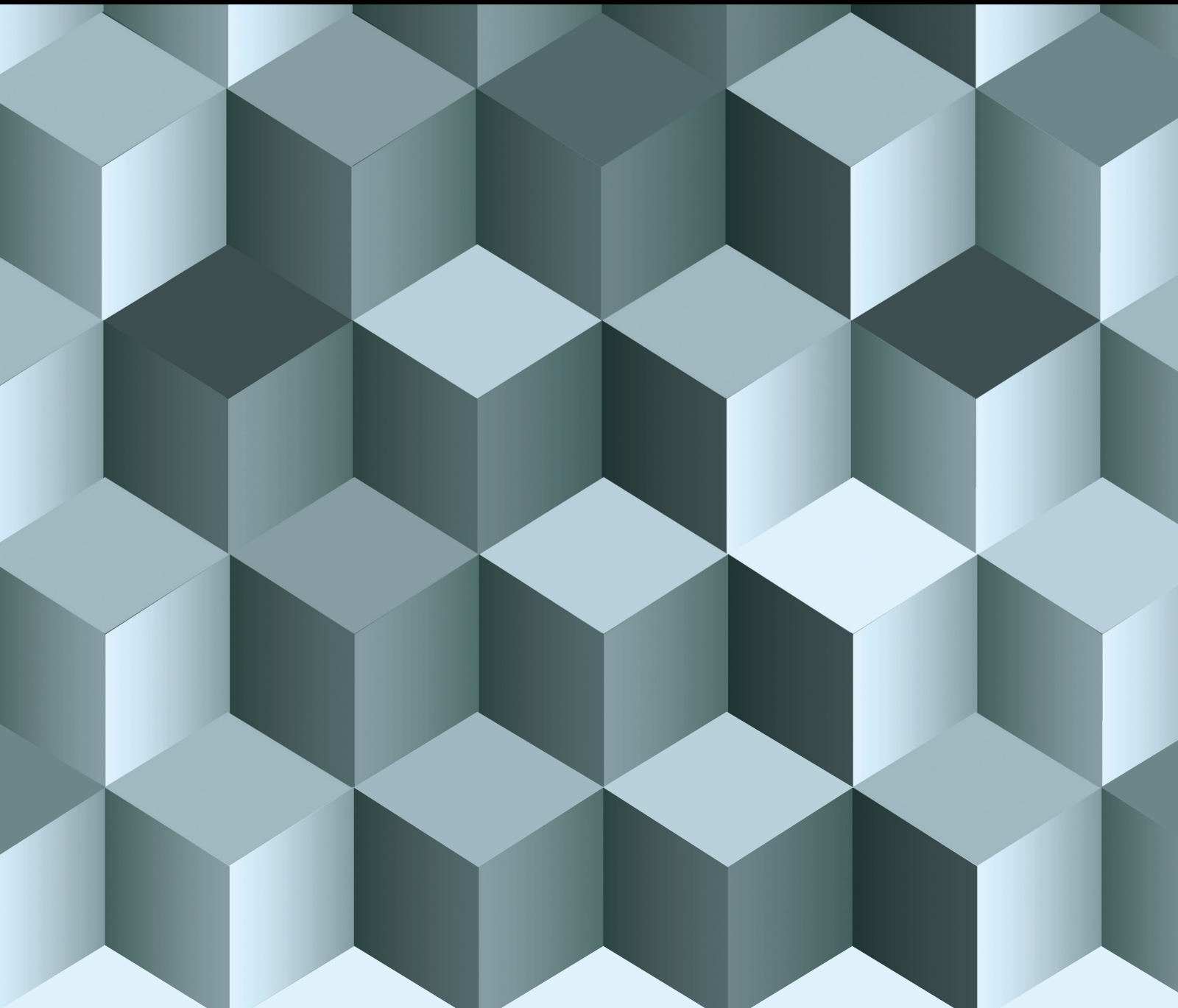


Mathematical Modeling for Next-Generation Big Data Technologies

Lead Guest Editor: Miaochoao Chen

Guest Editors: Zhiguo Zhou, Shengqi Lu, and Qilin Liu





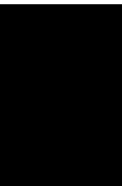
Mathematical Modeling for Next-Generation Big Data Technologies

Journal of Function Spaces

Mathematical Modeling for Next- Generation Big Data Technologies

Lead Guest Editor: Miaochoao Chen

Guest Editors: Zhiguo Zhou, Shengqi Lu, and Qilin
Liu






Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in "Journal of Function Spaces." 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

Maria Alessandra Ragusa, Italy

Associate Editors

Ismat Beg , Pakistan
Alberto Fiorenza , Italy
Adrian Petrusel , Romania

Academic Editors

Mohammed S. Abdo , Yemen
John R. Akeroyd , USA
Shrideh Al-Omari , Jordan
Richard I. Avery , USA
Bilal Bilalov, Azerbaijan
Salah Boulaaras, Saudi Arabia
Raúl E. Curto , USA
Giovanni Di Fratta, Austria
Konstantin M. Dyakonov , Spain
Hans G. Feichtinger , Austria
Baowei Feng , China
Aurelian Gheondea , Turkey
Xian-Ming Gu, China
Emanuel Guariglia, Italy
Yusuf Gurefe, Turkey
Yongsheng S. Han, USA
Seppo Hassi, Finland
Kwok-Pun Ho , Hong Kong
Gennaro Infante , Italy
Abdul Rauf Khan , Pakistan
Nikhil Khanna , Oman
Sebastian Krol, Poland
Yuri Latushkin , USA
Young Joo Lee , Republic of Korea
Guozhen Lu , USA
Giuseppe Marino , Italy
Mark A. McKibben , USA
Alexander Meskhi , Georgia
Feliz Minhós , Portugal
Alfonso Montes-Rodriguez , Spain
Gisele Mophou , France
Dumitru Motreanu , France
Sivaram K. Narayan, USA
Samuel Nicolay , Belgium
Kasso Okoudjou , USA
Gestur Ólafsson , USA
Gelu Popescu, USA
Humberto Rafeiro, United Arab Emirates

Paola Rubbioni , Italy
Natasha Samko , Portugal
Yoshihiro Sawano , Japan
Simone Secchi , Italy
Mitsuru Sugimoto , Japan
Wenchang Sun, China
Tomonari Suzuki , Japan
Wilfredo Urbina , USA
Calogero Vetro , Italy
Pasquale Vetro , Italy
Shanhe Wu , China
Kehe Zhu , USA

Contents

Retracted: Design and Communication of City Brand Image Based on Big Data and Personalized Recommendation System

Journal of Function Spaces

Retraction (1 page), Article ID 9898606, Volume 2023 (2023)

Retracted: Macro-Micro Study on Mechanical Properties of Frozen Fine Sandstone Based on DEM Mathematical Model

Journal of Function Spaces

Retraction (1 page), Article ID 9896372, Volume 2023 (2023)

Retracted: Prediction of Concrete Compressive Strength Based on the BP Neural Network Optimized by Random Forest and ISSA

Journal of Function Spaces

Retraction (1 page), Article ID 9892643, Volume 2023 (2023)

Retracted: Optimization of Ecological Mode of College Students' Innovation and Entrepreneurship Education Based on a Big Data Analysis Model

Journal of Function Spaces

Retraction (1 page), Article ID 9879764, Volume 2023 (2023)

Retracted: UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

Journal of Function Spaces

Retraction (1 page), Article ID 9863415, Volume 2023 (2023)

Retracted: Abnormal-Aware Multiperson Evaluation System with Improved Fuzzy Weighting

Journal of Function Spaces

Retraction (1 page), Article ID 9856068, Volume 2023 (2023)

Retracted: InSAR Phase Unwrapping Algorithm Based on Deep GAN

Journal of Function Spaces

Retraction (1 page), Article ID 9851623, Volume 2023 (2023)

Retracted: Data Analysis of Ceramic Material Color Matching Collection Points in Furniture Design Based on the Image Difference Prediction Model

Journal of Function Spaces

Retraction (1 page), Article ID 9848736, Volume 2023 (2023)

Retracted: Optimization of UAV Cooperative Path Planning Mathematical Model Based on Personalized Multigroup Sparrow Search Algorithm in Complex Environment

Journal of Function Spaces

Retraction (1 page), Article ID 9841681, Volume 2023 (2023)

Retracted: Simulation of Subject Coordination for a Smart Campus Based on Complex Network and Evolutionary Game Theory

Journal of Function Spaces

Retraction (1 page), Article ID 9814635, Volume 2023 (2023)

Retracted: Application of Optimized BP Neural Network In Financial Alert System

Journal of Function Spaces

Retraction (1 page), Article ID 9804696, Volume 2023 (2023)

Retracted: The Application of the Depth Model of Precise Matching between People and Posts Based on Ability Perception in Human Resource Management

Journal of Function Spaces

Retraction (1 page), Article ID 9801306, Volume 2023 (2023)

Retracted: Decompositions of Circulant-Balanced Complete Multipartite Graphs Based on a Novel Labelling Approach

Journal of Function Spaces

Retraction (1 page), Article ID 9789483, Volume 2023 (2023)

Retracted: Analysis of Option Butterfly Portfolio Models Based on Nonparametric Estimation Deep Learning Method

Journal of Function Spaces

Retraction (1 page), Article ID 9782789, Volume 2023 (2023)

Retracted: Application of Swarm Intelligence Optimization Algorithm in Logistics Delivery Path Optimization under the Background of Big Data

Journal of Function Spaces

Retraction (1 page), Article ID 9768576, Volume 2023 (2023)

Retracted: Analysis of the Impacts of Health Cost and Risk Preference on Farmers' Protective Behavior of Pesticide Application Based on the Autoregressive Threshold Model: A Case Study of Wuhu City in China

Journal of Function Spaces

Retraction (1 page), Article ID 9767353, Volume 2023 (2023)

Retracted: Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology

Journal of Function Spaces

Retraction (1 page), Article ID 9756926, Volume 2023 (2023)

Retracted: Exploration and Analysis of Collaborative Matching Algorithm Empowering Large Data Fiscal Budget

Journal of Function Spaces

Retraction (1 page), Article ID 9756857, Volume 2023 (2023)

Retracted: The Effect of Temperature on Intestinal Flora Imbalance Based on Time Series Prediction Model

Journal of Function Spaces

Retraction (1 page), Article ID 9753696, Volume 2023 (2023)

Contents

Retracted: Optimization of Digital Recommendation Service System for Tourist Attractions Based on Personalized Recommendation Algorithm

Journal of Function Spaces

Retraction (1 page), Article ID 9898754, Volume 2023 (2023)

Retracted: Construction of Mathematical Model of Logistics Delivering Based on Intelligent Mobilization

Journal of Function Spaces

Retraction (1 page), Article ID 9892364, Volume 2023 (2023)

Retracted: Construction of Holographic Immersion Chamber Based on Multisource Information Fusion and Interactive Virtual Reality Technology

Journal of Function Spaces

Retraction (1 page), Article ID 9863607, Volume 2023 (2023)

Retracted: Optimization Calculation Method and Mathematical Modeling of Big Data Chaotic Model Based on Improved Genetic Algorithm

Journal of Function Spaces

Retraction (1 page), Article ID 9837482, Volume 2023 (2023)

Retracted: Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles

Journal of Function Spaces

Retraction (1 page), Article ID 9830952, Volume 2023 (2023)

Retracted: Application of Internet of Things Based on Big Data Ecosystem in Factory Energy Consumption Analysis Model

Journal of Function Spaces

Retraction (1 page), Article ID 9805725, Volume 2023 (2023)

Retracted: Innovation Capacity, Industrial Structure, and Regional Development of High Quality: Empirical Research Based on the County Level

Journal of Function Spaces

Retraction (1 page), Article ID 9829467, Volume 2023 (2023)

Retracted: Application of Visual Guidance Design of Tourist City Combined with Deep Learning Computer Visual Symbols

Journal of Function Spaces

Retraction (1 page), Article ID 9872716, Volume 2023 (2023)

Retracted: Numerical Simulation of Particle Deposition in Arterial Bifurcation via Lattice Boltzmann Method

Journal of Function Spaces

Retraction (1 page), Article ID 9870187, Volume 2023 (2023)

Retracted: Analysis of Farmer Relocation Selection Behavior Based on Bayesian Network

Journal of Function Spaces

Retraction (1 page), Article ID 9863951, Volume 2023 (2023)

Retracted: Intelligent City 3D Modeling Model Based on Multisource Data Point Cloud Algorithm

Journal of Function Spaces

Retraction (1 page), Article ID 9850460, Volume 2023 (2023)

Retracted: A New Machine Learning Algorithm for Regional Low-Carbon Economic Development Analysis Based on Data Mining

Journal of Function Spaces

Retraction (1 page), Article ID 9848797, Volume 2023 (2023)

Retracted: Exercise Recommendation Model Based on Cognitive Level and Educational Big Data Mining

Journal of Function Spaces

Retraction (1 page), Article ID 9843983, Volume 2023 (2023)

Retracted: Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data

Journal of Function Spaces

Retraction (1 page), Article ID 9838451, Volume 2023 (2023)

Retracted: Prediction and Risk Assessment of Extreme Weather Events Based on Gumbel Copula Function

Journal of Function Spaces

Retraction (1 page), Article ID 9834132, Volume 2023 (2023)

Retracted: Application of Optimized Support Vector Machine Model in Tax Forecasting System

Journal of Function Spaces

Retraction (1 page), Article ID 9825379, Volume 2023 (2023)

Retracted: The Implementation of Unified Application Data Collection and Analysis in Tobacco Enterprise Platform Design

Journal of Function Spaces

Retraction (1 page), Article ID 9812574, Volume 2023 (2023)

Retracted: E-Commerce across Boarder Logistics Risk Evaluation Model Based on Improved Neural Network

Journal of Function Spaces

Retraction (1 page), Article ID 9797020, Volume 2023 (2023)

Retracted: Assisting Laboratory Management Based on Network Big Data Mining Technology

Journal of Function Spaces

Retraction (1 page), Article ID 9764513, Volume 2023 (2023)

Contents

Retracted: Exploration of Energy-Saving Chilling Landscape Design Based on Algo for Group Intelligence

Journal of Function Spaces

Retraction (1 page), Article ID 9757983, Volume 2023 (2023)

Retracted: Mathematical Models for Analysis the Impact of Institutional Quality and Multilateral Financial Institutions on the Success of Green PPP Projects

Journal of Function Spaces

Retraction (1 page), Article ID 9857587, Volume 2023 (2023)

Retracted: Optimization of College Students' Employment Prosperity Index System Based on Multiple Logit Models

Journal of Function Spaces





Retraction (1 page), Article ID 9846340, Volume 2023 (2023)

Retracted: Application of Wireless Sensor Network Model Based on Big Data Ecosystem in Intelligent Health Monitoring System

Journal of Function Spaces

Retraction (1 page), Article ID 9810980, Volume 2023 (2023)

[Retracted] Analysis of Option Butterfly Portfolio Models Based on Nonparametric Estimation Deep Learning Method

Xiangyu Ge , Xia Zhu , Gang Bi, Hao Zheng , and Qing Li 

Research Article (17 pages), Article ID 4989036, Volume 2023 (2023)

Retracted: Prediction of Urban Scale Expansion Based on Genetic Algorithm Optimized Neural Network Model

Journal of Function Spaces

Retraction (1 page), Article ID 9897063, Volume 2023 (2023)

Retracted: Designing Landscape of Urban Gardening Based on Optimized Artificial Intelligence Model

Journal of Function Spaces

Retraction (1 page), Article ID 9894653, Volume 2023 (2023)

Retracted: The Experimental Validation of a Digital Filter Mathematical Modelling for Nuclear Fuel Vibration Monitoring System

Journal of Function Spaces

Retraction (1 page), Article ID 9876837, Volume 2023 (2023)

Retracted: Construction of a Knowledge Map Based on Text CNN Algorithm for Maritime English Subjects

Journal of Function Spaces

Retraction (1 page), Article ID 9870920, Volume 2023 (2023)

Retracted: Modeling and Analysis of Multifocus Picture Division Algorithm Based on Deep Learning

Journal of Function Spaces

Retraction (1 page), Article ID 9869806, Volume 2023 (2023)

Retracted: Application of GA in Furniture Modeling Style Design

Journal of Function Spaces

Retraction (1 page), Article ID 9860741, Volume 2023 (2023)

Retracted: Mathematical Modeling and Simulation of Online Teaching Effect Evaluation Based on Decision Tree Algorithm

Journal of Function Spaces

Retraction (1 page), Article ID 9858792, Volume 2023 (2023)

Retracted: Optimization Algorithm of Communication Resource Allocation in a Complex Network Based on an Improved Neural Network

Journal of Function Spaces

Retraction (1 page), Article ID 9848901, Volume 2023 (2023)

Retracted: Optimization of Three-Dimensional Model of Landscape Space Based on Big Data Analysis

Journal of Function Spaces

Retraction (1 page), Article ID 9847536, Volume 2023 (2023)

Retracted: Participation in Targeted Poverty Alleviation and Enterprise Innovation Investment: Analysis of the Mediating Effect Test Model Based on Financing Constraints

Journal of Function Spaces

Retraction (1 page), Article ID 9846768, Volume 2023 (2023)

Retracted: Enterprise Management Decision and Financial Management Based on Dynamic Cost Volume Profit Model

Journal of Function Spaces

Retraction (1 page), Article ID 9835467, Volume 2023 (2023)

Retracted: Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model

Journal of Function Spaces

Retraction (1 page), Article ID 9834704, Volume 2023 (2023)

Retracted: Modeling of Energy Saving and Comfort of Building Layout in Extreme Weather Urban Residential Area under the Background of Spatial Structure and Form Evolution: Taking Yichun as an Example

Journal of Function Spaces

Retraction (1 page), Article ID 9825612, Volume 2023 (2023)

Retracted: Application of Neural Networks in Financial Time Series Forecasting Models

Journal of Function Spaces

Retraction (1 page), Article ID 9819013, Volume 2023 (2023)

Contents

Retracted: Prediction Model of Music Popular Trend Based on NNS and DM Technology

Journal of Function Spaces

Retraction (1 page), Article ID 9807252, Volume 2023 (2023)

Retracted: The Construction of Adaptive Learning for Sports Based on Aerobics Trajectory Recognition Model

Journal of Function Spaces

Retraction (1 page), Article ID 9804063, Volume 2023 (2023)

Retracted: Exploration of Flood Prediction in Watersheds Based on the Fusion Analysis of Remote Sensing Big Data with Multiple Physical Fields

Journal of Function Spaces

Retraction (1 page), Article ID 9794849, Volume 2023 (2023)

Retracted: Certificateless Batch Authentication Scheme and Intrusion Detection Model Based on the Mobile Edge Computing Technology NDN-IoT Environment

Journal of Function Spaces

Retraction (1 page), Article ID 9767396, Volume 2023 (2023)

Retracted: Analysis of Consumer Behavior Data Based on Deep Neural Network Model

Journal of Function Spaces

Retraction (1 page), Article ID 9767326, Volume 2023 (2023)

Retracted: Adjustment Method of "FAST" Active Reflector Based on Optimal Fitting Strategy

Journal of Function Spaces

Retraction (1 page), Article ID 9764289, Volume 2023 (2023)

Retracted: Optimization of the Intelligent Sports Health Management System Based on Big Data Technology and Internet of Things

Journal of Function Spaces

Retraction (1 page), Article ID 9763803, Volume 2023 (2023)

Retracted: Innovation and Research of Digital Artwork Design Based on Big Data and Integrated Media

Journal of Function Spaces

Retraction (1 page), Article ID 9757698, Volume 2023 (2023)

Retracted: Mathematical Model Construction of Teaching Evaluation in Colleges and Universities Based on Convolutional Neural Network under the Background of Big Data

Journal of Function Spaces

Retraction (1 page), Article ID 9757643, Volume 2023 (2023)

Retracted: The Evaluation Model of College English Diagnostic Exercises Based on Machine Learning

Journal of Function Spaces

Retraction (1 page), Article ID 9756390, Volume 2023 (2023)

[Retracted] UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

Gui Fu , Yang Gao , Liwen Liu , Mingye Yang , and Xinyu Zhu 

Research Article (11 pages), Article ID 9708143, Volume 2023 (2023)

[Retracted] Data Analysis of Ceramic Material Color Matching Collection Points in Furniture Design Based on the Image Difference Prediction Model

Cheng Zhou  and Yalan Li 


Research Article (10 pages), Article ID 4684904, Volume 2023 (2023)

[Retracted] InSAR Phase Unwrapping Algorithm Based on Deep GAN

Chenxia Wang , Pingli Sun , Zheng Li , and Linlin Tang 



Research Article (11 pages), Article ID 9230780, Volume 2023 (2023)

[Retracted] Application of Swarm Intelligence Optimization Algorithm in Logistics Delivery Path Optimization under the Background of Big Data

Guofu Zhao 



Research Article (12 pages), Article ID 3476711, Volume 2023 (2023)

[Retracted] Exploration and Analysis of Collaborative Matching Algorithm Empowering Large Data Fiscal Budget

Rui Gong  and Xu Zhang 



Research Article (11 pages), Article ID 8171318, Volume 2022 (2022)

[Retracted] Application of GA in Furniture Modeling Style Design

Heng Yu  and Chunjing Liu 

Research Article (11 pages), Article ID 5264185, Volume 2022 (2022)

[Retracted] Analysis of the Impacts of Health Cost and Risk Preference on Farmers' Protective Behavior of Pesticide Application Based on the Autoregressive Threshold Model: A Case Study of Wuhu City in China

Shukai Cai , Yaxing Gu, Runze Li, and Zhonghong Teng 


Research Article (9 pages), Article ID 9047754, Volume 2022 (2022)

[Retracted] The Application of the Depth Model of Precise Matching between People and Posts Based on Ability Perception in Human Resource Management

Hui Cao 

Research Article (9 pages), Article ID 5495437, Volume 2022 (2022)

[Retracted] Innovation and Research of Digital Artwork Design Based on Big Data and Integrated Media

Wenbin Zhai 

Research Article (10 pages), Article ID 6333688, Volume 2022 (2022)



[Retracted] Analysis of Farmer Relocation Selection Behavior Based on Bayesian Network

Song Li  and Yanbo Guo 

Research Article (8 pages), Article ID 9155510, Volume 2022 (2022)


Contents

[Retracted] Design and Communication of City Brand Image Based on Big Data and Personalized Recommendation System

Yan Lin  and Guofeng Hu 


Research Article (10 pages), Article ID 9378800, Volume 2022 (2022)

[Retracted] Assisting Laboratory Management Based on Network Big Data Mining Technology

Yonghua Xu 

Research Article (7 pages), Article ID 2087287, Volume 2022 (2022)

[Retracted] Analysis of Consumer Behavior Data Based on Deep Neural Network Model

Shijiao Yuan 

Research Article (10 pages), Article ID 4938278, Volume 2022 (2022)

[Retracted] Optimization of Digital Recommendation Service System for Tourist Attractions Based on Personalized Recommendation Algorithm

Yue Wang, Zhaoxiang Qin , Jun Tang , and Wei Zhang

Research Article (9 pages), Article ID 1191419, Volume 2022 (2022)

[Retracted] Construction of Mathematical Model of Logistics Delivering Based on Intelligent Mobilization

Haijun Liang  and Jing Guo 

Research Article (11 pages), Article ID 7386227, Volume 2022 (2022)

[Retracted] Application of Visual Guidance Design of Tourist City Combined with Deep Learning Computer Visual Symbols

Jiaying Cai  and Kai Liu 





Research Article (10 pages), Article ID 8612543, Volume 2022 (2022)

[Retracted] Innovation Capacity, Industrial Structure, and Regional Development of High Quality: Empirical Research Based on the County Level

Chuhao Zhao  and Hui Yao 


Research Article (13 pages), Article ID 4446292, Volume 2022 (2022)

[Retracted] Modeling of Energy Saving and Comfort of Building Layout in Extreme Weather Urban Residential Area under the Background of Spatial Structure and Form Evolution: Taking Yichun as an Example

Ming Lu , Guofei Sun , Endong Wang , and Zichong He 

Research Article (11 pages), Article ID 6892035, Volume 2022 (2022)

[Retracted] Application of Neural Networks in Financial Time Series Forecasting Models

Xinhui Li 

Research Article (9 pages), Article ID 7817264, Volume 2022 (2022)

[Retracted] Prediction Model of Music Popular Trend Based on NNS and DM Technology


Yichen Xu , Mingxun Wang , Hao Chen , and Fan Hu 

Research Article (9 pages), Article ID 6104056, Volume 2022 (2022)


[Retracted] Prediction of Concrete Compressive Strength Based on the BP Neural Network Optimized by Random Forest and ISSA

Gang Chen, Donglin Zhu , Xiao Wang , Changjun Zhou , and Xiangyu Chen
Research Article (20 pages), Article ID 8799429, Volume 2022 (2022)


[Retracted] A New Machine Learning Algorithm for Regional Low-Carbon Economic Development Analysis Based on Data Mining

Xinlei Liu 
Research Article (8 pages), Article ID 5692666, Volume 2022 (2022)




[Retracted] Designing Landscape of Urban Gardening Based on Optimized Artificial Intelligence Model

Long Liu 
Research Article (12 pages), Article ID 7963173, Volume 2022 (2022)


[Retracted] Optimization of Ecological Mode of College Students' Innovation and Entrepreneurship Education Based on a Big Data Analysis Model

Miao Yu , Yanju Lu , and Hongliang Wang 
Research Article (8 pages), Article ID 4357354, Volume 2022 (2022)

[Retracted] The Effect of Temperature on Intestinal Flora Imbalance Based on Time Series Prediction Model

Jing Zhu , Yafei Xue , and Shuyuan Yang 
Research Article (11 pages), Article ID 6122406, Volume 2022 (2022)

[Retracted] Application of Optimized Support Vector Machine Model in Tax Forecasting System

Yu Xin 
Research Article (10 pages), Article ID 6212579, Volume 2022 (2022)



[Retracted] The Implementation of Unified Application Data Collection and Analysis in Tobacco Enterprise Platform Design

Yong Jin , and Weidong Lou 
Research Article (9 pages), Article ID 3933465, Volume 2022 (2022)

[Retracted] Mathematical Model Construction of Teaching Evaluation in Colleges and Universities Based on Convolutional Neural Network under the Background of Big Data

Changming Tan 
Research Article (8 pages), Article ID 7064287, Volume 2022 (2022)

[Retracted] The Evaluation Model of College English Diagnostic Exercises Based on Machine Learning


Qi Wang , and Hao Wu 
Research Article (12 pages), Article ID 9185827, Volume 2022 (2022)

[Retracted] Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model

Shuli Shen 
Research Article (9 pages), Article ID 1078057, Volume 2022 (2022)



Contents

[Retracted] Optimization of Three-Dimensional Model of Landscape Space Based on Big Data Analysis

Cuixia Wang 


Research Article (11 pages), Article ID 7002983, Volume 2022 (2022)

[Retracted] Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles

Jijin Xu  and Zhen Liang 

Research Article (10 pages), Article ID 9295395, Volume 2022 (2022)

[Retracted] The Construction of Adaptive Learning for Sports Based on Aerobics Trajectory Recognition Model

Chaojie Xi 

Research Article (8 pages), Article ID 8339745, Volume 2022 (2022)

[Retracted] Adjustment Method of "FAST" Active Reflector Based on Optimal Fitting Strategy

Yu-Chen Xia 


Research Article (15 pages), Article ID 6994858, Volume 2022 (2022)

[Retracted] Participation in Targeted Poverty Alleviation and Enterprise Innovation Investment: Analysis of the Mediating Effect Test Model Based on Financing Constraints

Ting Gao  and Huaiming Wang


Research Article (16 pages), Article ID 7060462, Volume 2022 (2022)

[Retracted] Mathematical Modeling and Simulation of Online Teaching Effect Evaluation Based on Decision Tree Algorithm

Pingli Sun 


Research Article (11 pages), Article ID 7425196, Volume 2022 (2022)

[Retracted] Application of Optimized BP Neural Network In Financial Alert System

Qiaoyi Gao 



Research Article (10 pages), Article ID 1816315, Volume 2022 (2022)

[Retracted] Exploration of Energy-Saving Chilling Landscape Design Based on Algo for Group Intelligence

Zhuo Li 



Research Article (10 pages), Article ID 1851623, Volume 2022 (2022)

[Retracted] Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology

Yutang Liu  and Qin Zhang 


Research Article (10 pages), Article ID 4292063, Volume 2022 (2022)

[Retracted] Construction of a Knowledge Map Based on Text CNN Algorithm for Maritime English Subjects

Hui Wang  and Aimin Wei 

Research Article (9 pages), Article ID 6578682, Volume 2022 (2022)

[Retracted] Application of Wireless Sensor Network Model Based on Big Data Ecosystem in Intelligent Health Monitoring System

Jiaqing Tu 


Research Article (10 pages), Article ID 3179915, Volume 2022 (2022)

[Retracted] Optimization Algorithm of Communication Resource Allocation in a Complex Network Based on an Improved Neural Network

Haomiao Zhang  and Qing Liu


Research Article (8 pages), Article ID 7309612, Volume 2022 (2022)

[Retracted] Exercise Recommendation Model Based on Cognitive Level and Educational Big Data Mining

Yongming Pu  and Hongming Chen


Research Article (11 pages), Article ID 3845419, Volume 2022 (2022)

[Retracted] E-Commerce across Boarder Logistics Risk Evaluation Model Based on Improved Neural Network

Wei Qiao 



Research Article (10 pages), Article ID 2355298, Volume 2022 (2022)

[Retracted] Certificateless Batch Authentication Scheme and Intrusion Detection Model Based on the Mobile Edge Computing Technology NDN-IoT Environment

Jianzhao Sun 

Research Article (9 pages), Article ID 5926792, Volume 2022 (2022)

[Retracted] Intelligent City 3D Modeling Model Based on Multisource Data Point Cloud Algorithm

Youping Wu  and Zhihui Zhou 

Research Article (10 pages), Article ID 6135829, Volume 2022 (2022)

[Retracted] Prediction and Risk Assessment of Extreme Weather Events Based on Gumbel Copula Function

Peng-Hui Yang , Yao Yu , Feng Gu , Meng-Jie Qu , and Jia-Ming Zhu 

Research Article (13 pages), Article ID 1438373, Volume 2022 (2022)

[Retracted] Optimization Calculation Method and Mathematical Modeling of Big Data Chaotic Model Based on Improved Genetic Algorithm

Zhicheng Zhang  and Yan Zhang 

Research Article (9 pages), Article ID 6983242, Volume 2022 (2022)



Contents

[Retracted] Construction of Holographic Immersion Chamber Based on Multisource Information Fusion and Interactive Virtual Reality Technology

Chunying Wang  and Yin Liu 



Research Article (10 pages), Article ID 6406211, Volume 2022 (2022)

[Retracted] Decompositions of Circulant-Balanced Complete Multipartite Graphs Based on a Novel Labelling Approach

A. El-Mesady  and Omar Bazighifan 


Research Article (17 pages), Article ID 2017936, Volume 2022 (2022)

[Retracted] Exploration of Flood Prediction in Watersheds Based on the Fusion Analysis of Remote Sensing Big Data with Multiple Physical Fields

Minming Xu  and Yun Ouyang 


Research Article (10 pages), Article ID 9422553, Volume 2022 (2022)

[Retracted] Modeling and Analysis of Multifocus Picture Division Algorithm Based on Deep Learning

Hongxia You 




Research Article (10 pages), Article ID 8326626, Volume 2022 (2022)

[Retracted] Prediction of Urban Scale Expansion Based on Genetic Algorithm Optimized Neural Network Model

Hewu Kuang 


Research Article (11 pages), Article ID 5407319, Volume 2022 (2022)

[Retracted] Application of Internet of Things Based on Big Data Ecosystem in Factory Energy Consumption Analysis Model

Ang Li , Chen Zhang , and Lei Li 



Research Article (10 pages), Article ID 7631323, Volume 2022 (2022)

[Retracted] Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data

Ming Zhang 


Research Article (11 pages), Article ID 3367256, Volume 2022 (2022)

[Retracted] Optimization of College Students' Employment Prosperity Index System Based on Multiple Logit Models

Wei Xiang  and Weizhen Hu 


Research Article (8 pages), Article ID 9367675, Volume 2022 (2022)

[Retracted] Optimization of the Intelligent Sports Health Management System Based on Big Data Technology and Internet of Things

Jian Dang 


Research Article (19 pages), Article ID 5053150, Volume 2022 (2022)

[Retracted] Abnormal-Aware Multiperson Evaluation System with Improved Fuzzy Weighting

Shutong Ni 



Research Article (8 pages), Article ID 4899831, Volume 2022 (2022)

[Retracted] Numerical Simulation of Particle Deposition in Arterial Bifurcation via Lattice Boltzmann Method

Houhui Yi 

Research Article (6 pages), Article ID 3873484, Volume 2022 (2022)

[Retracted] Mathematical Models for Analysis the Impact of Institutional Quality and Multilateral Financial Institutions on the Success of Green PPP Projects

Wu-Yan Li , Ruo-Yu Chen , and Ming Ma

Research Article (12 pages), Article ID 3825211, Volume 2022 (2022)

[Retracted] Optimization of UAV Cooperative Path Planning Mathematical Model Based on Personalized Multigroup Sparrow Search Algorithm in Complex Environment

Shaoqiang Yan , Weidong Liu, Ping Yang , Fengxuan Wu , Donglin Zhu, and Gang Chen

Research Article (25 pages), Article ID 2521737, Volume 2022 (2022)

[Retracted] Simulation of Subject Coordination for a Smart Campus Based on Complex Network and Evolutionary Game Theory

Shaoyong Guo  and Ting Pang 

Research Article (9 pages), Article ID 4141475, Volume 2022 (2022)

[Retracted] Enterprise Management Decision and Financial Management Based on Dynamic Cost Volume Profit Model

Yanlin Guo 


Research Article (8 pages), Article ID 9016060, Volume 2022 (2022)

[Retracted] The Experimental Validation of a Digital Filter Mathematical Modelling for Nuclear Fuel Vibration Monitoring System

Han Yong , Huang Shiyong , Han MengXuan , Zhao Tao , Wen Guang , Gu Mingfei , and Zhang Qi 

Research Article (11 pages), Article ID 9741419, Volume 2022 (2022)

[Retracted] Macro-Micro Study on Mechanical Properties of Frozen Fine Sandstone Based on DEM Mathematical Model

Maoyan Ma , Min You, Shuguang Peng, Biao Zhang, and Yuan Lin

Research Article (10 pages), Article ID 7176665, Volume 2022 (2022)

Retraction

Retracted: Design and Communication of City Brand Image Based on Big Data and Personalized Recommendation System

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Lin and G. Hu, "Design and Communication of City Brand Image Based on Big Data and Personalized Recommendation System," *Journal of Function Spaces*, vol. 2022, Article ID 9378800, 10 pages, 2022.

Retraction

Retracted: Macro-Micro Study on Mechanical Properties of Frozen Fine Sandstone Based on DEM Mathematical Model

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Ma, M. You, S. Peng, B. Zhang, and Y. Lin, "Macro-Micro Study on Mechanical Properties of Frozen Fine Sandstone Based on DEM Mathematical Model," *Journal of Function Spaces*, vol. 2022, Article ID 7176665, 10 pages, 2022.

Retraction

Retracted: Prediction of Concrete Compressive Strength Based on the BP Neural Network Optimized by Random Forest and ISSA

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Chen, D. Zhu, X. Wang, C. Zhou, and X. Chen, "Prediction of Concrete Compressive Strength Based on the BP Neural Network Optimized by Random Forest and ISSA," *Journal of Function Spaces*, vol. 2022, Article ID 8799429, 20 pages, 2022.

Retraction

Retracted: Optimization of Ecological Mode of College Students' Innovation and Entrepreneurship Education Based on a Big Data Analysis Model

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Yu, Y. Lu, and H. Wang, "Optimization of Ecological Mode of College Students' Innovation and Entrepreneurship Education Based on a Big Data Analysis Model," *Journal of Function Spaces*, vol. 2022, Article ID 4357354, 8 pages, 2022.

Retraction

Retracted: UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Fu, Y. Gao, L. Liu, M. Yang, and X. Zhu, "UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment," *Journal of Function Spaces*, vol. 2023, Article ID 9708143, 11 pages, 2023.

Retraction

Retracted: Abnormal-Aware Multiperson Evaluation System with Improved Fuzzy Weighting

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Ni, "Abnormal-Aware Multiperson Evaluation System with Improved Fuzzy Weighting," *Journal of Function Spaces*, vol. 2022, Article ID 4899831, 8 pages, 2022.

Retraction

Retracted: InSAR Phase Unwrapping Algorithm Based on Deep GAN

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Wang, P. Sun, Z. Li, and L. Tang, "InSAR Phase Unwrapping Algorithm Based on Deep GAN," *Journal of Function Spaces*, vol. 2023, Article ID 9230780, 11 pages, 2023.

Retraction

Retracted: Data Analysis of Ceramic Material Color Matching Collection Points in Furniture Design Based on the Image Difference Prediction Model

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Zhou and Y. Li, "Data Analysis of Ceramic Material Color Matching Collection Points in Furniture Design Based on the Image Difference Prediction Model," *Journal of Function Spaces*, vol. 2023, Article ID 4684904, 10 pages, 2023.

Retraction

Retracted: Optimization of UAV Cooperative Path Planning Mathematical Model Based on Personalized Multigroup Sparrow Search Algorithm in Complex Environment

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Yan, W. Liu, P. Yang, F. Wu, D. Zhu, and G. Chen, "Optimization of UAV Cooperative Path Planning Mathematical Model Based on Personalized Multigroup Sparrow Search Algorithm in Complex Environment," *Journal of Function Spaces*, vol. 2022, Article ID 2521737, 25 pages, 2022.

Retraction

Retracted: Simulation of Subject Coordination for a Smart Campus Based on Complex Network and Evolutionary Game Theory

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Guo and T. Pang, "Simulation of Subject Coordination for a Smart Campus Based on Complex Network and Evolutionary Game Theory," *Journal of Function Spaces*, vol. 2022, Article ID 4141475, 9 pages, 2022.

Retraction

Retracted: Application of Optimized BP Neural Network In Financial Alert System

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Gao, "Application of Optimized BP Neural Network In Financial Alert System," *Journal of Function Spaces*, vol. 2022, Article ID 1816315, 10 pages, 2022.

Retraction

Retracted: The Application of the Depth Model of Precise Matching between People and Posts Based on Ability Perception in Human Resource Management

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Cao, "The Application of the Depth Model of Precise Matching between People and Posts Based on Ability Perception in Human Resource Management," *Journal of Function Spaces*, vol. 2022, Article ID 5495437, 9 pages, 2022.

Retraction

Retracted: Decompositions of Circulant-Balanced Complete Multipartite Graphs Based on a Novel Labelling Approach

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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] A. El-Mesady and O. Bazighifan, “Decompositions of Circulant-Balanced Complete Multipartite Graphs Based on a Novel Labelling Approach,” *Journal of Function Spaces*, vol. 2022, Article ID 2017936, 17 pages, 2022.

Retraction

Retracted: Analysis of Option Butterfly Portfolio Models Based on Nonparametric Estimation Deep Learning Method

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Ge, X. Zhu, G. Bi, H. Zheng, and Q. Li, “Analysis of Option Butterfly Portfolio Models Based on Nonparametric Estimation Deep Learning Method,” *Journal of Function Spaces*, vol. 2023, Article ID 4989036, 17 pages, 2023.

Retraction

Retracted: Application of Swarm Intelligence Optimization Algorithm in Logistics Delivery Path Optimization under the Background of Big Data

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Zhao, "Application of Swarm Intelligence Optimization Algorithm in Logistics Delivery Path Optimization under the Background of Big Data," *Journal of Function Spaces*, vol. 2023, Article ID 3476711, 12 pages, 2023.

Retraction

Retracted: Analysis of the Impacts of Health Cost and Risk Preference on Farmers' Protective Behavior of Pesticide Application Based on the Autoregressive Threshold Model: A Case Study of Wuhu City in China

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Cai, Y. Gu, R. Li, and Z. Teng, "Analysis of the Impacts of Health Cost and Risk Preference on Farmers' Protective Behavior of Pesticide Application Based on the Autoregressive Threshold Model: A Case Study of Wuhu City in China," *Journal of Function Spaces*, vol. 2022, Article ID 9047754, 9 pages, 2022.

Retraction

Retracted: Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Liu and Q. Zhang, "Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology," *Journal of Function Spaces*, vol. 2022, Article ID 4292063, 10 pages, 2022.

Retraction

Retracted: Exploration and Analysis of Collaborative Matching Algorithm Empowering Large Data Fiscal Budget

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Gong and X. Zhang, "Exploration and Analysis of Collaborative Matching Algorithm Empowering Large Data Fiscal Budget," *Journal of Function Spaces*, vol. 2022, Article ID 8171318, 11 pages, 2022.

Retraction

Retracted: The Effect of Temperature on Intestinal Flora Imbalance Based on Time Series Prediction Model

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Zhu, Y. Xue, and S. Yang, "The Effect of Temperature on Intestinal Flora Imbalance Based on Time Series Prediction Model," *Journal of Function Spaces*, vol. 2022, Article ID 6122406, 11 pages, 2022.

Retraction

Retracted: Optimization of Digital Recommendation Service System for Tourist Attractions Based on Personalized Recommendation Algorithm

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Wang, Z. Qin, J. Tang, and W. Zhang, "Optimization of Digital Recommendation Service System for Tourist Attractions Based on Personalized Recommendation Algorithm," *Journal of Function Spaces*, vol. 2022, Article ID 1191419, 9 pages, 2022.

Retraction

Retracted: Construction of Mathematical Model of Logistics Delivering Based on Intelligent Mobilization

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Liang and J. Guo, "Construction of Mathematical Model of Logistics Delivering Based on Intelligent Mobilization," *Journal of Function Spaces*, vol. 2022, Article ID 7386227, 11 pages, 2022.

Retraction

Retracted: Construction of Holographic Immersion Chamber Based on Multisource Information Fusion and Interactive Virtual Reality Technology

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Wang and Y. Liu, "Construction of Holographic Immersion Chamber Based on Multisource Information Fusion and Interactive Virtual Reality Technology," *Journal of Function Spaces*, vol. 2022, Article ID 6406211, 10 pages, 2022.

Retraction

Retracted: Optimization Calculation Method and Mathematical Modeling of Big Data Chaotic Model Based on Improved Genetic Algorithm

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Zhang and Y. Zhang, "Optimization Calculation Method and Mathematical Modeling of Big Data Chaotic Model Based on Improved Genetic Algorithm," *Journal of Function Spaces*, vol. 2022, Article ID 6983242, 9 pages, 2022.

Retraction

Retracted: Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xu and Z. Liang, "Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles," *Journal of Function Spaces*, vol. 2022, Article ID 9295395, 10 pages, 2022.

Retraction

Retracted: Application of Internet of Things Based on Big Data Ecosystem in Factory Energy Consumption Analysis Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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] A. Li, C. Zhang, and L. Li, "Application of Internet of Things Based on Big Data Ecosystem in Factory Energy Consumption Analysis Model," *Journal of Function Spaces*, vol. 2022, Article ID 7631323, 10 pages, 2022.

Retraction

Retracted: Innovation Capacity, Industrial Structure, and Regional Development of High Quality: Empirical Research Based on the County Level

Journal of Function Spaces

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 2023

Copyright © 2023 Journal of Function Spaces. 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. Zhao and H. Yao, "Innovation Capacity, Industrial Structure, and Regional Development of High Quality: Empirical Research Based on the County Level," *Journal of Function Spaces*, vol. 2022, Article ID 4446292, 13 pages, 2022.

Retraction

Retracted: Application of Visual Guidance Design of Tourist City Combined with Deep Learning Computer Visual Symbols

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Cai and K. Liu, "Application of Visual Guidance Design of Tourist City Combined with Deep Learning Computer Visual Symbols," *Journal of Function Spaces*, vol. 2022, Article ID 8612543, 10 pages, 2022.

Retraction

Retracted: Numerical Simulation of Particle Deposition in Arterial Bifurcation via Lattice Boltzmann Method

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Yi, "Numerical Simulation of Particle Deposition in Arterial Bifurcation via Lattice Boltzmann Method," *Journal of Function Spaces*, vol. 2022, Article ID 3873484, 6 pages, 2022.

Retraction

Retracted: Analysis of Farmer Relocation Selection Behavior Based on Bayesian Network

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Li and Y. Guo, "Analysis of Farmer Relocation Selection Behavior Based on Bayesian Network," *Journal of Function Spaces*, vol. 2022, Article ID 9155510, 8 pages, 2022.

Retraction

Retracted: Intelligent City 3D Modeling Model Based on Multisource Data Point Cloud Algorithm

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Wu and Z. Zhou, "Intelligent City 3D Modeling Model Based on Multisource Data Point Cloud Algorithm," *Journal of Function Spaces*, vol. 2022, Article ID 6135829, 10 pages, 2022.

Retraction

Retracted: A New Machine Learning Algorithm for Regional Low-Carbon Economic Development Analysis Based on Data Mining

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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, "A New Machine Learning Algorithm for Regional Low-Carbon Economic Development Analysis Based on Data Mining," *Journal of Function Spaces*, vol. 2022, Article ID 5692666, 8 pages, 2022.

Retraction

Retracted: Exercise Recommendation Model Based on Cognitive Level and Educational Big Data Mining

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Pu and H. Chen, "Exercise Recommendation Model Based on Cognitive Level and Educational Big Data Mining," *Journal of Function Spaces*, vol. 2022, Article ID 3845419, 11 pages, 2022.

Retraction

Retracted: Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Zhang, "Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data," *Journal of Function Spaces*, vol. 2022, Article ID 3367256, 11 pages, 2022.

Retraction

Retracted: Prediction and Risk Assessment of Extreme Weather Events Based on Gumbel Copula Function

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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.-H. Yang, Y. Yu, F. Gu, M.-J. Qu, and J.-M. Zhu, "Prediction and Risk Assessment of Extreme Weather Events Based on Gumbel Copula Function," *Journal of Function Spaces*, vol. 2022, Article ID 1438373, 13 pages, 2022.

Retraction

Retracted: Application of Optimized Support Vector Machine Model in Tax Forecasting System

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xin, "Application of Optimized Support Vector Machine Model in Tax Forecasting System," *Journal of Function Spaces*, vol. 2022, Article ID 6212579, 10 pages, 2022.

Retraction

Retracted: The Implementation of Unified Application Data Collection and Analysis in Tobacco Enterprise Platform Design

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Jin and W. Lou, "The Implementation of Unified Application Data Collection and Analysis in Tobacco Enterprise Platform Design," *Journal of Function Spaces*, vol. 2022, Article ID 3933465, 9 pages, 2022.

Retraction

Retracted: E-Commerce across Boarder Logistics Risk Evaluation Model Based on Improved Neural Network

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Qiao, "E-Commerce across Boarder Logistics Risk Evaluation Model Based on Improved Neural Network," *Journal of Function Spaces*, vol. 2022, Article ID 2355298, 10 pages, 2022.

Retraction

Retracted: Assisting Laboratory Management Based on Network Big Data Mining Technology

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xu, "Assisting Laboratory Management Based on Network Big Data Mining Technology," *Journal of Function Spaces*, vol. 2022, Article ID 2087287, 7 pages, 2022.

Retraction

Retracted: Exploration of Energy-Saving Chilling Landscape Design Based on Algo for Group Intelligence

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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, "Exploration of Energy-Saving Chilling Landscape Design Based on Algo for Group Intelligence," *Journal of Function Spaces*, vol. 2022, Article ID 1851623, 10 pages, 2022.

Retraction

Retracted: Mathematical Models for Analysis the Impact of Institutional Quality and Multilateral Financial Institutions on the Success of Green PPP Projects

Journal of Function Spaces

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 2023

Copyright © 2023 Journal of Function Spaces. 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. Li, R. Chen, and M. Ma, "Mathematical Models for Analysis the Impact of Institutional Quality and Multilateral Financial Institutions on the Success of Green PPP Projects," *Journal of Function Spaces*, vol. 2022, Article ID 3825211, 12 pages, 2022.

Retraction

Retracted: Optimization of College Students' Employment Prosperity Index System Based on Multiple Logit Models

Journal of Function Spaces

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 2023

Copyright © 2023 Journal of Function Spaces. 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. Xiang and W. Hu, "Optimization of College Students' Employment Prosperity Index System Based on Multiple Logit Models," *Journal of Function Spaces*, vol. 2022, Article ID 9367675, 8 pages, 2022.

Retraction

Retracted: Application of Wireless Sensor Network Model Based on Big Data Ecosystem in Intelligent Health Monitoring System

Journal of Function Spaces

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 2023

Copyright © 2023 Journal of Function Spaces. 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. Tu, "Application of Wireless Sensor Network Model Based on Big Data Ecosystem in Intelligent Health Monitoring System," *Journal of Function Spaces*, vol. 2022, Article ID 3179915, 10 pages, 2022.

Retraction

Retracted: Analysis of Option Butterfly Portfolio Models Based on Nonparametric Estimation Deep Learning Method

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Ge, X. Zhu, G. Bi, H. Zheng, and Q. Li, "Analysis of Option Butterfly Portfolio Models Based on Nonparametric Estimation Deep Learning Method," *Journal of Function Spaces*, vol. 2023, Article ID 4989036, 17 pages, 2023.

Research Article

Analysis of Option Butterfly Portfolio Models Based on Nonparametric Estimation Deep Learning Method

Xiangyu Ge ^{1,2}, Xia Zhu ², Gang Bi,² Hao Zheng ³, and Qing Li ²

¹Department of Finance, Wuhan Technology and Business University, Wuhan 430065, China

²School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China

³School of Management, Hubei University of Chinese Medicine, Wuhan 430065, China

Correspondence should be addressed to Xia Zhu; challengezhuzhu@163.com and Hao Zheng; 305743360@qq.com

Received 21 September 2022; Revised 19 October 2022; Accepted 13 April 2023; Published 28 August 2023

Academic Editor: Miaocho Chen

Copyright © 2023 Xiangyu Ge 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 option butterfly portfolio is the commonly option arbitrage strategy. In reality, because the distribution of the option state price density (SPD) function is not normal and unknown, so the nonparametric deep learning methods to estimate option butterfly portfolio returns are proposed. This paper constructs the single-index nonparametric option pricing model which contains multiple influencing factors and presents the nonparametric estimation form for option butterfly portfolio returns. The empirical analysis shows that the SPD function estimated by using single-index nonparametric option model can effectively calculate the option butterfly portfolio returns with the minimum option strike price interval and provide an effective reference tool for risk-averse investors with limited risk preferences.

1. Introduction

Since the listing of Shanghai Stock Exchange (SSE) 50ETF Fund options on February 9, 2015 has opened the prologue of options trading in mainland China, it becomes important to study the actual trading methods of various option spread combinations. Actually, options are of great significance to promote the healthy development of the capital market, which not only can help investors' hedge risks at a lower cost but also help different investors make appropriate investment choices based on their own risk preferences. How to reasonably test the effectiveness of option pricing has important practical significance. However, the existing methods to test the effectiveness of the option market mostly focus on posttest methods to verify the validity of the option parity formula (put-call parity) (see Stoll [1]) or the initial option fee and ending income of the spread portfolio model, and whether meet the no-arbitrage relationship put-call parity formula, few on option butterfly portfolio returns strategy method to test the effectiveness of the option market. Therefore, it would be a useful exploration to study how to use the option butterfly portfolio return strategy method to test the effectiveness of the option market.

In this paper, the butterfly portfolio returns have been constructed with the trading data in one minute from April 1, 2019 09:30:00 to May 31, 2019 15:00:00 of all SSE 50ETF European put options with expiration in June and July 2019, for a backtesting on the butterfly portfolio no-arbitrage return at maturity and actual return at maturity, including regression analysis, arbitrage the relationship between opportunities and arbitrage return and option value status, butterfly portfolio exercise price interval, the relationship of underlying asset price volatility rate, the intraday distribution of arbitrage opportunities, and the duration of the day, leading to a conclusion that although the SSE 50ETF option market has not been fully effective.

The butterfly portfolio return strategy is limited in terms of risks and returns, which is only constructed when the volatility of the future spot market is expected to be stable and benefits from the underlying price falling into the middle of the left and right strike prices of the butterfly portfolio. It is an arbitrage return strategy suitable for a market with flat fluctuations. The no-arbitrage return of this strategy is related to the state price density (SPD). Butterfly portfolio return strategy is featured with limited income and closed

risk exposure with low risk to achieve stable investment income. Importantly, the core of the option butterfly portfolio return strategy is to seize arbitrage opportunities to obtain profits. This paper will focus on option butterfly portfolio return strategy that generates returns when market volatility is stable.

In this paper, our contribution includes two aspects. Firstly, we extend finding the second derivative of option price of a single-indicator variable of option factor (the combination of the underlying asset price, the exercise price, exercise period, the risk-free rate and implied volatility, etc.) on the basis of nonparametric single-index option pricing model (see Li and Yang [2]). Secondly, we develop an efficient nonparametric estimation deep learning method to estimate the implied SPD function for single-index option pricing model and to present the numerical algorithm we use to estimate the option butterfly portfolio returns, with a comparison of results between parameter estimation method and nonparametric estimation method.

The remaining sections are compiled in the following manner. Section 2 provides a literature review. Section 3 proposes a nonparametric estimation deep learning method for single-index option price model with introducing the classical nonparametric estimation method of European option prices by performing kernel density estimation and local polynomial regression estimation by the classical nonparametric estimation method and the single-index nonparametric estimation method, respectively. Exactly, it is expected to find every minimum estimated mean square error values with these two methods. In Section 4, the empirical analysis of option butterfly portfolio returns based on parameter estimation method is attempted to find the relationship between the MSE of the classical nonparametric method and the MSE of the single-index nonparametric estimation deep learning method. The values of nonparametric estimation deep learning algorithm for option butterfly portfolio returns may generate. A conclusion of this paper is given in Section 5. The fundamental properties for the option pricing, the SPD function and the butterfly portfolio returns, and the calculation program codes are reviewed in the Appendix.

2. Literature Review

The time-state preference model proposed by Arrow and Debreu [3] has promoted the development of uncertainty investing theory and introduced Arrow-Debreu securities (i.e., underlying securities) whose prices are determined by the SPD function to define each Arrow-Debreu security that generates a payment at state x . The option price can be obtained indirectly by estimating the SPD. The information of SPD not only can be used to derive more than derivatives prices but also to measure the size of financial risks of commercial banks, investment banks, securities companies, and other financial institutions. In order to overcome the shortcomings of traditional VaR risk measurement, Ait-Sahalia and Lo [4] have proposed a new risk measurement method (E-VaR) based on SPD, which has two important characteristics; on the one hand, it contains all relevant economic

information, such as investors' risk appetite, asset price dynamics, and market clearing; it can be derived from the preference-based equilibrium model or the measurement on the basis of the Black-Scholes-Merton (BSM) model. Therefore, of risk based on SPD is more attractive than that in the traditional statistical sense. Regarding the estimation of SPD, the method of neural network is used to make nonparametric estimation of option price to estimate SPD by Hutchinson et al. [5]; and the method of binary tree is given to estimate SPD by Rubinstein [6]. The estimation of SPD is obtained by taking the second derivative of the option pricing formula with respect to the strike price on the basis of the BSM model by Ait-Sahalia and Lo [4]. Yang [7] has proposed a new semiparametric estimation method by combining mathematical models and nonparametric estimation methods to estimate the SPD function and has verified that the effect of option price estimation based on the semiparametric estimation method is better than adapted BSM estimation, direct nonparametric estimation, and semiparametric BSM estimation.

As mentioned earlier, the problem of derivatives pricing can be transformed into the problem of estimating SPD function. Breeden and Litzenberger [8] with strong assumptions about the underlying asset have proposed the analytical solution of SPD by the BSM model; that is, if the underlying asset price obeys geometric Brownian motion and the risk-free interest rate is unchanged, SPD obeys logarithmic positive state distribution. However, the underlying asset price is a more complex random process, and the analytical solution cannot be obtained; therefore, the estimation of SPD can only be performed by numerical approximation. Rubinstein [6] and Jackwerth and Rubinstein [9] have minimized the gap between the SPD and the prior distribution by using the method of prior distribution of the SPD. Because of the classical BSM model with too many assumptions, the estimation accuracy of the SPD method is proved to be insufficient by nonparametric estimation method. Therefore, nonparametric estimation methods without any presupposition requirement to estimate the SPD are crucial and necessary. The existing nonparametric estimation method for estimating the SPD function is mainly divided into fully nonparametric Nadaraya-Watson kernel estimation, semi-parametric Nadaraya-Watson kernel estimation relying on the BSM formula, and local polynomial estimation methods. However, all nonparametric estimation methods are very dependent on the quality of sample data. If the sample set is sparse in a certain area, the nonparametric point estimation near the area is not effective. Then, it is very meaningful to explore a nonparametric estimation deep learning method to estimate SPD that the performance of point estimation will not be greatly influenced when the sample points are sparse, leading to obtain the nonarbitrage income of option butterfly portfolio returns by estimating SPD.

The estimation method of option butterfly portfolio returns was first proposed by Breeden and Litzenberger [8] who have successfully induced the SPD function by the price of European call option price, and the first derivative of the option price related to the strike price is the distribution function of the asset state price underlying the option price,

while the second derivative is SPD function, which has opened the door to estimate option butterfly portfolio returns by nonparametric estimation method. Since then, more and more researches have turned the direction to estimate the SPD function. The performance of parametric estimation method of the BSM formula derivation has been presented with a strong premise, but option price state function actually is hard to meet the assumption of normal distribution form. Based on the analysis of Breeden and Litzenberger [8], one study reported by Ait-Sahalia [10] is mainly about the nonparametric estimation method for the asset SPD function of the underlying option price. On this basis, Kiesel [11] has made a full explanation for the nonparametric estimation of the asset SPD function of the underlying option price under the butterfly arbitrage principle. It has been proved to be arbitrage-free nature of butterfly portfolio return strategy by Carr and Madan [12]. The conditions of arbitrage opportunities in the butterfly portfolio return strategy that are available have been provided by Davis and Hobson [13] with the finite probability space method.

The estimation of option butterfly portfolio return centers on the SPD function of the underlying asset of the option. The method proposed by Breeden and Litzenberger [8] will not be ensured unless the assumptions that state price obeys normal distribution are satisfied. It also assumes that the SPD function is an unknown nonparametric form by Ait-Sahalia [10] and Kiesel [11]. According to the nonparametric estimation theory of Ait-Sahalia [10] and Kiesel [11], the focus of the estimation of the asset SPD function of the underlying option is put on the nonparametric estimation of the option price. It is supposed to first estimate the nonparametric form of the option price, jumping to the nonparametric form of the SPD function after the first-order derivative.

However, previous studies of nonparametric estimation deep learning methods are carried out with larger sample size, while data of delivery option price in a given time (e.g., a trading day) is limited from the number of 20 to 50. It is useful to aggregate the data over time to increase the number of samples for nonparametric pricing methods. For example, option data set for 1 year was adopted by Ait-Sahalia and Lo [14], and two-dimensional rolling model of cross section and time was supported by Fan and Mancini [15] who chose option data set for 3 years. Ludwig [16] has argued that although the option pricing model of aggregating data over time is effective to solve the problem of sample size, it is simple to combine option contracts with different term structures to make nonparametric regression: price contract of different rights under the same term, ignoring the influence of the term structure of option pricing, which is vulnerable to nonstationary and calendar effect.

The founding discovered by Breeden and Litzenberger [8] is that the first derivative of the exercise price of a call option is less than zero (monotonicity constraint), and the second derivative is greater than zero (convexity constraint). The monotonicity and the constraint condition of convexity are called as shape constraints or no-arbitrage constraints, which the pricing model is named as a nonparametric

option pricing model no-arbitrage constraints by Ait-Sahalia and Duarte [17]. Compared with the genuine nonparametric option pricing method of Ait-Sahalia and Lo [14], option pricing model based on a nonparametric estimation deep learning without arbitrage constraints is shown as follow aspects. Firstly, it is to ensure that the risk-neutral probability density function is positive value to have arbitrage opportunities. Secondly, the model has no calendar arbitrage effect with no need for scrolling the data set by time (see Ludwig [16]). Lastly, only a small sample of a single term structure is needed without rolling over a larger number of data sets (there are only dozens of option contracts for a single term structure).

Followed by the no-arbitrage constraint pricing model of Ait-Sahalia and Duarte [17], there have been many studies on nonparametric regression under no-arbitrage constraints. For example, the idea of Yatchew and Hardle [18] is to use nonparametric least squares method and Bootstrap method to consider tail constraints under the condition of the call option price to find better the effect of getting the tail constraint. Hardle and Hlavka [19] and Birke and Pilz [20] have further studied the nonparametric estimation method under call option pricing no-arbitrage constraints but only for the different estimation method. Monteiro and Santos [21] have established a nonparametric regression model with both call and put option data, which has been transformed into a quadratic programming model to solve it.

Regarding the nonparametric estimation of the SPD function, local polynomial estimation is a genuine nonparametric estimation method to overcome the boundary effect. Because the least squares method is used to estimate the regression function and the reciprocal in this method, it is easy to calculate the estimated value of the explained variable and its partial derivative. Ait-Sahalia and Duarte [17] have estimated the SPD by using a local polynomial approach, without a proof of the asymptotic nature of the estimator. Based on the research of Ait-Sahalia and Duarte [17], many scholars have discussed the local polynomial method to estimate the SPD. After the discussion of the convergence of nonparametric estimation, Li et al. [22] have obtained the estimator of SPD and the deviation and variance by the method of local polynomial estimation, with the analysis of the convergence of local polynomial estimation and the speed of the convergence of that.

3. The Single-Index Nonparametric Estimation Method for the SPD Function

According to the nonparametric estimation method for the SPD function raised by Ait-Sahalia and Lo [14], it is related to option price concerning on the strike price. Actually, the butterfly portfolio return is influenced not only by the SPD function but also by the exercise price and other factors such as volatility. This main idea of this paper is to estimate the option butterfly returns by nonparametric estimation—the local polynomial method. Specially, putting forward a nonparametric estimation method for the SPD function has taken multiple factors into consideration, which is one of the machine learning method.

As shown in Appendix A.1, as to GBM Equation (A.1) for nondividend stocks, it can be extended to the case of dividends as Equation (A.9). If the stock price $S(t)$ satisfies the GBM stochastic differential (Equation (A.9)), the solution to $S = S(t)$ is

$$S(t) = S(0) \exp \left[\sigma(W(t) - W(0)) + \left(\mu - q - \frac{1}{2}\sigma^2 \right) t \right]. \quad (1)$$

Therefore

$$S(T) = S(t) \exp \left[\sigma(W(T) - W(t)) + \left(\mu - q - \frac{1}{2}\sigma^2 \right) (T - t) \right]. \quad (2)$$

Now defining the drift transformation $\tilde{W}(t)$ of Brownian motion $W(t)$ as

$$\tilde{W}(t) = W(t) + \frac{\mu - r}{\sigma} t, t \leq t \leq T, \quad (3)$$

so

$$\tilde{W}(T) - \tilde{W}(t) = W(T) - W(t) + \frac{\mu - r}{\sigma} (T - t). \quad (4)$$

Substituting Equation (4) into Equation (2) to get the new expression of $S(T)$ as

$$S(T) = S(t) \exp \left[\sigma(\tilde{W}(T) - \tilde{W}(t)) + \left(r - q - \frac{1}{2}\sigma^2 \right) (T - t) \right], \quad (5)$$

let

$$\begin{aligned} X &= -\frac{W(T) - W(t)}{\sqrt{T - t}}, \\ Y &= -\frac{\tilde{W}(T) - \tilde{W}(t)}{\sqrt{T - t}}, \text{ or} \\ Y &= X - \frac{\mu - r}{\sigma} \sqrt{T - t}, \end{aligned} \quad (6)$$

then

$$S(T) = S(t) \exp \left[-\sigma\sqrt{\tau}y + \left(r - q - \frac{1}{2}\sigma^2 \right) \tau \right]. \quad (7)$$

Note that the difference between Equation (7) and Equation (2) is that μ is replaced by r . It is known that under the original probability measure \mathbb{P} , the random variable X obeys the standard normal distribution, and its density function is $f_X(x) = (1/\sqrt{2\pi})e^{-(x^2/2)}$. However, the random variable Y in \mathbb{P} does not obey the standard normal distribution. Now, we

define the Radon-Nikodym derivative as follows:

$$Z(t, T) = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \exp \left[\int_t^T \theta dW(s) - \frac{1}{2} \int_t^T \theta^2 ds \right] = e^{\theta\sqrt{\tau}x - (1/2)\theta^2\tau}, \quad (8)$$

where $\theta = (\mu - r)/\sigma$. Then, we achieve a new probability measure $\tilde{\mathbb{P}}$, such that

$$\begin{aligned} d\tilde{\mathbb{P}} &= \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} d\mathbb{P} = e^{\theta\sqrt{\tau}x - (1/2)\theta^2\tau} d\mathbb{P} = \frac{1}{\sqrt{2\pi}} e^{\theta\sqrt{\tau}x - (1/2)\theta^2\tau} e^{-(x^2/2)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-(x - \theta\sqrt{\tau})^2/2} = \frac{1}{\sqrt{2\pi}} e^{-(y^2/2)} = f_Y(y) \end{aligned} \quad (9)$$

That is, $y = x - \theta\sqrt{\tau}$, Y obeys the standard normal distribution under $\tilde{\mathbb{P}}$; that is, the random process $\{\tilde{W}(t), 0 \leq t \leq T\}$ is a standard Brownian motion under $\tilde{\mathbb{P}}$. Let $\tilde{\mathbb{E}}(\cdot)$ represent the mathematical expectation about the new measure $\tilde{\mathbb{P}}$ and $\tilde{\mathbb{E}}_t(\cdot)$ represent the conditional expectation about the new measure $\tilde{\mathbb{P}}$ up to time t , under the condition of known information. Based on Equation (8) and Equation (9), Equation (A.16) (see Appendix A.2) can be written as

$$V(S_t, t) = \tilde{\mathbb{E}}_t \left[e^{-r(T-t)} \max \{S_T - K, 0\} \right]. \quad (10)$$

Furthermore, under the assumption that $y = B_t$ is a random variable with unknown distribution over time with a dividend, that is, $B_t \sim f_Y(B_t)$ represents a random variable, and $f_Y(B_t)$ is specified as probability density function, such as a normal distribution, or a nonparametric probability density function, so the price of the underlying asset of the option is proved to be an equation of the random variable for info set $\mathfrak{F} = (S_t, K, r, q, \tau)$. Changing the time from $\tau = T - t$ to $\tau = t - 0 = t$; that is, considering the time τ from 0 'to t , the state price of the underlying asset is described as S_0 , and the random variable B_t can be expressed as a state price function S_t as

$$S_t = S_0 \exp [\mu(\mathfrak{F}) - \sigma(\mathfrak{F})B_t], B_t = -\frac{\ln(S_t/S_0) - \mu(\mathfrak{F})}{\sigma(\mathfrak{F})}, \quad (11)$$

where $\mu(\mathfrak{F}) = (r - q - \sigma^2/2)\tau$ refers to the drift coefficient, and $\sigma(\mathfrak{F}) = \sigma\sqrt{\tau}$ means the diffusion coefficient, while the continuous dividend yield is defined as q .

According to the above Formula (11) and Formula (A.18) (see Appendix A.2), the antiderivative of the European call option price with random variable B_t can be deduced as follows:

$$\begin{aligned} c(S_0, \mathfrak{F}) &= e^{-r\tau} \tilde{\mathbb{E}}_0 [\max \{S_t - K, 0\}] \\ &= e^{-r\tau} \int_D^{+\infty} (S_0 e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F})B_t} - K) f_Y(B_t) dB_t, \end{aligned} \quad (12)$$

where the notation $\tilde{\mathbb{E}}_t$ (here, $t=0$) for conditional expectations, $D = (\ln(K/S_0) - \mu(\mathfrak{F}))/\sigma(\mathfrak{F})$, which is identified as factor single indicator variable of option price in this paper. On the analysis of Equation (12), the option price is subject to variable transformation to become an anti-derivative of random variables B_t , and the lower limit of integral turns to be a single-index D . Thus, all the influencing factors of the option price are combined to achieve model dimension reduction. At the same time, it is noted that the SPD function $f_X(S_t)$ of the underlying asset of the option has a connection with the probability density function $f_Y(B_t)$ of the random variable as follows:

$$f_X(S_t) = \frac{f_Y(B_t)}{S_0 \sigma(\mathfrak{F}) e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F}) B_t}}, \quad (13)$$

when the properties of probability density function $f_Y(B_t)$ are the standard normal distribution, namely, $f_Y(B_t) = (1/\sqrt{2\pi}) e^{-(B_t^2/2)}$, then

$$\begin{aligned} c(S_0, \mathfrak{F}) &= e^{-r\tau} \int_D^{+\infty} (S_0 e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F}) B_t} - K) f_Y(B_t) dB_t \\ &= e^{-r\tau} \int_D^{+\infty} S_0 e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F}) B_t} f_Y(B_t) dB_t - e^{-r\tau} K \int_D^{+\infty} f_Y(B_t) dB_t \\ &= e^{-r\tau} \int_D^{+\infty} S_0 e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F}) B_t} \frac{1}{\sqrt{2\pi}} e^{-(B_t^2/2)} dB_t - e^{-r\tau} K \int_D^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(B_t^2/2)} dB_t \\ &= e^{-r\tau} S_0 e^{\mu(\mathfrak{F}) + \sigma^2(\mathfrak{F})/2} \int_D^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(1/2)[B_t + \sigma(\mathfrak{F})]^2} dB_t - e^{-r\tau} K [1 - N(D)] \\ &= e^{-r\tau} S_0 e^{\mu(\mathfrak{F}) + \sigma^2(\mathfrak{F})/2} \int_{D - \sigma(\mathfrak{F})}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(Z^2/2)} \\ &\quad \cdot dZ - e^{-r\tau} K [1 - N(D)] (Z = B_t + \sigma(\mathfrak{F})) \\ &= e^{-q\tau} S_0 [1 - N(D - \sigma(\mathfrak{F}))] - e^{-r\tau} K [1 - N(D)] \\ &= e^{-q\tau} S_0 N(-D + \sigma(\mathfrak{F})) - e^{-r\tau} KN(-D) \\ &= e^{-q\tau} S_0 N\left(\frac{\ln(S_0/K) + (r - q + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right) \\ &\quad - e^{-r\tau} KN\left(\frac{\ln(S_0/K) + (r - q - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right) \\ &= e^{-q\tau} S_0 N(d_{10}) - e^{-r\tau} KN(d_{20}) = c_{BSM}(S_0, \tau), \end{aligned} \quad (14)$$

where $Z = B_t + \sigma(\mathfrak{F})$, $r - q = (\mu(\mathfrak{F}) + \sigma^2/2)/\tau$, $d_{10} = (\ln(S_0/K) + (r - q + \sigma^2/2)\tau)/(\sigma\sqrt{\tau})$ and $d_{20} = (\ln(S_0/K) + (r - q - \sigma^2/2)\tau)/(\sigma\sqrt{\tau}) = d_{10} - \sigma\sqrt{\tau}$. Specifically, in Equation (14), when the volatility σ is equal as the classic BSM option pricing formula. When the volatility is supposed to be implied, it appears to be the semiparametric BSM option pricing formula (see Ait-Sahalia and Lo [14]).

Similarly, it is easy to derive the pricing formula of European put options with dividends. When the volatility σ in the formula is the historical volatility, the formula is the classic BSM option pricing formula; when the volatility σ changes to the implied volatility, the formula turns to be the semiparametric BSM option pricing formula. The put option price formula of the single-index model can be

deduced as the following (Formula (15)):

$$\begin{aligned} p(S_0, \mathfrak{F}) &= e^{-r\tau} \int_{-\infty}^D (K - S_0 e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F}) B_t}) f_Y(B_t) dB_t \\ &= e^{-r\tau} K \int_{-\infty}^D f_Y(B_t) dB_t - e^{-r\tau} \int_{-\infty}^D S_0 e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F}) B_t} f_Y(B_t) dB_t \\ &= e^{-r\tau} K \int_{-\infty}^D \frac{1}{\sqrt{2\pi}} e^{-(B_t^2/2)} dB_t - e^{-r\tau} \int_{-\infty}^D S_0 e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F}) B_t} \frac{1}{\sqrt{2\pi}} e^{-(B_t^2/2)} dB_t \\ &= e^{-r\tau} KN(D) - e^{-r\tau} S_0 e^{\mu(\mathfrak{F}) + \sigma^2(Z)/2} \int_{-\infty}^D \frac{1}{\sqrt{2\pi}} e^{-(1/2)[B_t + \sigma(\mathfrak{F})]^2} dB_t \\ &= e^{-r\tau} KN(D) - e^{-r\tau} S_0 e^{\mu(\mathfrak{F}) + \sigma^2(\mathfrak{F})/2} \int_{-\infty}^{D - \sigma(\mathfrak{F})} \frac{1}{\sqrt{2\pi}} e^{-(Z^2/2)} dZ (Z = B_t + \sigma(\mathfrak{F})) \\ &= e^{-r\tau} KN(D) - e^{-q\tau} S_0 N(D - \sigma(\mathfrak{F})) \\ &= e^{-r\tau} KN\left(-\frac{\ln(S_0/K) + (r - q - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right) \\ &\quad - e^{-q\tau} S_0 N\left(-\frac{\ln(S_0/K) + (r - q + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right) \\ &= e^{-r\tau} KN(-d_{20}) - e^{-q\tau} S_0 N(-d_{10}) = p_{BSM}(S_0, \tau). \end{aligned} \quad (15)$$

Based on the existing research of the single-index nonparametric option pricing model proposed by Li and Yang [2], the five influencing factors of option price $\mathfrak{F} = (S, K, r, q, \tau)$ are transformed into integral function of random variables B_t by the process of variable transformation, leading to the regression equation of put option price to single-index variable D :

$$\begin{aligned} p(S_0, \mathfrak{F}) &= e^{-r\tau} \int_{-\infty}^D (K - S_0 e^{\mu(\mathfrak{F}) - \sigma(\mathfrak{F}) B_t}) f_Y(B_t) dB_t \\ &= e^{-r\tau} \int_{-\infty}^D K \sigma(\mathfrak{F}) F_Y(B_t) dB_t = e^{-r\tau} K \sigma(\mathfrak{F}) \int_{-\infty}^D F_Y(B_t) dB_t. \end{aligned} \quad (16)$$

The above formula can also be written as

$$p(S_0, \mathfrak{F}) \frac{e^{r\tau}}{K \sigma(\mathfrak{F})} = \int_{-\infty}^D F_Y(B_t) dB_t. \quad (17)$$

We can see that Equation (17) provides us with a method for estimating the SPD function. $F_Y(B_t)$ is the distribution function of the random variable B_t . The first derivative of D is the probability density function of the random variable B_t . When Formula (17) is used to estimate the option price, the integral part of Formula (17) can be regarded as a whole, such as the function $G(D)$ about D to estimate the whole; that is, the above Formula (17) can be expressed as

$$Y_i = G(X_i) + \varepsilon_i, \quad (18)$$

where $Y_i = p(S_0, \mathfrak{F}_i)(e^{r\tau}/K\sigma(\mathfrak{F}_i))$, $X_i = D_i = \ln(K_i/S_0) - \mu(\mathfrak{F}_i)/\sigma(\mathfrak{F}_i)$. Therefore, in order to obtain the SPD function, it is only necessary to obtain the second-order partial derivative of the estimated function $G(D)$ with respect to the single-index D , such as the following formula.

$$\left. \frac{\partial^2 G(D)}{\partial D^2} \right|_{D=B_t} = f_Y(B_t). \quad (19)$$

TABLE 1: The historical dividend situation of the SSE 50ETF.

Years	2019	2018	2017	2016	2014	2013
Equity registration date	29/11/2019	30/11/2018	27/11/2017	28/11/2016	14/11/2014	14/11/2013
Dividends per share (RMB)	0.047	0.049	0.054	0.053	0.043	0.053
Years	2012	2012	2010	2008	2006	2006
Equity registration date	12/11/2012	15/5/2012	15/11/2010	18/11/2008	15/11/2006	18/5/2006
Dividends per share (RMB)	0.037	0.011	0.026	0.060	0.037	0.024

4. Empirical Analysis to Nonparametric Estimation Deep Learning Method for Butterfly Portfolio Models

Since the focus of this paper is put on the stock options of the SSE 50ETF, the underlying asset of which contains dividends, it depends on the BSM formula with dividends. Underlying assets are selected from January 4, 2016 to July 24, 2018 while the options come from the closing price of the day at each strike price for each contract month. The information of SSE 50ETF option contract terms is recently associated with four kinds of contracts on the market such as July, August, September, and December, and all of contracts expire on the fourth Wednesday of each month while July contract is about to expire. All of the data are source from strike price options in circulation in the market from July 3, 2018 to July 24, 2018 for contracts in August and December as well as June 1, 2018 to July 24, 2018 for contracts in September.

In this paper, the risk-free interest rate is replaced by the 20-day average of SHIBOR interest rate on July 24, 2018. More precisely, there are 21 working days left to expire for the August contracts, and SHIBOR2W is used as the risk-free interest rate. However, SHIBOR1M is selected for the September contracts which have 46 days left. For the December contract, which has 111 days to remain, we use SHIBOR3M as the risk-free rate. In addition, the dividend has been regarded as a constant q_i that stands for dividend per share at i times. Fortunately, the data of dividend per share of each time since the establishment of China Shanghai 50ETF Fund on December 30, 2004 are available. As for the dividend yield, the historical average dividend rate under continuous compound interest is more effective to calculate.

In order to determine the dividend rate q , this paper has inquired the historical dividend situation of SSE 50ETF Fund, which has paid out 12 dividends by 2019, as shown in Table 1.

In this paper, the historical average dividend rate under continuous compounding is used to represent the dividend rate q . Details are below:

$$q = \ln \left(1 + \frac{1/n \sum_{i=1}^n q_i}{S_0} \right). \quad (20)$$

Among them, q_i is the dividend of each dividend at the i -th dividend, and $S_0 = 2.938$ refers to the closing price of the SSE 50ETF on the day of equity registration in

2019. Therefore, the calculation result of q obtained by Formula (20) is $q = 0.0139$.

According to the above assumption about risk-free interest rate and dividend yield, the historical annual volatility σ_{year} of the underlying SSE 50ETF is specified as the volatility σ_{day} as follows:

$$R_t = \ln \left(\frac{S_t}{S_{t-1}} \right),$$

$$\mu = \frac{1}{T} \sum_{t=1}^T R_t, \quad (21)$$

$$\sigma_{\text{day}}^2 = \frac{1}{T-1} \sum_{t=1}^T (R_t - \mu)^2,$$

$$\sigma_{\text{year}} = \sigma_{\text{day}} \cdot \sqrt{252},$$

where S_t means the closing price of the SSE 50ETF on t day, and R_t refers to the daily yield, while T stands for the sample length of SSE 50ETF. The calculated annual historical volatility is $\sigma_{\text{year}} = 0.20657$.

4.1. The Classical Nonparametric Estimation Method for Option Price. From Equation (A.27) (see Appendix A.3), the SPD function is easy to be deduced as long as the option price function is valid. The parametric form results are obtained from the direct derivation of the classical BSM formula. Much more previous literatures have discussed the limitations of the parametric form method. For example, the calculation of the option price formula has been provided by Ait-Sahalia and Lo [14] depending on the nonparametric kernel estimation method, which estimate the nonparametric estimator $\hat{p}(\cdot)$ of the put option price very intuitively with the financial market data to obtain the second-order partial differential $\partial^2 \hat{p} / \partial K^2$. Under appropriate regular conditions, when $\hat{p}(\cdot)$ converges to the real's put option price $p(\cdot)$ according to the probability, A also probabilistically converges to $\partial^2 p / \partial K^2$ proportional to the SPD. When with a call option price market data set $\{p_i\}$, accompanying with the feature sets $\{\mathfrak{F}_i = (S_i, K_i, r_{t_i, \tau_i}, q_{t_i, \tau_i}, \tau_i)\}$, the Nadaraya-Watson kernel estimator of the put option price function is estimated as

$$\hat{p}(\cdot) = \mathbb{E}(p | \mathfrak{F}) = \frac{\sum_{i=1}^n \mathbb{K}_h(\mathfrak{F} - \mathfrak{F}_i) p_i}{\sum_{i=1}^n \mathbb{K}_h(\mathfrak{F} - \mathfrak{F}_i)}, \quad (22)$$

where $\mathbb{K}_h(x_i - x_0) = \mathbb{K}((x_i - x_0)/h)$ with bandwidth h , since Formula (22) involves 5 variables, and in the case of a limited number of samples data, the accuracy of the put option price function estimation will decrease with the increase of the number of variables.

Ait-Sahalia and Lo [14] used two methods to reduce the number of variables. First, it is assumed that the option pricing formula is not a function of the underlying asset price, risk-free interest rate, and dividend rate but depends on the future price of these variables $F_t = S_t e^{(r-q)\tau}$. Under this assumption, the number of regressors is reduced from 5 to 4. The put option pricing formula is reformulated as the following:

$$\hat{p}(F_{t,\tau}, K, r_{t,\tau}, \tau) = \frac{\sum_{i=1}^n \mathbb{K}_{h_F}(F_{t,\tau} - F_{t_i,\tau}) \mathbb{K}_{h_K}(K - K_i) \mathbb{K}_{h_r}(\tau - \tau_i) \mathbb{K}_{h_\tau}(r - r_{t_i,\tau}) P_i}{\sum_{i=1}^n \mathbb{K}_{h_F}(F_{t,\tau} - F_{t_i,\tau}) \mathbb{K}_{h_K}(K - K_i) \mathbb{K}_{h_r}(\tau - \tau_i) \mathbb{K}_{h_\tau}(r - r_{t_i,\tau})}. \quad (23)$$

The second method is semiparametric methods. Here, the option pricing formula is still given by the BSM model, but the volatility σ is based on the result of a nonparametric estimation $\hat{\sigma}(F_{t,\tau}, K, r_{t,\tau}, \tau)$, that is

$$\hat{p}(F_{t,\tau}, K, r_{t,\tau}, \tau) = p_{BSM}(F_{t,\tau}, K, r_{t,\tau}, \tau, \hat{\sigma}(F_{t,\tau}, K, r_{t,\tau}, \tau)). \quad (24)$$

Then, the kernel estimator of volatility $\hat{\sigma}(F_{t,\tau}, K, r_{t,\tau}, \tau)$ can be written as

$$\hat{\sigma}^2(F_{t,\tau}, K, r_{t,\tau}, \tau) = \mathbb{E}(\sigma^2 | F_{t,\tau}, K, r_{t,\tau}, \tau) = \frac{\sum_{i=1}^n \mathbb{K}_{h_F}(F_{t,\tau} - F_{t_i,\tau}) \mathbb{K}_{h_K}(K - K_i) \mathbb{K}_{h_r}(\tau - \tau_i) \mathbb{K}_{h_\tau}(r - r_{t_i,\tau}) \sigma_i^2}{\sum_{i=1}^n \mathbb{K}_{h_F}(F_{t,\tau} - F_{t_i,\tau}) \mathbb{K}_{h_K}(K - K_i) \mathbb{K}_{h_r}(\tau - \tau_i) \mathbb{K}_{h_\tau}(r - r_{t_i,\tau})}. \quad (25)$$

Suppose that X_1, X_2, \dots, X_n are independent and identically distributed samples extraded through the future price from a one-dimensional population X , and the probability density functions $f(x), x \in R$ of X are unknown, then, the kernel density of $f(x)$ is estimated as

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n \mathbb{K}_h(x - X_i), \quad (26)$$

where $\hat{f}(x)$ is the estimation of the probability density function and n means the number of samples. Here, h is the bandwidth, and $\mathbb{K}_h(x)$ is represented as the kernel function.

Two problems in using kernel density estimation are the choice of kernel function and the choice of bandwidth. Firstly, taking the one-dimensional case as an example, there are six commonly used kernel functions. Generally speaking, when the amount of data is large enough, the choice of the kernel function is not important. According to theoretical calculations, the kernel density estimation is very similar by using different kernel functions, which means that kernel density estimation is not sensitive to the choice of kernel function. Relatively speaking, kernel density estimation is

more sensitive to the choice of bandwidth with a total of 9901 pieces of data screened for the empirical study in this paper, and the sample size should be classified into a sufficient category. Therefore, it is more appropriate to select the Gaussian kernel function with the bandwidth h , which is $\mathbb{K}_h(x) = (1/\sqrt{2\pi})e^{-x^2/(2h^2)}$. Secondly, in theory, the bandwidth should decrease as the sample size increases, when $n \rightarrow \infty, h \rightarrow 0$. Based on Formula (26), the bandwidth h controls the degree of smoothness. If the bandwidth h is smaller, the influence of randomness will increase, and the kernel density function $f(x)$ will become an irregular shape. The important features of kernel density may be concealed, causing the estimated value of the kernel density function to fluctuate greatly and resulting in overfitting; and if the bandwidth h is larger, the sample information will be averaged by $(x - X_i)/h$, and the participation of each sample point will be reduced. The estimated result will be very smooth and accompanied by a large deviation. So under a given sample, the choice of bandwidth is crucial. Therefore, the cross-validated (CV) method (see Li and Racine [23]) is applied to obtain the empirically optimal bandwidth h in the nonparametric regression model as follows:

$$h_{\text{opt}} = \left(\frac{4\hat{\sigma}^5}{3n} \right)^{1/5} = 1.06\hat{\sigma}n^{-1/5}. \quad (27)$$

Because the purpose of this paper is mainly to use a nonparametric estimation deep learning method to estimate the butterfly portfolio strategy return, the use of the multivariate Nadaraya-Watson estimation method with a Gaussian kernel function or bandwidth matrix follows the general form in nonparametric estimation. If the heteroscedasticity is considered, the bandwidth h can be selected by using a nonparametric estimation model with variable bandwidth, but the focus of this paper is to estimate the butterfly portfolio returns, so there is no special requirement for the selection of the method.

Therefore, this paper uses the usual training mean squared error (MSE) method to select the bandwidth h , which assigned values in turn from 0.1 to 1 with an interval of 0.1 to select the bandwidth that minimizes MSE. More details are showing in

$$\mathbb{K}_h(x - X_i) = \frac{1}{\sqrt{2\pi}} e^{-(1/2)((x-X_i)/h)^2}, \quad (28)$$

$$\text{MSE} = 1/n \sum_{i=1}^n (p_i - \hat{p}_i)^2.$$

As shown in Table 2, when the bandwidth is selected as $h_{\text{opt}} = 0.1$, the MSE results of the kernel density nonparametric estimation for the SSE 50ETF put option price are the smallest. That is, the nonparametric estimation method can take the test see to the smallest MSE of option price estimation, e.g., $\text{MSE} = 0.0005965$. After determining the optimal bandwidth $h_{\text{opt}} = 0.1$, the data set is divided into two equal parts by random sampling, which are used as training set and test set, respectively. Next, it is designed to observe

TABLE 2: The MSE results of the kernel density nonparametric estimation for put option prices.

Bandwidth	0.1	0.2	0.3	0.4	0.5
MSE	0.0005965	0.0021889	0.0039184	0.0051961	0.0061221
Bandwidth	0.6	0.7	0.8	0.9	1
MSE	0.0068078	0.0073195	0.0077041	0.0079962	0.0082210

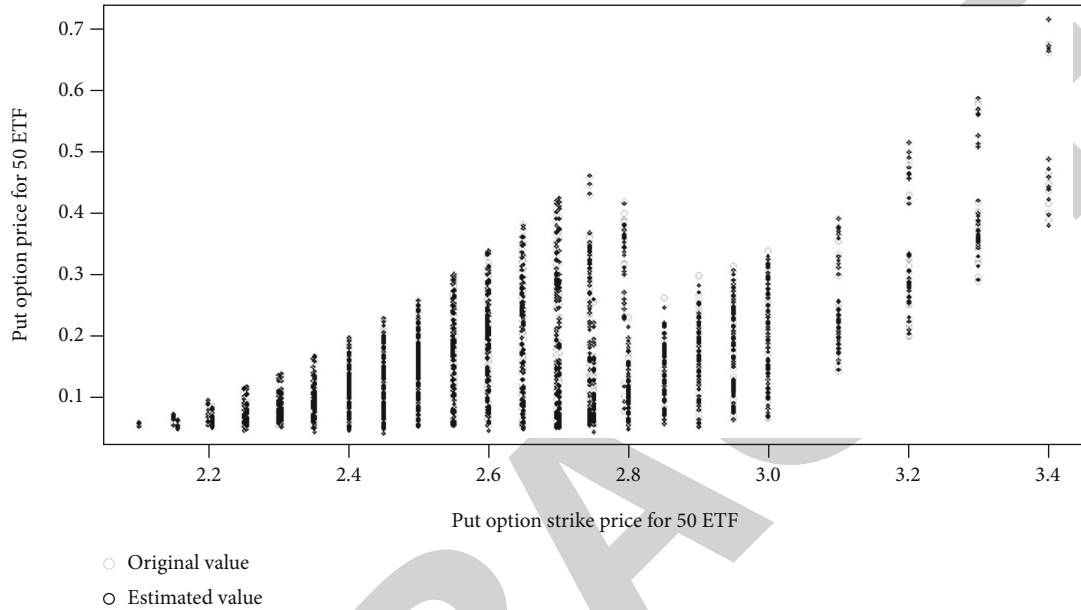


FIGURE 1: The values of kernel density nonparametric estimation for SSE 50ETF put option prices.

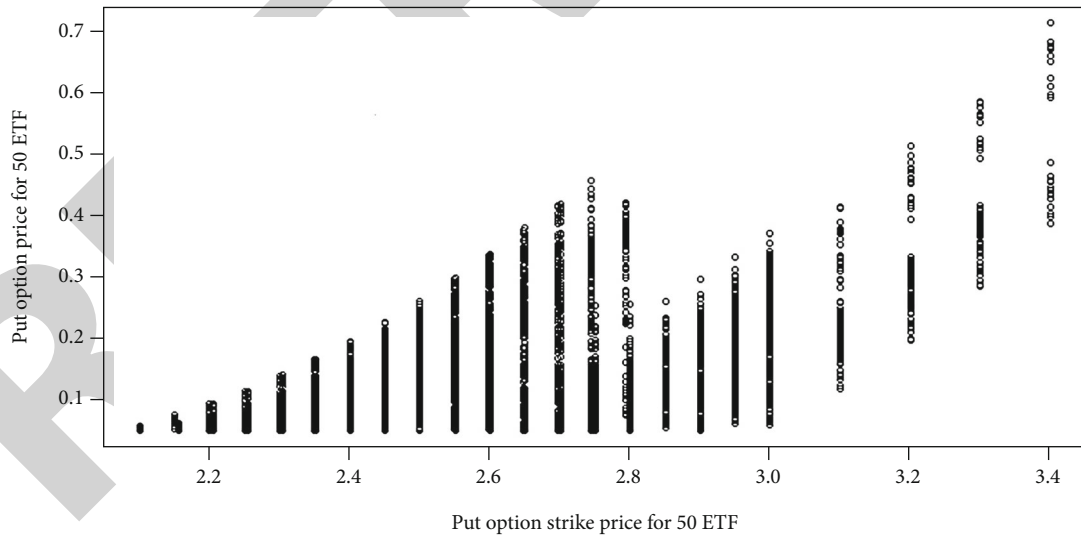


FIGURE 2: The relationship of all SSE 50ETF put option prices and corresponding strike prices.

the degree of deviation of the option price estimated according to Formula (23) from the original price series, as shown in Figure 1.

In order to observe the difference between the original values and the estimated values more clearly, it is necessary to plot the relationship between all SSE 50ETF European put option prices with expiration time equal to 27 days on

2019-05-30, as well as the original value and their corresponding strike prices during the sample period as shown in Figure 2.

Based on the estimated option price, it is easy to derive the strike price. In practice, an approximate solution to the SPD can be obtained according to the finite difference between the market-observed strike price and the discrete

TABLE 3: The provisions of the SSE's execution price and exercise price spacing.

Exercise price (RMB)	$K \leq 3$	$3 < K \leq 5$	$5 < K \leq 10$	$10 < K \leq 20$
Exercise price spacing (RMB)	0.05	0.1	0.25	0.5
Exercise price (RMB)	$20 < K \leq 50$	$50 < K \leq 100$	$K > 100$	
Exercise price spacing (RMB)	1.0	2.5	5.0	

option price. Assuming that there are N different strike prices with expiration time τ , K_1 represents the lowest strike price, and K_n represents the highest strike price. Here, three options with sequential exercise prices are K_{n-1} , K_n , and K_{n+1} . Generally speaking, $K_n - K_{n-1}$ and $K_{n+1} - K_n$ are not necessarily equal. The state price distribution function value $F(K_n)$ centered on K_n can be estimated by the following formula (29), namely

$$F(K_n) \approx e^{r\tau} \frac{P_{n+1} - P_{n-1}}{K_{n+1} - K_{n-1}}. \quad (29)$$

Therefore, the estimation formula of the SDP function is

$$f(K_n) \approx e^{r\tau} \frac{P_{n+1} - 2P_n + P_{n-1}}{(K_n - K_{n-1})^2}. \quad (30)$$

Formula (30) involves the exercise price interval of different options. It is noted that the provisions of SSE are shown in Table 3.

More importantly, the curve concerning the second derivative of the strike price is the SPD. It is found that the slope of the curve raises along with the increasing of the execution price, which is in line with the conclusion of Equation (A.9) (see Appendix A.1).

The idea of the Nadaraya-Watson kernel density estimation is that the option price at a given point \mathfrak{F} is obtained by the weighted average of the option price observations in the neighborhood of point \mathfrak{F} . When there are fewer points on both sides of the characteristic variable \mathfrak{F} , the results show a relatively large errors by the method, especially near boundary points, where there are no observations on one side of the boundary point. Therefore, the local polynomial method is useful to estimate the SPD.

4.2. The Nonparametric Estimation Deep Learning Method for Single-Index Option Price. Ait-Sahalia and Duarte [17] have estimated the SPD using a local polynomial estimation approach, without the shortcomings of Nadaraya-Watson kernel density estimation method, which overcomes the problem of large deviations on boundary points. The local polynomial estimation has the same order of magnitude on boundary points and interior points. At the same time, by the least squares method, the regression function and derivative can be estimated, without taking partial derivatives of the nonparametric estimators of the function.

The basic idea of local polynomial estimation is to let the value range of the independent variable D be \mathfrak{D} , for $\forall D_0 \in \mathfrak{D}$, and to select a certain neighborhood of D_0 . Those observation values of the dependent variable correspond to the obser-

vations of the independent variable, which is fitted in some way in this neighborhood, and the value of the curve obtained by this local fitting at D_0 is used as the estimated value $\hat{G}(D_0)$ of the regression function $G(\cdot)$ at G_0 . Supposing the model as follows

$$Y = G(D) + \varepsilon, E(\varepsilon) = 0, \text{Var}(\varepsilon) = \sigma^2 < \infty. \quad (31)$$

Let $G^{(k)}(D)$ be the k derivative of the regression function in the model (31), and then, $G(D) = G^{(0)}(D)$. Assuming that $G(D)$ has $p + 1$ order derivative, for $\forall D_0 \in \mathfrak{D}$, $G(D)$'s Taylor-expanded in the ε -neighborhood of $D = D_0$ as

$$\begin{aligned} G(D) &\approx G(D_0) + G'(D_0)(D - D_0) \\ &\quad + \frac{G''(D_0)}{2!}(D - D_0)^2 + \dots + \frac{G^{(p)}(D_0)}{p!}(D - D_0)^p. \end{aligned} \quad (32)$$

Set $G^{(k)}(D)/k! = \beta_k(D)$, the above Equation (32) can be rewritten as

$$\begin{aligned} G(D) &\approx \beta_0(D_0) + \beta_1(D_0)(D - D_0) \\ &\quad + \beta_2(D_0)(D - D_0)^2 + \dots + \beta_p(D_0)(D - D_0)^p \\ &= \sum_{j=1}^p \beta_j(D_0)(D - D_0)^j. \end{aligned} \quad (33)$$

In order to draw the SPD function curve derived from the single-index model, after obtaining the estimation of the function $G(D)$, according to Formula (31), the second-order derivative of $G(D)$ with respect to the single-index D is the SPD function, which is effective to estimate this second derivative by the local polynomial estimation method, and the parameter $\beta_j(D_0)$ is selected to minimize Formula (34) as

$$\min_{\beta_0, \dots, \beta_p} \sum_{i=1}^n \left[G(D_i) - \sum_{j=0}^p \beta_j(D_0)(D_i - D_0)^j \right]^2 \mathbb{K}_h(D_i - D_0), \quad (34)$$

where h is the bandwidth that controls the local neighborhood, $\mathbb{K}_h(D_i - D_0) = \mathbb{K}(D_i - D_0/h)$, and $\mathbb{K}(\cdot)$ represents the kernel function. The kernel function $\mathbb{K}(\cdot)$ and bandwidth h are still selected by the Gaussian kernel and the method of Formula (27), respectively. Since the inversion matrix part in the local polynomial estimation solution process is usually a singular matrix after substituting the actual data, the inversion matrix

TABLE 4: MSE of $G(D)$'s local polynomial estimation.

Bandwidth	0.1	0.2	0.3	0.4	0.5
MSE	0.0033341	0.0033006	0.0032892	0.0032945	0.0033008
Bandwidth	0.6	0.7	0.8	0.9	1
MSE	0.0033142	0.0033707	0.0034453	0.0036354	0.0057721

cannot be solved normally. Therefore, the Moore-Penrose generalized inverse matrix solution method is also used as an alternative method in the actual calculation. Let

$$\begin{aligned}
 X(D_0) &= \begin{pmatrix} 1 & D_1 - D_0 & (D_1 - D_0)^2 \\ 1 & D_2 - D_0 & (D_2 - D_0)^2 \\ \vdots & \vdots & \vdots \\ 1 & D_n - D_0 & (D_n - D_0)^2 \end{pmatrix}, \\
 Y &= \begin{pmatrix} G(D_1) \\ G(D_2) \\ \vdots \\ G(D_n) \end{pmatrix}, \\
 \beta(D_0) &= \begin{pmatrix} \beta_0(D_0) \\ \beta_1(D_0) \\ \vdots \\ \beta_p(D_0) \end{pmatrix},
 \end{aligned} \tag{35}$$

$$W(D_0) = \text{diag}(\mathbb{K}_h(D_1 - D_0), \dots, \mathbb{K}_h(D_n - D_0)).$$

Therefore, the estimated value of $\beta(D_0)$ can be obtained by the weighted least squares method as follows:

$$\widehat{\beta}(D_0) = (X^T W X)^{-1} X^T W Y. \tag{36}$$

In particular, when $p = 0$, we can get

$$\widehat{\beta}_0(D_0) = \widehat{G}(D_0) = \frac{\sum_{i=1}^n \mathbb{K}_h(D_i - D_0) Y_i}{\sum_{i=1}^n \mathbb{K}_h(D_i - D_0)}, \tag{37}$$

where $Y_i = G(D_i)$, Formula (37) changes form as the same as Formula (22). Formula (37) is the Nadaraya-Watson kernel estimation of $G(D_0)$; that is, the Nadaraya-Watson estimation is a zero-order local polynomial estimation. At the same time, the estimated value of the second-order partial derivative of the estimated function $G(D)$ with respect to a single-index D is

$$2\widehat{\beta}_2(D_0) = \left. \frac{\partial^2 G(D)}{\partial D^2} \right|_{D=D_0}, \tag{38}$$

which derivative $\widehat{\beta}_2(D_0)$ is the SPD function in the new sense proposed in this paper.

The distance between D_i and D_0 is designed to measure the weight of D_i when the estimation the density is D_0 .

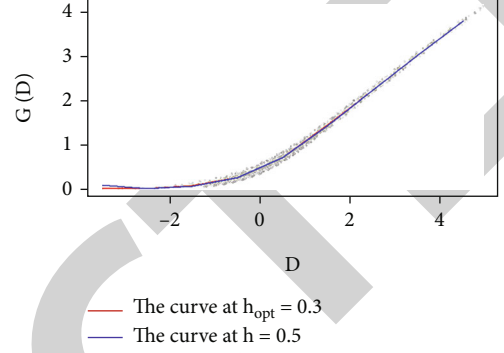


FIGURE 3: The curve of local polynomial fitting for the $G(D)$ function.

Besides, h has defined the size of D_0 in the process of estimation as the bandwidth. Here, Y for $G(D)$ the same as X for D . Similarly, the Gaussian kernel function is also applied. Specifically, the CV method is formulated to select the bandwidth of minimum MSE on the test set, with assigning 10 numbers from 0.1 to 1 with an interval of 0.1 as shown in Table 4.

As shown in Table 4, it can be seen that the selection of optimal bandwidth by using the CV method that minimizes MSE on the test set is $h_{\text{opt}} = 0.3$ (the optimal curve of local polynomial fitting for the $G(D)$ function represented by the red curve). The curve of local polynomial fitting for the $G(D)$ function can be seen in Figure 3. For comparison, the curve at $h = 0.5$ is included in Figure 3, in which real data are successful to be shown in the gray scattered points, and the blue curve is identified as the suboptimal curve of local polynomial fitting for the $G(D)$ function.

The put option price on the test set can be estimated by using the previously estimated Formula (19) which can be calculated after estimating $G(D)$. It can be seen that the comparison of all put option strike prices and the estimated value of the put option prices with their true values during the sample period has been presented in Figure 4.

By comparison, it is clear to see that the results estimated by the method in this section are more approximate to the actual data than the method provided in previous Section 4.1. It is evident that the best MSE of the put option price calculated by the nonparametric estimation method for single-index put option prices is 0.0001599, which is more superior than that calculated by the common nonparametric estimation method for put option prices as 0.0005965.

From Figure 4, the estimated value of the SPD function curve by using the single-index model is distributed around the value of Formula (38). Although the center of these values is not as close to the curve of Formula (38) as in the

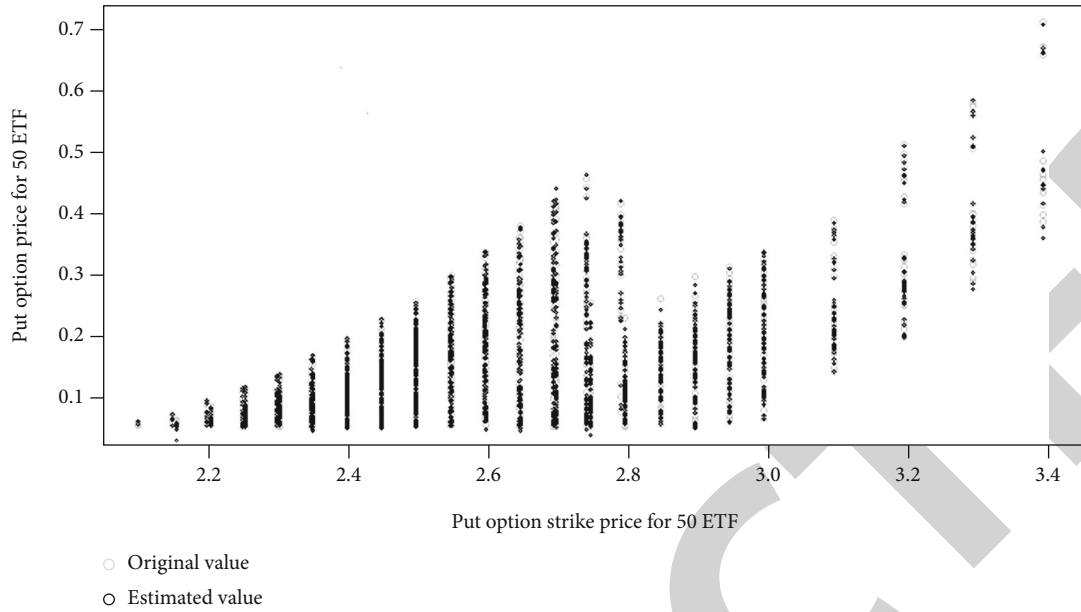


FIGURE 4: Original and estimated values of single-index model for SSE 50ETF put option price.

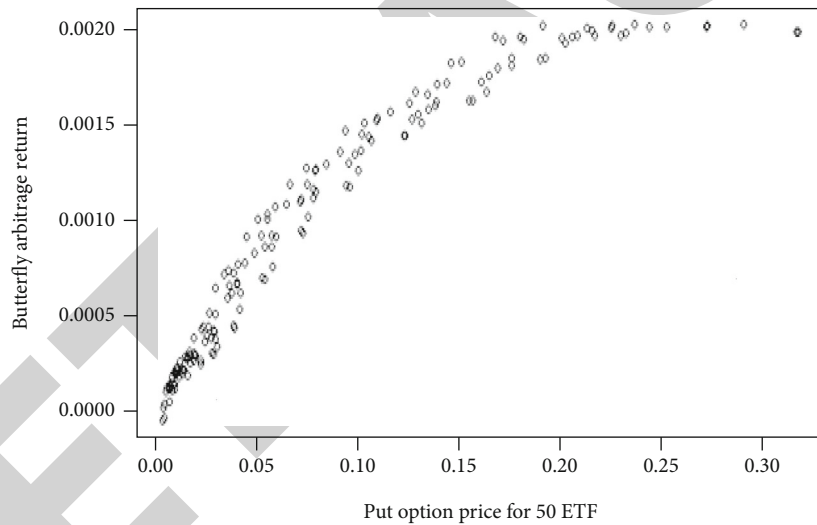


FIGURE 5: Put option contract price—butterfly portfolio returns of SSE 50ETF option in August.

previous Section 4.1, the curve estimated by the single-index model is smoother. The strike price near the sample boundary point of 3.4 RMB does not happen to wild volatility and breakpoint problems like the estimated value in the previous Section 4.1, which is especially important to obtain as much information as possible about the put option prices when the sample is limited or some points are missing. In practice, estimating the SPD can be as accurate as possible by combining models, such as using a nonparametric local polynomial estimation method at the nonboundary points, or a nonparametric single-index model method near the boundary points. Therefore, such a combined model has certain requirements on the sample size. Under small samples are methods such as biasing local polynomial estimation greatly. Similarly, the SPD function estimated by the single-index model estimation method may also show the characteristic

that the SPD value decreases with the increase of the expiration time.

4.3. The Nonparametric Estimation Deep Learning Method for Butterfly Portfolio Returns. With the proof of optimality, it has access to the single-index option price nonparametric estimation method provided in this paper to estimate the possible butterfly portfolio returns. After the estimation of $G(D)$, it is known that the first-order derivative of D has the property of SPD function, but Formula (13) is supposed to be valid only when it is satisfied with the condition of $\varepsilon \rightarrow 0$. As a matter of fact, the equation is established only when the strike price lies in the minimum interval in the real market.

Indeed, the value of the derivative required in the Formula (13) can be found by numerical solution. Furthermore,

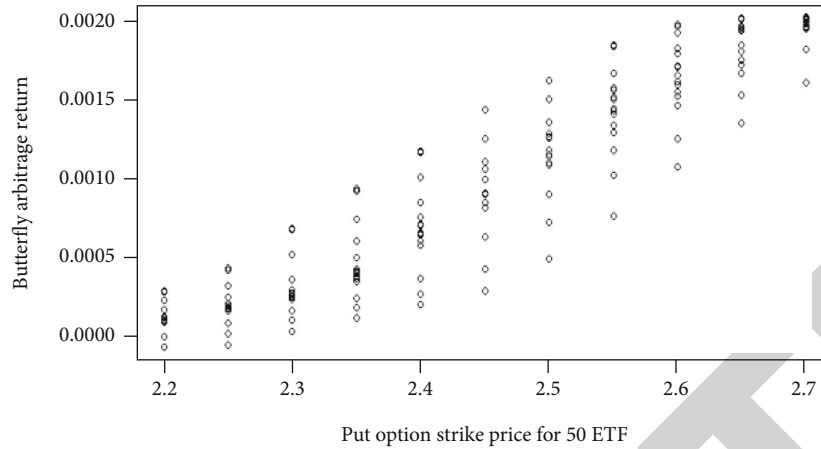


FIGURE 6: Put option strike price—butterfly portfolio returns of SSE 50ETF option in August.

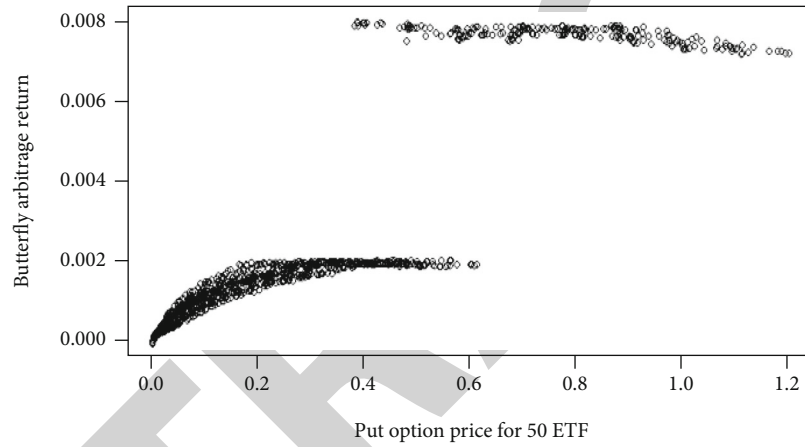


FIGURE 7: The chart of put option contract price—butterfly portfolio returns of SSE 50ETF option.

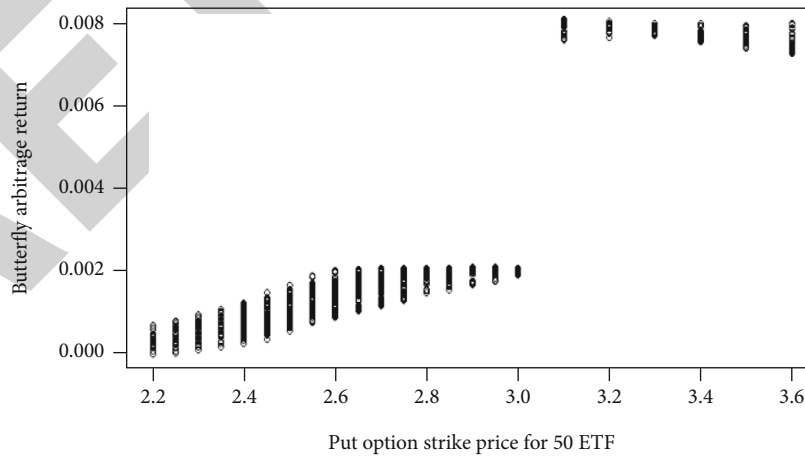


FIGURE 8: The chart of put option strike price—butterfly portfolio returns of SSE 50ETF option.

it is necessary to work out the result of $\Delta G(D)$ within the minimum change of D , such as $\Delta D = 0.00001$ so that the slope with a tiny move approximates to derivative value. Finally, substituting them into the formula to obtain parameter estimations of possible return from a short or long but-

terfly portfolio at minimum, the strike price interval of $\mathfrak{F} = (S, K, r, q, \tau)$. After estimating the optimal function $G(D)$, it is easy to obtain the possible butterfly portfolio returns at each point \mathfrak{F} in the sample period, taking the August contract as an example.

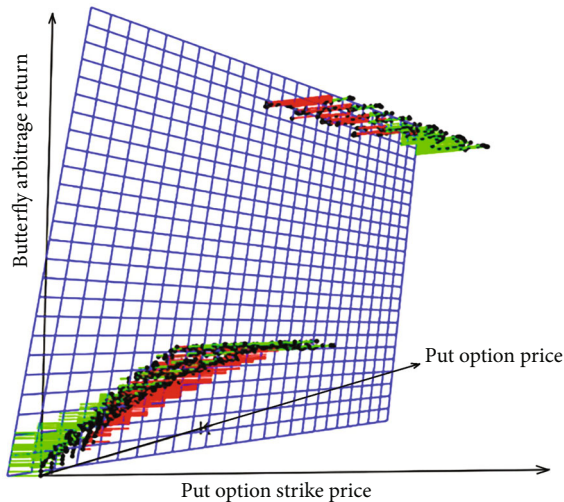


FIGURE 9: Three-dimensional graph of butterfly portfolio returns—strike prices—option prices.

Generally, the risk-free interest rate, volatility, and dividend yield are considered to be constant in the analysis. From Figure 5, it can be observed that the butterfly portfolio returns raise along with the increase of option price, taking the change of expiration time and the difference of strike price into consideration. On account of the change of expiration time and option price, it can be seen that there is a growth trend in butterfly portfolio returns with the growth of the strike price from Figure 6. Therefore, all of option contracts in August, September, and December in the market are summarized as follows.

As shown in Figures 7 and 8, it is feasible to extend the analysis period from the August contracts to the full-contract month put options listed in the market. The jump gap in the figure is attributed to the change of the strike price interval as specified in Table 4. In addition, this paper also draws three-dimensional graphs of nonparametric estimation for butterfly portfolio returns, strike prices, and option prices for butterfly portfolio construction at all points \mathfrak{F} in the selected sample period (see Figure 9).

For more discussion on the issue of put option butterfly portfolio returns, the corresponding put option price is calculated based on the bullish and bearish parity relation. The nonparametric estimation deep learning method in this paper can be used for the same research.

5. Conclusions

In this paper, more emphases have been put on the nonparametric estimation method of option price, and the practical application has been added to obtain the possible butterfly portfolio return at the minimum strike price interval. Firstly, the evidence of classical definition of SPD and the origin of butterfly portfolio return formula have been reviewed. Secondly, the nonparametric estimation model of single-index option price is proposed to deduce another form of SPD function calculated by European put option price, which provides a new model to calculate the important index of

financial risk measurement. As a result, the classic nonparametric estimation method of European option price is supposed to arrive at the minimum MSE values with approach of kernel density estimation, while the single-index nonparametric method of that to the minimum MSE values is required with local polynomial regression estimation. In comparison, it is reasonable to explain that the MSE of single-index method is smaller than classical nonparametric estimation method by the comparison, which illustrates that the method proposed has the superiority. Finally, the single-index nonparametric model is performed to estimate a new form of the SPD function, making it possible to obtain parameter estimations of possible butterfly portfolio returns at the minimum of the strike price interval, which has provided a powerful reference for investors taking butterfly positions in the options market. In future, the advanced method to estimate the option-implied state price density (SPD) will be proposed, such as sieve method which is one of the seminonparametric models.

Appendix

A. Fundamental Properties for Option Pricing, SPD and Butterfly Portfolio

A.1. Black-Scholes-Merton Option Pricing. Assuming that the change in the stock price $S(t)$ obeys the generalized Wiener process $W(t)$, or the stock price respects geometric Brownian movement (GBM), that is, with constant drift rate and variance rate, the model is derived as the following; the stochastic differential equation (SDE) shows

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \quad (\text{A.1})$$

where $S(t)$ denotes the stock price, σ means the stock price volatility, and μ is the stock return expectation. Assuming that the stock price $S(t) = S$ respects the Itô process, the assumption about $V = V(S, t)$ is the derivative price related to S , and the variable V is a function of S and t by Itô's lemma

$$dV = \left(\frac{\partial V}{\partial S} \mu S + \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial V}{\partial S} \sigma S dW(t). \quad (\text{A.2})$$

It is appropriate to choose the combination to remove uncertainty, such as taking a short position in a derivative and regarding a $\partial V/\partial S$ amount of stock as a combination. After defining Π as the value of the combination, then

$$\begin{aligned} \Pi &= V - \frac{\partial V}{\partial S} S, \\ d\Pi &= dV - \frac{\partial V}{\partial S} dS. \end{aligned} \quad (\text{A.3})$$

Substituting Equation (A.1) and Equation (A.2) into Equation (A.3), the equation in discrete form is

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt. \quad (\text{A.4})$$

It is found that the $W(t)$ term in the equation has been eliminated, so the investment portfolio must be risk-free within dt time, and at the same time, it can be discovered that an instantaneous rate of return is equal to the market risk-free rate that can be obtained; therefore

$$\frac{d\Pi}{\Pi} = rdt, \quad r\Pi dt = \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt, \quad (\text{A.5})$$

where r is the risk-free interest rate. Substituting Equations (A.3) and (A.4) into Equation (A.5), we can get

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0. \quad (\text{A.6})$$

Equation (A.6) is the Black-Scholes-Merton (BSM) equation.

When $t = T$, the boundary conditions for European call option are $C(S, T) = \max(S_T - K, 0)$, where K is the option strike price. When $S_T = S(T) = 0$, the call option yield is 0, so when $S(t) = S = 0$, there is $C(0, t) = 0$. When $S \rightarrow \infty$, the value of the call option becomes the value of the stock, that is, $C(S, t) \sim S$. The boundary conditions for European put option are $P(S, T) = \max(K - S_T, 0)$, if $S \equiv 0$, then, the terminal payoff of the put option is K , assuming the interest rate r is constant, and the boundary condition for $S = 0$ is $P(0, t) = Ke^{-r(T-t)}$. When $S \rightarrow \infty$, the put option cannot be exercised, so $P(S, t) \rightarrow 0$. Substituting the boundary conditions into the BSM (Equation (A.6)) to solve the pricing formulas for European call option price and European put option price, then

$$C(S, t) = SN(d_1) - Ke^{-r(T-t)}N(d_2), \quad (\text{A.7})$$

$$P(S, t) = Ke^{-r(T-t)}N(-d_2) - SN(-d_1), \quad (\text{A.8})$$

where $d_1 = (\ln(S/K) + (r + \sigma^2/2)(T-t))/(\sigma\sqrt{T-t})$ and $d_2 = (\ln(S/K) + (r - \sigma^2/2)(T-t))/(\sigma\sqrt{T-t}) = d_1 - \sigma\sqrt{T-t}$.

When the underlying common stock pays dividends, $D(S, t)$ is denoted as the dividends paid by one stock per unit time, and the expected rate of return in the stochastic differential Equation (A.1) that the stock price obeys goes to become $\mu - D(S, t)/S$, and then, the value of the option with dividends is defined as $H(S, t)$, and when r, σ are constants, $D(S, t) = qS$, on account of further proof in the book by Hull [24], the stock price satisfies the stochastic differential equation.

$$dS(t) = (\mu - q)S(t)dt + \sigma S(t)dW(t). \quad (\text{A.9})$$

The corresponding option value equation is

$$\frac{\partial H}{\partial t} + (r - q)S \frac{\partial H}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 H}{\partial S^2} - rH = 0. \quad (\text{A.10})$$

In Formula (A.10), making variable substitution $\tilde{S} = Se^{-q(T-t)}$, we can get

$$\frac{\partial H}{\partial t} + r\tilde{S} \frac{\partial H}{\partial \tilde{S}} + \frac{1}{2} \sigma^2 \tilde{S}^2 \frac{\partial^2 H}{\partial \tilde{S}^2} - rH = 0. \quad (\text{A.11})$$

The terminal condition for a call option with a dividend payment remains as $H(S, T) = \max(S_T - K, 0)$, when $S = 0$, there still is $H(0, t) = 0$. When $S \rightarrow \infty$, the value of the call option becomes the value of the stock; that is, $H(S, t) \sim Se^{-q(T-t)}$, we can get

$$H(S, t) = C(\tilde{S}, t) = e^{-q(T-t)}SN(d_{10}) - Ke^{-r(T-t)}N(d_{20}), \quad (\text{A.12})$$

where $d_{10} = (\ln(S/K) + (r - q + \sigma^2/2)(T-t))/(\sigma\sqrt{T-t})$ and $d_{20} = (\ln(S/K) + (r - q - \sigma^2/2)(T-t))/(\sigma\sqrt{T-t}) = d_{10} - \sigma\sqrt{T-t}$. It also shows that stock holders have dividend income, while option holders have no dividend income.

A.2. State Price Density (SPD) Function. Because of the difficulty in solving the BSM differential equation, the most commonly used derivative pricing method is the equivalent martingale measure. The solution of BSM option pricing (Equation (A.6)) has probabilistic expression as

$$V(S, t) = e^{-r(T-t)} \int_0^\infty \max\{S_T - K, 0\} f(S_T, T; S, t) dS_T, \quad (\text{A.13})$$

where $f(S_T, T; S, t)$ is the risk-neutral transfer density function or SPD function of the stock taking the value of S_T at time T under the condition of S . Equation (A.13) means that the price of the option at time t is the discount of its value at time T under the risk-neutral probability. Defining $\max\{S_T - K, 0\} = \Phi(S_T)$ and differentiating Formula (A.13), we can get

$$\begin{aligned} \frac{\partial V}{\partial t} &= rV + e^{-r(T-t)} \int_0^\infty \Phi(S_T) \frac{\partial f}{\partial t} dS_T, \\ \frac{\partial V}{\partial S} &= e^{-r(T-t)} \int_0^\infty \Phi(S_T) \frac{\partial f}{\partial S} dS_T, \\ \frac{\partial^2 V}{\partial S^2} &= e^{-r(T-t)} \int_0^\infty \Phi(S_T) \frac{\partial^2 f}{\partial S^2} dS_T. \end{aligned} \quad (\text{A.14})$$

Since the geometric Brown motion S satisfies Equation (A.1), its transfer density function $f(S_T, T; S, t)$ satisfies the Kolmogorov equation

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = 0. \quad (\text{A.15})$$

In particular, for Equation (A.6), with boundary conditions $V(0, t) = 0$ and $V(S, t) < S$ and a terminal payoff of

$V(S, T) = \Phi(S)$, the value of an undetermined interest at time t is

$$V(S, t) = e^{-r(T-t)} \int_0^\infty \Phi(S_T) f(S_T, T; S, t) dS_T = e^{-r(T-t)} \tilde{\mathbb{E}}_t(\Phi(S_T)), \quad (\text{A.16})$$

where the notation $\tilde{\mathbb{E}}_t$ for conditional expectations is explained in Section 4.

Under the assumption that the market is complete, assuming that there is a market in which bonds and stocks can be traded freely, and the risk-free interest on bonds is fixed as r , that is, when there is a unique risk-neutral probability measure in the market, at any given time t , for a derivative security that generates a payment on the maturity date, its price can be presented with the risk-neutral pricing formula provided by Harrison and Kreps [25] and Pliska [26] for a derivative security with a payment of $K = \Phi(S_T)$ at maturity T . When the risk-free interest rate r in the above formula is constant, $f(S_T, T; S, t)$ also refers to the SPD function, which summarizes all the required information in the course of pricing derivative securities. The SPD function $f(S_T, T; S, t)$ is the solution of the backward Kolmogorov Equation (A.15), which is derived from the European call option by Breeden and Litzenberger [8]:

$$f(S_T, T; S_t, t) = \frac{1}{S_t \sqrt{2\pi\sigma^2(T-t)}} \cdot \exp \left[-\frac{[(r-q-\sigma^2/2)(T-t) - \ln S_T + \ln S_t]^2}{2\sigma^2(T-t)} \right]. \quad (\text{A.17})$$

If noting $S(t) = S = S_t$, $\tau = T - t$, and substituting Formula (A.17) into Formula (A.16), we can get

$$V(S_t, t) = e^{-r\tau} \int_0^\infty \max\{S_T - K, 0\} \frac{1}{S_t \sqrt{2\pi\sigma^2\tau}} \cdot \exp \left[-\frac{[(r-q-\sigma^2/2)\tau - \ln S_T + \ln S_t]^2}{2\sigma^2\tau} \right] dS_T, \quad (\text{A.18})$$

making $y = ((r-q-\sigma^2/2)\tau - \ln S_T + \ln S_t)/(\sigma\sqrt{\tau})$, and then, Formula (A.18) becomes

$$V(S_t, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-r\tau} \max\{S_t e^{-\sigma\sqrt{\tau}y + (r-q-\sigma^2/2)\tau} - K, 0\} e^{-(y^2/2)} dy. \quad (\text{A.19})$$

For a no-dividend European call option, the process of the first-order derivative of Equation (A.16) of the strike

price is designed to obtain the SPD function as follows:

$$\frac{\partial c(S_t, \tau)}{\partial K} = \frac{\partial (e^{-r\tau} \int_K^\infty (S_T - K) f(S_T, T; S_t, t) dS_T)}{\partial K} = -e^{-r\tau} \int_K^\infty f(S_T, T; S_t, t) dS_T. \quad (\text{A.20})$$

Since the density function is definitely greater than 0 in the value range, corresponding to the European call option, when $S_T > K$, then $f(S_T, T; S_t, t) > 0$

$$\frac{\partial^2 c(S_t, \tau)}{\partial K^2} \Big|_{K=x} = e^{-r\tau} f(x, S_t, \tau) > 0. \quad (\text{A.21})$$

To be more specific, the call option price $c(S_t, \tau)$ is a convex function of the strike price K , and there is

$$c(S_t, \lambda K_1 + (1-\lambda)K_3, \tau) < \lambda c(S_t, K_1, \tau) + (1-\lambda)c(S_t, K_3, \tau), \forall K_1 < K_3, 0 < \lambda < 1. \quad (\text{A.22})$$

Let

$$\lambda K_1 + (1-\lambda)K_3 = K_2 \longrightarrow \lambda = \frac{K_3 - K_2}{K_3 - K_1}, \quad (\text{A.23})$$

Then

$$c(S_t, K_2, \tau) < \lambda c(S_t, K_1, \tau) + (1-\lambda)c(S_t, K_3, \tau), \forall K_1 < K_2 < K_3, 0 < \lambda < 1. \quad (\text{A.24})$$

Theoretically, the price of European call options without arbitrage is supposed to conform with the above equation.

A.3. Butterfly Portfolio Model. It is intuitive to construct an option butterfly portfolio returns to the explained Equation (A.17). Here, we can see that a call option butterfly portfolio returns are designed to buy a call option with a strike price of $K - \varepsilon$ and $K + \varepsilon$, respectively, with the same expiration date, and sell two calls with a strike price of K . Then, the gain and loss of the butterfly portfolio on the option premium are shown as

$$2c(S_t, K, \tau) - c(S_t, K - \varepsilon, \tau) - c(S_t, K + \varepsilon, \tau). \quad (\text{A.25})$$

The profit of call option butterfly portfolio returns is shown in Figure 10.

When $\varepsilon \rightarrow 0$, the return function form of the butterfly portfolio is similar to Dirac δ function ($\delta(x) = 0, x \neq 0$ and $\int_{-\infty}^{+\infty} \delta(x) dx = 1$). More precisely, only allowing the underlying stock price to be K , in this case, the return is shown as the area ε^2 of the shaded part in Figure 11.

However in other cases, it is virtually believed to be 0. Geometrically, regarding the combination at this time as a single security, it turns out to be $\varepsilon^2 f(x) e^{-r\tau}$ at (t) the moment, and $f(x)$ serves as the SPD function. In the principle of no arbitrage, theoretically, the profit and loss for option of constructing butterfly portfolio returns are

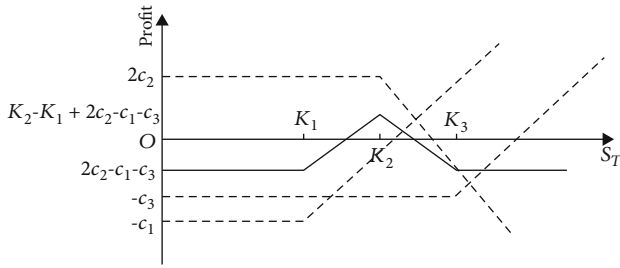


FIGURE 10: The profit of call option butterfly portfolio returns.

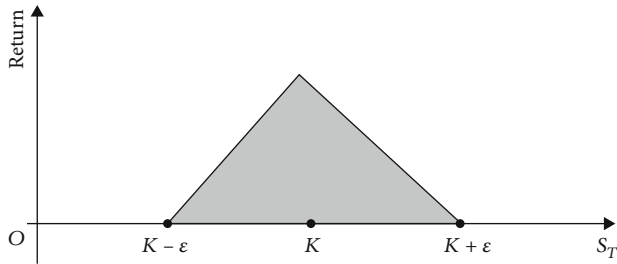


FIGURE 11: The butterfly portfolio returns.

possible to be equal to the return from the portfolio, so when $\epsilon \rightarrow 0$:

$$c(S_t, K - \epsilon, \tau) + c(S_t, K + \epsilon, \tau) - 2c(S_t, K, \tau) = \epsilon^2 f(K, S_t, \tau) e^{-r\tau}. \tag{A.26}$$

Substituting Formula (A.21) into the above Formula (A.26), then

$$\frac{c(S_t, K - \epsilon, \tau) + c(S_t, K + \epsilon, \tau) - 2c(S_t, K, \tau)}{\epsilon^2} = \frac{\partial^2 c(S_t, K, \tau)}{\partial K^2}. \tag{A.27}$$

Equation (A.27) has express the intuitive interpretation of option butterfly portfolio returns of Equation (A.21). It is acknowledged that the option butterfly portfolio returns are capable of being calculated by SPD function. However, it has been assumed that the SPD function is normal density function with the normal distribution in most existing researches. Actually, it is not. Therefore, the nonparametric estimation method comes into being.

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 is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant 71974204. This work was supported in part by the Humanities and Social Science Foundation of the Ministry of Education of China under grant 21YJC790065 and “the Fundamental Research Funds for the Central Universities,” Zhongnan University of Economics and Law under grant 2722022AK001.

References

- [1] H. R. Stoll, “The relationship between put and call option prices,” *The Journal of Finance*, vol. 24, no. 5, pp. 801–824, 1969.
- [2] Q. Li and Q. L. Yang, “Single-index nonparametric option pricing guided by model,” *Mathematical Statistics and Management*, vol. 37, pp. 1086–1094, 2018, (In Chinese).
- [3] D. Arrow and G. Debreu, *Theory of Value*, Science Direct Topics, 1959.
- [4] Y. Ait-Sahalia and A. W. Lo, “Nonparametric risk management and implied risk aversion,” *Journal of Econometrics*, vol. 94, no. 1-2, pp. 9–51, 2000.
- [5] J. M. Hutchinson, A. W. Lo, and T. Poggio, “A nonparametric approach to pricing and hedging derivative securities via learning networks,” *The Journal of Finance*, vol. 49, no. 3, pp. 851–889, 1994.
- [6] M. Rubinstein, “Implied binomial trees,” *The Journal of Finance*, vol. 49, no. 3, pp. 771–818, 1994.
- [7] Y. Yang, *Research of Option Pricing Based on the Extended Black Scholes Model and Nonparametric Method*, PhD thesis, Capital University of Economics and Business, 2015, (In Chinese).
- [8] D. T. Breeden and R. H. Litzenberger, “Prices of state-contingent claims implicit in option prices,” *Journal of Business*, vol. 51, no. 4, pp. 621–651, 1978.
- [9] J. C. Jackwerth and M. Rubinstein, “Recovering probability distributions from option prices,” *The Journal of Finance*, vol. 51, no. 5, pp. 1611–1631, 1996.
- [10] Y. Ait-Sahalia, “Nonparametric pricing of interest rate derivative securities,” *Econometrica*, vol. 64, no. 3, pp. 527–560, 1996.
- [11] R. Kiesel, “Nonparametric statistical methods and the pricing of derivative securities,” *Advances in Decision Sciences*, vol. 6, no. 1, Article ID S1173912602000019, p. 22, 2002.
- [12] P. Carr and D. B. Madan, “A note on sufficient conditions for no arbitrage,” *Finance Research Letters*, vol. 2, no. 3, pp. 125–130, 2005.
- [13] M. H. A. Davis and D. G. Hobson, “The range of traded option prices,” *Mathematical Finance*, vol. 17, no. 1, pp. 1–14, 2007.
- [14] Y. Ait-Sahalia and A. W. Lo, “Nonparametric estimation of state-price densities implicit in financial asset prices,” *Journal of Finance*, vol. 53, no. 2, pp. 499–547, 1998.
- [15] J. Fan and L. Mancini, “Option pricing with model-guided nonparametric methods,” *The Journal of American Statistical Association*, vol. 104, no. 488, pp. 1351–1372, 2009.
- [16] M. Ludwig, “Robust estimation of shape-constrained state price density surfaces,” *The Journal of Derivatives*, vol. 22, no. 3, pp. 56–72, 2015.
- [17] Y. Ait-Sahalia and J. Duarte, “Nonparametric option pricing under shape restrictions,” *Journal of Econometrics*, vol. 116, no. 1-2, pp. 9–47, 2003.

Retraction

Retracted: Prediction of Urban Scale Expansion Based on Genetic Algorithm Optimized Neural Network Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Kuang, "Prediction of Urban Scale Expansion Based on Genetic Algorithm Optimized Neural Network Model," *Journal of Function Spaces*, vol. 2022, Article ID 5407319, 11 pages, 2022.

Retraction

Retracted: Designing Landscape of Urban Gardening Based on Optimized Artificial Intelligence Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Liu, "Designing Landscape of Urban Gardening Based on Optimized Artificial Intelligence Model," *Journal of Function Spaces*, vol. 2022, Article ID 7963173, 12 pages, 2022.

Retraction

Retracted: The Experimental Validation of a Digital Filter Mathematical Modelling for Nuclear Fuel Vibration Monitoring System

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Yong, H. Shiyong, H. MengXuan et al., "The Experimental Validation of a Digital Filter Mathematical Modelling for Nuclear Fuel Vibration Monitoring System," *Journal of Function Spaces*, vol. 2022, Article ID 9741419, 11 pages, 2022.

Retraction

Retracted: Construction of a Knowledge Map Based on Text CNN Algorithm for Maritime English Subjects

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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 and A. Wei, "Construction of a Knowledge Map Based on Text CNN Algorithm for Maritime English Subjects," *Journal of Function Spaces*, vol. 2022, Article ID 6578682, 9 pages, 2022.

Retraction

Retracted: Modeling and Analysis of Multifocus Picture Division Algorithm Based on Deep Learning

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. You, "Modeling and Analysis of Multifocus Picture Division Algorithm Based on Deep Learning," *Journal of Function Spaces*, vol. 2022, Article ID 8326626, 10 pages, 2022.

Retraction

Retracted: Application of GA in Furniture Modeling Style Design

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Yu and C. Liu, "Application of GA in Furniture Modeling Style Design," *Journal of Function Spaces*, vol. 2022, Article ID 5264185, 11 pages, 2022.

Retraction

Retracted: Mathematical Modeling and Simulation of Online Teaching Effect Evaluation Based on Decision Tree Algorithm

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Sun, "Mathematical Modeling and Simulation of Online Teaching Effect Evaluation Based on Decision Tree Algorithm," *Journal of Function Spaces*, vol. 2022, Article ID 7425196, 11 pages, 2022.

Retraction

Retracted: Optimization Algorithm of Communication Resource Allocation in a Complex Network Based on an Improved Neural Network

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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 and Q. Liu, "Optimization Algorithm of Communication Resource Allocation in a Complex Network Based on an Improved Neural Network," *Journal of Function Spaces*, vol. 2022, Article ID 7309612, 8 pages, 2022.

Retraction

Retracted: Optimization of Three-Dimensional Model of Landscape Space Based on Big Data Analysis

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Wang, "Optimization of Three-Dimensional Model of Landscape Space Based on Big Data Analysis," *Journal of Function Spaces*, vol. 2022, Article ID 7002983, 11 pages, 2022.

Retraction

Retracted: Participation in Targeted Poverty Alleviation and Enterprise Innovation Investment: Analysis of the Mediating Effect Test Model Based on Financing Constraints

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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] T. Gao and H. Wang, "Participation in Targeted Poverty Alleviation and Enterprise Innovation Investment: Analysis of the Mediating Effect Test Model Based on Financing Constraints," *Journal of Function Spaces*, vol. 2022, Article ID 7060462, 16 pages, 2022.

Retraction

Retracted: Enterprise Management Decision and Financial Management Based on Dynamic Cost Volume Profit Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Guo, "Enterprise Management Decision and Financial Management Based on Dynamic Cost Volume Profit Model," *Journal of Function Spaces*, vol. 2022, Article ID 9016060, 8 pages, 2022.

Retraction

Retracted: Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Shen, "Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model," *Journal of Function Spaces*, vol. 2022, Article ID 1078057, 9 pages, 2022.

Retraction

Retracted: Modeling of Energy Saving and Comfort of Building Layout in Extreme Weather Urban Residential Area under the Background of Spatial Structure and Form Evolution: Taking Yichun as an Example

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Lu, G. Sun, E. Wang, and Z. He, "Modeling of Energy Saving and Comfort of Building Layout in Extreme Weather Urban Residential Area under the Background of Spatial Structure and Form Evolution: Taking Yichun as an Example," *Journal of Function Spaces*, vol. 2022, Article ID 6892035, 11 pages, 2022.

Retraction

Retracted: Application of Neural Networks in Financial Time Series Forecasting Models

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Li, "Application of Neural Networks in Financial Time Series Forecasting Models," *Journal of Function Spaces*, vol. 2022, Article ID 7817264, 9 pages, 2022.

Retraction

Retracted: Prediction Model of Music Popular Trend Based on NNS and DM Technology

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xu, M. Wang, H. Chen, and F. Hu, "Prediction Model of Music Popular Trend Based on NNS and DM Technology," *Journal of Function Spaces*, vol. 2022, Article ID 6104056, 9 pages, 2022.

Retraction

Retracted: The Construction of Adaptive Learning for Sports Based on Aerobics Trajectory Recognition Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xi, "The Construction of Adaptive Learning for Sports Based on Aerobics Trajectory Recognition Model," *Journal of Function Spaces*, vol. 2022, Article ID 8339745, 8 pages, 2022.

Retraction

Retracted: Exploration of Flood Prediction in Watersheds Based on the Fusion Analysis of Remote Sensing Big Data with Multiple Physical Fields

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xu and Y. Ouyang, "Exploration of Flood Prediction in Watersheds Based on the Fusion Analysis of Remote Sensing Big Data with Multiple Physical Fields," *Journal of Function Spaces*, vol. 2022, Article ID 9422553, 10 pages, 2022.

Retraction

Retracted: Certificateless Batch Authentication Scheme and Intrusion Detection Model Based on the Mobile Edge Computing Technology NDN-IoT Environment

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Sun, "Certificateless Batch Authentication Scheme and Intrusion Detection Model Based on the Mobile Edge Computing Technology NDN-IoT Environment," *Journal of Function Spaces*, vol. 2022, Article ID 5926792, 9 pages, 2022.

Retraction

Retracted: Analysis of Consumer Behavior Data Based on Deep Neural Network Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Yuan, "Analysis of Consumer Behavior Data Based on Deep Neural Network Model," *Journal of Function Spaces*, vol. 2022, Article ID 4938278, 10 pages, 2022.

Retraction

Retracted: Adjustment Method of “FAST” Active Reflector Based on Optimal Fitting Strategy

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xia, “Adjustment Method of “FAST” Active Reflector Based on Optimal Fitting Strategy,” *Journal of Function Spaces*, vol. 2022, Article ID 6994858, 15 pages, 2022.

Retraction

Retracted: Optimization of the Intelligent Sports Health Management System Based on Big Data Technology and Internet of Things

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Dang, "Optimization of the Intelligent Sports Health Management System Based on Big Data Technology and Internet of Things," *Journal of Function Spaces*, vol. 2022, Article ID 5053150, 19 pages, 2022.

Retraction

Retracted: Innovation and Research of Digital Artwork Design Based on Big Data and Integrated Media

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Zhai, "Innovation and Research of Digital Artwork Design Based on Big Data and Integrated Media," *Journal of Function Spaces*, vol. 2022, Article ID 6333688, 10 pages, 2022.

Retraction

Retracted: Mathematical Model Construction of Teaching Evaluation in Colleges and Universities Based on Convolutional Neural Network under the Background of Big Data

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Tan, "Mathematical Model Construction of Teaching Evaluation in Colleges and Universities Based on Convolutional Neural Network under the Background of Big Data," *Journal of Function Spaces*, vol. 2022, Article ID 7064287, 8 pages, 2022.

Retraction

Retracted: The Evaluation Model of College English Diagnostic Exercises Based on Machine Learning

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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 H. Wu, "The Evaluation Model of College English Diagnostic Exercises Based on Machine Learning," *Journal of Function Spaces*, vol. 2022, Article ID 9185827, 12 pages, 2022.

Retraction

Retracted: UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Fu, Y. Gao, L. Liu, M. Yang, and X. Zhu, "UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment," *Journal of Function Spaces*, vol. 2023, Article ID 9708143, 11 pages, 2023.

Research Article

UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

Gui Fu ^{1,2}, Yang Gao ², Liwen Liu ³, Mingye Yang ² and Xinyu Zhu ¹

¹Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Guanghan 618307, China

²School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China

³College of Flight Technology, Civil Aviation Flight University of China, Guanghan 618307, China

Correspondence should be addressed to Liwen Liu; llw@cafuc.edu.cn

Received 5 August 2022; Revised 15 October 2022; Accepted 20 March 2023; Published 28 March 2023

Academic Editor: Miaochoao Chen

Copyright © 2023 Gui Fu 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 information technology, various products used in information technology are also constantly optimized. Among them, the task and path planning of UAV in the high-end robot industry has always been the focus of relevant researchers. In the high-end robot industry, in addition to the research and development of UAVs, they also continue to learn and strengthen the task and path planning of UAVs. Nowadays, using unmanned aerial vehicles for real-time shooting has become the trend of this era. Drones have brought great convenience to people's lives, and more and more people are willing to use drones. Based on the above situation, this paper studies the task and path planning of UAV based on reinforcement learning in dynamic environment. In the case of unpredictable scene parameters, reinforcement learning method can be established by value function. Thus, a more reasonable path can be given to realize the reconnaissance and detection of points of interest. MATLAB simulation experiments show that the algorithm can effectively detect targets in complex terrain composed of terrain restricted areas, and return to the designated end point to complete communication. Firstly, the development of unmanned aerial vehicles in various countries and the social status of unmanned aerial vehicles are discussed. By making UAV build threat model and task allocation in dynamic environment. The path planning and optimization of UAV in dynamic environment is studied, and the path planning algorithm and Hungarian algorithm are added. The optimized UAV has the fastest data transmission and calculation speed, while the other two types of UAVs have slower data transmission and calculation speed. In particular, ordinary UAVs also have data transmission failures, resulting in incomplete experimental results. The results show that the optimized UAV system is better in data calculation and transmission, which also shows that the UAV can quickly plan and process flight paths, which is suitable for practical applications.

1. Introduction

In today's era, the high-end robot industry is developing more and more rapidly. In the development of UAV, the wrong task and path of UAV often occur [1]. In the process of UAV working, due to the environmental factors, the sudden change of the environment cannot be avoided, which leads to the UAV system failure, and failure fall problems often occur when the UAV is in the working state [2]. Therefore, the reinforcement learning of UAV mission and path planning in dynamic environment has always been the research direction of relevant researchers [3]. Because

the working environment of UAV is outdoor, UAV needs more powerful internal system than other camera equipment. From the perspective of UAV machine itself, during the execution of work, the chip in the system will be short circuited due to long-time continuous work, and the original work task will not be completed eventually [4]. From the perspective of environment, due to the variable environmental impact, UAV will also be separated from the originally planned path and mission. Extreme environment cannot avoid varying degrees of damage to UAVs. Therefore, we should not only consider the internal system construction of UAV but also try our best to ensure that UAV can work

normally in different environments [5–8]. Therefore, in the process of summarizing the UAV mission planning path, we should not only focus on whether the UAV system is complete but also improve the working state of UAV in the dynamic environment.

In order to make the UAV better adapt to the dynamic environment, we should purposefully improve and optimize the internal tasks and planning paths of the UAV. Because the UAV itself is relatively light, it also needs to meet the principle of simplifying the internal system of the UAV, which is more practical. With the development of the times, the application range of UAV is more and more extensive, and UAV can be found in all fields. At first, when unmanned aerial vehicles were first introduced, their endurance and shooting clarity were very poor [9]. Later, the performance of UAV was continuously upgraded. UAV is mainly used to complete complex tasks with high difficulty coefficient, and the task and path planning of UAV is the key to solve complex tasks. Path planning mainly refers to the existence of threatening obstacles in the environment of UAV [10]. Then, planning the optimal route for the UAV to avoid obstacles from the starting point to the destination is also one of the main factors for the UAV to realize autonomous flight. In the task allocation of UAV, the main purpose is to enhance the time performance and environmental adaptability of UAV. For the path planning level of UAV, modifying and improving the algorithm can improve the overall performance of UAV and then improve the selection of smooth flight path in the working process of UAV, so as to improve work efficiency [11].

As the core of artificial intelligence, machine learning has three main classifications. They are supervised learning, unsupervised learning, and reinforcement learning. The purpose of reinforcement learning is to make agents autonomous in an environment and get the maximum reward. At the same time, the concept of reinforcement learning is very broad, which is called reinforcement learning or reinforcement learning in the field of artificial intelligence. In cybernetics, it is called dynamic programming; although strictly speaking, the concept of dynamic programming was put forward long before reinforcement learning [12]. At the same time, the two are the relationship between inheritance and development. However, at present, reinforcement learning, which is generally considered to be approximately equivalent in concept, is one of the three branches of machine learning. At the same time, reinforcement learning algorithm is sequential decision-making. In order to maximize the overall benefits of the system, the action strategies taken by agents in reinforcement learning are allowed to have a long-term impact. At the same time, the return is delayed, and the previous action will affect the return after multiple steps. Finally, at the expense of immediate return, the agents in the decision-making system can obtain better long-term return [13].

The innovative contribution of the research lies in the analysis of the problem of UAV mission path planning based on reinforcement learning in a dynamic environment. Adding the correlation function to the flying altitude of the UAV makes the UAV more invisible. Then, the optimized Hungarian algorithm is added to the UAV

system to make the data communication flow within the system faster. Finally, the improved artificial potential field algorithm is added to the UAV system, which also accelerates the ability of computing data in the system. The optimized UAV has the fastest data transmission and calculation speed, while the other two types of UAVs have slower data transmission and calculation speed. Compared with ordinary UAV, the UAV after reinforcement learning can better adapt to the environment and greatly improve the internal system performance of the whole UAV. From the above, it can be seen that the UAV based on reinforcement learning in the dynamic environment has a very good market prospect.

2. Development Status of Reinforcement Learning UAV in Various Countries

With the advent of the era of big data, ordinary UAVs have also been upgraded into new UAVs that can adapt to the environment [14]. Nowadays, new unmanned aerial vehicles are widely used in various industries and fields, and the performance of internal systems is constantly optimized. This paper is based on the new UAV and then studies the UAV mission and path planning [15]. Among them, in the path planning of UAV, adding optimization algorithm to the UAV system can make a simple judgment on the path more accurately [16]. Only when the path of UAV flight is correct can it improve the efficiency of UAV to complete the task and then provide convenience for people's life. Using the optimized new UAV, the multiangle picture of the captured object can be collected in real time, and the collected picture can also be saved and automatically projected in the UAV system. Only if the UAV itself does not have a fault, it can normally take flight photos according to the planned flight path. However, the performance requirements of UAV vary in different application fields. The new UAV system studied in this paper has improved its internal stability and endurance compared with the ordinary UAV system. At the same time, how to fly normally in different dynamic environments is also the main problem. Only by constantly integrating the algorithm with the internal system of UAV can the ability of UAV to adapt to the environment be greatly improved, which is unmatched by ordinary UAV. The deep reinforcement learning (DRL) method can solve the problem of creating data sets by letting UAVs collect data by themselves in the training environment. Using sac algorithm can realize the action space of obstacle avoidance scheme based on continuous UAV so that UAV can make more accurate and smooth action choices. Using depth map as input, sac is combined with variational automatic encoder (VAE). The UAV is trained to complete the obstacle avoidance task in a simulated environment composed of multiple wall obstacles.

China is the country where drones are most widely used, and most of them are used in shooting, for example, photography, video recording, and follow-up. Drones are also occasionally used by police, firefighters, the military, or geological monitoring [17]. The use of drones is convenient for observation. By taking advantage of the small size of the

UAV, the target can be photographed or tracked in the surveillance dead corner that cannot be observed by people [18]. The picture information stored inside is automatically sent to the computer terminal through the network to achieve the work purpose and complete the layout task. Compared with ordinary UAVs, it reduces human resources, saves working time, and greatly reduces the difficulty of work. Moreover, compared with the new type of UAV, the ordinary UAV is larger in size, faster in power consumption, and higher in cost, and its production and use are not a small amount [19]. The emergence of new unmanned aerial vehicles not only reduces human and material resources but also greatly reduces expenditure [20]. It is also in line with China's concept of environmental protection and energy conservation. Moreover, the service performance and service life of the internal system of the entire UAV have also been improved so that the whole UAV can give full play to its greatest advantages and shoot clearer and better images in its working state.

In order to improve the efficiency of UAV and improve the disadvantages of UAV in a single environment, the internal system of UAV is the main research object in the UK. Under the premise of ensuring that the normal working state of UAV is not affected, the internal system of UAV is actually applied to different environments, and the internal system is continuously improved, which breaks the problem that ordinary UAV cannot work in a dynamic environment. Based on the dynamic environment, this paper also studies the reinforcement learning of UAV mission and path planning.

3. UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

3.1. Assignment of Reinforcement Learning UAV Tasks and Threats in a Dynamic Environment. Reinforcement learning is an important method in path planning. When the scene parameters are unpredictable, reinforcement learning method can be established by the value function. Thus, a more reasonable path can be given to realize the reconnaissance and detection of points of interest. In the case of unpredictable scene parameters, reinforcement learning method can be established by value function. Thus, a more reasonable path can be given to realize the reconnaissance and detection of points of interest. MATLAB simulation experiments show that the algorithm can effectively detect targets in complex terrain composed of terrain-restricted areas and return to the designated end point to complete communication. In the process of UAV working under the dynamic environment, the task of strengthening UAV and the allocation of existing threats in the task are the focus of this paper. In this paper, the obstacles are divided according to the classification method and the robot path planning. The main threat sources in the path planning part of UAV mission planning are divided into the following three categories for description: static threat, dynamic threat, and pop-up threat. (1) Static

threat: it mainly refers to the threat that is known during mission planning and will not change in the actual flight process. (2) Dynamic threat: it is a known type of threat with a very high probability of occurrence within a certain area identified by the UAV's onboard sensors. In the working state of UAV, it is very important to model the threat of obstacles. In the process of threat modeling of obstacles, the distribution position and situation of obstacles are mainly detected. In the threat modeling, the threat source should be transmitted to the UAV so that the UAV can analyze the threat of objects. In specific practical applications, obstacles are detected first, and then, appropriate responses are made to obstacles. Therefore, no matter where the UAV is in any environment, only by avoiding damage to the fuselage can it perform all tasks normally. In the process of UAV performing tasks, due to the different difficulty of tasks, a UAV may not be able to complete. At this time, multiple UAVs are required to perform a task at the same time. Multiple unmanned aerial vehicles can perform tasks better through information transmission between them. However, whether it is a UAV or multiple UAVs, its core is to manage and implement the internal system of the UAV. The internal framework of the UAV system is shown in Figure 1.

It can be seen from Figure 1 that the control system of UAV is the above-mentioned management part. The control system of UAV manages other UAVs that perform tasks together, including the overall cognition of terrain and the analysis and judgment of obstacles. Among them, the amount of two poles in the threat potential field generated by each obstacle is as follows:

$$\begin{aligned} U_i(x, y) &= 0, \\ U_i(x, y) &= 1. \end{aligned} \quad (1)$$

From the above formula, we can judge whether there is a threat of obstacles. When the result is 0, the UAV can fly normally, and when the result is 1, the UAV needs to make an evasive action and then return to the normal flight path. In practical application, there may be more than one obstacle. After multiple obstacles are superimposed, the relevant two terminal quantities are as follows:

$$\begin{aligned} U_1 &= \frac{2y_1 + y_2 - 2y_2}{2} \\ U_2 &= \frac{y_2}{2} \end{aligned} \quad (2)$$

Before the UAV enters the working state, it usually needs to evaluate and analyze the received tasks, that is, to analyze the tasks accurately. The relevant formula is as follows:

$$f_r = \lim_{m \rightarrow \infty} \sum_{j=1}^m \left(\frac{r_{\text{safe}}}{r_j} \right)^2. \quad (3)$$

According to the above formula, the task can be accurately analyzed and evaluated. Due to the different types of

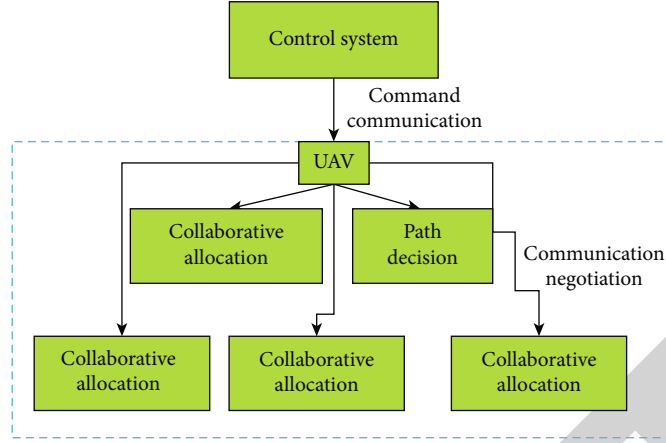


FIGURE 1: Internal framework structure management flow chart of UAV system.

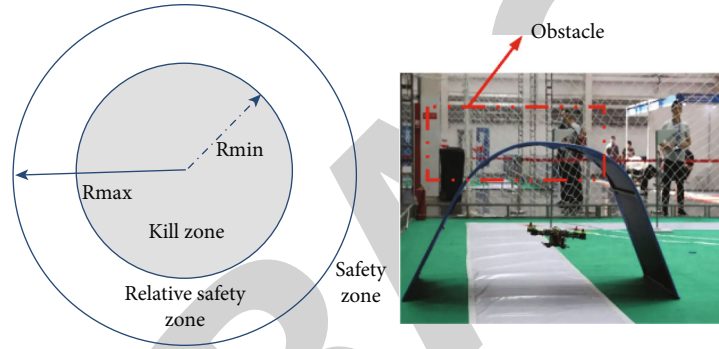


FIGURE 2: Obstacle signal threat model.

obstacles encountered by UAV in the process of mission execution, this paper analyzes and designs several types of threat sources. The first is the same UAV obstacle, which can also be called signal obstacle threat. The relevant formula for obtaining the signal of this type of obstacle is as follows:

$$U = \begin{cases} 1 & \\ \frac{|R_{\min} - r|}{e^{(R_{\max} - R_{\min})}} & \\ 0 & \end{cases} \quad (4)$$

According to the above formula, the signal received by the UAV can be calculated, and then, the threat judgment of the transmitted signal can be carried out. The obstacle signal threat model is shown in Figure 2.

It can be seen from Figure 2 that the UAV analyzes and judges the threat area of the signal sent by the obstacle, the relative safety area, and the safety area and then makes the flight route change action. In the process of UAV entering the working state, it often meets the threat of terrain obstacles as well as signal obstacles. Because in the process of carrying out tasks, it is impossible for unmanned aerial vehicles to have a smooth flow. Terrain factors should always be considered. If you pay little attention, drones will fall, causing

great damage to drones. The calculation formula of obstacle terrain model is as follows:

$$U = \begin{cases} U = \sum_{j=1}^m U_j & \\ U_j = \text{peak}_j \times \frac{1}{e^{\left(\frac{x-h_{xj}}{h_{rxj}}\right)^2 + \left(\frac{y-h_{yj}}{h_{ryj}}\right)^2}} & \\ \frac{4(r - R_{\min}) * (R_{\max} - r)}{(R_{\max} + R_{\min})^2}, R_{\min} \leq r \leq R_{\max} & \\ 0 & \end{cases} \quad (5)$$

From the above formula, the sum of the number of threats of the obstacle terrain model and the minimum and maximum distance from the terrain obstacle can be calculated. Only by obtaining the value of this sum and the distance length can the drone ensure its own flight safety. The collision rate and mission success rate of the above two different obstacles transmitted to the computer during the operation of the UAV are shown in Figure 3.

It can be seen from Figure 3 that although there is no collision during the execution of the task, the probability of the UAV successfully completing the task is gradually

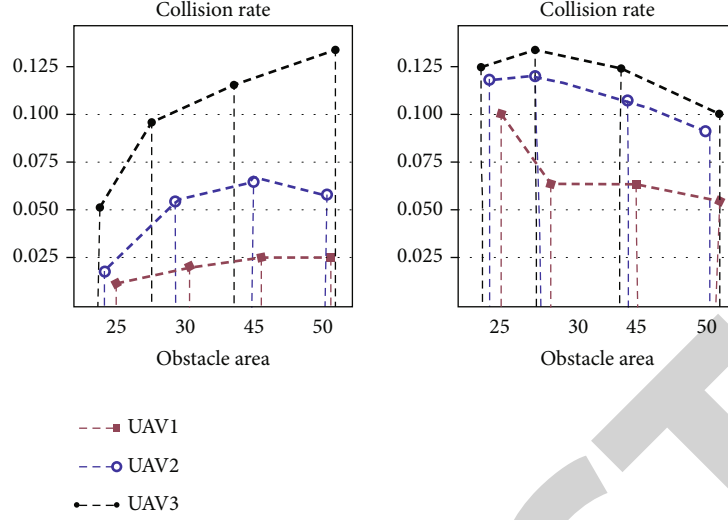


FIGURE 3: Working collision rate and mission success rate of UAV.

decreasing with the increase of obstacles. This also shows that the threat of obstacles to UAVs is an inevitable factor.

In the process of task allocation when UAV is working, the principle of task allocation has always been that task efficiency is the first, and cost is the least. In the problem of task allocation of UAV, this paper simply modeled the task allocation, and the relevant formula is as follows

$$\begin{aligned}
 A_{ij} &= \begin{cases} 1, & \text{Implementation} \\ 0, & \text{Non-implementation} \end{cases}, \\
 C &= \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix} \text{ and } D = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \cdots & \cdots & \cdots \\ d_{m1} & \cdots & d_{mn} \end{bmatrix}, \\
 &\left\{ \begin{array}{l} \text{Benefit}(A) = \lim_{m \rightarrow \infty} \sum_{i=1}^m \sum_{j=1}^m A_{ij} b_{ij} \\ \text{Cost}(A) = \lim_{m \rightarrow \infty} \sum_{i=1}^m \sum_{j=1}^m A_{ij} c_{ij} \end{array} \right\}, \\
 &\left\{ \begin{array}{l} \text{Max}(\text{Benefit}(A)) \\ \text{Min}(\text{Cost}(A)) \end{array} \right\}.
 \end{aligned} \quad (6)$$

According to the above formula, the UAV can process the cost of the received tasks and then process the task allocation time and sequencing. In the process of task processing, the task time allocation data transmitted to the computer is shown in Figure 4.

It can be seen from Figure 4 that the UAV is flexible in solving tasks in the face of different task assignments. It is also more intuitive to clarify the allocation of UAV tasks at different times.

3.2. Path Planning and Optimization of UAV Based on Reinforcement Learning in Dynamic Environment. The above content describes the relevant research on reinforce-

ment learning of UAV tasks and existing threats in a dynamic environment. Next, the path planning and optimization performance of UAV in dynamic environment are studied in detail. The path planning problem of UAV mainly refers to the optimal flight path of UAV. The flight route selected by UAV must ensure the least obstacles and meet the mission requirements. Minimize the risk of UAV as much as possible and complete the tasks delivered at the least cost. In the process of path planning, the new UAV first judges whether the path is reasonable and whether the algorithm is complete and then decides the next specific operation. The second is to screen the path and analyze the performance parameters of the UAV itself. Thirdly, the route planning should be full of security, and the path planning must be hidden. Finally, in the process of executing the task, the algorithm inside the system can quickly respond to the task and make modifications and adjustments to the existing problems in time. Among them, the cost function of UAV flight altitude is as follows:

$$f_H = \begin{cases} \frac{1}{\sum_{k=1}^n H_{\max} - h_k}, & h_k \leq H_{\max} \\ \infty, & h_k > H_{\max} \end{cases}. \quad (7)$$

According to the above function, we can get the flight altitude of the UAV, better give the flight path, and realize the feature of concealment. In the process of UAV path planning, a variety of path planning algorithms can be added to the whole UAV system.

The control method of UAV formation transformation is the premise of realizing formation flight of multiple UAVs. The formation reconstruction of clustered UAVs is an important problem that we need to consider so that each UAV can reach the final position from the initial position without collision, thus ensuring the minimum cost or optimal energy consumption in the formation reconstruction process. The target allocation problem is solved by Hungary algorithm at most. It is the most common algorithm for

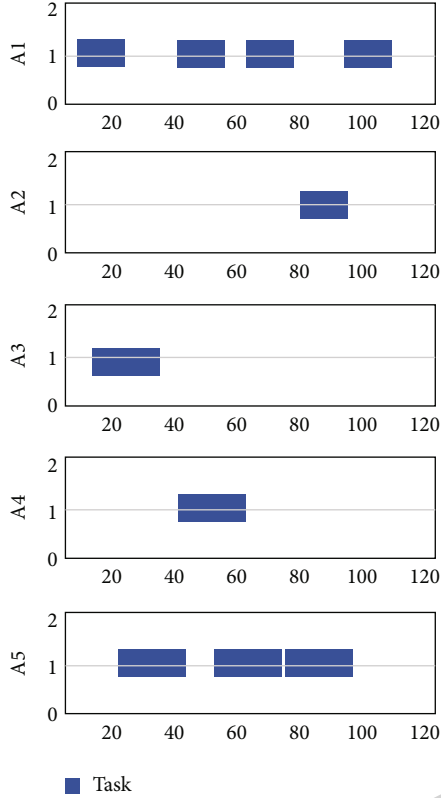


FIGURE 4: Instantaneous assignment data diagram of UAV mission.

partial graph matching. The core of this algorithm is to find the augmented path. It is an algorithm for finding the maximum matching of bipartite graph by using the augmented path. Problems with low calculation difficulty, short planning time, and high planning efficiency can meet the actual needs. Hungarian algorithm is a combinatorial optimization algorithm that solves the task allocation problem in polynomial time and promotes the later primal dual method.

The problems suitable for dynamic programming must have the following characteristics. (1) Optimal substructure: if the optimal solution of the parent problem contains the optimal solution of its subproblem, we say that the problem has an optimal substructure. That is to say, when the subproblem is optimal, the parent problem must be able to obtain the optimal solution through optimization. (2) Overlapping subproblems are essentially the same as the parent problem, except that the input parameters of the problem are different, which can be called overlapping subproblems, which is the essence of the efficiency of dynamic programming in solving problems. (3) The problem has a boundary. The subproblem does not exist under certain circumstances. We call this situation one in which the problem has a boundary. For the top-up and bottom-down methods, the boundary is the exit and entrance of the problem, respectively. (4) The subproblems are independent of each other. The subproblems are independent of each other when solving the optimal solution, that is, the solution of the self-problem is irrelevant to other parallel subproblems. Since there are few applications at present, it will not be intro-

duced in detail. Its basic idea is the same as that of dynamic programming. It also adopts the method of strategy estimation, strategy improvement, and strategy iteration to obtain the optimal strategy. However, in policy estimation, it takes the value function record of the first access to the state in a cycle. After countless rounds, the strategy is estimated by approaching the real value. It has three main characteristics: (1) the algorithm can obtain new decision experience from the past decision experience without modeling the world where the agent is located; (2) the estimation of the state value function by the algorithm is independent of each other; (3) the algorithm can only deal with the problem of episode task mode.

The two algorithms added in this paper are Hungarian algorithm and artificial potential field method, which are further optimized. The Hungarian algorithm mainly improves the communication transmission speed of the UAV internal system, while the artificial potential field method improves the calculation speed of the whole UAV internal system. Only when the speed of communication and calculation data is accelerated, the UAV can better plan the flight path. The specific process of implementing the Hungarian algorithm in the UAV is shown in Figure 5.

Figure 5 shows the internal implementation process of the Hungarian algorithm. First initialize the data, then sort the data and calculate the weight so that the idle rows are filled with data for sorting. Finally, the UAV selects the optimal flight path through the specific value calculated internally. Within the Hungarian algorithm, this paper mainly integrates the data of each gradient. The traditional Hungarian data are calculated separately, which not only wastes a lot of time but also does not improve the accuracy of calculation. The calculation formula of each gradient after integration is as follows:

$$\lim_{n \rightarrow \infty} \left(\sum_{j=1}^n \frac{(E-1) \cdot b_j}{E^j} + \frac{b_n}{E^n} \right). \quad (8)$$

Using Hungarian algorithm to integrate and process the data can greatly shorten the time of data allocation and calculation. In the implementation of Hungarian algorithm within the UAV system, it is also necessary to calculate the relevant weights. This paper also integrates the weights, and the relevant formula of the comprehensive weight of relevant statistics is as follows:

$$SC_i = \lim_{n \rightarrow \infty} \left(\sum_{j=1}^n \frac{(E-1) \cdot Cb_j}{E^j} + \frac{cb_n}{E^n} \right). \quad (9)$$

From the above formula, the calculated weight can continue the overall operation of the data. According to the above Hungarian algorithm optimized for the UAV internal system, the UAV has achieved good communication transmission performance in the practical application in the dynamic environment. In order to more intuitively see the specific situation of path planning of UAV in practical application, the flight trajectory data of UAV is transmitted to the

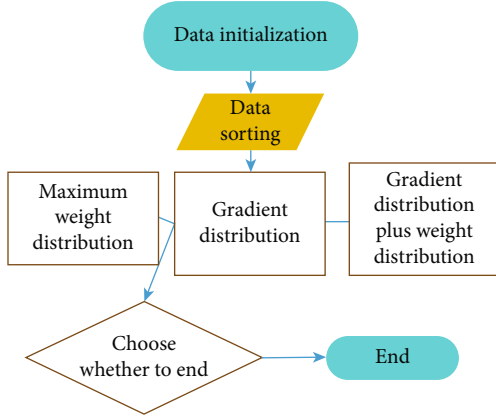


FIGURE 5: Specific process of implementing Hungarian algorithm in UAV.

computer, and the flight data trajectory of UAV is obtained, as shown in Figure 6.

It can be seen from Figure 6 that the data trajectory of the UAV flying without obstacles and the route selected by the UAV are the fastest routes to complete the task. After setting obstacles in the flight process, the UAV also automatically analyzes the data of obstacles and finally takes the most appropriate flight route for planning. From the feedback data trajectory diagram, the optimized and integrated Hungarian algorithm added to the UAV system can indeed improve the system performance of the UAV and greatly improve the communication and transmission ability of the UAV to data. In order to see the overall change of the UAV after adding the optimization algorithm in many aspects and angles, the data comparison of the total energy consumption generated by the UAV during the task execution is also carried out, as shown in Figure 7.

It can be seen from Figure 7 that the UAV without the optimization algorithm has a high energy consumption at the beginning of the task, and the energy consumption required has also reached a very high value due to the growth of working hours. The UAV with the optimized algorithm has halved the overall energy consumption compared with the UAV without the optimization algorithm during the flight mission, which can also enable the UAV to better complete the assigned tasks in a limited time.

The Hungarian algorithm and its optimization are described in detail in the above content, and the artificial potential field method is understood and optimized below. Artificial potential field is the most widely used algorithm for unmanned aerial vehicles, because the mathematical principle of artificial potential field method and its simple and easy to understand characteristics make the artificial potential field algorithm possible to change. However, the artificial potential field algorithm also has defects. The algorithm is weak in self-regulation, which is easy to minimize the processing of data information, so that the data information obtained by the UAV is wrong, and there is a misjudgment on the path selection and planning, which will

eventually cause serious losses. In order to optimize the algorithm, the internal formula for calculating the safety distance is improved. The formula is as follows:

$$F_{\text{rep}} \left(X_r \times \left(\frac{1}{(d(X_r, X_0))^2} - \frac{1}{(d_m - d_0)^2} \right) \right) d < d_m. \quad (10)$$

From the above formula, the safe flight distance of UAV can be accurately calculated, and then, the reasonable flight path can be analyzed and planned in the UAV system. In addition to calculating the safe distance, it is also necessary to add a cost function inside the algorithm so that the overall artificial potential field algorithm can give full play to the maximum performance of UAV path planning. The relevant formula is as follows:

$$J = \lim_{n \rightarrow \infty} \sum_{i,j=1}^n (w_k f_{ij} + w_d D_{ij}). \quad (11)$$

After using the optimized artificial potential field algorithm for the UAV, the actual task simulation is carried out, and the feedback path planning data is shown in Figure 8.

It can be seen from Figure 8 that the UAV is planning the route for different obstacles. By adding the optimized artificial potential field algorithm, the UAV can plan and design the path more quickly. The safety of UAV is ensured to a greater extent.

4. Results and Effect Analysis

4.1. Assignment of UAV Tasks and Threats in a Dynamic Environment. In this paper, a three-dimensional point cloud map of the environment is established by visual slam. Then, a two-dimensional mesh map is established by the three-dimensional point cloud of feature points proposed by SLAM algorithm. The height of each grid is calculated by projecting the map points of the graph into the corresponding grid. Then, an image segmentation algorithm based on mean shift is used to smooth the height of the mesh map, divide the obstacles and the ground, and combine the image blocks with similar height. The algorithm calculates the spatial distance between the landing area and the obstacle and selects the area farthest from the obstacle as the filtered landing area. In this way, the area suitable for UAV landing is selected. The UAV finally lands in a safe area according to the descent procedure.

This article further validates the ability of the new UAV to detect obstacles during mission completion. According to research on task and threat allocation of reinforcement learning UAVs in dynamic environments, the same tasks are assigned to UAVs. Three different mission environments were selected to simulate the UAV mission. In order to ensure the accuracy of the experiment, we repeat the operation for many times and finally take the average value to evaluate the experimental results. On the way of executing the mission, the UAV first receives the mission, assigns and processes the mission, and drives according to the

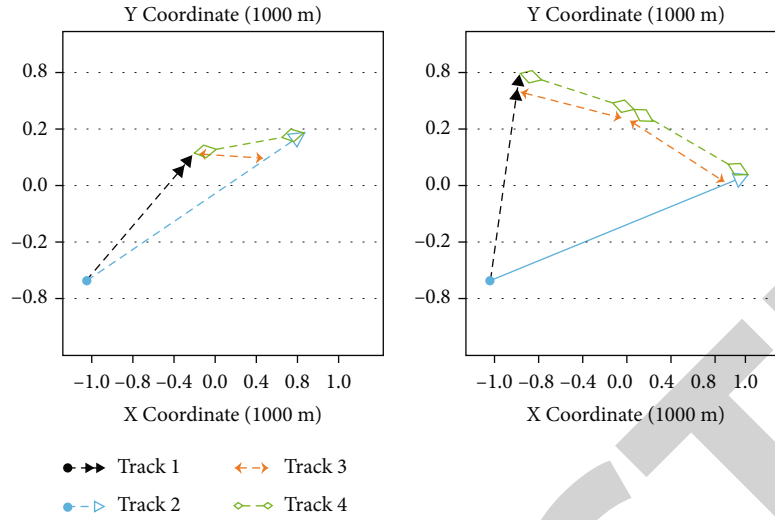


FIGURE 6: UAV flight data trajectory.

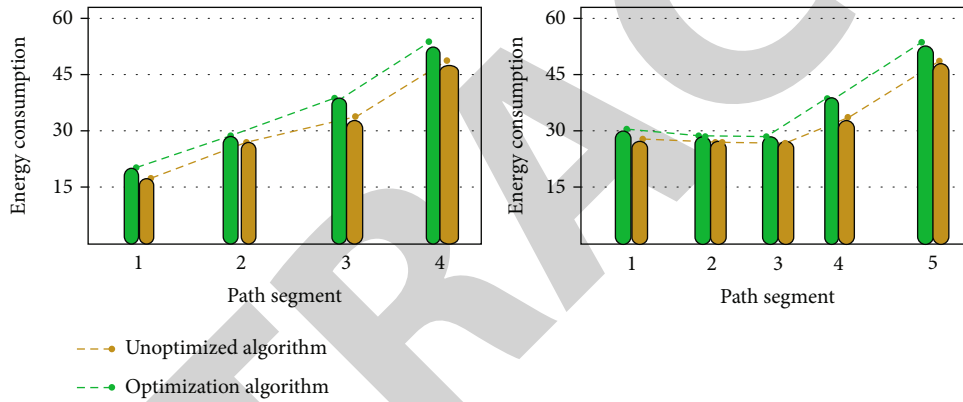


FIGURE 7: Comparison diagram of flight energy consumption during UAV operation.

selected flight path. Then, in the process of UAV flight, it monitors the possible obstacles, starts to judge whether the obstacle target is a threat to itself, and further makes action feedback. The purpose of this is that UAVs can save energy consumption. If obstacles are avoided, UAVs will often complete tasks significantly. After no one analyzes the target obstacles, it is necessary to focus on completing the tasks received within the system, analyze and process the tasks, and then complete the assigned tasks one by one. In the whole experiment, we mainly focus on whether the UAV can accurately perceive the existence of target obstacles under the simulated working state. The final experimental results are transmitted to the data formed by the computer, as shown in Figure 9.

It can be seen from Figure 9 that the data of obstacles fed back by UAV is different when it processes the same task in three different working environments. In the first environment, due to less obstacle model settings, the UAV has less changes to modify the original planned route when performing tasks, and the completion time is faster. In the third environment, although there are many target obstacle models, UAVs can accurately detect the existence of obstacles. To

sum up, the process of task and threat allocation of reinforcement learning UAV in dynamic environment studied in this paper is more suitable for practical application and has better detection performance.

4.2. UAV Path Planning and Optimization in Dynamic Environment. In the research of path planning and optimization of UAV based on reinforcement learning in dynamic environment, the problem of task path planning in UAV system is addressed. Firstly, a correlation function is added to the flight altitude of the UAV to make the UAV more invisible. Then, an optimized Hungarian algorithm is added to the UAV system to make the data communication flow faster within the system. Finally, an improved artificial potential field algorithm is added to the UAV system, which also accelerates the ability of calculating data in the system. In order to further verify the research results and practical application effects of UAV path planning and optimization, three kinds of UAVs—ordinary UAV, nonoptimized UAV, and optimized UAV—are compared for system efficiency. Considering the accuracy of the experimental results, the three UAVs are sent to the same task to test the system

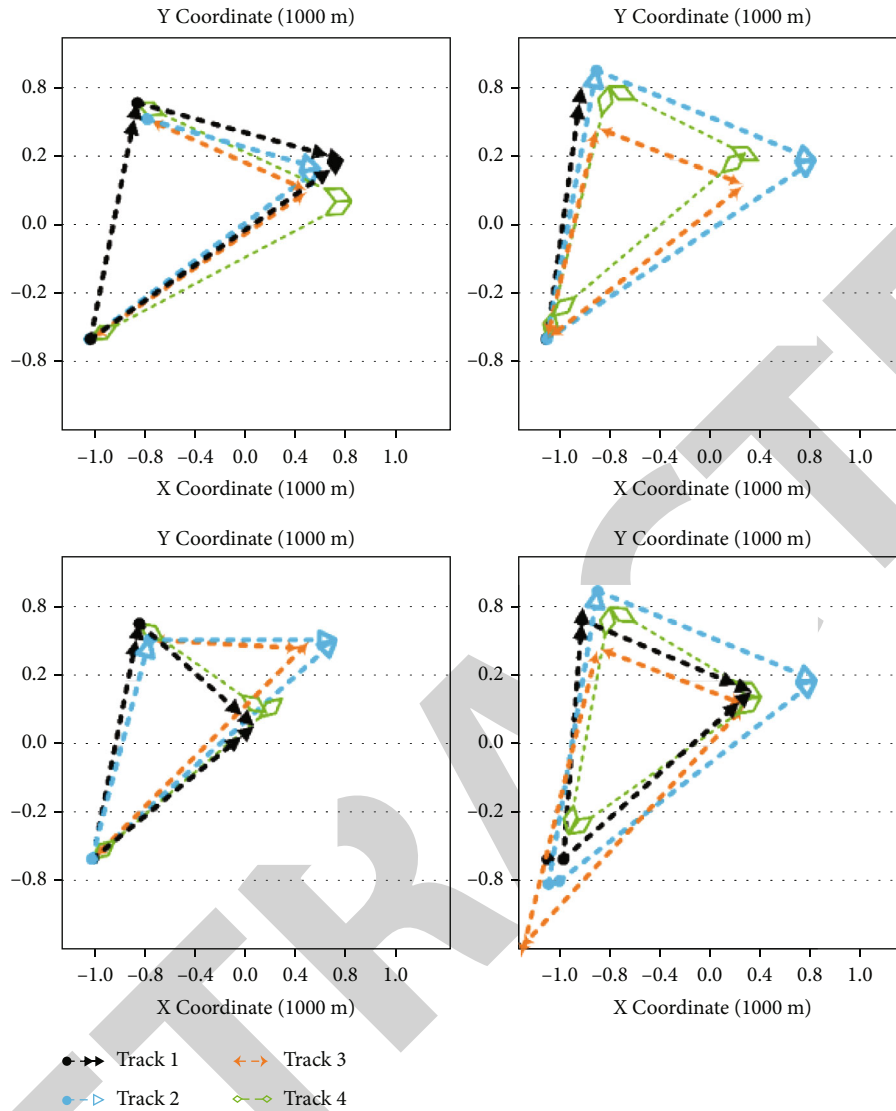


FIGURE 8: Path planning data diagram of actual task simulation feedback.

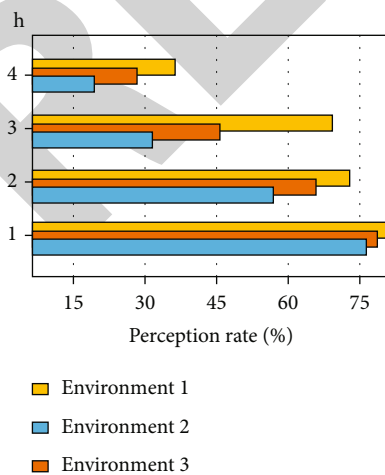


FIGURE 9: UAV perceived target obstacle ratio diagram.

performance under the same environmental state. However, due to the relatively backward system performance of ordinary UAV, less environmental obstacles are selected in this experiment, which is mainly based on the fluctuation of data processing wavelength in the system. The fluctuation amplitude of the internal data processing and calculation of the UAV system generated during the experimental test in this paper is shown in Figure 10.

It can be seen from Figure 10 that the results of the research on reinforcement learning UAV path planning and optimization in a dynamic environment are the data processing status of three types of UAVs. The optimized UAV has the fastest speed in the process of data transmission and calculation, while the data transmission and calculation express of the other two types of UAVs are slow. In particular, ordinary UAVs also have data transmission failures, resulting in incomplete experimental results. The results show that the optimized UAV system is better at data calculation and transmission, which also shows that the

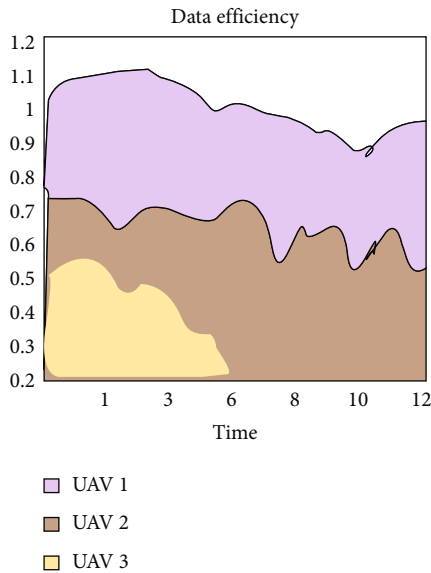


FIGURE 10: Fluctuation range of data processing and calculation in UAV system.

UAV can quickly plan and process the flight path, which is suitable for practical application.

This paper introduces the background of using reinforcement learning algorithm and points out that reinforcement learning algorithm is suitable for grid modeling. At the same time, the main parameters to measure the scene are given, such as the definition of state and grid coordinates. All of these lay a theoretical foundation for the introduction of subsequent algorithms, further verify the detection ability of the new UAV to detect obstacles in the process of completing tasks, and study the assignment of UAV tasks and threats based on reinforcement learning in a dynamic environment. Assigning the same task to the UAV and selecting three different task environments for the UAV task simulation are in order to ensure the accuracy of the experiment.

5. Conclusion

This paper studies the problem of UAV mission path planning based on reinforcement learning in dynamic environment and further verifies the detection ability of the new UAV to detect obstacles in the process of completing tasks. According to reinforcement learning in dynamic environment, the task and threat assignment of UAV are studied. The optimized Hungarian algorithm is added to the UAV system to make the data communication flow within the system faster. Finally, the improved artificial potential field algorithm is added to the UAV system, which also accelerates the ability of computing data in the system. The system efficiency of three kinds of UAVs—ordinary UAV, nonoptimized UAV, and optimized UAV—is compared. The optimized UAV has the fastest data transmission and calculation speed, while the other two types of UAVs have slower data transmission and calculation speed. In particular, ordinary UAVs also have data transmission failures, resulting in incomplete experimental results. The results

show that the optimized UAV system is better in data calculation and transmission, which also shows that the UAV can quickly plan and process flight paths, which is suitable for practical applications. However, there are still many shortcomings. For example, in a dynamic environment, the UAV encounters too many obstacles at the same time. The data collected by the UAV system will be mixed together, resulting in disorder of the internal system. Solving this situation is still a big problem. Further analysis is needed in future research and analysis.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by a fund project for basic scientific research expenses of central universities (nos. J2022-024 and J2022-07) and an independent research project of Key Laboratory of Flight Techniques and Flight Safety (no. FZ2021ZZ04).

References

- [1] L. Zhang and N. Ansari, "Optimizing the operation cost for UAV-aided mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 6085–6093, 2021.
- [2] Y. Choi, Y. Choi, S. Briceno, and D. N. Mavris, "Energy-constrained multi-UAV coverage path planning for an aerial imagery mission using column generation," *Journal of Intelligent & Robotic Systems*, vol. 97, no. 1, pp. 125–139, 2020.
- [3] B. López, J. Muñoz, F. Quevedo, C. A. Monje, S. Garrido, and L. E. Moreno, "Path planning and collision risk management strategy for multi-UAV systems in 3D environments," *Sensors*, vol. 21, no. 13, p. 4414, 2021.
- [4] T. M. Cabreira, C. Di Franco, P. R. Ferreira, and G. C. Buttazzo, "Energy-aware spiral coverage path planning for UAV photogrammetric applications," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3662–3668, 2018.
- [5] F. Causa, G. Fasano, and M. Grassi, "Multi-UAV path planning for autonomous missions in mixed GNSS coverage scenarios," *Sensors*, vol. 18, no. 12, p. 4188, 2018.
- [6] L. Huo, J. Zhu, G. Wu, and Z. Li, "A novel simulated annealing based strategy for balanced UAV task assignment and path planning," *Sensors*, vol. 20, no. 17, p. 4769, 2020.
- [7] J. Kok, L. F. Gonzalez, and N. Kelson, "FPGA implementation of an evolutionary algorithm for autonomous unmanned aerial vehicle on-board path planning," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 2, pp. 272–281, 2013.
- [8] M. Popović, T. Vidal-Calleja, G. Hitz et al., "An informative path planning framework for UAV-based terrain monitoring," *Autonomous Robots*, vol. 44, no. 6, pp. 889–911, 2020.

Retraction

Retracted: Data Analysis of Ceramic Material Color Matching Collection Points in Furniture Design Based on the Image Difference Prediction Model

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Zhou and Y. Li, "Data Analysis of Ceramic Material Color Matching Collection Points in Furniture Design Based on the Image Difference Prediction Model," *Journal of Function Spaces*, vol. 2023, Article ID 4684904, 10 pages, 2023.

Research Article

Data Analysis of Ceramic Material Color Matching Collection Points in Furniture Design Based on the Image Difference Prediction Model

Cheng Zhou ^{1,2} and Yalan Li ^{3,4}

¹School of Art and Design, Jingdezhen Ceramic University, Jiangxi 333403, China

²School of Fine Arts and Design, Xiangnan University, Hunan 423000, China

³School of Physics and Electronic Electrical Engineering, Xiangnan University, Hunan 423000, China

⁴Hunan Engineering Research Center of Advanced Embedded Computing and Intelligent Medical Systems, China

Correspondence should be addressed to Yalan Li; liyalan@xnu.edu.cn

Received 12 August 2022; Revised 28 September 2022; Accepted 12 October 2022; Published 28 March 2023

Academic Editor: Miaocho Chen

Copyright © 2023 Cheng Zhou and Yalan 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.

The manufacturing process of ceramic materials and exquisite patterns is not simply artistic acts. At the same time, it condensed the excellent traditional culture of ancient China. The application of ceramic materials in modern furniture can not only enhance the taste of living room culture but also show the charm of China's ceramic culture. *Objective.* To optimize the scientific data model required by traditional ceramic materials. On the basis of image difference prediction calculation, the data of ceramic material color matching collection points were analyzed and processed, and the image difference prediction model was established. This paper introduces the application and development of ceramic materials in furniture industry. On the basis of the traditional image difference prediction algorithm, the consideration of image difference and difference prediction expansion is added, and the image difference prediction calculation model of ceramic material color matching is established. The survey results show that 83% of customers are very satisfied with the color matching of new ceramic materials calculated by image difference prediction. This experiment uses image difference prediction model to remove the redundant pixels of ceramic material color and make multiple judgments on color display. According to the initial color data of the collection points, this experiment finally obtained a new combination color matching of ceramic wardrobe, which is more suitable for the "minimalist" and "quiet" styles.

1. Introduction

Thousands of years of human history is essentially the history of human use and invention of materials. Materials are the most basic material needs for human survival and the cornerstone of human civilization [1]. Cohen, a British material scientist, wrote in the preface of fundamentals of material science and engineering that "we are surrounded by materials. They not only exist in our real life in the physical form of entities, but also have been deeply rooted in our culture and ideological field." This further shows the importance of materials to human beings. Every scientific and technological leap of mankind is inseparable from the

understanding and application of materials [2]. At the same time, with the development of science and technology, mankind has developed a variety of new materials, indicating that the application of materials has entered a higher stage from the backward era [3]. The application of these new materials in various fields has fundamentally changed the traditional way of material selection and product design and has had an incredible impact in the fields of industry and art [4]. Because of this, new materials are not limited to the field of industrial manufacturing but are more and more closely related to product modeling and design. Every material that can be mass produced will greatly change and promote the change of product structure, thus affecting the

product modeling design, and may eventually develop a new design style [5]. However, the application of new materials is a new challenge for product designers. Designers need to fully understand the physical and chemical parameters of materials, meet the design requirements and the emotional needs of users, and truly reflect the humanistic care required by products. Through the collected literature on the application of ceramic furniture, we found the following shortcomings in the theoretical research on the application of ceramic furniture. The development speed of the combination of ceramic materials and modern home decoration is relatively slow. This is due to the lack of relevant theoretical research and theoretical works. Relevant scholars should pave the way for the latecomers and do more theoretical research with practical significance. To lay a certain foundation, because of the educational background and aesthetic perspective, different consumer groups have different attitudes towards ceramic furniture, and the deviation between them is large, which is difficult to balance.

In a broad sense, materials are defined as all substances except human ideological activities. The research object materials of this experiment refer to the collection of substances of articles owned and used by human beings in a narrow sense [6]. In the scope of industrial production and design, materials are usually divided into traditional materials and new materials. Traditional materials usually refer to substances that have been used for a long time and have been used in most fields of human activities, mainly including metal, wood, ceramics, glass, and plastics [7]. There are two kinds of new materials. The first one refers to special materials that are emerging and in the early stage of development and have one or more physical and chemical properties better than traditional materials [8]. The other refers to materials that greatly improve traditional materials or have some excellent properties through new technologies or preparation processes [9]. From the definition, we can know that new materials are the inevitable product of human scientific and technological development, so they have the characteristics of timeliness, adaptability, and excellence. The timeliness refers to that with the preparation and application of new materials; they are finally extended to all aspects of human life and finally gradually become traditional materials. Today's plastics have evolved from new materials to traditional materials through this path [10].

In the context of people's life, furniture is conventionally defined as an appliance used in human daily life and social activities, which has the functions of sitting, lying, chair, storage, spacing, and so on. It is generally assembled by several parts and components according to a certain joint mode [11]. With the application of new materials, this definition is also changing dynamically. The definition of furniture is becoming larger and larger, including more and more types. For example, with the invention of cars, aircraft, and ships, the seats and other supporting facilities of these vehicles also belong to the field of furniture [12]. With the improvement of urban infrastructure and the rise of the concept of outdoor furniture, such as park seats, cement stone platforms and decorative ornaments have been included in the category of furniture [13]. Only by embracing and storing, con-

stantly updating ideas and adapting to the emergence and use of new materials, can it be possible to design furniture products that do not fall behind the times.

Modern furniture materials are not limited to the use of wood, metal, and plastic. In addition to the use of new materials, the addition of traditional materials and ceramics has also evolved the furniture design style [14]. Moreover, with the aid of image difference prediction, ceramic materials have also got rid of people's impression of its inherent white and cyan. More colorful ceramics have been added and applied to furniture design. This new visual design will deeply affect people's aesthetics of furniture.

2. Introduction and Application of Ceramic Materials

2.1. Overview of Ceramic Materials. Ceramic material is a kind of inorganic nonmetallic material made by forming and sintering [15, 16]. As an artificial synthetic material first manufactured and used by primitive humans, it has also made two major scientific and technological leaps along with human development [17]. With the control of fire and full understanding of combustion, ceramic materials have evolved from pottery with lower sintering temperature to porcelain with higher sintering temperature and denser structure, realizing the first leap in the development history of ceramic materials. The second leap took place in modern times. The breakthroughs in human physics and chemistry have made great breakthroughs in ceramic materials in light, sound, electricity, magnetism, and heat. Therefore, the evolution from porcelain to special ceramics is the second leap in the development of ceramic materials. With the two leaps, ceramic materials have developed from the original utensils containing water or food to the fields of decoration, daily necessities, and even military industry. Stronger hardness, richer color glaze, and unique warm temperament of ceramic materials make it also favored by furniture designers [18].

2.2. Application of Ceramic Materials in Furniture. Furniture is not only an object with use value but also a historical and cultural carrier of various countries and nationalities. The application of ceramic materials in furniture adds a unique artistic beauty and endows this cultural carrier with richer connotation. When ceramics no longer only appear in people's daily life as tableware, decorative vases, and other objects as furniture, this transformation reflects and conforms to the pursuit of diversified lifestyles of modern mankind. Moreover, as a pollution-free and completely degradable material, ceramic materials are also suitable for the current human concept of environmental protection. They are a favorable tool for the furniture industry to achieve the goal of sustainable development. As the origin and manufacturing country of porcelain, in the long historical development, ceramic craftsmen and scholars have endowed ceramic materials with profound cultural heritage. For the Chinese people, every piece of porcelain is the embodiment of the breadth and depth of Chinese civilization. Ceramic furniture also appeared in the life of ancient Chinese people [19]. The four-animal foot three-color glazed

pottery cabinet unearthed in the Tang tomb in Xi'an has a beautiful and generous shape. It is not only a daily furniture but also a rare art.

Ceramic itself has the characteristics of high hardness, wear resistance, and high chemical stability, so that it can participate in the manufacture of furniture. However, the problems of fragility, difficult processing after forming, and low yield also limit its application in the field of furniture. Except for a few all porcelain furniture, porcelain in the Ming and Qing Dynasties mostly appeared in the form of patches in furniture design. However, with the progress of modern technology, high-strength ceramics have emerged. This ceramic has the performance of preventing falling and scratching, greatly expanding the application of ceramic materials in furniture [20]. The technical breakthrough of ceramic glaze makes the ceramic get rid of the white and cyan in the traditional impression. The bolder and bright colors combined with the texture of ceramic as warm as jade have brought new use elements to the furniture design industry [21]. At the 2019 Milan International Furniture Exhibition, a complete set of furniture with ceramics as the main body appeared. A large number of colored dots of different sizes are hand painted on the white ceramic background as decoration, which makes the whole set of furniture stable and reveal a different vitality. In recent years, nanotechnology has overcome the shortcomings that ceramic materials have always been difficult to change the shape after firing. New nanoceramics have achieved a great breakthrough in plasticity and greatly strengthened the machinability and flexibility of ceramic materials. If it reaches the standard of industrialized mass production and enters the furniture industry, it will inevitably bring broader space for the development of ceramic furniture.

2.3. Reasons for Ceramic Materials Entering Furniture Industry. This paper analyzes the development process of furniture materials and obtains three reasons for entering the ceramic industry. Firstly, it is because of the continuous progress of materials and furniture manufacturing technology. As mentioned above, the progress of ceramic materials enables them to enter the field of furniture, but this is not enough. Only with the common progress of different materials and the innovation of furniture manufacturing technology can ceramic materials be better and more organically integrated into furniture design. Nowadays, ceramics can not only be used as the main structure of furniture but also beautify and decorate the human indoor and outdoor environment together with other materials. The second reason is the human demand for individuality. With the progress of human science and technology, the vast majority of human demand rises from the low end of Maslow's demand pyramid to the high. The demand for beauty and personality different from most people can give birth to the emergence and development of ceramic furniture. The application of ceramic materials in the hands of different designers has produced wonderful chemical reactions and produced a wealth of furniture categories. Finally, earth shaking changes have taken place in human life style, and the functionality of furniture has been gradually weakened, which is more the

embodiment and display of the taste of the owner of the house. As a scarce and precious material, it is natural for ceramics to be used in the furniture industry. With its changeable texture and pattern of glaze, ceramics exudes unique charm in beautifying residents' life and cultivating cultural spirit.

3. Prediction and Calculation of Image Difference

The technology of prediction and calculation based on image difference is a research hotspot in image processing in recent years. The predictable algorithm based on difference expansion proposed by some scholars has attracted extensive attention. This algorithm can provide a large amount of embedded values and has good plasticity. The algorithm can be adjusted by using integer wavelet coefficients and image prediction error to meet different needs. Based on the traditional image difference prediction method, this study makes full use of the correlation between color components, improves the traditional method, reduces the amount of calculation of the algorithm, and significantly improves the accuracy of prediction calculation, so as to select the color of ceramic materials. This article compares the two images with Python using the Structural Similarity Index (SSIM). Using this method, we can easily determine whether the two images are the same or judge the difference due to the intentional tampering of slight image processing compression. The SSIM method will be extended so that we can use OpenCV and Python to visualize the differences between images. Specifically, we will draw bounding boxes around areas in two different input images.

3.1. Traditional Image Difference Prediction Method. The traditional image difference method can predict the adjacent color pixels. The mean and difference of two color pixels can be expressed by the following formula:

$$l = \left\lfloor \frac{x+y}{2} \right\rfloor, \quad (1)$$

$$h = x - y.$$

x and y in the formula are the gray values of two adjacent color pixels, respectively, and their value range is 0 to 255. l represents the mean value, and h represents the difference value. The integer wavelet inverse transform of formula (1) can be obtained as

$$\begin{cases} x = l + \left\lfloor \frac{h+1}{2} \right\rfloor, y = l - \left\lfloor \frac{h}{2} \right\rfloor & |h| \geq 0, \\ x = l + \left\lfloor \frac{h}{2} \right\rfloor, y = l - \left\lfloor \frac{h-1}{2} \right\rfloor & |h| < 0, \end{cases} \quad (2)$$

where $\lfloor \cdot \rfloor$ is rounding down symbol.

Then, the predicted difference h_w of the embed 1 pixel between two adjacent pixels is obtained by

$$h_w = 2h + w, \quad (3)$$

where $w \in [0, 1]$ and it is the information to be embedded between two adjacent pixels. And h_w is the difference after expansion with h . h_w would be substituted into formula (2) to obtain the new embedded 1 pixel's value, which is the predicted pixel value. In order to ensure that the new predicted pixel value does not overflow, the formula (4) should be satisfied.

$$|h_w| \in [0, \min \{510 - 2l, 2l + 1\}]. \quad (4)$$

3.2. Single Component Image Prediction Error. There is a strong correlation between the color pixels of ceramic materials, which is the basis of color prediction and linear compression. With the increase of the distance between pixels, this correlation is negatively correlated and gradually weakened. Therefore, it can be predicted by adjacent pixels, and its accuracy is higher than that of distant pixels. Therefore, this experiment takes any designated pixel on the color of ceramic materials, determines the surrounding pixels, and then calculates the prediction error.

$$p_w = 2p + w. \quad (5)$$

Suppose a is the value of one pixel in an image and \bar{a} is the average value of the adjacent pixels surrounding a . \bar{a} can be used as the predicted value of a . In formula (5), p is the difference between the pixel value a and the predicted value \bar{a} . It is also called prediction error. Combine the embedded information w and the original prediction error p into formula (5) to obtain the prediction error p_w .

Then, the predicted pixel value could be calculated by

$$a_w = \bar{a} + p_w, \quad (6)$$

where a_w is the new embedded pixel value.

According to the principle that the components of RGB in the hue triangle are independent, the pixels on the color of ceramic materials also have their own independent components. In order to improve the accuracy of image prediction and make use of the strong correlation of adjacent pixels, take the upper, lower, left, and right pixels adjacent to the pixel to be predicted as the reference points, and their average value can be calculated by the temple as

$$\bar{S} = \frac{1}{4} \times \begin{vmatrix} & 1 & \\ 1 & \cdot & 1 \\ & 1 & \end{vmatrix}, \quad (7)$$

where the black dot in the center of the prediction template in the above formula represents the pixel to be predicted. According to equation (7), $\bar{a} = \lfloor a_{\text{upper}} + a_{\text{lower}} + a_{\text{left}} + a_{\text{right}} \rfloor$ can be obtained as the prediction value of the embedded point pixel.

3.3. Single Component Image Prediction Difference Expansion. When predicting the color components of any two pixels in the color of ceramic materials, the mean and difference of errors are as follows:

$$\begin{aligned} l &= \left\lfloor \frac{p_1 + p_2}{2} \right\rfloor, \\ h &= p_1 - p_2. \end{aligned} \quad (8)$$

Referring to the traditional image difference prediction method, the integer wavelet inverse transform of the above formula can be obtained:

$$\begin{cases} p_1 = l + \left\lfloor \frac{h+1}{2} \right\rfloor, p_2 = l - \left\lfloor \frac{h}{2} \right\rfloor & |l \geq 0, h \geq 0, \\ p_1 = l + \left\lfloor \frac{h}{2} \right\rfloor, p_2 = l - \left\lfloor \frac{h-1}{2} \right\rfloor & |l \geq 0, h < 0, \\ p_1 = l + \left\lfloor \frac{h}{2} \right\rfloor, p_2 = l - \left\lfloor \frac{h+1}{2} \right\rfloor & |l < 0, h \geq 0, \\ p_1 = l + \left\lfloor \frac{h-1}{2} \right\rfloor, p_2 = l - \left\lfloor \frac{h}{2} \right\rfloor & |l < 0, h < 0. \end{cases} \quad (9)$$

Then, utilize the original difference h of two adjacent pixels, the embed information w , and the adjacent pixel values p_1 and p_2 to compute the expansion difference h_w of the prediction pixel by

$$h_w = (p_1 + p_2)\sqrt{2h + w}. \quad (10)$$

In order to suppress errors caused by rounding down, the l and h are updated by formula (11) and marked as l' and h' .

$$l' = \left\lfloor \frac{(p_1 + p_2)^2 + (p_1 - p_2)^2}{2} \right\rfloor, \quad (11)$$

$$h' = 2\sqrt{(p_1 + p_2)^2 + (p_1 - p_2)^2}.$$

The corresponding integer wavelet inverse transform is transformed into two equations:

$$\begin{cases} p_1 = l' \pm \left\lfloor \frac{h'+1}{2} \right\rfloor, p_2 = p_2 \pm \left\lfloor \frac{h'}{2} \right\rfloor & |l' \geq 0, h' \geq 0, \\ p_1 = l' \pm \left\lfloor \frac{h'}{2} \right\rfloor, p_2 = p_2 \pm \left\lfloor \frac{h'-1}{2} \right\rfloor & |l' \geq 0, h' < 0, \end{cases} \quad (12)$$

$$\begin{cases} p_1 = l' \pm \left\lfloor \frac{h'}{2} \right\rfloor, p_2 = p_2 \pm \left\lfloor \frac{h'+1}{2} \right\rfloor & |l' < 0, h' \geq 0, \\ p_1 = l' \pm \left\lfloor \frac{h'-1}{2} \right\rfloor, p_2 = p_2 \pm \left\lfloor \frac{h'}{2} \right\rfloor & |l' < 0, h' < 0. \end{cases} \quad (13)$$

The least effective pixel function can be further obtained by calculating

$$w = LSB(a_w). \quad (14)$$



FIGURE 1: Schematic diagram of the closing and opening of the Keramos cabinet door with ceramic furniture wardrobe combination.

According to (6), the combined pixel gray value should be between 0 and 255; the gray value range of pixels to be predicted is also limited:

$$p_w \in [-\bar{a}, -255 - \bar{a}]. \quad (15)$$

In this range, considering the error obtained above, the color component of any pixel to be predicted is calculated, and finally, the ceramic material color image difference prediction model is obtained.

$$\begin{cases} |p_{nw}| = |p_n| - 1 |p_n > T_n, \\ p_{nw} = p_n + 1 | \bar{a} < 128, p_n < T_n, \\ p_{nw} = p_n - 1 | \bar{a} \geq 128, p_n < T_n. \end{cases} \quad (16)$$

4. Case Analysis of Ceramic Furniture

4.1. Color Ceramic Wardrobe Combination Keramos. The limited ceramic wardrobe combination designed and launched by two Italian designers in 2016 is named after Keramos in Greek, which means “pottery.” With its lively and flexible color and flawless luster of ceramic glaze, it was very popular once launched and later won several design awards one after another.

Figure 1 is the schematic diagram of the closing and opening of the Keramos cabinet door of the ceramic furniture wardrobe combination. When you see this group of ceramic wardrobe, three bright colors of red, yellow, and blue first come into sight, leaving an unforgettable impression. After careful observation, it will be found that solid wood is used for the cabinet door, interior, and cabinet legs, and the log color is retained, which is perfectly and harmoniously combined with ceramics without dominating. However, the bright color of ceramic materials usually carries a young and lively label in the design field, and at present, a large number of customers prefer the aesthetic style of “minimalism” and “silence.” Therefore, in order to meet the aesthetic preferences of these customers, this study uses the image difference prediction model established earlier to predict the color of ceramic materials and get the required results.

4.2. Prediction Using Image Difference Model. This paper uses Python as the programming language, calls OpenCV and other libraries, and builds and designs the program system. The difference image sequence is filtered to remove the

noise interference in the image and get more accurate moving targets. After the moving object is extracted and segmented, the Camshift algorithm is used to calculate the candidate window, and the gray histogram distribution feature and area feature of the moving area are extracted for target tracking. Firstly, the ceramic material surface is photographed and sampled, and the image processing model is used to analyze the effective ceramic color pixels and redundant ceramic color pixels. The results are shown in Figure 2.

In Figure 2, the white part is the effective pixel area, which is the pixel block that can be used by the image difference model, while the green part is the redundant pixel area, which needs to be removed during prediction calculation. Redundant pixels include pixels overlapping with effective pixel information and pixels containing useless information. Both pixels will interfere with the calculation of the model and slow down the calculation speed, which must be eliminated.

The space and memory occupied by the extracted ceramic material color matching original pixels and the predicted pixels also need to be calculated and considered. The spatial complexity of the original pixels increases smoothly and approaches a straight line, which shows that the original pixel data extracted in this experiment is stable. As the abscissa pixel value increases, the spatial complexity of the ordinate increases gradually. The pixel data of the color to be predicted shows exponential growth, which shows that the predicted data contains large information, and the higher the spatial complexity, the more accurate the prediction result is. The detailed and intuitive data structure is shown in Figure 3.

Select a circular area with radius r on the surface of ceramic material. If n sampling points are included, $2n$ modes will be generated by using the image difference prediction model. Therefore, with the increase of sampling points, the generated patterns will also double. Too many binary patterns are not conducive to the extraction, recognition, classification, and prediction of color information. Traditional statistical methods usually use histogram to express the information of image pixels, but too many patterns will cause the sparsity of histogram and weaken its expression function. In order to solve this problem and improve the efficiency of statistics, it is necessary to reduce the dimension of data in order to reduce the amount of data and complete the image expression information. In this study, dynamic uniform pattern is used to reduce the dimension of too many binary patterns generated by the image difference

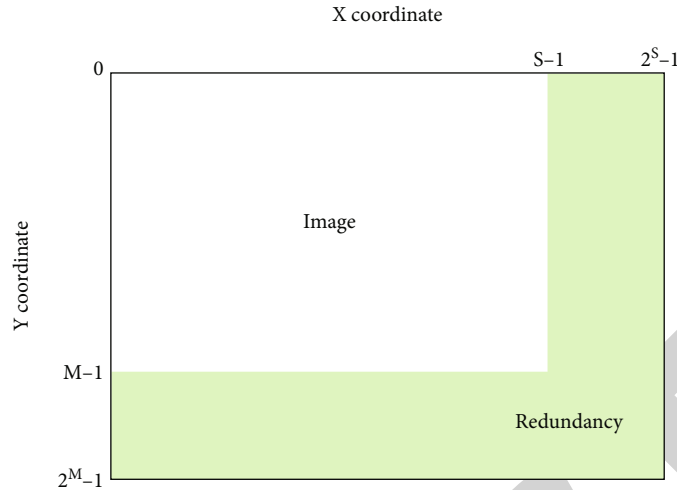


FIGURE 2: Effective ceramic color pixel and redundant ceramic color pixel map obtained by model processing.

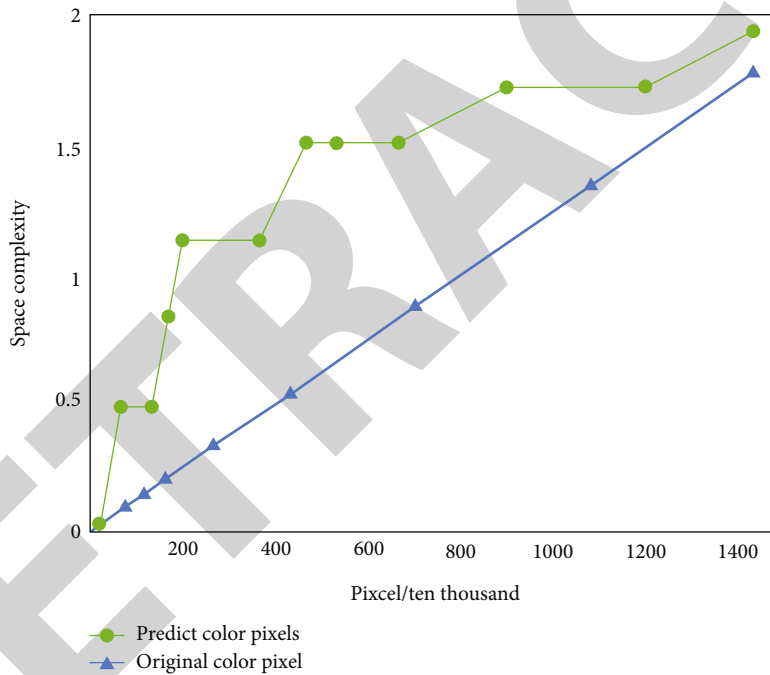


FIGURE 3: The spatial complexity of the original color matching pixels of ceramic materials and the pixels obtained after prediction.

prediction model. The dynamic equivalence pattern holds that when a cyclic binary number jumps from 0 to 1 or from 1 to 0 at most twice, its corresponding binary is an equivalence pattern class. For example, 000000 (0 jumps) and 110011 (2 jumps) are dynamic equivalent mode classes. Through this improvement, the types of binary modes are greatly reduced. Therefore, in this study, 8 sampling points are taken from a 3 by 3 circular area, as shown in Figure 4. The dimension reduction of binary mode can reduce the final eigenvectors greatly by using dynamic equivalent mode. Noise is an important cause of image interference. An image may have a variety of noises in practical applications. These noises may be generated in transmission, or in quantization and other processing. Although these model-based methods

have strong mathematical derivation, the performance of texture structure restoration under heavy noise will be significantly reduced. In addition, because of the high complexity of iterative optimization, they are usually time-consuming. After removing the interference of high-frequency noise, the accuracy of experimental results will be improved.

This HSV model is widely used to predict the color matching required by human vision. It represents the color through the brightness, hue, and saturation of the color. Its structural model is shown in Figure 5.

It can be clearly seen from the figure that the HSV color space structure is an equilateral hexagonal pyramid; black is located at the vertex of the equilateral hexagonal pyramid,

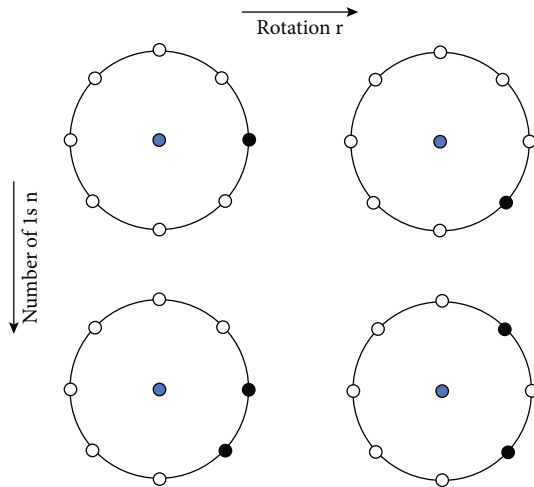


FIGURE 4: Four modes of color pixel collection points of ceramic materials after dimensionality reduction using dynamic equivalent mode.

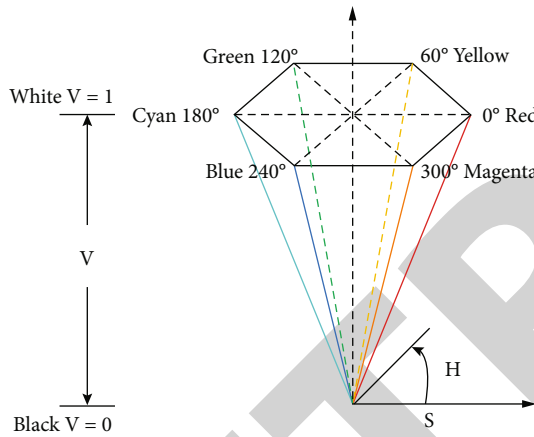


FIGURE 5: Hue, saturation, and value color space model structure diagram.

because the brightness of this point is 0, and white is located at the center of the bottom surface of the equilateral hexagonal pyramid, because the brightness of this point is 1, which is the point with the highest brightness. As indicated by the central axis in Figure 6(a), V represents the brightness of the color, and the color on the central axis represents all the colors of this gradient from black to gray to white. H is the hue of the color. The six edges of the equilateral hexagonal pyramid represent the six main colors of the model, respectively, of which 0° is red, 60° is yellow, 120° is green, 180° is cyan, 240° is blue, and 300° is magenta. These edges are arranged equidistantly around the central axis and contain all colors visible to human vision from 0° to 360° . S refers to the saturation of color. This parameter is used to characterize the purity of color. The point in the figure represents the color. The closer the point is to the central axis, the lower the saturation of the color. The brightness, hue, and saturation represented by S and V are three independent color parameters. One of them can be adjusted arbitrarily without affecting the other two. Therefore, the color prediction of

ceramic materials can be selected in a larger range without considering the interaction between various color parameters, which effectively improves the efficiency of model calculation.

The color display of ceramic materials follows the general rules of color. The lower the saturation of the color, the lower the purity of the color. When the saturation is low, it is more gray in human visual perception. When the brightness is close to 0 in Figure 6, it is also shown when the brightness is near 0 in Figure 6. When the saturation is still 0 and the hue is close to 0, the color near the horizontal axis will be perceived from black to gray and then to white with the increase of brightness. For furniture made of ceramics, white glaze and colored glaze are mostly applied on the surface as the base color, and large decorative patterns are decorated with blue and white, multicolored and pastel glaze on the base color. Placed in other wooden furniture, it is eye-catching and bright, playing a very bright decorative effect. It is an important ornament of high-end home environment in the bright and cool period. If the saturation of a color is low, the hue information will become unstable and unreliable. This is because the three components of red, green, and blue are almost equal in low saturation colors. This makes it difficult to determine the exact color; “achromatic” refers to scaling the edge color channel (or subtracting the partially scaled edge channel). From Figure 6(b), we can know to design the S value of the color and take the overlapped part, not the values on both sides, so that the human eye can see the real color. Only when the three are in an appropriate range can human beings perceive the original appearance of this color in the traditional sense. Therefore, in the image difference prediction model, the fuzzy method is used to truncate the saturation. And Figure 6(b) will be obtained. Because of the subjectivity of correction in the perception of color and noncolor by the human eyes, the two areas in the picture overlap.

As shown in Figure 7(a), when the saturation and color values are 0, the reduction interval of brightness is 0 to 1, showing a gradual process from black to gray and then from gray to white. From this display, it can be inferred that the most intuitive difference of noncolor for human perception is the difference of brightness. Therefore, use the fuzzy method to blur the noncolor trapezoid again to get Figure 7(b). Similar to the trapezoidal blur result in Figure 6, there are overlapping areas of black and gray and gray and white due to the human eye itself.

By removing the pixels of the color of ceramic materials, making a number of decisions on the color display, and according to the data collected from the initial color, the experiment finally predicts a new combined color matching of ceramic wardrobe that is more suitable for the “minimalist” and “quiet” style. The results are shown in Figure 8.

It can be seen from Figure 8 that the color saturation of the changed ceramic combined wardrobe is greatly reduced after image difference prediction. Compared with the bright red, yellow, and blue of the original color, the color value orientation of the new color is lower, but the predicted color of ceramic materials still retains flexibility and elegance because of the permeability of ceramic glaze. The predicted

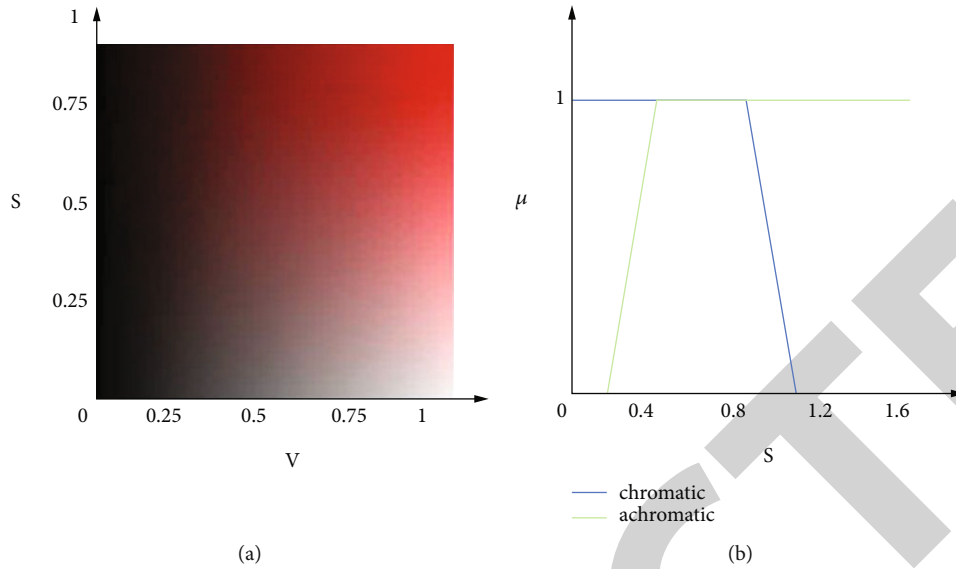


FIGURE 6: Variation of color with S and V at $H = 0$ (a) and S-component fuzzy set partition (b).

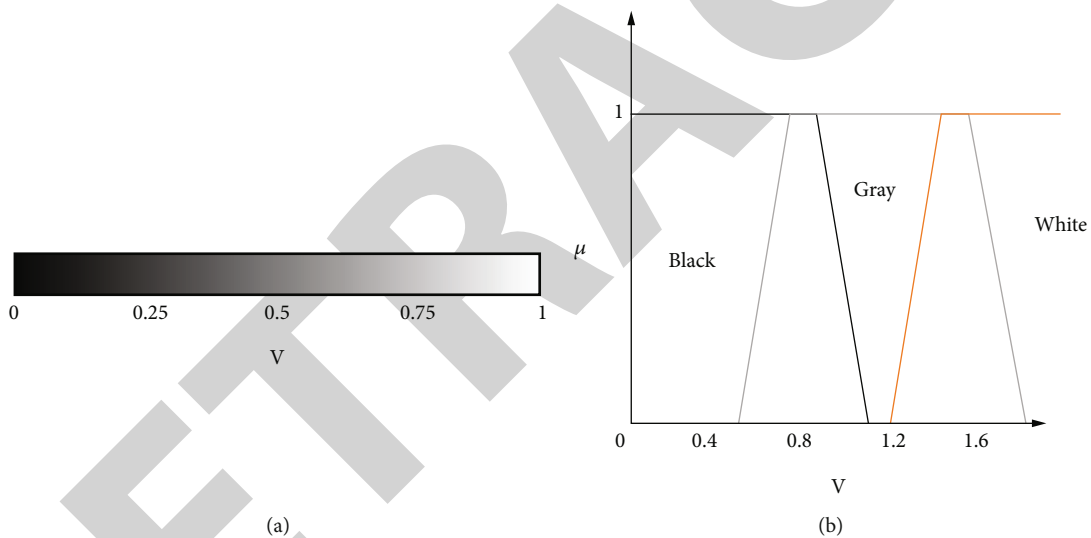


FIGURE 7: Plot of ceramic material color as a function of V (a) and V-component fuzzy set partition (b).

colors are dark green and black cyan with low brightness and ivory white with high brightness. The first two are in line with the popular Morandi color style, and the latter is a classic color, which can be matched with any interior decoration style. The predicted color finally needs the market and customers to test the popularity. Therefore, this study sent a questionnaire on the sales website. The final statistical results are shown in Figure 9.

The survey results show that 83% of furniture customers are very satisfied with the color matching of new ceramic materials calculated by image difference prediction, 16% are basically satisfied with this color, and about 1% of customers express their dislike. From the results, the color matching of new ceramic materials calculated by image difference prediction has achieved the expected effect. Based on the original bright color matching and young customer groups, the color matching of ceramic wardrobe more suit-



FIGURE 8: Keramos schematic diagram of ceramic furniture wardrobe combination after image difference prediction.

able for the current popular style is predicted, which greatly expands the market.

The change and evolution of traditional furniture production are of epochal significance. The application of modern furniture design to ceramic decorative materials is not only a continuous line of traditional Chinese culture but also

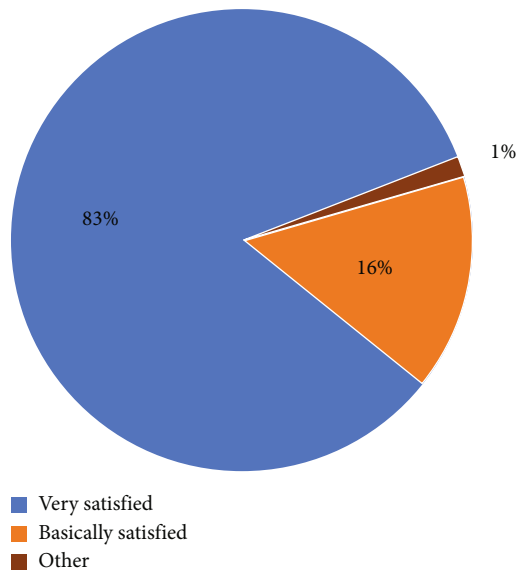


FIGURE 9: The result of the Keramos customer satisfaction questionnaire survey of ceramic furniture wardrobe combination after image difference prediction.

an important driving force for expanding the industry with new elements. From this point of view, we should adjust the artistic style of contemporary furniture design according to the development of national spirit connotation. While summarizing the experience of applying ancient and modern ceramic decoration to furniture design, it can also provide a new unified way of comfort and beauty for home design.

5. Conclusion

This paper describes the combination of ceramic furniture and wardrobe—Keramos. The problem of bright color matching and low age of audience group is solved. The redundant pixels of ceramic material color are removed by image difference prediction model, and multiple judgments are made for color display. The new color matching is predicted according to the data of the collected points on the initial color. After image difference prediction, the color saturation of the changed ceramic combination wardrobe is greatly reduced. Compared with the bright red, yellow, and blue of the primary color, the color value orientation of the new color is lower. However, due to the permeability of ceramic glaze, the predicted color of ceramic materials still maintains flexibility and elegance. By removing the pixels of ceramic material color, some decisions are made on color display. Based on the data collected from the initial color, the experiment finally predicts a new color matching of ceramic wardrobe combination, which is more suitable for the “minimalist” and “quiet” styles. From the results, the color matching of the new ceramic materials calculated by image difference prediction has achieved the expected results. Based on the original bright color matching and young customer base, it is predicted that the color matching of ceramic wardrobe will be more suitable for the current popular style,

thus greatly expanding the market. In the future home decoration, more attention should be paid to the use of ceramic art; the embellishment of ceramic art should be better reflected, so that ceramic art decoration can be better displayed in the indoor space, bringing better and more aesthetic living space to the residents.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the Applied Characteristic Disciplines of Electronic Science and Technology of Xiangnan University (XNXY20221210).

References

- [1] S. G. Lee, J. Y. Seo, J. W. Lee, W. B. Park, K. S. Sohn, and M. Pyo, “Composition-tuned lithium aluminosilicate as a new humidity-sensing ceramic material with high sensitivity,” *Sensors and Actuators B: Chemical*, vol. 339, p. 129928, 2021.
- [2] L. Makarichi, K. A. Techato, and W. Jutidamrongphan, “Material flow analysis as a support tool for multi-criteria analysis in solid waste management decision-making,” *Resources Conservation & Recycling*, vol. 139, pp. 351–365, 2018.
- [3] Z. Niu and K. Cheng, “Investigation on the material removal and surface roughness in ultraprecision machining of Al/B4C/50p metal matrix composites,” *The International Journal of Advanced Manufacturing Technology*, vol. 105, no. 7–8, pp. 2815–2831, 2019.
- [4] A. Sg, A. Sn, B. Kss, and C. Sprc, “Fabrication and experimental study to optimize the recast layer and the material removal in electric discharge machining (edm) of aa6061-b 4 c composite,” *Materials Today: Proceedings*, vol. 19, no. 12, pp. 448–454, 2019.
- [5] M. P. Lebedev, V. N. Tagrov, and E. S. Lukin, “Ceramic composition material with mullite crystals obtained from the clay raw material of Yakutia,” *Materials Science Forum*, vol. 992, pp. 253–258, 2020.
- [6] X. Lian, J. Chen, Y. Zhang, G. Wu, and H. Chen, “Inverted perovskite solar cells based on small molecular hole transport material C8-dioctylbenzothienobenzothiophene,” *Chinese Journal of Chemistry*, vol. 37, no. 12, pp. 1239–1244, 2019.
- [7] Z. Li, G. Dong, R. Guan, and J. Kong, “Electrochemical properties of Co-doped rod-like Li4Ti4.92Co0.08O12 as anode material for lithium-ion batteries,” *Journal of Materials Science: Materials in Electronics*, vol. 30, no. 8, pp. 20588–20595, 2019.
- [8] M. H. R. Khan, S. Hossain, and M. M. Rahman, “Study on the physical and mechanical properties of Al-powder reinforced bioactive glass: ceramic composite material,” *Advances in*

Retraction

Retracted: InSAR Phase Unwrapping Algorithm Based on Deep GAN

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Wang, P. Sun, Z. Li, and L. Tang, "InSAR Phase Unwrapping Algorithm Based on Deep GAN," *Journal of Function Spaces*, vol. 2023, Article ID 9230780, 11 pages, 2023.

Research Article

InSAR Phase Unwrapping Algorithm Based on Deep GAN

Chenxia Wang ¹, Pingli Sun ¹, Zheng Li ¹, and Linlin Tang ²

¹Department of Education and Teaching, Zhengzhou Preschool Education College, Zhengzhou, Henan 450000, China

²School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, Guangdong 518055, China

Correspondence should be addressed to Chenxia Wang; 2003003@zzpec.edu.cn

Received 26 July 2022; Revised 13 September 2022; Accepted 23 September 2022; Published 27 March 2023

Academic Editor: Miaochao Chen

Copyright © 2023 Chenxia 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.

At present, the traditional phase unwrapping algorithm is difficult to balance the accuracy and unwrapping efficiency. The traditional phase unwrapping algorithm is difficult to balance the accuracy and efficiency in the phase unwrapping experiments of simulated and measured topographic interferograms. In this paper, the phase unwrapping technology will be studied under the framework of deep learning theory according to the development trend of InSAR intelligence. A phase unwrapping algorithm based on deep GAN is proposed. The model structure includes dense convolution layer blocks and series structure, so that the network can achieve a better balance between feature extraction and detail preservation of interferogram. This is helpful to improve the phase unwrapping accuracy and training efficiency of the network. The experimental results show that the algorithm has a good expansion effect on the interferogram with high signal-to-noise ratio. The synthesis algorithm makes full use of the advantages of the branch cutting method and the finite element method. The phase reliable region and unreliable region are determined, and the transmission of phase error from the unreliable region to reliable region is effectively avoided. The accuracy of the phase unwrapping results in the reliable region is ensured, and the overall phase unwrapping convergence accuracy is greatly improved.

1. Introduction

Since the 1950s, the theory and technology of synthetic aperture radar remote sensing have been in a state of rapid development. Compared with traditional visible light or infrared remote sensing technology, SAR has all-weather and all-day observation capabilities [1]. It does not rely on sunlight as an illumination source and can penetrate clouds, mist, and dust [2]. Now, it has gradually become an important object. Interferometric synthetic aperture radar is a remote sensing mapping technology that combines synthetic aperture radar and radio astronomy interferometry [3]. In the past few decades, InSAR theory and technology have been continuously developed. InSAR mainly uses the phase information of two or more SAR images to obtain target elevation and has been widely used in geographic information system construction, environmental monitoring, and surface deformation monitoring, seismic activity, volcanic activity, and other fields [4]. Interferogram phase unwrapping is a key step in InSAR elevation measurement technology, which

directly affects the accuracy of InSAR elevation measurement [5].

At present, there are many research achievements in SAR image segmentation processing technology. However, due to the complexity of SAR terrain scene, all kinds of segmentation algorithms have great pertinence and poor universality. This paper summarizes these research results, summarizes the existing segmentation methods, and segments the SAR image according to the characteristics of using prior knowledge. Methods are mainly divided into two categories: data-driven and model-driven. SAR image segmentation based on data-driven directly operates on the current image data. Although it also uses the prior knowledge, it does not depend on the prior knowledge. Model-driven SAR image segmentation is directly based on prior knowledge. In addition, the performance evaluation index of SAR image segmentation algorithm is given. And use Radarsat data to verify the performance of the above segmentation algorithm. Deep learning (DL) is a machine learning method that simulates the neural structure of the

human brain, and it is also a breakthrough technology in the field of computer vision for decades. It can effectively extract rich low-level and high-level feature information from sample data. It has achieved the best performance in many problems in different fields, such as speech recognition, text data mining, text translation, face recognition, image classification and recognition, image segmentation, etc [6]. With the expansion of the application of DL technology by domestic and foreign scholars, it has been gradually applied to synthetic aperture radar (SAR) image classification and segmentation, SAR target detection, interferometric SAR image segmentation, SAR image registration, and inverse synthetic aperture radar (ISAR) imaging. InSAR interferogram phase unwrapping and other fields effectively promote the development of related technologies in these fields [7, 8].

This paper analyzes the phase unwrapping problem in synthetic aperture radar interferometry. The study is divided into five parts. The first section describes the phase unwrapping algorithm based on full convolution and the improved U-net phase unwrapping algorithm. This paper proposes to use the above expansion algorithm to study the interferometric SAR interferogram data set. Section 2 describes the history and current situation of phase unwrapping algorithm and deep learning. The importance of phase unwrapping derived from InSAR height measurement principle is introduced. Section 3 analyzes the technical principle of InSAR and the basic theory of GAN. The traditional phase unwrapping algorithm is studied. Three classical GAN neural networks used in phase unwrapping are introduced. The model architecture is introduced and analyzed in detail. Section 4 simulates the experimental analysis of interferogram. Finally, an improved U-net phase unwrapping algorithm is proposed, and the model structure is divided into modules. Compared with the deep learning phase unwrapping algorithm, the unwrapping results of the proposed algorithm are analyzed.

2. State of the Art

2.1. Phase Unwrapping Algorithm Research History and Present. Phase unwrapping is an important step in InSAR data, and its results directly affect the accuracy of target elevation measurements. At present, phase unwrapping algorithms are mainly divided into three categories, including path-tracking algorithms represented by branch-cut and quality-guided algorithms, minimum-norm algorithms represented by least squares, and state-of-the-art algorithms represented by Kalman filtering [9]:

- (1) Path tracking algorithms use various strategies to define a suitable path and integrate along this path to obtain its unwrapped phase to minimize or avoid the cumulative effect of errors in the phase unwrapping process, including the classic branch proposed by Kumar et al. [10]. In recent years, the tangent method, which detects the residue points and places the branch tangents according to the polarity balance rule, effectively avoids the transmission of unwrapping errors [11]. If the residue points are too dense,

it will lead to an “island” effect in the unwrapping results. The quality map-guided method uses the precomputed quality map as auxiliary information to make the phase unwrapping proceed in the direction of high-quality pixels to low-quality pixels; Flynn equals the mask-cut method proposed in 1996. This algorithm combines the advantages of the branch-cut method and the quality-map-guided algorithm and uses the quality map to guide the placement of branch tangents. Tang et al. improved the quality of graph-guided method by combining the minimum discontinuity method to shorten the running time of the algorithm [12]

- (2) The state estimation algorithm transforms the interferogram phase unwrapping problem into a state estimation problem under nonlinear conditions and performs phase noise suppression and phase unwrapping almost simultaneously [13]. These include the phase unwrapping algorithm improved by using the extended Kalman filter model in 2008, which has better unwrapping results; the unscented Kalman filter phase unwrapping algorithm proposed by Y et al. and a series of nonlinear filter phase unwrapping algorithms. The unwrapping algorithm optimizes the performance of the algorithm by combining the gradient estimator and the fast path tracking strategy to ensure the accuracy and speed of the unwrapping. Then, on this basis, the particle filter phase unwrapping algorithm and the information filter phase unwrapping algorithm are proposed [14]

2.2. History and Current Situation of Deep Learning Research. The concept of deep learning originates from artificial neural network, which is a technology for feature learning on a large amount of data. Its essence is a deep neural network that simulates the human brain for learning and judgment. The convolutional neural network LeNet-5 by combining the convolutional layer and the downsampling layer, which can realize the task of handwritten digit recognition, and this network became the pioneer of the modern structure of the convolutional neural network (CNN). Since then, Li et al. took the lead in using the nonlinear activation function and the Dropout method to prevent overfitting and proposed a new CNN structure AlexNet based on the LeNet-5 model [15]. This network has become an important turning point in the development of deep learning and has paved the way for in-depth research on convolutional neural networks; in 2019, Ma and Li proposed VGGNet, which improves the overall performance of the network by increasing the number of layers of the model [16]. The VGG16 and VGG19 convolutional neural networks have been successfully constructed [17]. Compared with the previous network structure, the error rate of VGGNet in the classification task is greatly reduced. The concept of deep learning originates from artificial neural networks and is a feature learning technology for large amounts of data [18]. The essence is a deep neural network that simulates the human brain for learning and judgment [19]. Since then, J et al took the lead in using

the nonlinear activation function ReLu and Dropout method to prevent overfitting and proposed a new CNN structure AlexNet based on the LeNet-5 model. This network has become an important turning point in the development of deep learning and has paved the way for in-depth research on convolutional neural networks; which improves the overall performance of the network by increasing the number of layers of the model. The VGG16 and VGG19 convolutional neural networks were successfully constructed. Compared with the previous network structure, the error rate of VGGNet in the classification task was greatly reduced [20].

In addition, GAN neural network has gradually begun to be applied to the field of interferogram phase disentanglement. This method transforms the interferogram phase unwrapping problem into a multiclassification problem. The VGG16 network model is used in the first half of the network model to downsample the input image, and the convolutional layer is used instead of the fully connected layer. The feature map enters the decoding path after going through the convolution layer and the pooling layer, and finally, the feature map is upsampled and the classification result is output. Based on the DeeplabV3+ network, the algorithm similarly transforms the unwrapping problem into the classification problem of phase ambiguities and divides, stacks, and corrects the interferogram. After processing, better disentanglement results can be obtained. After the winding phase interferogram is passed through the trained network, the unwrapped phase interferogram can be directly obtained. The network establishes the relationship between the winding phase and the real. The nonlinear mapping relationship between the phases has a good unwrapping effect.

3. Methodology

3.1. InSAR Technology Principle

3.1.1. InSAR Elevation Measurement Principle. The basic idea of InSAR technology is to use two antennas installed on the same flight carrier to observe the same target and obtain the complex radar image pair of the target area. The digital elevation model (DEM) of the target area is reconstructed by calculating the digital elevation information of the target point by calculating the phase information difference contained in the pixels in the two interferograms (i.e., the main and auxiliary images of the interference pair). InSAR error propagation can be used for three-dimensional surface reconstruction by using the phase data of primary differential interference and satellite orbit data. The surface deformation can be detected by the second differential interference processing. This interference analysis needs to use radar system parameters, radar platform attitude (baseline) data phase observation, terrain data (used for terrain phase deduction in secondary difference), etc. Obviously, the uncertainty or error of these data will propagate to the interference elevation or deformation results. This section will introduce the area elevation calculation method through the geometric schematic diagram of InSAR elevation measurement.

In Figure 1, A_1 and A_2 represent radar satellite antennas; assuming that the distance between the two antennas is the baseline length of B , the horizontal angle of the baseline is α , and the downward viewing angle of the antenna A_1 relative to point P is β . The height of the radar antenna A_1 is H . P is a point in the observation area; the distances from the two radar satellite antennas A_1 and A_2 to the observation point P are R_1 and R_2 , respectively; and the height of the observation point P from the ground is h . Due to the difference in the viewing angles of the antennas A_1 and A_2 , the two images cannot be completely overlapped. In order to make the pixels in the two images correspond to each other, an image registration operation is required. Assuming that the signals received by the antennas A_1 and A_2 are S_1 and S_2 , respectively, the following formula is the complex conjugate operation of the registered image:

$$S_1 \times S_2^* = A \exp \left[j \frac{4\pi}{\lambda} (R_2 - R_1) \right]. \quad (1)$$

Among them, “*” is the conjugate symbol, A is the amplitude of the signal, and λ is the wavelength of the radar. From this, the phase difference of the two SAR images can be obtained:

$$\varphi = \varphi_2 - \varphi_1 = \frac{4\pi}{\lambda} (R_2 - R_1). \quad (2)$$

According to the cosine law,

$$R_2^2 = R_1^2 + B^2 - 2R_1B \cos \left(\frac{\pi}{2} - \theta + \alpha \right). \quad (3)$$

The height of the observation point P is derived from the above formula:

$$h = H - R_1 \sin \alpha - \sec \left(\frac{-\lambda\varphi}{4\pi B} \right). \quad (4)$$

In formula (4), H , R_1 , and α are all known quantities, so the above formula is only related to the real phase φ . Since the range of the main phase value limited by the trigonometric function operation is between $[-\pi, \pi)$, $2k\pi$, $k = 0, 1, 2, \dots$:

$$\phi = \varphi + \phi_c = \varphi + 2k\pi \quad (k = 0, \pm 1, \pm 2, \dots). \quad (5)$$

The k in equation (5) is called the phase ambiguity number, which can be classified to obtain the phase ambiguity number distribution map in the deep learning phase unwrapping. The main value of the phase is the winding phase. The process of recovering the real phase of the interferogram from the winding phase is phase unwrapping. Only by obtaining the real phase can the observation target elevation be obtained. It can be seen that phase unwrapping is an important step in the InSAR elevation measurement link.

3.1.2. InSAR Interferometric Phase Extraction Process of Surface Elevation Information. InSAR applications mainly

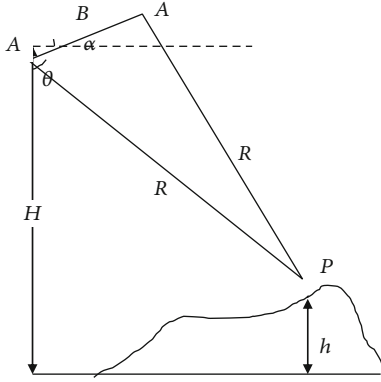


FIGURE 1: InSAR geometric schematic diagram.

include surface deformation measurement and 3D reconstruction, including the following steps:

Interferometric image pair prefiltering and image pair registration processing: during the actual measurement process, the interferogram will be spectral shifts that occur in both the range and azimuth directions, resulting in phase noise in the interference image pair. Therefore, the interferogram needs to be prefiltered to improve the coherence of the interferogram. The topography is more sensitive to the influence of the phase. If the corresponding pixels of the two images are deviated, it will cause a large error in the measurement results. Therefore, it is necessary to perform registration processing on the interference image pair. The registered interferogram can be subjected to the next step of conjugate multiplication and interference processing.

Elimination of the flat ground effect: the flat ground effect refers to the periodic changes of the interferogram fringes in the azimuth and distance directions when characterizing the flat terrain, which cannot reflect the real topographic changes. Therefore, eliminating the flat-earth effect is an important step before the phase unwrapping operation.

Interference filtering processing: the unfiltered interferogram usually has a lot of noise. If the phase unwrapping is performed directly, it will lead to a large deviation in the result. Before unwrapping, the interferogram needs to be filtered. The filtering methods include spatial filtering method and filtering method based on wavelet transform.

Phase unwrapping: phase unwrapping is the process of restoring the interference phase after the flat ground effect has been eliminated from $(-\pi, \pi)$ to the real phase, so that the interference phase corresponds to the terrain elevation, and the result of phase unwrapping will directly affect the to the accuracy of the digital elevation reconstruction.

DEM reconstruction: after the real phase is obtained through phase unwrapping, the digital elevation model can be reconstructed according to the InSAR elevation measurement principle introduced in this chapter in combination with orbital parameters.

3.2. The Basic Theory of GAN. In GAN, the basic architecture of both the generator and the discriminator are neural networks. The input of the generation network is a random noise vector that conforms to a certain distribution, the out-

put is generally a random image, and the input of the discriminator is a real image from the data set or an image generated by the generator and the discriminator. When GAN is training, the generator and the discriminator will alternately train, and the samples generated by the generator will be transformed according to the output of the discriminator during the training process. Therefore, the optimization objective function of GAN is shown in the following formula:

$$\mathcal{L}_{\text{GAN}}(G, D) = \mathbb{E}_x[\lg D(x)] + \mathbb{E}_z[\lg (1 - D(G(z)))], \quad (6)$$

where x represents the random noise vector input to the generator, z is the output of the generator, G represents the generator network, and D represents the discriminator network. The generation countermeasure network (GAN) has two networks, a generator network and a discriminator network. These two networks can be neural networks, from convolutional neural networks, recursive neural networks to automatic encoders. In this configuration, the two networks participate in the competitive game and try to surpass each other while helping them complete their tasks. In the original GAN, the architectures of both the generator network and the discriminator network are multilayer perceptrons, so the generation ability of the generator and the discriminative ability of the discriminator are greatly limited. In DCGAN, the authors used convolutional neural network as the basic architecture of generator and discriminator to enhance the generative and discriminative capabilities of GAN. At the same time, in order to make it differentiable in network training, DCGAN removes the pooling layer in the general CNN and replaces the fully connected layer of the discriminant network with a global pooling layer to reduce the amount of computation.

3.2.1. GAN Optimization and Generation Control. Although DCGAN improves the generation ability and discrimination ability of GAN, it does not fundamentally solve the problems of GAN training difficulties, unstable generator training, and lack of diversity of generated samples during the training process. Wasserstein Generative Adversarial Network (Wasserstein GAN, WGAN) fundamentally analyzed these existing problems and gave a series of solutions. WGAN analyzes the reasons for the problems of the original GAN, mainly including the following points: (1) the better the discriminator, the more serious the generator gradient disappears; (2) the original generator loss function is unreasonable, and the generator loss is minimized during training. When the operation is performed, it is equivalent to minimizing an unreasonable measurement distance between the generated sample and the target. Doing this leads to two problems: (1) unstable gradients and (2) insufficient generative diversity for the generator. At the same time, WGAN also provides a series of solutions: (1) remove the sigmoid activation function of the last layer of the discriminator; (2) do not perform logarithmic transformation when calculating the loss functions of the generator and discriminator; (3) each time after updating the parameters of the discriminator, truncate their absolute values to be less than or equal

to a fixed constant; (4) do not use momentum-based optimization algorithms (including momentum and Adam) in GAN training. It is recommended to use RMSProp, SGD, and other excellent primitives. The GAN not only has the above-mentioned problems, but the most important thing is that the generation results of the generator are not controllable. The main reason is that the input of the generator is a random noise vector, so the output of the generator will completely depend on the input random noise vector, making the output of the generator unpredictable. cGAN controls the generation results of the generator by adding generative and discriminative conditions. Compared with the original GAN model, the optimization objective function of cGAN is shown in formula (7), where y is the added constraint vector.

$$\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}_x[\lg D(x|y)] + \mathbb{E}_z[\lg (1 - D(G(z|y)))]. \quad (7)$$

3.2.2. The Loss Function Mechanism of GAN. The main function of the discriminator network in GAN is to guide the training of the generator by judging the authenticity of the input samples, so that the output of the generator is closer to the real samples. Specifically, the role of the discriminator is equivalent to first extracting the features of the generated samples or real samples and then discriminating the gap between the generated samples and the real samples at the feature level. The general supervised learning network will use a predefined loss function to calculate the error between the output of the model and the real sample, so as to guide the training of the model. Therefore, the discriminative mechanism of GAN is equivalent to providing a trainable loss function to supervise the training of the generator. Figure 2 is a graph of the GAN loss function. If the generative network structure is replaced by an end-to-end segmentation model structure, GAN is quite a more powerful segmentation model at this time, because it has a better trainable loss function. The pix model perfectly solves the image translation problem by applying the cGAN mechanism for the first time. The cGAN optimization objective function of Pix2pix is shown in the following formula (8):

$$\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}_{x,y}[\lg D(x, y)] + \mathbb{E}_{x,z}[\lg (1 - D(x, G(x, z)))]. \quad (8)$$

In Pix2pix, the input x of the generator is changed from a random vector to an image, and the input of the discriminator is the original input image x spliced in the original generated image z or marked image y as a condition for discrimination. Inspired by the work of Pix2pix, this paper constructs different GAN-based segmentation models in different medical image segmentation tasks and achieves very good segmentation results.

3.3. Research on Traditional Phase Unwrapping Algorithms

3.3.1. Phase Unwrapping Algorithms. Based on path tracking, the concept of residue point in 1988 and proposed the classical branch-cut method for phase unwrapping (BranchCut-PhaseUnwrapping, BCPU). The algorithm generates

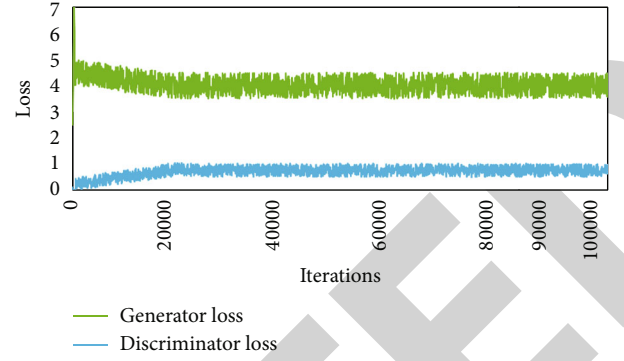


FIGURE 2: GAN loss function graph.

branch tangents by identifying positive and negative residue points. According to the principle that the integral path does not cross the branch tangents, the error transmission caused by unwrapping of discontinuous points is avoided. The definition of the residual point is as follows: calculate the adjacent phase gradient winding value, as follows:

$$\begin{aligned} \Delta_1 &= W[\psi(m+1, n) - \psi(m, n)], \\ \Delta_2 &= W[\psi(m+1, n+1) - \psi(m+1, n)], \\ \Delta_3 &= W[\psi(m, n+1) - \psi(m+1, n+1)], \\ \Delta_4 &= W[\psi(m, n) - \psi(m, n+1)]. \end{aligned} \quad (9)$$

Adding the above formula can get

$$S = \sum_{i=1}^4 \Delta_i. \quad (10)$$

If s is greater than 0, that is, the polarity of the residual point charge is “positive,” it means that the pixel $\Psi(m, n)$ in the upper left corner of the closed path is a positive residual point, otherwise, $\psi(m, n)$ is called a negative residual point. Count the points.

3.3.2. Quality Graph Guidance Algorithm. In addition to the above-mentioned method of determining the integral path using the residual point distribution, other auxiliary information can also be used to plan the integral path. The QualityGuidePhaseUnwrapping algorithm (QGPU) analyzes the phase quality, so that the winding phase is unwrapped in the direction of high-quality pixels to low-quality pixels. The core of the mass map-guided method is to use the phase quality of the interferogram to guide the interferometric pixels for diffusion. The following are the algorithm steps:

Step 1. Take the pixel with the highest quality in the winding interferogram as the starting point of unwrapping, mark it as the unwrapped pixel, and store the adjacent pixels around the starting pixel in the adjacency list.

Step 2. Sort the pixels in the adjacency list according to the quality value, select the pixel with the highest quality, mark the pixel as the unwrapped point, remove the unwrapped

pixel from the adjacency list, and then, put this pixel. The adjacent unwrapped pixels are stored in the adjacency list.

Step 3. Repeatedly select the pixel with the highest quality in the adjacency list to perform Step 2, until the unwrapped pixels in the adjacency list are empty, and the phase unwrapping is completed.

The mass map-guided algorithm references the mass map as auxiliary information to avoid error propagation due to unwrapping discontinuous phases. In the case of a relatively reliable phase quality map, the unwrapping effect of the algorithm is better than that of the Goldstein branch-cut method. Otherwise, the unwrapping results may be inaccurate due to the existence of unbalanced residual points. In this paper, we use weighting to avoid the error caused by residual points, which leads to the accumulation and propagation of errors. The residual points and nonresidual points can be distinguished by solving the quality map of the wrapped phase, so as to obtain better expansion results. At present, there are mainly four methods to evaluate the quality of interferograms: coherence coefficient map, pseudocoherence map, phase derivative change map, and maximum phase gradient map.

- (1) The coherence coefficient map is defined as follows:

$$\gamma = \frac{\lim_{M,N \rightarrow \infty} \sum_{n=1}^N \sum_{m=1}^M |\mu_1(n, m)| |\mu_2(n, m)|}{\sqrt{\lim_{M,N \rightarrow \infty} \sum_{n=1}^N \sum_{m=1}^M |\mu_1(n, m)|^2 \lim_{M,N \rightarrow \infty} \sum_{n=1}^N \sum_{m=1}^M |\mu_2(n, m)|^2}} \quad (11)$$

In the formula, γ is the coherence coefficient value; M and N are the data size for calculating the coherence; m and n are the row and column numbers in the data.

- (2) The pseudocoherence coefficient map is defined as follows:

$$Z_{m,n} = \frac{\sqrt{(\sum \cos \phi_{i,j})^2 + (\sum \sin \phi_{i,j})^2}}{k^2} \quad (12)$$

- (3) The phase derivative change diagram is defined as follows:

$$Z_{m,n} = \frac{\sqrt{\sum (\Delta_{m,n}^x - \bar{\Delta}_{m,n}^x)^2} + \sqrt{\sum (\Delta_{m,n}^y - \bar{\Delta}_{m,n}^y)^2}}{k^2} \quad (13)$$

- (4) The maximum phase gradient map is defined as follows:

$$Z_{m,n} = \begin{cases} \max [|\Delta_{m,n}^x|] \\ \max [|\Delta_{m,n}^y|] \end{cases} \quad (14)$$

In the above four evaluation methods, the coherence coefficient map indicates the quality of the corresponding position of the interference image through the brightness and darkness, so this is the most intuitive interference phase quality evaluation standard. When the terrain is relatively flat and the phase change is not obvious, the pseudocoherence coefficient map is often used. Similar to the maximum phase gradient, it also appears as low-quality data in areas with steep changes in terrain, which is not conducive to guiding the unwrapping path. In the case where the coherence map cannot be obtained, the effect of the phase derivative change map is relatively more reliable.

3.4. Phase Unwrapping Algorithm Based on Deep Learning

3.4.1. Phase Unwrapping Network Based on FC-DenseNet.

According to this network, the excellent performance on the data set confirms the potential of this network in dealing with semantic segmentation problems. Traditional neural networks such as LeNet, VGGNet, FC-DenseNet can reuse the feature information of each layer through dense block (DB) with its unique multilayer cascade structure, and the input of each layer of network includes all previous layer learning. The obtained image features can improve the utilization rate of feature information by the network. In order to deal with the complex phase unwrapping problem, this paper deepens the network on the basis of the original model and proposes a phase unwrapping model based on FC-DenseNet. The schematic diagram of the network model is as follows. As shown in Figure 3, (a) is the initial model, and (b) is the deepened model.

It can be seen from Figure 3 that the structure of the deepened FC-DenseNet model is similar to that of U-net, with the downsampling path on (a) and the upsampling path on (b). The network is composed of two convolution layers, nine dense blocks, two transition layers, layers and jump connections. In the phase unwrapping model based on the FC-DenseNet network, the black arrow represents the flow direction of the network, the blue box represents the convolution module (convolutional layer + batch normalization BN + activation function ReLu + Dropout layer); TD is denoted as the downsampling transition layer; TU is denoted as the upsampling transition layer; and the gray arrows are skip connections. The higher resolution information in the encoding path is passed to the decoding path through skip connections. From the input layer, the current output is concatenated into the input of the next layer. If the operation is repeated four times, the final dense block output is the concatenation result of the output features of the four convolutional layers. Therefore, such a multilayer dense structure combines the feature information contained in the remaining layers to ensure the efficient use of information.

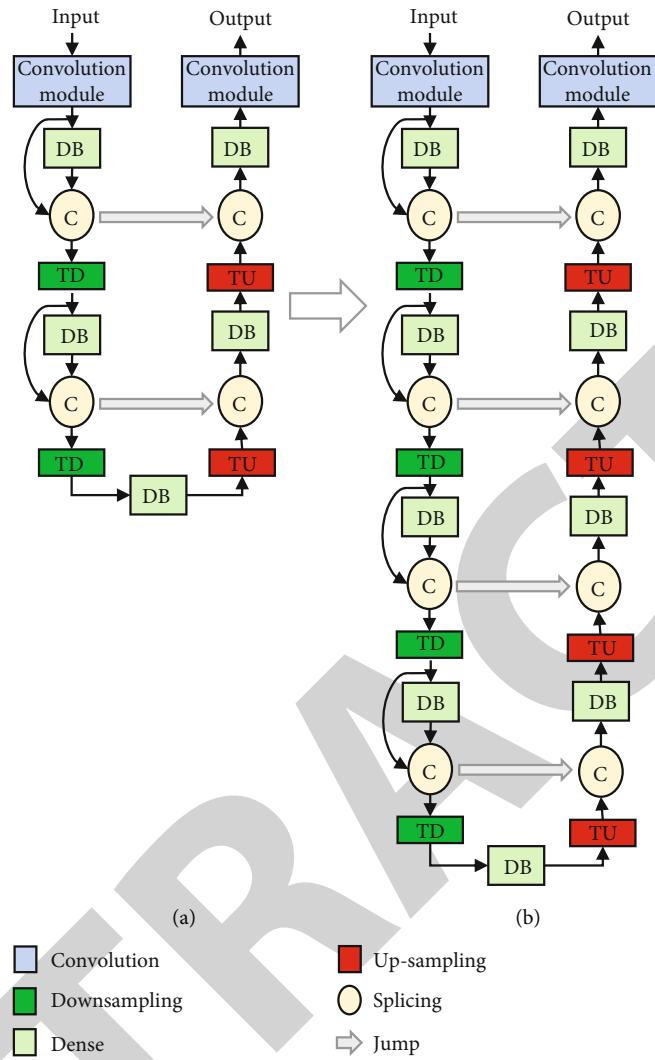


FIGURE 3: Initial model and phase unwrapping model based on FC-DenseNet network.

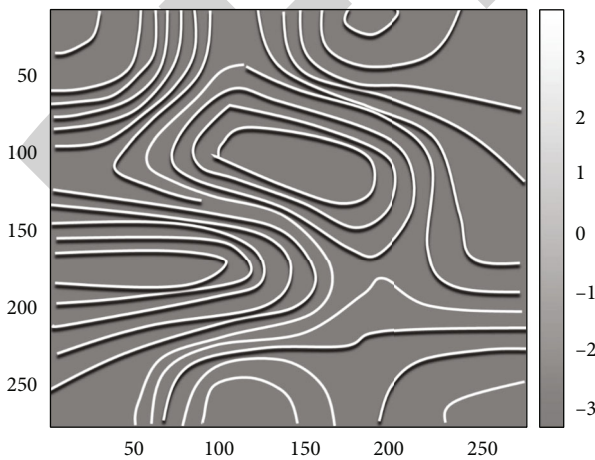


FIGURE 4: Noise winding phase diagram.

4. Result Analysis and Discussion

4.1. *Experimental Analysis of Simulated Interferograms.* Interferogram denoising is a key step in InSAR data processing. If the noise pollution in the interferogram is serious. This section will use simulated and measured interferometric maps to analyze the branch-cut method, quality map-guided method, FFT-based least squares method (LSPU), and iterative least squares method (ILSPU) introduced in this chapter. Based on the unscented Kalman filter algorithm (UKFPU), five classical phase unwrapping algorithms are tested, respectively, and the performance of each algorithm is analyzed by comparing the unwrapping results of each algorithm. The unit of horizontal and vertical coordinates in the figure is pixels, and the unit of the color label on the right is radians.

- (1) Simulation experiment 1: Figure 4 shows the simulated terrain winding interferogram (signal-to-noise ratio is 8 dB). The phase unwrapping of Figure 4 is carried out using the branch-cut method, the quality

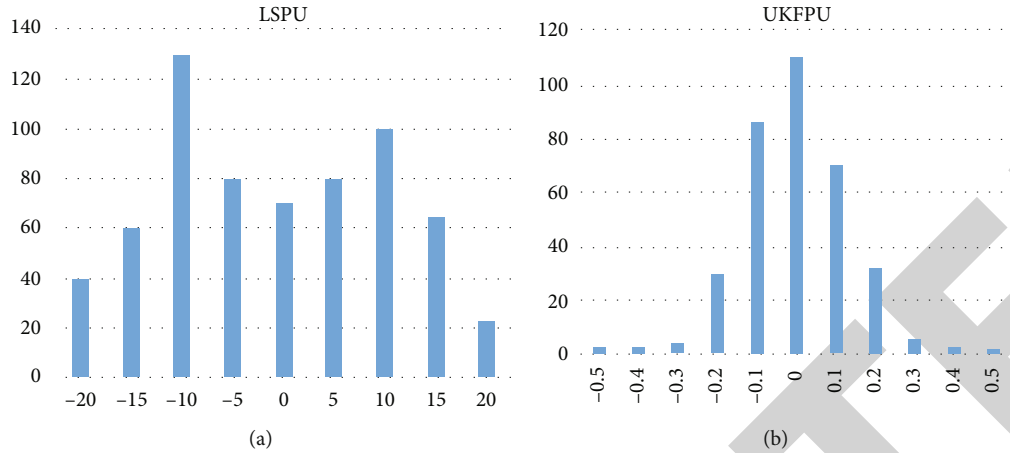


FIGURE 5: The statistical histogram of the unwrapping error of each algorithm: (a) LSPU algorithm; (b) UKFPU algorithm.

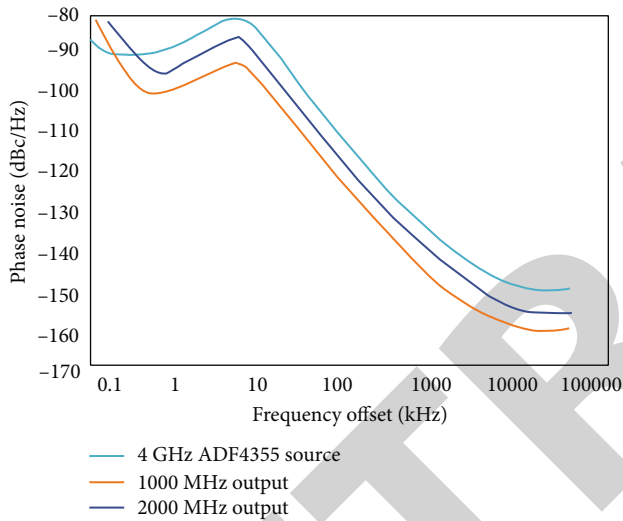


FIGURE 6: Interferogram segmentation results after adding noise.

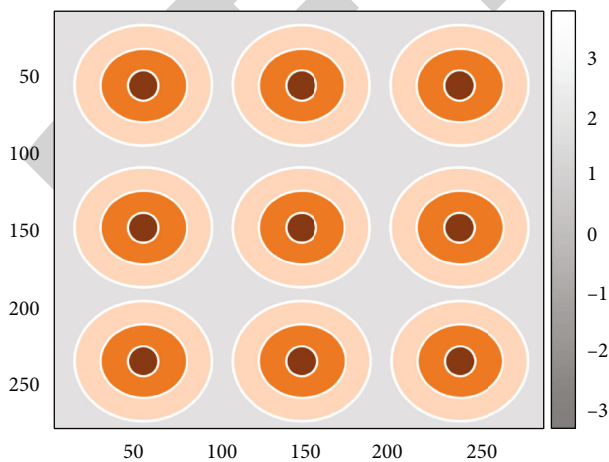


FIGURE 7: Noise winding interference diagram.

map-guided method, LSPU, ILSPU, and UKFPU, respectively, and the unwrapping results of each algorithm are shown in Figure 4

It can be clearly seen from Figure 5 that the unwrapping results of the LSPU algorithm have a large error, and the error values are distributed between -20 and 20, while the unwrapping errors of the other four algorithms are mostly distributed between -1 and +1. The experimental results show that other algorithms except the LSPU algorithm have good unwrapping effect when dealing with interferograms with high signal-to-noise ratio.

(2) Simulation experiment 2: Figure 6 shows the interferogram with added noise. It can be seen from the figure that when dealing with interferograms with lower signal-to-noise ratios. Too many points are left in the noisy area, which leads to the “islanding” effect, and a large number of unwrapped black spots appear in the unwrapped result, which cannot effectively unwrap all the interferograms. The unwrapping error is mainly distributed around 0, but there are still areas in the unwrapping result that are inconsistent with the original real phase map

(3) Simulation experiment 3: Figure 7 is the interferogram of winding. In order to compare the unwrapping efficiency of each algorithm and the unwrapping accuracy under different signal-to-noise ratios, different degrees of noise are added to Figure 7 before unwrapping. Since the statistical standard of the branch-cut method is different from other algorithms, the statistics and analysis are not carried out in this experiment. As the signal-to-noise ratio of the interferogram gradually decreases, the mean square error of the unwrapping results of each algorithm is also increasing

Figure 7 shows the distribution of residual points of the measured optical interferogram. From the unwrapping results of the measured terrain, it can be seen that when the branch tangent method unwrapped the dense area of the residual points of the interferogram, the branch tangent

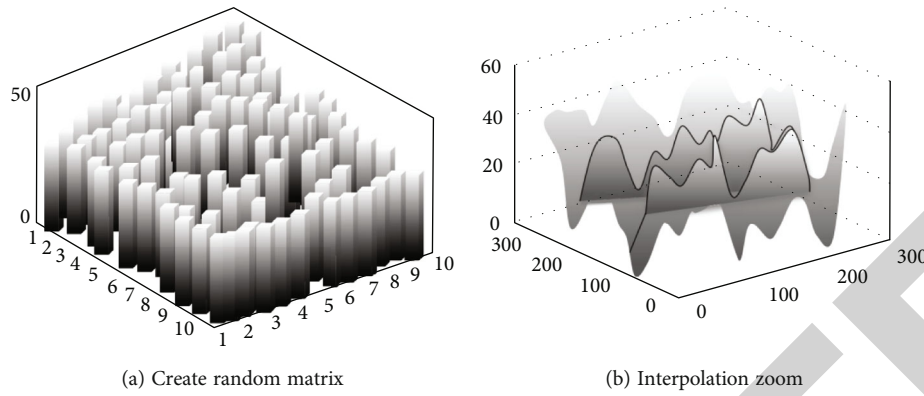


FIGURE 8: A simulated InSAR data set.

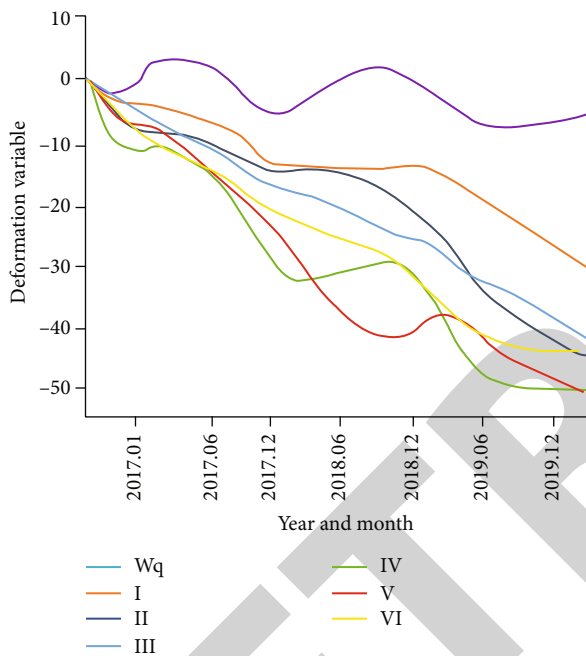


FIGURE 9: Phase unwrapping feature points obtained by the correlation coefficient method.

lines formed a closed loop, and the complete interferogram cannot be unwrapped; the unwrapped result of the mass-map guided method is basically consistent with the real phase, but it can be seen from the rewinding image that there are still unresolved edges of the interference fringes. For filtering noise, the fringe distortion in unwrapped phase rewind images of LSPU and ILSPU is not obvious, and there is burr noise, indicating that although the minimum norm class algorithm has relatively high phase unwrapping efficiency, it is easy to perform in the unwrapping process. The details of the interferogram are ignored, causing the unwrapped results to deviate too much from the true phase.

4.2. Research on GAN-Based Deep Learning Phase Unwrapping Algorithm

4.2.1. InSAR Interferogram Data Set. Construction of data set is an important step in supervised learning, and high-quality

data sets can often improve the quality of network model training and the accuracy of prediction. Data sets with better quality often have the following two characteristics:

(1) The amount of sample data is moderate. If the amount of data is too large, the neural network will lead to overfitting after generalizing a large amount of data, and if the amount of data is too small, it will lead to underfitting. Therefore, in the actual training process, data enhancement is generally adopted to expand the data

(2) The sample data types are diverse. If the sample data features are single, the trained network has poor robustness and weak generalization ability, and the data set with diverse features is more representative. In the confrontation training (as a classical algorithm to improve the robustness of the model), the author effectively combines it with the training process of GAN, and names it as the combination model rob GAN. Experiments show that rob GAN can not only make the training of GAN more stable and the generation results more realistic but also reduce the performance gap between the training set and the test set. Currently, there are no publicly available InSAR data sets in the field of deep learning phase unwrapping. In this paper, the simulated InSAR data, quasimeasured InSAR data, and DEM terrain data of $256 \text{ pixels} \times 256 \text{ pixels}$ are constructed, respectively, and the three types of data are mixed to form a complete InSAR data set. Label units are radians. In phase unwrapping, the relationship between the real phase φ and the corresponding wrapping phase φ can be expressed as

$$\varphi = \text{angle}[\exp(j\phi)]. \quad (15)$$

Among them, the winding phase $\varphi \in (-\pi, +\pi)$. When creating an InSAR data set, first generate the real phase value of InSAR, calculate the winding interferogram through the above formula, and construct a simulated InSAR data set as shown in Figure 8. Columns in Figure 8(a) are the

randomly created 10×10 initial matrices; Figure 8(b) is the interpolation and enlargement of the original initial matrix; in the network training process, the real phase map generated above is used as its corresponding noise label image of the wrapped phase map.

Figure 9 is a phase unwrapping feature point diagram obtained by the correlation coefficient method. The larger the correlation coefficient, the smaller the interference and the better the reliability; the smaller the correlation coefficient, the greater the interference, and the worse the reliability. Therefore, when the unit phase unwrapping is performed, the area with the largest average correlation coefficient is selected as the object of the network planning algorithm, and the remaining three areas in the unit are used as the unwrapping object of the FFT-based least squares algorithm. Similarly, in the process of image correction and fusion, the area with the largest average correlation coefficient should be used as the benchmark for correction and fusion.

5. Conclusion

This paper applies the deep learning method to phase unwrapping to expand and explore the potential of GAN neural network in the field of InSAR interferogram phase unwrapping. The main research work is as follows: the InSAR data set for deep learning phase unwrapping is proposed. The mixture of simulated data and quasimeasured data improves the diversity of sample data features in the training set and increases the generalization ability of the neural network. The experimental results show that although the structure level is deep enough, the FC Dense Net network model cannot process interferograms with low SNR and complex interference fringes. The experimental results of simulated and measured terrain interferometric unwrapping show that, compared with other types of deep learning phase unwrapping methods, the proposed method achieves better results in both simulated and measured data unwrapping experiments, with a relatively small mean square error and a better solution. The winding precision is relatively high, and the noise resistance is relatively strong. This paper has some limitations. When the interference filtering of the original interferogram suppresses a lot of noise, it will also lead to the loss of a lot of useful information. The irregular window needs further research and analysis. The research also needs to improve the ability of adaptive median filter to remove phase particle noise and keep it unclear.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] K. D. Murray, R. B. Lohman, and D. Bekaert, "Cluster-based empirical tropospheric corrections applied to InSAR time series analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 3, no. 99, pp. 1–9, 2020.
- [2] L. Pulvirenti, M. Chini, and N. Pierdicca, "InSAR multitemporal data over persistent scatterers to detect floodwater in urban areas: a case study in Beletweyne, Somalia," *Remote Sensing*, vol. 13, no. 1, p. 37, 2020.
- [3] Z. Yuan, Z. Lu, L. Chen, and X. Xing, "A closed-form robust cluster-analysis-based multibaseline InSAR phase unwrapping and filtering algorithm with optimal baseline combination analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 6, pp. 4251–4262, 2020.
- [4] M. Wang, H. Lu, S. Liu, and Z. Zhu, "How to mislead AI-assisted network automation in SD-IPoEONs: a comparison study of DRL- and GAN-based approaches," *Journal of Lightwave Technology*, vol. 38, no. 20, pp. 5574–5585, 2020.
- [5] K. S. Gill, S. Nguyen, and M. Thein, "Three-way deep neural network for radio frequency map generation and source localization," 2021, <https://arxiv.org/abs/2111.12175>.
- [6] K. B. Park, S. H. Choi, and J. Y. Lee, "M-GAN: retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks," *IEEE Access*, vol. 8, no. 99, pp. 146308–146322, 2020.
- [7] X. Y. Lim, K. B. Gan, and N. A. A. Aziz, "Deep ConvLSTM network with dataset resampling for upper body activity recognition using minimal number of IMU sensors," *Applied Sciences*, vol. 11, no. 8, p. 3543, 2021.
- [8] S. Krylov and V. Krylov, "Inverse problem solving approach using deep network trained by GAN simulated data," *Journal of Physics: Conference Series*, vol. 2021, no. 1, article 012002, 12 pages, 2021.
- [9] A. Amyar, S. Ruan, P. Vera, P. Decazes, and R. Modzelewski, "RADIOGAN: deep convolutional conditional generative adversarial network to generate PET images," 2020, <https://arxiv.org/abs/2003.08663>.
- [10] A. Kumar, A. Alsadoon, P. W. C. Prasad et al., "Generative adversarial network (GAN) and enhanced root mean square error (ERMSE): deep learning for stock price movement prediction," *Papers*, vol. 17, no. 20, pp. 115–123, 2021.
- [11] R. Nandhini Abirami, P. M. Durai Raj Vincent, K. Srinivasan, U. Tariq, and C. Y. Chang, "Deep CNN and deep GAN in computational visual perception-driven image analysis," *Complexity*, vol. 2021, no. 25, Article ID 5541134, 30 pages, 2021.
- [12] T. W. Tang, W. H. Kuo, J. H. Lan, C. F. Ding, H. Hsu, and H. T. Young, "Anomaly detection neural network with dual auto-encoders GAN and its industrial inspection applications," *Sensors*, vol. 20, no. 12, 2020.
- [13] N. Shivsharan and S. Ganorkar, "Diabetic retinopathy detection using optimization assisted deep learning model: outlook on improved grey wolf algorithm," *International Journal of Image and Graphics*, vol. 21, no. 3, article 2150035, 2021.
- [14] Y. Hua, R. Li, Z. Zhao, X. Chen, and H. Zhang, "GAN-powered deep distributional reinforcement learning for resource management in network slicing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 2, pp. 334–349, 2020.
- [15] R. Li, X. Lv, J. Yuan, and J. Yao, "A triangle-oriented spatial-temporal phase unwrapping algorithm based on irrotational constraints for time-series InSAR," *IEEE Transactions on*

Retraction

Retracted: Application of Swarm Intelligence Optimization Algorithm in Logistics Delivery Path Optimization under the Background of Big Data

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Zhao, "Application of Swarm Intelligence Optimization Algorithm in Logistics Delivery Path Optimization under the Background of Big Data," *Journal of Function Spaces*, vol. 2023, Article ID 3476711, 12 pages, 2023.

Research Article

Application of Swarm Intelligence Optimization Algorithm in Logistics Delivery Path Optimization under the Background of Big Data

Guofu Zhao ^{1,2}

¹School of Business, Zhengzhou Sias University, Xinzheng, Henan 451150, China

²School of Economics & Trade, Zhengzhou University of Technology, Zhengzhou, Henan 450044, China

Correspondence should be addressed to Guofu Zhao; 14999@sias.edu.cn

Received 2 September 2022; Revised 21 October 2022; Accepted 24 November 2022; Published 24 February 2023

Academic Editor: Miaochao Chen

Copyright © 2023 Guofu Zhao. 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.

Traditional logistics delivery route optimization algorithm has some problems such as long time to find the optimal route. Based on this, this paper discusses swarm intelligence optimization algorithm and logistics delivery route optimization. To solve the logistics vehicle routing problem, considering that the basic ACO (ant colony optimization) has the disadvantages of slow convergence speed and easy to fall into local optimum, this paper proposes a new hybrid population optimization algorithm and applies it to VRPTW (vehicle routing problem with time windows). In addition, the concept of crowding degree in AFSA (artificial fish swarm algorithm) is introduced into ACO. In the early stage of the optimization process, a strong crowding degree limit is set to ensure that most ants are not affected by pheromone concentration to conduct random optimization. The simulation results show that the AC (accuracy rate) of this algorithm is 95.08%, which is higher than the traditional PSO (particle swarm optimization) algorithm and general heuristic algorithm. The hybrid algorithm can effectively improve the optimization efficiency of VRPTW, lay a foundation for solving large-scale VRPTW, and provide new research ideas and methods. At the same time, the results fully show that the algorithm in this paper has certain advantages in performance, and it can be applied to logistics delivery route optimization.

1. Introduction

The distribution route is very important in the growth of urban economy. Under certain conditions, the most appropriate distribution route can be improved and optimized to effectively reduce the distribution cost and material loss [1, 2]. Shortening the distribution path reduces the transportation expenditure for enterprises and the consumption loss of the masses [3]. IA is an important service industry in the national economy, the logistics industry is developing rapidly all over the world and has gradually become the artery of basic industry and national economic development [4]. At the same time, urban logistics delivery channels are related to the people's livelihood of the whole city. Therefore, it is necessary to plan transportation routes reasonably and improve logistics delivery efficiency

while consuming the lowest cost. VRP (vehicle routing problem) is an important content in logistics system research. Based on this, this paper discusses swarm intelligence optimization algorithm and logistics delivery path optimization. A new optimization algorithm of logistics delivery path is proposed. Now that we have entered the era of big data, we have formed a "big group" space for network interaction under this support. Its collaboration and reliability need to be solved by means of swarm intelligence. The swarm intelligence algorithms represented by particle swarm optimization and ant colony algorithm have good robustness and flexibility, which is the key to solving complex problems. In view of this, the article takes the network swarm intelligence in the age of big data as the research base point, combines the application principle and characteristics of swarm intelligence, explains its application

advantages in NP problems, and provides effective support for the security of network information interaction. Through intelligent route planning algorithm, automatically match vehicle data and route information, customer location and driver information, commodity specification and quantity, and loading and unloading points, and combine a large amount of data. Second, calculate the vehicle arrangement route, and finally, formulate an optimized distribution route that meets the transportation and distribution objectives, for example, shorter mileage, less cost, shorter time, and fewer vehicles.

Big data is an abstract concept, but at present, the academic circles have not yet formed an exact and unified definition. In this paper, traffic big data is defined as a data set composed of a large quantity of electronic maps, roads, vehicles, and other types of traffic information [5]. According to the vehicle route model, VRP can be converted into TSP (traveling salesman problem). ACO has been applied effectively in solving famous problems such as TSP, but when solving large-scale problems, its convergence speed is slow, and it takes a long time. In this paper, the traditional ACO is improved to build an optimization model. Big data can be used for positioning when building an optimization model. The data needs to be collected from the route collection center about the specific distribution requirements and time of each logistics branch, and based on this information, the data will be uniformly distributed and the distribution route will be planned. In this way, the efficiency of each distribution route can be ensured, while the needs of each customer can be met. The research innovation contribution lies in the discretization of the hybrid algorithm and the introduction of the local path optimization operator. The improved algorithm is applied to solve the VRPTW problem. In the early stage of the optimization process, a strong congestion limit is set. It ensures that most ants are not affected by pheromone concentration, to conduct random optimization. A HSIA model for logistics distribution route optimization is proposed, and its basic principle, mathematical description, parameter analysis, and algorithm flow are analyzed and studied. At the same time, set the distribution target weight, and find the optimal distribution path in the target function according to the different needs of logistics, to complete the optimization of logistics distribution path.

According to the research on the application of swarm intelligence optimization algorithm in logistics delivery route optimization and the need of this paper structure, this paper will be divided into five parts; the specific contents are as follows.

The first section introduces the research background and significance of logistics distribution path optimization and explains the content of this paper and the organizational structure of the full text. The second section is related work. This section expounds the research status of the research topic of this paper and puts forward the research work of this paper. In the third section, the swarm intelligence optimization algorithm and logistics delivery path optimization are analyzed. A new optimization algorithm of logistics delivery path of mixed population is proposed. In the fourth section, a large quantity of experiments are carried out to explore the performance of the algorithm. The fifth section is summary and prospect. This section makes a comprehensive summary of this research;

finally, the shortcomings of the research and the research direction in the future are given.

2. Related Work

Shukla et al. established a logistics delivery model with rigid requirements for the travel time of distribution vehicles [6]. The model studies how the total cost of delivery will change if the hard time window is considered to be relaxed to a soft time window and takes a delivery task as an example to optimize the delivery route with a heuristic algorithm. Mousavi and Vahdani designed a tabu search algorithm for the most basic model of logistics delivery path optimization in the B2C e-business environment and conducted numerical tests and comparisons at the same time [7]. Ouyang added fuzzy constraints to the close-range open VRPTW and solved it with a hybrid ACO [8]. Raad et al. believe that information technology is a logistics capability that third-party logistics users are particularly concerned about, and users expect third-party logistics to make breakthroughs in services in areas such as integrated supply chain management and e-business. Third-party logistics users have higher and higher requirements for the quality of logistics services [9]. Mardaneh et al. proposed an adaptive multimodal continuous ACO with minimum habitat and extended ACO to solve multimode optimization problems [10]. The algorithm adjusts the ant colony pheromone update strategy and uses the differential evolution operator to construct the initial solution of the ant colony to speed up the convergence rate. To enhance the search, a Gaussian distribution-based local search scheme is adaptively performed around the seeds of the niche, switching between global and local searches. Rodríguez et al. utilized a hybrid algorithm based on tabu search and simulated annealing algorithms to solve VRP with backhaul and time windows [11]. Chu et al. proposed the column generation method to solve the idea of VRP [12]. The idea is to transform the original problem into a simplified problem, and the range that needs to be considered is a subset of all possible feasible solutions, and the shortest path is found by repeatedly solving on this basis. In order to reduce logistics delivery costs and improve customer satisfaction, Oliveira and Hamacher established a distribution system under changing demand and various components of operating costs and formed an integrated model [13]. Kuznietsov et al. proposed a nonlinear programming optimization model aiming at the node construction cost and operating cost of cold chain logistics and used the quantum PSO to solve the model [14, 15]. At the same time, a GA (genetic algorithm) is designed for this problem from the aspects of genetic coding, genetic operator, algorithm termination conditions, etc., to effectively solve the exponential explosion phenomenon when solving combinatorial optimization problems. Rooeinfar et al. explored the logistics delivery system and proposed corresponding development strategies [16]. These include building logistics alliances; third-party logistics models; introducing fourth-party logistics models; and accelerating lean supply chain management. Lang and Shen constructed a multiobjective route optimization model considering customer satisfaction and delivery costs and used an improved GA to simulate the model [17]. Niknejad and

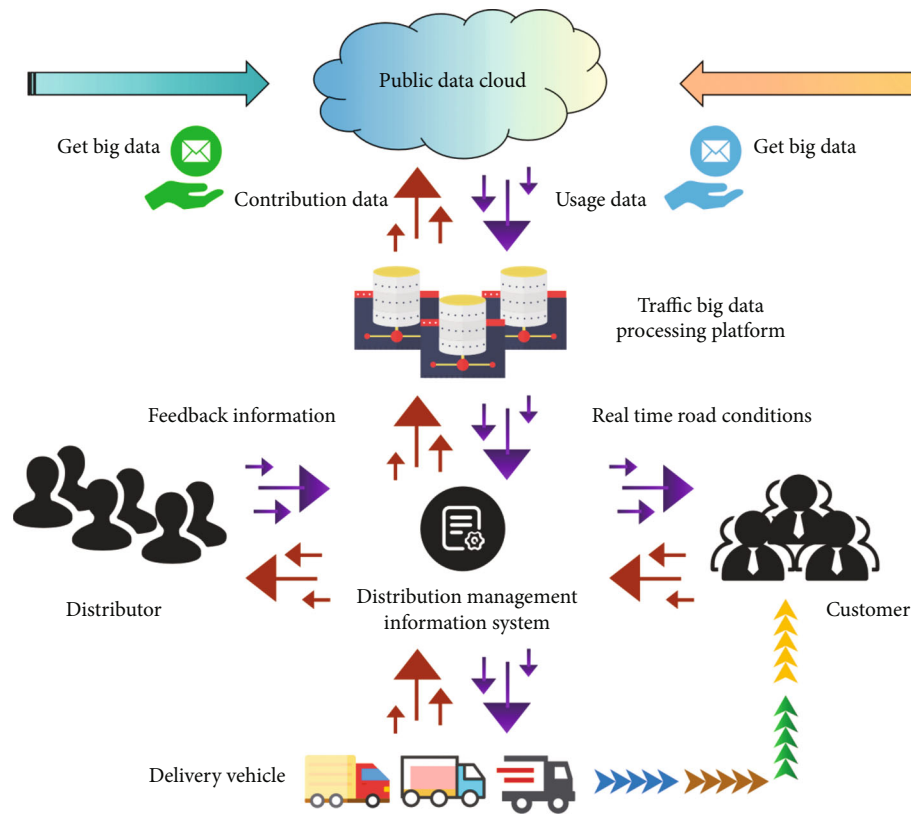


FIGURE 1: Application of big data in logistics delivery.

Petrovic studied VRP considering customer preference in dynamic environment and gave a solution idea [18]. Xiao and Rao believe that there are many uncertain factors in the real environment, which brings great difficulties to the location selection of logistics delivery [19]. The traditional two-layer objective planning cannot meet the requirements of uncertain conditions. The improved planning model is more suitable for the changing external environment, and then, an example is verified. Gharbi et al. further studied the selection of connection points and the optimization of cold chain logistics delivery paths based on connection points and established a cold chain distribution path optimization model based on transportation big data with connection points [20, 21]. Finally, an example is used to verify the validity of the above model.

By summarizing the above algorithms for solving VRP, we can see that there is no best algorithm, only the most suitable one. However, with the increasing scale of VRP and more and more complicated constraints, heuristic algorithm based on swarm intelligence is the main development trend to solve this kind of problems. This paper mainly discusses swarm intelligence optimization algorithm and logistics delivery path optimization. Based on the improvement of ACO, a new optimization algorithm of mixed population logistics delivery path is proposed. At the same time, in order to apply the improved algorithm to solve VRPTW, the hybrid adaptive algorithm is discretized, and a local path optimization operator is introduced. Finally, an example of path optimization test is used to verify the effectiveness and better effect of the algorithm in solving VRPTW.

3. Methodology

3.1. Logistics Distribution Management Information System. Whether the distribution path is reasonable or not directly affects the speed and cost of logistics delivery, selecting the path optimization goal is the premise of path planning. According to the specific distribution problems of customers, we can design a reasonable distribution scheme to optimize the route. This kind of scheduling belongs to the logistics delivery path optimization problem [22]. According to the different logistics constraints in reality, there are many models of VRP, among which: VRPTW is the most typical problem, and it is also a representative constrained multi-objective problem with the most research value at present. In order to reduce the operating cost of service providers, the logistics system will optimize the vehicle distribution route. Logistics delivery with high timeliness requirements can use the distribution management information system to obtain information such as roads and real-time road conditions through the traffic big data platform to manage and direct vehicles in transit for distribution; at the same time, customers can also inquire about vehicles and goods and feedback information through this platform. The application of big data in logistics delivery is shown in Figure 1.

The application of big data analysis in the freight field is the most common. It is mainly reflected in site selection optimization, inventory scale, supply route, and other activities. Data analysis can group customers of enterprises. Transportation and route selection are the most widely used fields of big data analysis in logistics management. Many enterprises use

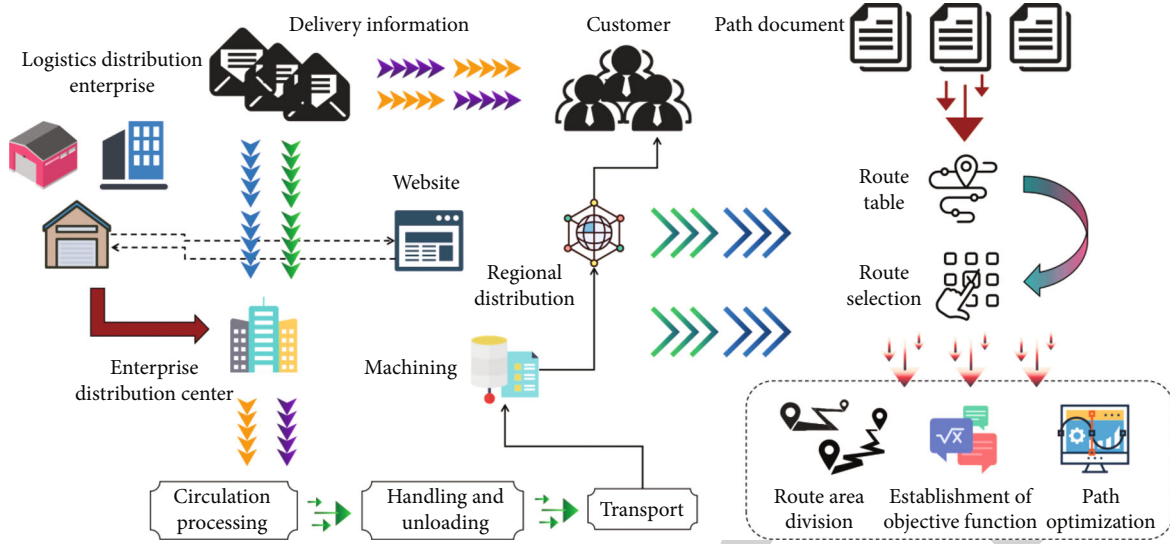


FIGURE 2: Logistics delivery and route design process.

remote big data information processing technology loaded with GPS navigation to optimize freight transportation routes. The limitations of traditional optimization methods make it impossible to meet the complex requirements of multiconstraints and multiobjectives of path planning problems. At present, heuristic algorithm based on swarm intelligence is the main development trend to solve this kind of problems. Swarm intelligent algorithm is distributed, individual intelligence is not controlled by a unified subject, and it has strong robustness. Moreover, each individual can only perceive local information, so it is relatively simple to follow the rules and easy to implement the algorithm [23]. The intelligent group has the characteristics of self-adaptability, and the group can change its behavior at an appropriate time when the required cost is not too high. Heuristic algorithm is highly sensitive to VRP path, so its solution efficiency is also high. Its VRP is basically the same for different logistics companies. Therefore, heuristic algorithm can be used to plan the path, reduce the cost, consume the least resources, and get the maximum profit. AFSA is a new intelligent bionic algorithm based on the characteristics of fish activities.

The optimization of logistics distribution path requires that the vehicle distribution route be arranged quickly and reasonably, so that the goods can be delivered to the customers in time and accurately, and the distribution cost is the lowest. Soft time window means that if the delivery vehicle cannot start service within the time required by the customer, it must be punished accordingly. The meaning of hard time window is that if the delivery vehicle cannot start service within the time required by the customer, the solution is not feasible. The optimization goal is to quickly find a distribution strategy by obtaining real-time road congestion information, building a mathematical model and adopting appropriate algorithms according to the existing resources and customer demand. For the processing of optimization problems, because each logistics delivery route optimization problem faces different situations, it is necessary to design a general optimization algorithm to solve it.

Therefore, based on big data, we can simulate the route distribution status in various situations, integrate these distribution statuses, and establish a tracking optimization algorithm to calculate the specific loss of distribution.

3.2. Optimization Algorithm of Mixed Population Logistics Delivery Route. In this paper, the time window of logistics delivery problem is set as a special time window, which is a time interval, and the distribution vehicles need to deliver the materials to the material demand points in advance or on time within this time window. At the same time, this paper adopts the two-population strategy, and the combination of ACO and differential evolution in the first population can solve the parameter optimization problem well. Secondly, the ant colony hybrid PSO is adopted in the population, which enhances the global search ability in a wide range and solves the problems of slow convergence and easy falling into local optimum. The flow chart of logistics and route design is shown in Figure 2.

A single objective optimization model is established to minimize the total vehicle transportation cost, as shown in

$$Z = \min \sum_i \sum_j \sum_k C_{ij} x_{ijk} (i \neq j), \quad (1)$$

where Z stands for the total cost; i and j represent any two points; k represents any transport vehicle; and C_{ij} represents the transportation cost from point i to point j . The variable x_{ijk} is defined as follows:

$$x_{ijk} = \begin{cases} 1, & \text{The transportation task from } i \text{ to } j \text{ is completed by } k, \\ 0, & \text{Otherwise.} \end{cases} \quad (2)$$

Considering the actual situation, this paper puts forward three optimization objectives. That is, the shortest delivery time, the shortest driving distance of all delivery vehicles,

TABLE 1: Experimental environment.

Hardware environment	Pentium D
Operating system	Windows
Implementation software	Matlab
Web server	IIS5.1
Programming language	C#
Database system	Access
Hard disc	1 T
Display card	512 G

and the least cost, and the established optimization model is shown in

$$\begin{cases} \min Z_1 = \sum_i \sum_j \sum_k T_{ij} x_{ijk} (\text{The shortest time}), \\ \min Z_2 = \sum_i \sum_j \sum_k D_{ij} x_{ijk} (\text{The shortest distance}), \\ \min Z_3 = \sum_i \sum_j \sum_k C_{ij} x_{ijk} (\text{The least cost}). \end{cases} \quad (3)$$

The establishment of the model is based on certain assumptions: the travel time, distance, and the demand of consumers between any two points can be known, that is, the distribution model under certain conditions.

Let m be the quantity of ants in the ant colony; d_{ij} is the distance between the city i and j ; $\eta_{ij}^k(t) = 1/d_{ij}$ is the heuristic function; τ_{ij} is the amount of information on the path (i, j) at the moment t ; $p_{ij}^k(t)$ means that the ants k move from city i to city j at time t the probability of j . The formula is as follows:

$$p_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}(t)]^\beta}{\sum_{s \in \text{allowed}_k} [\tau_{is}(t)]^\alpha \cdot [\eta_{is}(t)]^\beta}, & j \in \text{allowed}_k, \\ 0, & \text{Otherwise,} \end{cases} \quad (4)$$

$$\text{tabu}_k (k = 1, 2, \dots, m). \quad (5)$$

Among them, j is the city that has not been visited yet; $\text{allowed}_k = \{C - \text{tabu}_k\}$ is the city that the ant k is allowed to choose next; α and β are heuristic factors; tabu_k is used to record the city that the ant k has walked through at the moment of t . The ant is not allowed to repeatedly pass through this cycle, and the taboo table is cleared after the end of this cycle. The ant completes a cycle and updates each path pheromone:

$$\tau_{ij}(t+n) = (1-\rho) \cdot \tau_{ij}(t) + \Delta\tau_{ij}(t), \quad (6)$$

$$\Delta\tau_{ij}(t) = \sum_{k=1}^m \Delta\tau_{ij}^k(t). \quad (7)$$

Among them, $\rho \in [0, 1]$ represents the pheromone volatilization coefficient; $(1-\rho)$ represents the pheromone resid-

ual factor; $\Delta\tau_{ij}$ represents the information amount increment on the path ij in this cycle. Initially, $\Delta\tau_{ij} = 0$; τ_{ij}^k represents the amount of information left by the ant k between cities i and j in this cycle. The formula is as follows:

$$\Delta\tau_{ij}^k(t) = \begin{cases} \frac{Q}{L_k}, & \text{If the } k\text{th ant passes through } (i, j) \text{ in this cycle,} \\ 0, & \text{Otherwise.} \end{cases} \quad (8)$$

Among them, L_k represents the path length of the ant k to travel around; Q is a constant. In the ACO, the parameter selection method and selection principle directly affect the computational efficiency and convergence of the ACO.

In the process of e-business logistics delivery, the weight of the goods is a factor that must be considered. In general, goods with heavy weight should be dispatched first to reduce fuel consumption during the delivery process. The weight index is expressed as:

$$S_g = \sum_i f(g-i)^{tu}. \quad (9)$$

In the formula, S_g represents the quantity of locations that need to be delivered; $\sum_i f$ represents the weight of the delivered goods; g is the order of locations for logistics delivery; i is the weight of the goods to be delivered at the i location; and tu is the weight index calculation factor. The establishment of the timeliness index reflects the timeliness requirements of the goods, calculated as follows:

$$S_t = \frac{1}{N} \left(\frac{t_a - t_i}{T_{fs}} \right). \quad (10)$$

In the formula, t_i represents the preservation time of the goods at the i delivery point; T_{fs} represents the time required to deliver the goods; t_a is the delivery departure time; $1/N$ represents the impact factor in the delivery process; and S_t is the arrival time of the goods. The customer importance index is expressed as:

$$S_j = \frac{1}{N} \left(\frac{N-i}{m_j} \right). \quad (11)$$

In the formula, $1/N$ represents the order of important customer goods; m_j is the priority selection factor; N represents the importance degree division factor; and i is the quantity of priority customers.

The hybrid intelligent algorithm proposed in this paper adopts a combination of deterministic selection and random selection in the strategy of determining the transfer path. That is, the ants can dynamically adjust the transition probability of the current state during the path search process. The calculation formula of the transition probability is as

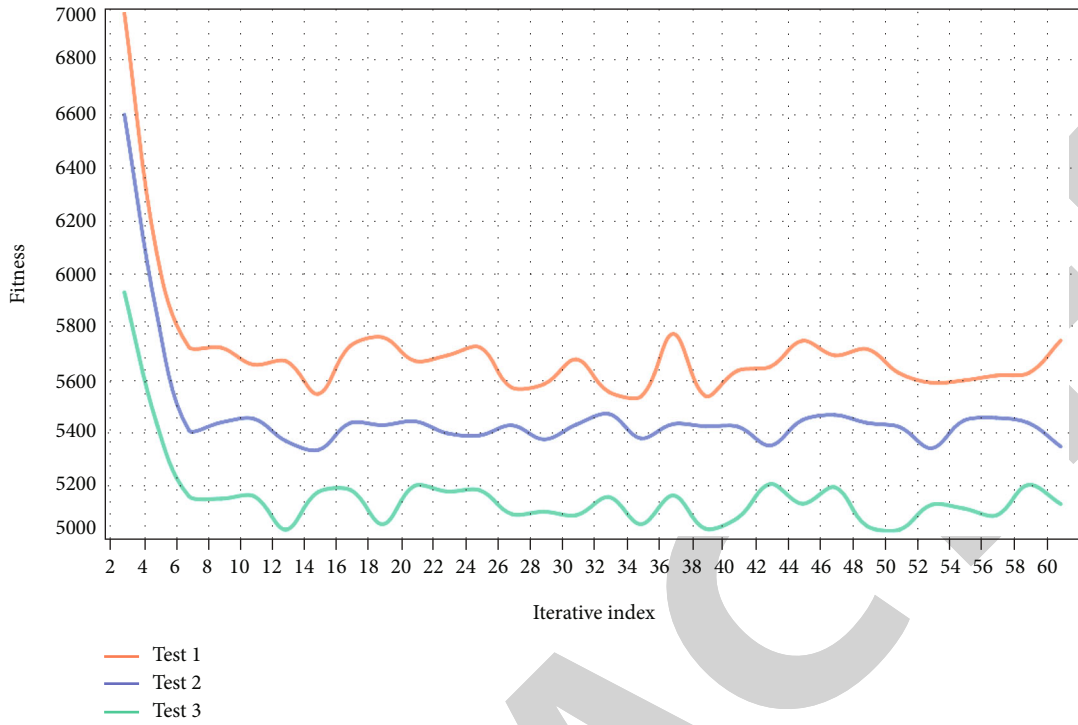


FIGURE 3: Comparison of convergence curves of algorithms.

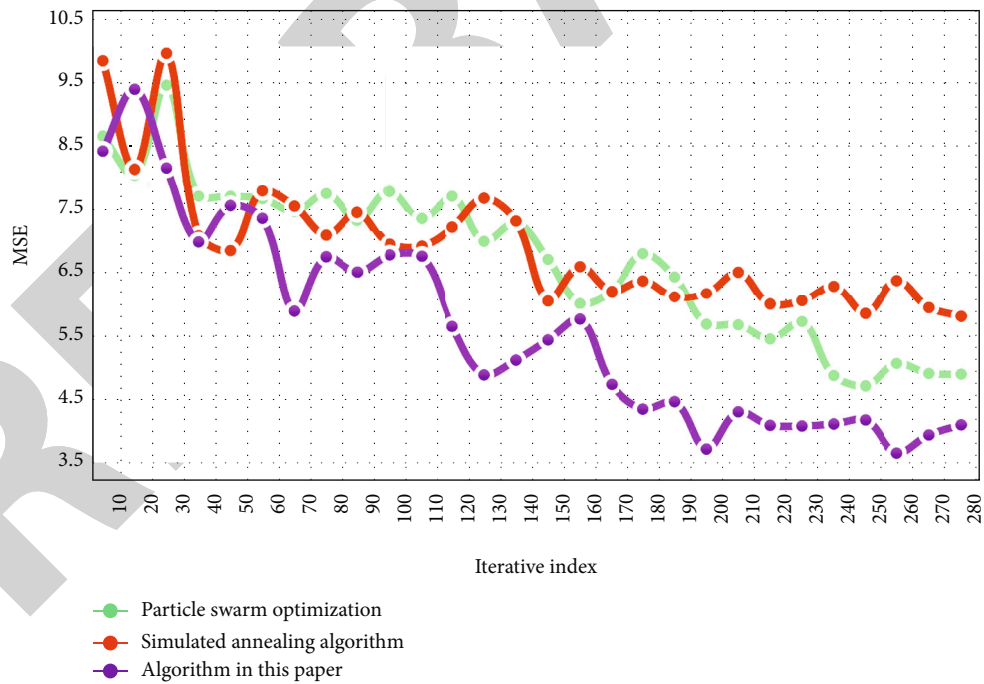


FIGURE 4: MSE test results.

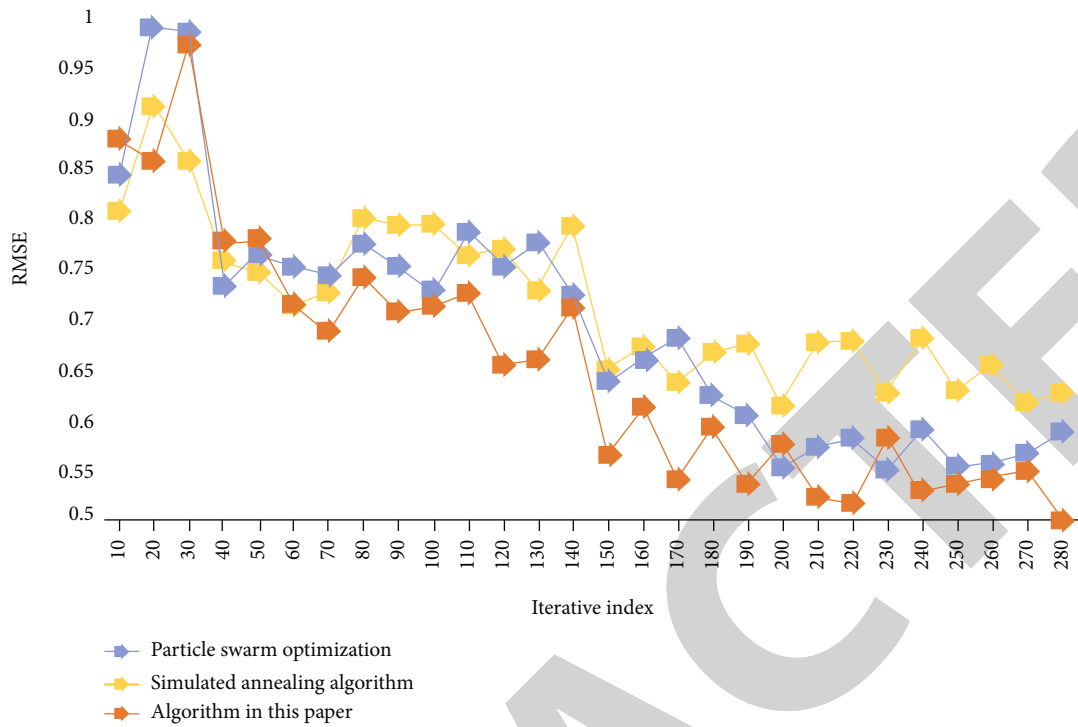


FIGURE 5: RMSE test results.

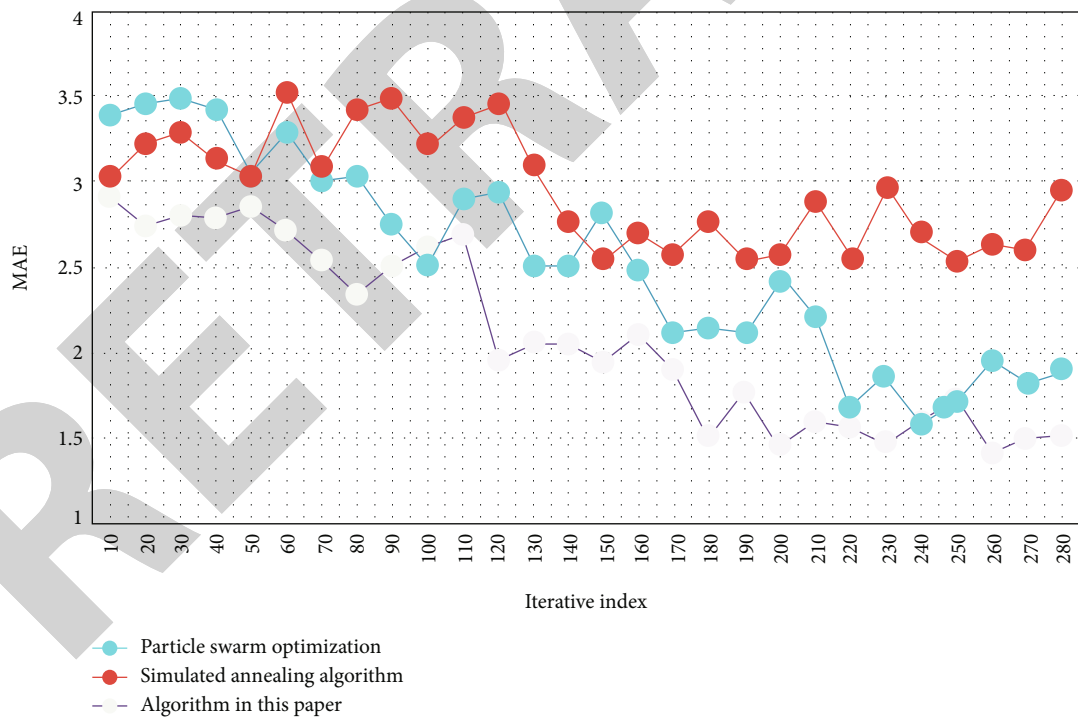


FIGURE 6: MAE test results.

TABLE 2: Comparison of errors of each algorithm.

Algorithm	MSE	RMSE	MAE
PSO	5.189	0.551	1.578
Tabu search algorithm	6.254	0.768	2.754
Simulated annealing algorithm	6.192	0.651	2.657
Algorithm in this paper	4.054	0.509	1.465

follows:

$$J = \begin{cases} \max_{h \in U_i^k} \{ [\tau_{ih}(t)]^\alpha, [\eta_{ih}(t)]^\beta \}, & \text{If } q < q_0, \\ p_{ij}^k(t), & \text{Else.} \end{cases} \quad (12)$$

In the formula, a random number with a uniform distribution between (0, 1) is randomly generated before the path selection in

$$q_0 < [0, 1]. \quad (13)$$

If the random value is less than the threshold q_0 , the largest one in $\{ \tau_{ih}(t)^\alpha, \eta_{ij}(t)^\beta \}$ is selected from the set of feasible cities. The city where this maximum value is located is the next transfer path; otherwise, it is calculated according to the method of calculating the transfer probability in the basic ACO. The process is shown in the following formula:

$$p_{ij}^k(t) = \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}(t)]^\beta}{\sum_{h \in U_i^k} [\tau_{ih}(t)]^\alpha \cdot [\eta_{ih}(t)]^\beta}. \quad (14)$$

After determining the transition probability, this paper also introduces the concept of crowding degree in AFSA. The calculation formula of crowding degree q_{ij} is shown in the following formula:

$$q_{ij} = \frac{2\tau_{ij}}{\sum_{i \neq j} \tau_{ij}}. \quad (15)$$

The key to the introduction of the congestion degree is to determine a threshold according to the actual situation of the problem and use $\delta(t)$ to represent the congestion degree threshold at the moment of t , if the following formula is satisfied:

$$q_{ij} < \delta(t). \quad (16)$$

It means that the current path is not too crowded, and the ant k selects this path as the path to move in the next step. Otherwise, the ants reselect a random path within the feasible neighborhood to transfer. Among them, the congestion degree $\delta(t)$ is updated as follows:

$$\delta(t) = 1 - e^{-ct}. \quad (17)$$

At the same time, in the process of evolution, by introducing an information exchange mechanism, the information can be transmitted between the two populations, which helps individuals avoid wrong information judgment and fall into the local optimum point. A nonlinear dynamic adaptive inertia weight strategy is adopted to improve the performance of the algorithm. Its update status is as follows:

$$w(t) = w_{\text{end}} + (w_{\text{start}} - w_{\text{end}}) \times \exp\left(-k \times \left(\frac{t}{t_{\text{max}}}\right)^2\right). \quad (18)$$

Among them, k is the control factor; it controls the smoothness of the change curve of w and t . Calculate the total path index:

$$S_k = \frac{1}{F} \left(\frac{n+1}{h_y \cdot j} \right). \quad (19)$$

In the formula, S_k represents the distance between the k distribution points; $1/F$ represents the return distance after the delivery is completed; $h_y \cdot j$ is the adjustment coefficient of the total path; and n is the total path index value.

Build an ant colony of a certain scale. From the starting point, each ant chooses the path to move to the next node according to the pheromone concentration of each path. The advantages and disadvantages of each path are reflected by the amount of pheromone released. Every ant's transfer process is a solution, repeated and circulated until the best solution is found. In this paper, the pheromone concentration is updated after all ants have completed a complete optimization process. The pheromone-exerting mechanism weakens the influence of ants' experience earlier in time. Due to the different emphasis on the distribution objectives of goods in logistics delivery centers, the distribution objectives are weighted. Before setting, the obtained objective function is dimensionless.

VRP focuses on the path planning between a supplier and K sales points, which can be briefly described as: given one or more centers (central garages), a vehicle set, and a customer set, vehicles and customers have their own attributes, each vehicle has capacity, and the goods carried cannot exceed its capacity. In this paper, the "extreme difference" dimensionless processing method is adopted. By embedding the obtained data into the evaluation function, the comprehensive evaluation value can be obtained. For the problem of logistics delivery route, the smaller the comprehensive evaluation value, the better. However, this method is suitable for the situation where there are many alternatives. In the case of few schemes, the continuous elimination method can be implemented by comparing two schemes, and finally the best scheme can be selected. Minimizing the quantity of vehicles used is the first optimization goal, which has a higher priority. Therefore, the solution with less vehicles is always better than the solution with more vehicles, although this may lead to the increase of vehicle running costs. The delivery service of key users should try to ensure punctuality, so the penalty coefficient of

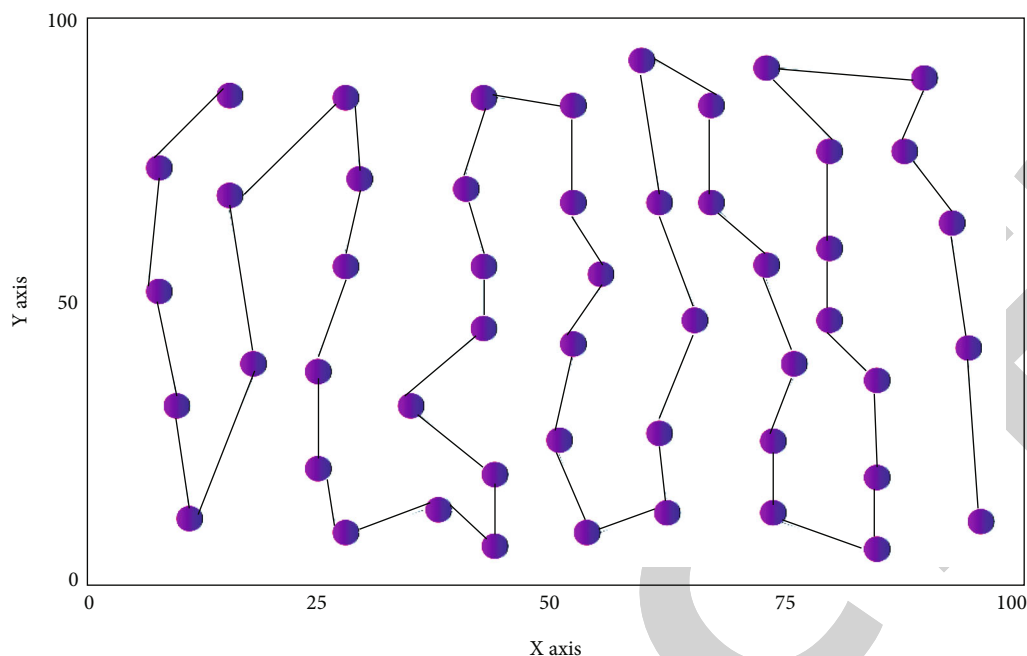


FIGURE 7: Calculation results of the algorithm.

deviation between service time and customer target time in delivery should be higher than that of ordinary users. The introduction of customer importance factors can highlight the important customer value in logistics delivery decision, give priority to ensuring its distribution service, and achieve the balance of customer satisfaction.

4. Result Analysis and Discussion

The urban logistics distribution path planning problem is essentially a vehicle path optimization problem, which can be defined as: under the condition that the distribution center and the customer point are known, find an optimal vehicle distribution path scheme, and deliver the goods to the customer within the specified time, usually with the goal of minimizing the distribution cost. In order to verify the rationality of the proposed model and the effectiveness of the solution algorithm, this section simulates the proposed model and algorithm. Test question bank by simulating VRPTW standard. In VRPTW, test data including 25, 50, and 100 customer nodes are set according to the scale of problem solving, and it is agreed that each demand point of goods has corresponding time window constraint and corresponding demand of goods. Moreover, it is assumed that each cargo transportation task point in the distribution network has corresponding time window constraints and certain demand, and the maximum load capacity of each vehicle is given. The actual running time of vehicles between customers is completely determined by the physical distance between customers. Firstly, the performance of the two-population hybrid algorithm is tested by using 10 city outlets. The quantity of ants is 100, the quantity of iterations is 500, the heuristic factors are 1 and 2, and the pheromone volatilization coefficient is 0.3. In this section, under the Windows operating system, the simulation experiment of model realization and algorithm solution is carried out with

Matlab software. The specific experimental environment is shown in Table 1.

The soft time window means that if the delivery vehicle fails to start service within the time required by the customer, a certain amount of compensation will be paid; the longer the delay, the more compensation you have to pay. If the delivery vehicle is earlier than the earliest start time required by the customer, additional waiting time will be generated, which will affect the rest of the delivery staff's work arrangement and their salary. In this paper, the total transportation area is divided according to the constraints such as vehicle load, and then, the optimal transportation route is designed in the divided subareas. This can reduce the space of single optimization search, greatly reduce the amount of calculation, and improve the speed and accuracy of solution. The convergence curve pairs of the algorithm are shown in Figure 3 below.

The distance between customer nodes in the distribution network is calculated by the Euclid distance formula of two-point coordinates, and it is assumed that the running time of vehicles between two points is equal to the transportation distance. The purpose of the test is to test whether the HSIA proposed in this paper can achieve the optimal distribution path and test the related performance of the HSIA. The MSE (mean-squared error) of the algorithm is shown in Figure 4. RMSE (root mean square error) of the algorithm is shown in Figure 5. The MAE (mean absolute error) of the algorithm is shown in Figure 6.

In order to reflect the test results more intuitively, this paper draws the test results of various errors in the above figure into tables. Table 2 shows the error comparison of each algorithm.

In this paper, a delivery route optimization model based on differentiated satisfaction is proposed. Through customer set division, key customers, ordinary customers, and both

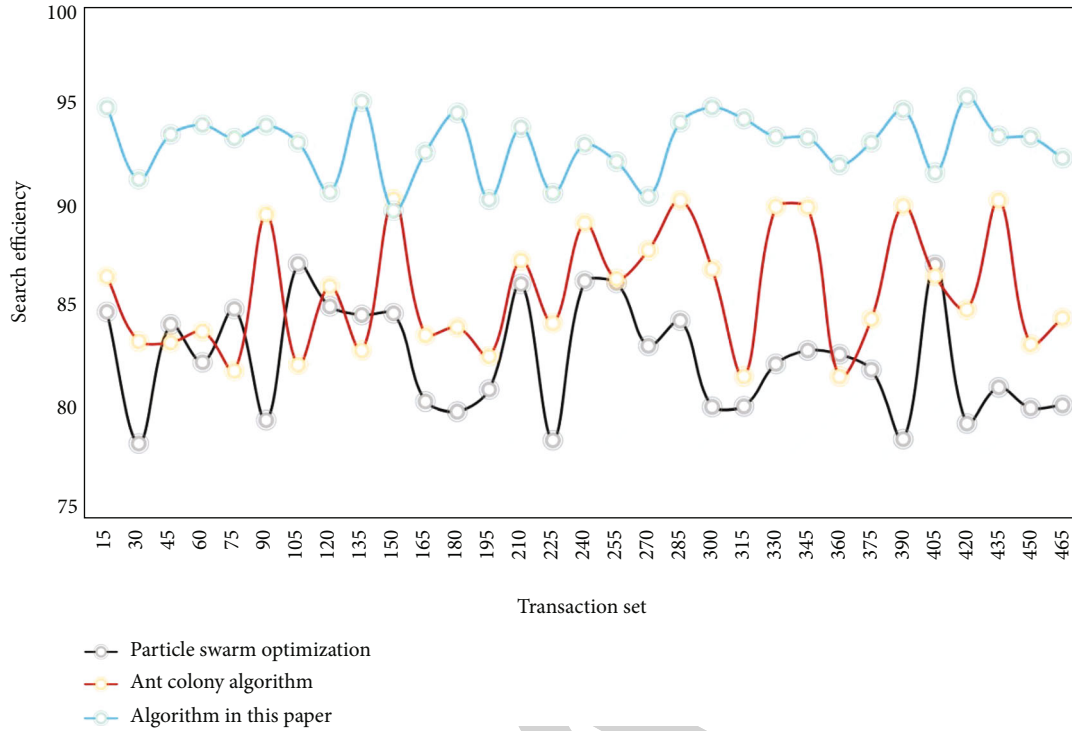


FIGURE 8: Comparison of search efficiency of algorithms.

TABLE 3: Comparison of time consumption of different algorithms.

Index	PSO	Tabu search algorithm	Simulated annealing algorithm	Algorithm in this paper
Best solution	676	662	669	672
Worst solution	689	697	696	681
Average value	681	690	688	678
The quantity of times to reach the best solution	15	5	9	21

customers are distinguished. The differentiated satisfaction of different customer roles is used to participate in the decision-making in route optimization to improve the satisfaction of key customers of enterprises, and a balanced route scheme with lower transportation cost is produced. The calculation results of the algorithm are shown in Figure 7.

In this paper, the hybrid population algorithm absorbs the advantages of different algorithms, and it reaches the best solution the most times, which increases the probability of the best solution. In addition, because of the communication mechanism in different populations, the algorithm improves the search efficiency of populations.

The algorithm proposed in this paper is carefully tested on other examples in Solomon standard test data source. In order to make the test results more accurate and reliable, each instance is operated 20 times under the same hardware and software configuration, and the average of these operation results is taken as the running result of the algorithm and compared with the settlement results of other algorithms. The comparison of search efficiency of the algorithm is shown in Figure 8.

It can be seen that there is a big difference in running time between HSIA and PSO for path optimization. At the same time, the PSO has some shortcomings, such as slightly insufficient search precision and accuracy, and high-dependence on parameter setting. As a result, the total length of the optimal path searched by the PSO is farther than the search result of the hybrid algorithm.

An example is used to test the performance of the HSIA, and the results are compared with those of other algorithms. The experiment was conducted independently for 50 times, and the statistical results are shown in Table 3.

The results show that, although other algorithms can improve the convergence speed, there are many worst solutions obtained by the algorithm, which indicates that the algorithm is easy to converge to the local minimum solution. This paper can improve the global optimization ability of the algorithm, get a better solution, and be relatively stable. The AC comparison of the algorithm is shown in Figure 9.

Test results show that the AC of this algorithm is 95.08%, which is higher than that of traditional PSO and general heuristic algorithm. In this paper, the advantages of different

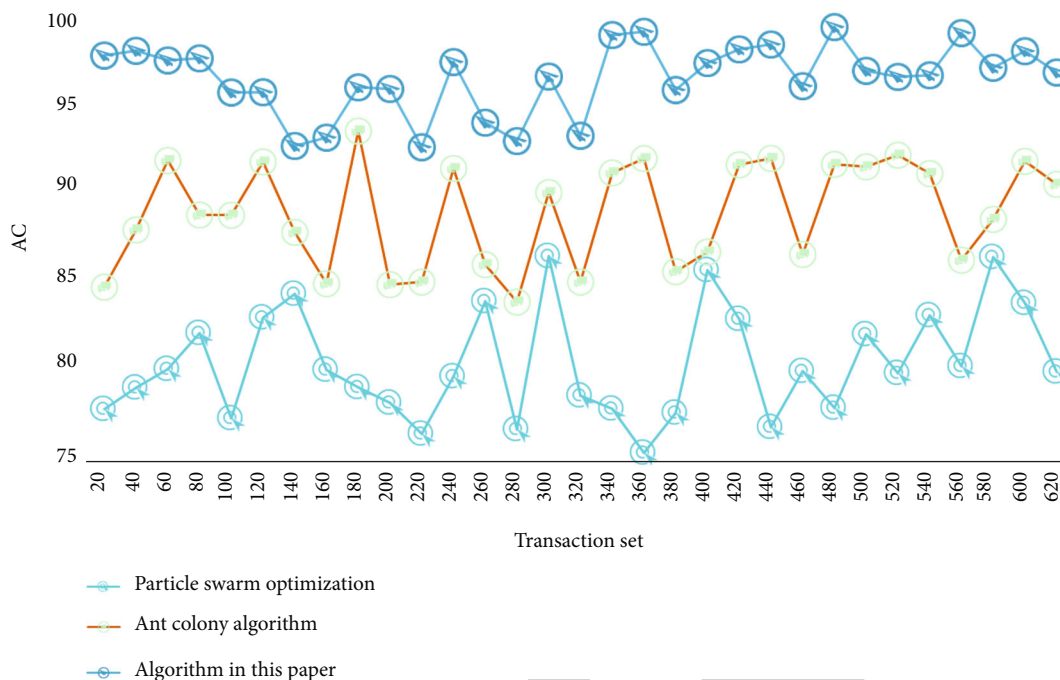


FIGURE 9: AC comparison of algorithms.

algorithms are taken into account at the same time, and the crowding degree of fish school is introduced into the iterative process of ACO. When the initial feasible solution is obtained by using artificial fish school, the ability of the algorithm to obtain the global optimal solution is improved by gradually changing the crowding degree. Therefore, the HSIA proposed in this paper does have strong global optimization ability. The test results in this section fully demonstrate that our algorithm has certain advantages in performance. It can be completely applied to the optimization of logistics delivery path.

Based on the above experimental results, it can be seen that the e-commerce logistics distribution path optimization algorithm under the big data background designed this time can save a lot of logistics distribution costs compared with the traditional algorithm. It conforms to the economic effectiveness of e-commerce logistics distribution path optimization. Compared with the actual optimal path, the error values of the path optimized by the e-commerce logistics distribution path optimization algorithm are lower than those of the traditional algorithm. The above experimental results show that the path optimized by the e-commerce logistics distribution path optimization algorithm is more consistent with the actual optimal path. The accuracy of the e-commerce logistics distribution path optimization algorithm is verified.

5. Conclusions

With the growth of the Internet, the growth of e-business has also been advanced by leaps and bounds. In the current environment, the logistics industry must keep pace with the growth of the Internet and actively innovate and shorten the distribution route, which can significantly reduce the logistics cost. This plays an important role in reducing material loss and improving benefits. Based on this, this paper deeply discusses

the application of swarm intelligence optimization algorithm in logistics delivery path optimization under the background of big data. In this paper, under the constraints of time window, delivery vehicles, and customer demand, the specific delivery requirements and delivery time of each logistics branch are obtained through big data, and based on this information, the distribution is unified, and the distribution route is planned. In this paper, in order to apply the improved algorithm to solve VRPTW, the hybrid algorithm is discretized, and a local path optimization operator is introduced. At the same time, the concept of crowding degree in the AFSA is introduced into the ACO. In the early stage of the optimization process, a strong crowding degree limit is set to ensure that most ants are not affected by the pheromone concentration to conduct random optimization. Test results show that the AC of this algorithm is 95.08%, which is higher than that of traditional PSO and general heuristic algorithm. The algorithm achieves the best solution the most times and increases the probability of the best solution. The test results of this paper fully show that the algorithm in this paper has certain advantages in performance. It can be completely applied to the optimization of logistics delivery path. However, when the algorithm is applied in practice, it should be constantly updated with the growth of urban roads.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Retraction

Retracted: Exploration and Analysis of Collaborative Matching Algorithm Empowering Large Data Fiscal Budget

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Gong and X. Zhang, "Exploration and Analysis of Collaborative Matching Algorithm Empowering Large Data Fiscal Budget," *Journal of Function Spaces*, vol. 2022, Article ID 8171318, 11 pages, 2022.

Research Article

Exploration and Analysis of Collaborative Matching Algorithm Empowering Large Data Fiscal Budget

Rui Gong ¹ and Xu Zhang ²

¹Enrollment and Employment Division, Changde Vocational Technical College, Changde 415000, China

²School of Public Finance and Taxation, Zhejiang University of Finance and Economics, Hangzhou 310018, China

Correspondence should be addressed to Xu Zhang; 20150039@zufe.edu.cn

Received 2 August 2022; Revised 15 September 2022; Accepted 19 September 2022; Published 10 October 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Rui Gong and Xu 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.

At present, many budget companies have realized high-level computerized accounting and have high-speed and advanced database management systems. In order to adapt to the growth of budget entity informatization, the financial department must improve its own financial budget level. Introduce information technology, and develop corresponding financial budget software to better perform its own financial calculation function. This paper explores and analyzes the full collaborative matching algorithm of enabling big data financial budget and builds a financial big data financial budget platform through offline and online data collection. Combined with the matching algorithm and linear regression prediction algorithm in data analysis, the financial data is analyzed in depth. Finally, the PSO-SA algorithm runs 10 rounds on the datasets with matching scales $d = 10$ and 30, respectively. It is found that the fluctuation is large when $d = 10$ and relatively stable when $d = 30$. When $d = 10$, the maximum value is 7.2354, and the minimum value is 6.9969. When $d = 10$, the maximum value is 26.6403, and the minimum value is 23.9599. It can be concluded that when $d = 0$, PSO-SA can obtain a relatively good matching scheme, but it is easy to fall into a local optimal solution. The superposition effect of “data + computing power + algorithm + scenario” can help enterprises make better decisions. Embed complex analysis into daily management and trading scenarios to build a financial empowerment platform. Make the increasingly complex work more automated and intelligent, and improve financial efficiency.

1. Introduction

With the popularization of technologies such as Internet+, mobile Internet, and cloud computing, the amount of data in human society has experienced explosive growth and has entered the era of large data. The rapid increase in the amount of data and the increasing competition between countries and enterprises require governments and enterprises to use a large amount of data to provide customers with products and services more accurately, quickly, and individually. Large data tech is a comprehensive data engineering application tech that includes statistical analysis, data mining, artificial intelligence, parallel computing, natural language processing, and data storage [1]. Big data audit refers to the construction of a data-based audit work mode based on the original data of the auditor’s database. It forms an audit intermediate table by collecting, converting, sorting,

analyzing, and verifying the underlying data. And use query analysis, multidimensional analysis, data mining, and other technical methods to build a model for data analysis. Discover trends, anomalies, and errors; grasp the overall situation; and highlight the key points and accurately extend, so as to collect audit evidence and achieve audit objectives [2]. Today, large data-related technologies have profoundly changed and affected the growth of the entire society, and even the way of thinking. The transformation of human beings in turn affects the value system of human beings. Large data tech has attracted much attention. Therefore, it is necessary to centrally store large data and build a unified financial budget system; it is necessary to increase the intensity of financial budgeting, innovate financial budgeting methods, improve the efficiency and quality of financial budgeting, and gradually explore the application of large data tech in financial budgeting work. Budgeting in a timely

manner is conducive to macro-control and timely detection of problems, thereby improving the capacity, quality, and efficiency of budgetary work. In recent years, the Commissioner's Office has made some attempts in the supervision of fiscal revenue and the comprehensive fiscal supervision of central budget units. However, using data to analyze and predict the results is still superficial, because only historical data is used for simple linear analysis. . However, in the current financial budget information systems, due to the lack of a unified caliber, it is impossible to compare and coordinate, and it is difficult to carry out data analysis. (1) Project budget audit: Audit is a key work of the National Audit Office. The number of units within the central budget is huge and widely distributed and has a lot of data, making auditing difficult. In this process, the use of ABC classification method has played a great role in the audit of project budget. By establishing a database of key projects of local budget units, compiling ABC analysis charts, scientifically classifying project budget amounts, and comparing them with the data of the previous and next years, the quality and work efficiency of the audit are guaranteed. (2) Supervision progress data: use progress analysis to supervise the budget implementation of central basic budget units. Every half year, the third quarter, and the whole year collect statistics on the overall execution rate, basic expenditure, project expenditure, and "three public" funds execution rate of each department and compare them with the work progress, annual progress, and annual progress of each department. Analyze the budget execution of each department. Communicate the budget implementation and problems of financial departments at all levels so that financial departments at all levels can better understand and grasp the use of financial funds. (3) Payment audit data analysis process: First, by comparing the number of audits and the number of audited accounts over the years, we learned about the payment audit over the years. Secondly, the change of audit workload is analyzed. The third is to compare the changes in the total amount of audit funds over the years and reflect the overall progress of financial funds. On this basis, it analyzes and discusses the problems in the implementation of direct financial payment in my country. (4) Local government regulation and policy analysis: (1) Analysis of the scale and structure of central government revenue. Using the linear analysis method, this paper forecasts and analyzes the changes in the scale and structure of local government revenue. The various tax and non-tax revenues of the year are compared, and the changes in their scale and structure are analyzed and forecasted. . It is necessary to implement the research on the impact of the new corporate income tax law on local finance.

The core significance of big data audit lies not only in the reform of audit mode, but also in the breakthrough of audit ideas. In the era of big data, auditors need to make good use of big data thinking and awareness to improve the methods and ways to find clues to problems. Only by maximizing the use of audit results and improving audit efficiency can we achieve our audit goal—to find problem clues, evaluate risks, and reveal system defects [3]. Making good use of large data tech to carry out financial work is a way for the country to promote data transparency, sharing, and openness and

improve national data capabilities. . Secondly, with the continuous growth and improvement of Internet tech and the popularization of informatization of the financial budget system, various business and management information are gradually transformed into paperless, information flow, and automation. It has become more and more complex, getting rid of the business trajectory of paper finance of traditional financial budget and extending the scope of financial budget from internal data to external data, from traditional financial budget data to business data, resulting in an explosion of financial information. Growth, which poses a great challenge to the processing capacity of traditional financial budget analysis systems [4]. The traditional data analysis process is that professional data analysts discover possible doubts by observing and analyzing data, combined with past experience, repeatedly observe and analyze them, find problems, and provide corresponding suggestions for specific responsible business personnel, or It is the business personnel who put forward the corresponding requirements, and the analysts conduct related query and analysis according to the requirements [5]. All in all, it starts with the details of business data and draws relatively accurate conclusions with the professional knowledge and years of experience of financial personnel. Through large data tech, it is possible to obtain and comprehensively analyze the entire financial-related data, conduct analysis and mining, and then discover the internal connection hidden between the data, improve the financial system and personnel's insight into problems, and gradually strengthen the entire industry. The growth direction of the industry and the introduction of relevant systems and other macro-level understandings make predictive thinking and macro-level strategic deployment of the growth strategy of the entire industry and put forward specific growth goals that are staged, feasible, and assessable [6]. Therefore, this paper proposes a research on collaborative matching algorithm to empower large data financial budget. Collaborative matching algorithm and large data tech can gradually realize automatic financial budget and comprehensive financial budget, save labor costs, and help improve the quality and improvement of financial budget. The efficiency of financial budget and the research on collaborative matching algorithm to empower large data financial budget are of great significance.

In recent years, the function of the local regional budget to execute the financial budget has been continuously expanded, and it has played an active and effective supervisory role in many fields such as budget preparation management, non-tax revenue collection, and performance management [7]. In the era of large data, the level of informatization of budget execution units is increasing day by day, and financial and budget organs at all levels have begun to actively explore the application of emerging technologies such as large data financial budgets in budget execution, and have achieved initial results. However, at present, there are still some problems in the implementation of financial budgets of local regional budgets in China, such as imperfect laws and regulations, relatively lack of financial budget resources, and single technical means of financial budgets, which lead to limited functions [8]. For future work, the

Ministry of Finance and Budget has put forward the goals of integrating financial budget resources, focusing on the budget implementation of financial budget departments, conducting comprehensive financial budgets for budget units and subordinate units of budget units, and expanding the scope of financial budgets. At present, many budgeted units have achieved a high level of accounting calculating and have high-speed and advanced database management systems. In order to adapt to the growth of the informatization level of the budgeted units, the budgetary department must improve its own financial budget level and introduce information [9]. The platform uses large data tech to realize an all-round financial budget including overall budget, departmental budget, and special budget. Combined with related algorithms such as collaborative matching, association rules, and regression prediction in data analysis, it conducts in-depth analysis of financial budget data and discovers data. Therefore, the innovation of this paper is as follows:

- (1) According to the characteristics and existing problems of the current financial budget, the collaborative matching algorithm and big data technology are used to integrate the financial and business data of the budget unit and its related units. Combine manual labor with on-site financial budget execution system, and use structured query language and other relevant analysis languages.
- (2) The fiscal budget model reflects the advantages of “unified analysis, discovery of doubts, and decentralized verification,” thus realizing the transformation of the fiscal budget model from loose to joint, which helps to improve the work efficiency of the fiscal budget department

This paper is divided into five parts. The first section describes the research background and significance of collaborative matching algorithm enabling big data financial budget. Section 2 makes a multi-angle, multilevel, dynamic, and efficient statistical analysis of historical data from the perspective of big data matching algorithm technology. Section 3 analyzes the key technologies of big data audit and data mining. At the same time, the matching association rules are analyzed. This means that in the frequent itemset, some data can be derived from other data and reach the lowest confidence level. Section 4 is the experimental analysis. This part carries out experimental verification on the dataset to analyze the performance of the model. Section 5 is the conclusion and prospect. This part mainly reviews the main contents and results of this study. Finally, the full text is summarized. This paper expands the scope, breadth, and depth of financial budget. It also greatly shortened the time of members of the financial team on the audit site and effectively improved the efficiency of financial budget work.

2. Related Work

Large data matching algorithm tech is a multi-angle, multilevel, dynamic, and efficient statistical analysis of historical

data based on the use of efficient data analysis and matching tech to provide support for prediction, judgment, matching, and decision-making [10]. Since the 1990s, this tech has been used in various fields such as science and tech, business, economy, and public management and has continuously innovated management models, effectively reducing management costs and improving management levels. The application of large data in financial budget supervision has broad prospects and is of great significance.

Kilby et al. found that a hybrid approach of professional judgment and data mining can produce more accurate financial budget forecasts [11]. Singh et al. went further and argued that the analysis should be combined with qualitative data mining methods from financial budget management databases [12]. Aragonés et al. used neural networks and case-based reasoning, as well as the selection of two markets and the choice of passive or active trading strategies, to generate significantly more prediction of financial budgets holding returns [13]. Moghaddam et al. employ neural networks and use financial ratios and macroeconomic variables to predict market returns [14]. Wang et al. propose a deep learning approach to neural networks to construct financial distress prediction model [15]. Ahmadi et al. used the CART optimization algorithm to study the prediction performance and significance test of the short-term capital adequacy ratio of Chinese-listed companies [16]. Yu et al. believe that the combination of professional equipment and big data technology can provide a good help for professional budget analysis [17]. Koyuncugil proposed a financial risk early warning system model based on data mining [18]. Zhao examined the behavioral impact of large data on auditors' judgment and discussed issues such as information overload, information relevance, pattern recognition, [19]. Ming et al. argue that large data provides a complementary source of evidence for the budget function and should be budgeted based on adequacy, reliability, and relevance. The evidence standards framework assesses its use [20].

However, matching algorithms and large data tech will bring new financial thinking and directions in financial budgeting and give birth to new financial budgeting methods. Financial personnel need to keep pace with the times and actively learn new financial budgeting thinking modes and methods. The impact of large data on the growth of financial budgets. Fundamentally solve the loopholes and problems of sampling budget. Financial personnel collect all the data of the objects subject to the financial budget and use the complete and comprehensive business data information to conduct multidimensional analysis of the data to discover hidden financial budget problems. All in all, large data-related technologies provide professional financial personnel with new financial budgeting methods, can control the overall data, and help professional financial personnel find overlooked problems from a macro and more comprehensive perspective.

3. Methodology

3.1. Key Technologies of Large Data Auditing. Big data audit refers to the construction of a data-based audit work mode

based on the original data of the auditor's database. It forms an audit intermediate table by collecting, converting, sorting, analyzing, and verifying the underlying data. And use query analysis, multidimensional analysis, data mining, and other technical methods to build a model for data analysis. By discovering trends, anomalies, and errors; grasping the overall situation; highlighting the key points; and accurately extending, we can collect audit evidence and achieve audit objectives. Hadoop is an open source distributed storage computing platform of the Apache Foundation, which can store and process large amounts of data through simple program patterns. The infrastructure of Hadoop includes Hadoop Digital File System (HDFS) and MR framework. HDFS is a data storage system containing named nodes and data nodes. MR is responsible for the processing of data, including tracking, tracking, tracking, and working. The HDFS and MR processes are performed on the same computer, allowing direct communication. In the Hadoop ecosystem, there are many other elements, basically based on HDFS and MR. HBase is a distributed key-value database with randomness and range search performance. The software uses HDFS as the base storage, but is augmented by specific indexes and storage structures. All Hadoop clusters are built on a resource manager, and YARN is the default resource manager for Apache Hadoop. These components provide efficient parallel data processing services, enabling fast and reliable analysis of structured and complex data. The data processing of the Hadoop architecture is organized according to the operation, including three elements: data input, configuration, and program. This project is divided into two parts: One is Map and the other is Reduce (Figure 1).

3.2. Data Mining. Data mining refers to analyzing massive data and extracting useful information from it. In the past, data mining was only for knowledge discovery in databases, that is, to convert data into useful knowledge. The KDD process involves a series of data preprocessing to transform the raw data into a suitable format for subsequent analysis. Data mining technology is to convert existing data into patterns or models, then post-process them to evaluate the correctness and practicability of the extracted patterns and models, and integrate them into DSS using appropriate methods. This provides end users (business analysts, scientists, planners, etc.) with relevant information. Today, data mining represents the entire process of knowledge discovery in databases (KDD) and has been widely recognized as a powerful and general data analysis tool. The complete data mining process generally includes evaluating and specifying business objectives, data sources, data processing, and building analytical models using data mining algorithms such as logistic regression or neural networks Figure 2.

Figure 3 shows the overall architectural design of the data mining system, where data mining includes many tasks that can be used according to the needs of a specific application context, or even combined. Data mining tasks are usually divided into prediction tasks and description tasks. A prediction task refers to building a useful model for predicting the future behavior or value of some feature. These

include classification and prediction, i.e., from a set of data objects with known class labels, predicting the classification of objects with unknown class labels; in descriptive data mining tasks, built-in models are designed to be understandable, effective, and efficient. Form to describe the data. A related example of a descriptive task is data representation, whose main purpose is to summarize general characteristics of the target data class.

The overall design of the data mining project needs to divide the functional modules of the system based on the clear system requirements. Divide the work of system development and define the interface between each module. To prepare for the later detailed design and implementation. The data needs to be read first, that is, the data is read from the log file, and then the data put into the memory set is matched into the required data, and the matched set is sent to the server. The server receives the data, saves the data to the database, and the data enters the database for integration.

3.3. Matching Algorithm. Matching algorithm, also known as matching mining, is a task to find interesting relationships between data and data such as frequent patterns, associations, correlations, or causal structures, including two aspects: frequent itemsets and matching association rules. Frequent itemsets indicate that in a dataset, some data items appear frequently and exceed a set threshold. These data itemsets are called frequent itemsets. Matching association rules means that in the frequent itemset, some data can be derived from other data, and the lowest confidence level is reached, suggesting that there may be a strong relationship between different data, such as the implication of $X > Y$.

3.3.1. Apriori Algorithm. Support: Measure frequent itemsets; the support of an itemset is the proportion of records that contain this set in the dataset.

$$\text{sup port}(X, Y) = P(XY) = \text{number}(XY) / \text{number}(\text{AllSamples}). \quad (1)$$

Confidence or credibility: defined for association rules, for example, the confidence of the rule $YZ \Rightarrow X$ is the conditional probability of X appearing when YZ appears. According to the definition of support, it can be transformed into the following formula:

$$\text{confidence}(YZ \Rightarrow X) = P(X|YZ) = \text{sup port}(X, Y, Z) / \text{sup port}(Y, Z). \quad (2)$$

Principle: If an itemset is frequent, then all its subsets are also frequent; conversely, if an itemset is infrequent, then all its supersets are also infrequent. For example, $\{1\}$ is an infrequent item, then $\{0, 1\}$, $\{1, 2, 3\}$, etc. are infrequent; using this principle can avoid the exponential increase in the number of itemsets and reduce the time complexity of the algorithm. In the subsequent description, C_k is used to represent the list of candidate itemsets, and L_k is used to represent the list of frequent itemsets.

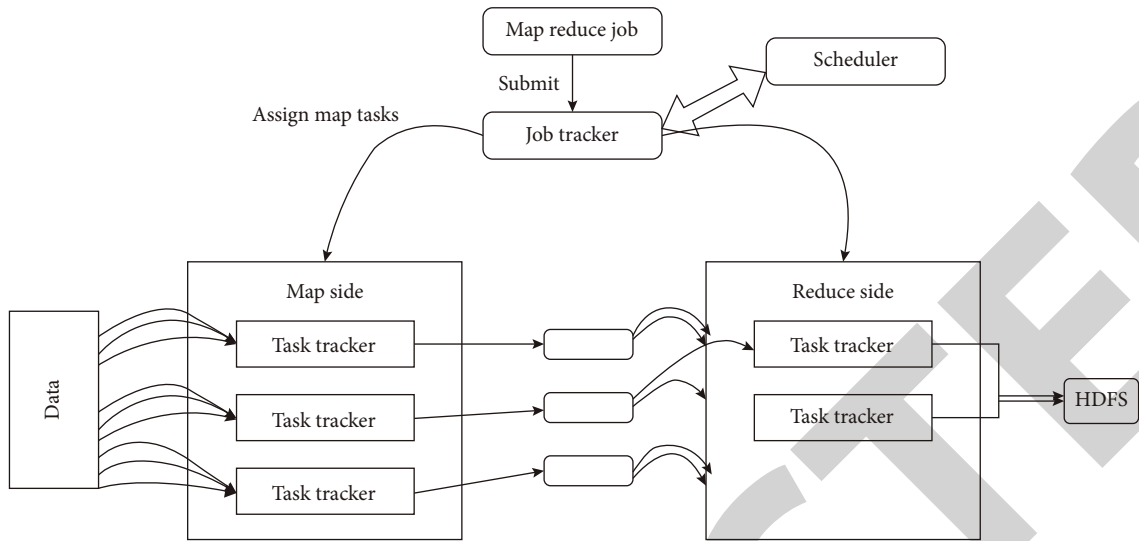


FIGURE 1: Map Reduce distributed workflow.

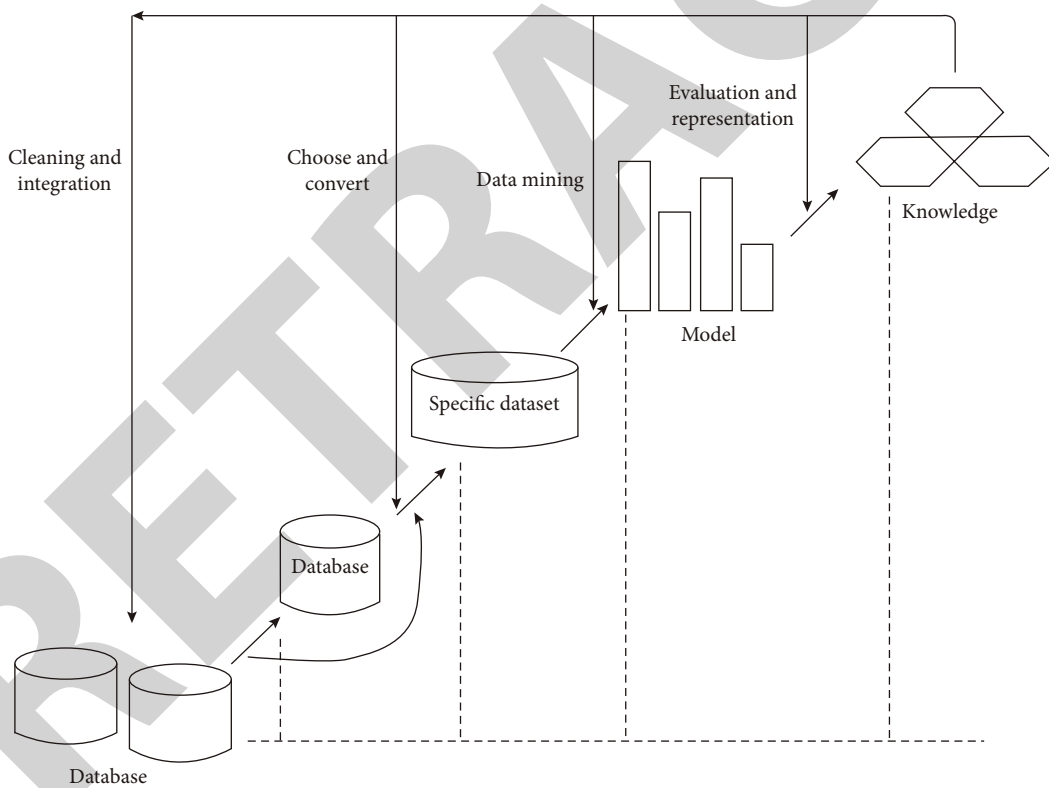


FIGURE 2: Schematic diagram of the data mining process.

3.3.2. *Apriori Algorithm Improvement (Table 1)*. The candidate itemset of Apriori algorithm is generated by the upper-level frequent itemset. The method of generation is to compare whether the first $k - 1$ items of the two elements in L_k are the same. If they are the same, then combine the two items to generate a candidate of $k + 1$ elements. Items, the advantage of doing so is to reduce the number of tra-

versals to get duplicate itemsets. For example, it is known that $\{0, 1\}$, $\{0, 2\}$, and $\{1, 2\}$ are frequent itemsets, and now they generate ternary candidate itemsets. If each set is merged in pairs, it will be merged three times, and the same result $\{0, 1, 2\}$ is obtained, but if the merge is performed only when the elements of the previous item are the same, only one merge is required, which reduces the number of

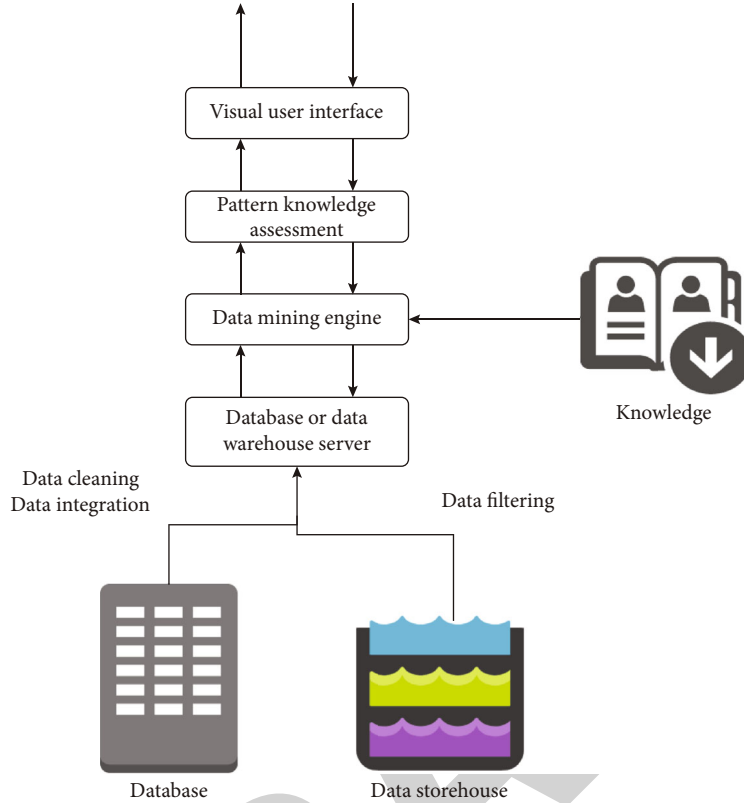


FIGURE 3: Overall architecture of the data mining system.

TABLE 1: Algorithm performance comparison.

	Apriori	Improved Apriori
C_1	514	514
L_1	95	95
C_2	4370	841
L_2	130	130
C_3	586	0
L_3	0	0
Scan dataset (number of times)	3	2
Time (S)	185	68

merges to obtain duplicate values. This method can effectively reduce the size of the candidate itemset and improve the efficiency of the algorithm. But not all candidate itemsets generated by frequent items are frequent itemsets, and the data of candidate itemsets need to be further reduced so that fewer comparisons are made when scanning the dataset. In order to further reduce the number of candidate itemsets and make full use of known infrequent itemsets, a filter set can be added during the generation process of candidate itemsets. Supersets are removed directly. The filter set generation formula can be expressed by the formula:

$$filter = C_k - L_k. \quad (3)$$

TABLE 2: Time complexity analysis of PSO-SA.

Step name	Time complexity
Update particle velocity and displacement	$O(Sn)$
Adjust solution by constraints	$O(Sn)$
Update local and global optimal particles	$O(Sz(n))$

3.3.3. PSO-SA Algorithm. Cooperative matching evolution algorithm is easy to implement and has a faster convergence speed, but it is easy to fall into the local optimal solution, which is characterized by strong convergence and low population diversity. The PSO evolution strategy used by PSO-SA is similar to the standard PSO strategy. Both the optimal position of the group and the historical optimal position of the individual are used to control the current speed of the particle. Table 2 shows the results of PSO-SA time complexity analysis. λ_1 and λ_2 Influence on inertia factor. (Table 2).

3.4. Regression Analysis. Regression analysis is a set of techniques and tools used in statistics to explore relationships between variables. In its simplest form (simple linear regression), with one variable considered as the dependent variable and one variable as the independent variable, ordinary least squares (OLS) is used to estimate the linear regression line. In regression analysis, you need to collect data on many variables and then determine if there is an actual relationship between those variables, and if so, the equation can be used

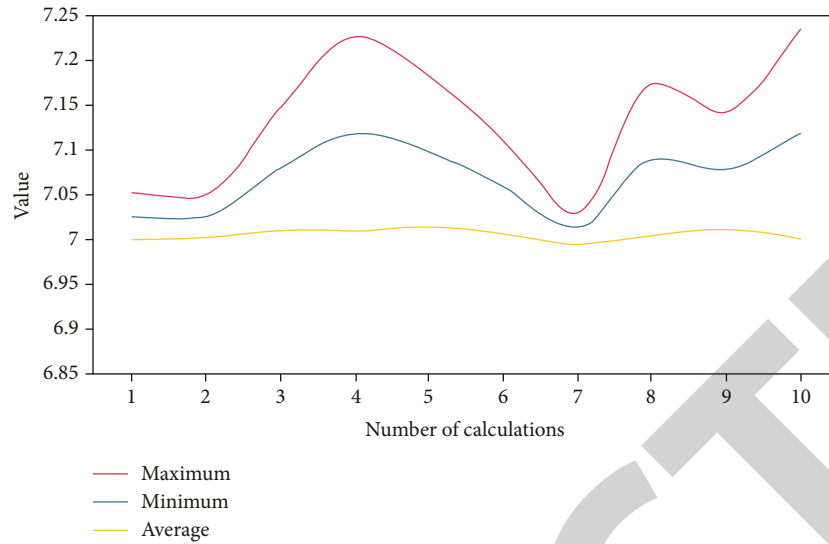


FIGURE 4: The statistics of the matching scheme solved by PSO-SA ($d = 10$).

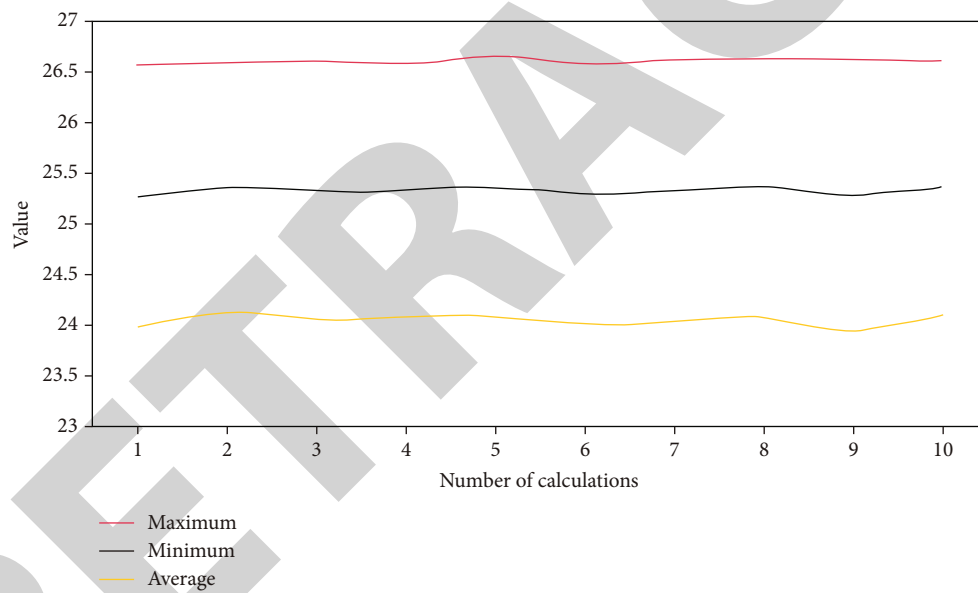


FIGURE 5: Statistics of PSO-SA solving matching scheme ($d = 30$).

to predict the value of the dependent variable if the independent variable takes on a specific value. Among them, the relationship model of these variables needs to be assumed first, and then a regression model is established, and the correlation coefficient is calculated according to the predicted value to determine whether the model is correct. Regression analysis is usually used to predict numerical targets. It is a predictive modeling technique that studies the relationship between independent variables and dependent variables. It is a very practical and common prediction algorithm.

The linear regression algorithm refers to the output of a regression equation composed of multiple input characteristic variables and uses the most suitable straight line (regression line) to establish the relationship between the

dependent variable Y and one or more independent variables X . Multiple linear regression has many methods (>1) independent variables, while simple linear regression has only one independent variable.

Simple linear regression is also called univariate linear regression. The relationship between the independent variable and the dependent variable is represented by a straight line, and the linear correlation between the two variables is studied; X represents the independent variable, and Y represents the dependent variable. The regression equation is as follows:

$$Y = aX + b. \quad (4)$$

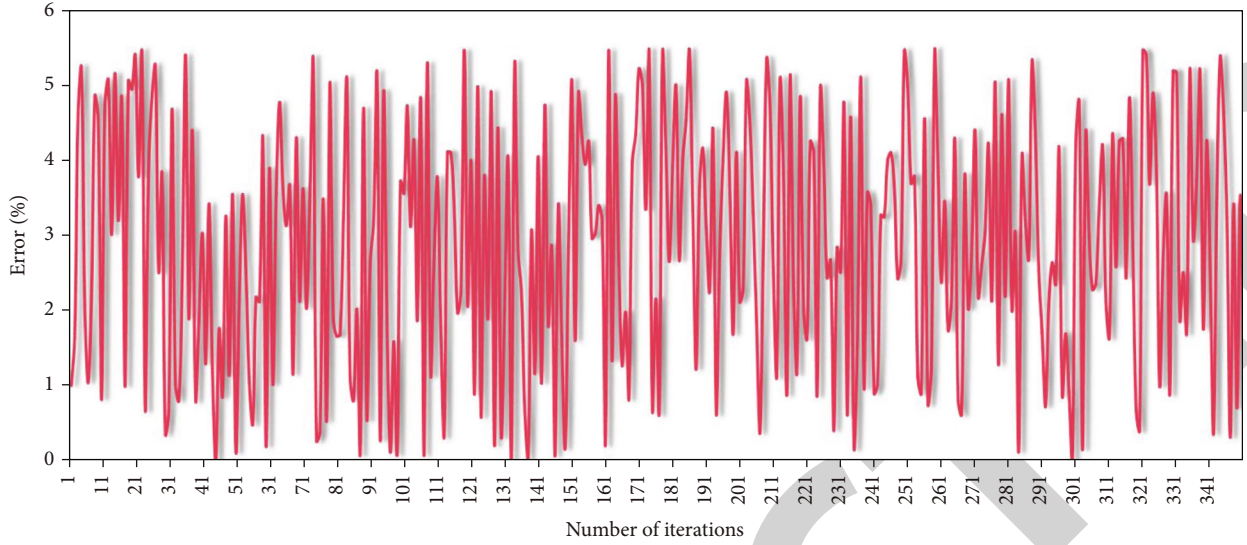


FIGURE 6: The relationship between the number of iterations and the error.

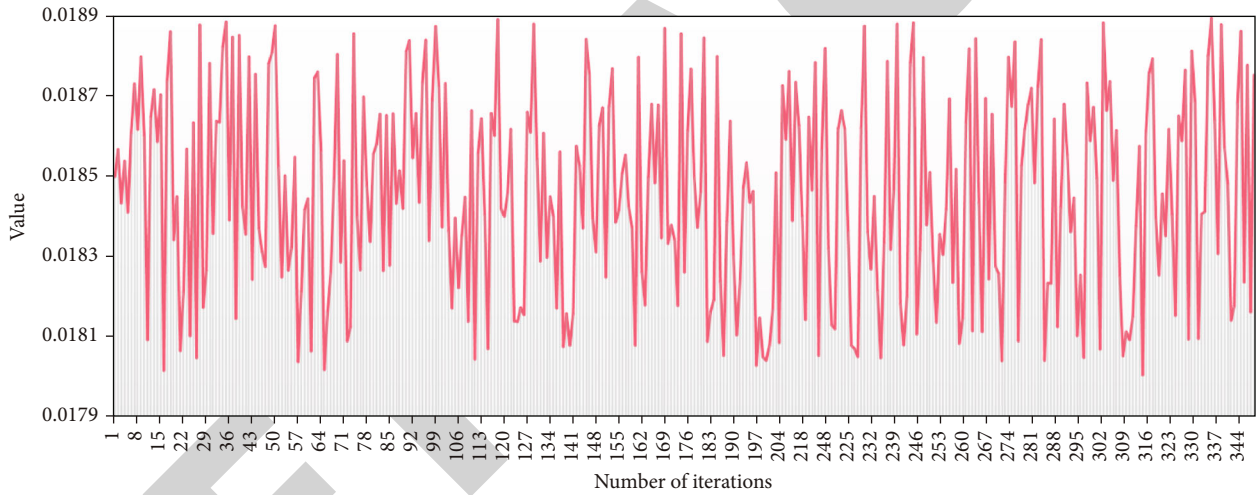


FIGURE 7: The relationship between the number of iterations and variance.

Among them, a and b are the regression coefficients required; the regression coefficients are solved by minimizing the sum of squares of the error between the actual value and the predicted value. The equation of the sum of squares of the error is as follows:

$$Q(a, b) = \sum_{i=1}^n (Y_i - (aX_i + b))^2. \quad (5)$$

Expand and simplify the above equation to get

$$Q(a, b) = n\bar{Y}^2 - 2an\bar{X}\bar{Y} - 2bn\bar{Y} + a^2n\bar{X}^2 + 2aBn\bar{X} + nb^2. \quad (6)$$

Partial derivatives with respect to a and b , respectively, and making them equal to zero, we get

$$a = \frac{\bar{X} \cdot \bar{Y} - \bar{X}\bar{Y}}{(\bar{X})^2 - \bar{X}^2}, \quad (7)$$

$$b = \bar{Y} - a\bar{X}.$$

Multiple linear regression has multiple independent variables, and the regression equation is as follows:

$$y_i = w_1x_{i1} + w_2x_{i2} + w_nx_{in}. \quad (8)$$

X represents an input data, Y represented by a vector, n represents an input data with n independent variables, and the regression coefficient is represented by a W vector.

The establishment of the regression model is mainly to solve the regression coefficient W vector. The least squares

method can be used to find the value that minimizes the error of w . In fact, the least squares method is the loss function of the linear regression model, which minimizes the sum of squares subtracted from the actual value and the predicted value, as long as the loss function takes the minimum value. Parameter is the desired parameter.

4. Result Analysis and Discussion

In the experiment, the horizontal comparison algorithm includes the PSO-SA algorithm. The vertical comparison group is serial collaborative matching and parallel collaborative matching. HPSOSA adopts SA strategy to try to jump out of the local area only when PSO-SA mutates many times, but the optimal value remains unchanged. The results of parallel collaboration are few, mainly represented by PSO-SA, which adopts PSP for collaborative matching (Figures 4 and 5).

On the datasets with question matching scales of $d = 10$ and 30, the PSO-SA algorithm was run for 10 rounds, respectively, and it was found that the fluctuation was large when $d = 10$, while it was relatively stable when $d = 30$; when $d = 10$, the maximum value was 7.2354, and the minimum value is 6.9969; when $d = 30$, the maximum value is 26.6403, and the minimum value is 23.9599. It can be seen that when $d = 10$, PSO-SA can obtain a relatively good matching scheme, but it is easy to fall into the local optimal solution, which is characterized by convergence.

It can be seen from Figures 6 and 7 that as the number of iterations increases, the error and variance are also different, changing all the time, and there are errors.

5. Conclusions

Data is the basic support for financial auditors to carry out audit work and effectively exert the function of the “immune system.” The financial budget execution financial budget data comprehensive analysis system developed in this paper collects the backup data of all financial business systems and financial budget software; uses early warning, query, and multidimensional models for unified analysis; and further analyzes and confirms the doubts queried by each model. Each audit team conducts decentralized inspections, which expands the coverage of audits, transforms the traditional single department budget execution budget to reflect individual problems into a centralized reflection of general problems in the budget execution process of all budget units, and expands the scope of the financial budget. The breadth and depth also greatly shorten the time of the members of the financial team at the audit site and effectively improve the efficiency of financial budget work.

5.1. Collect and Process Data to Build a Solid Foundation for Data Analysis

- (a) Collect data and establish a financial database. The acquisition and storage of massive data is a prerequisite for data analysis. Therefore, when the commissioner’s office uses big data for budget supervision,

the first thing to do is to gradually expand the collection of data, the collection of data, the collection of data, the collection of data, the collection of data, the management of data, the management and management of data, etc., aspects of work. First, we must do a good job in the collection of financial data. In order to meet the requirements of “interconnection,” this paper promotes the research on treasury information system, department budget system, and government procurement system of the Commissioner’s Office. Second, gradually collect social and economic data. Gradually collect local economic and social data, and establish a working mechanism for data collection with local finance, taxation, statistics, People’s Bank, and other departments. Only on the basis of data sharing and openness can big data-related technologies be used for data mining and analysis

- (b) Strengthen the processing of financial data, and build a quality system for financial information. Financial data is an important carrier of the financial department, and its problems will have a significant negative impact on the financial decision-making of the financial department. In order to ensure the accuracy and integrity of financial data, the authenticity of data collection, analysis, and other links, it is necessary to build a financial data quality management system: first, to build a financial data structure and standardize the content of financial data and second, to build financial data. The quality management system uses a variety of methods such as grouping technology, comparative analysis technology, and model technology to construct the quality management process of financial data

5.2. *Introducing and Cultivating Talents and Providing Method and Technical Support.* The use of big data for financial budget management requires relevant departments to have a certain understanding of financial work and have certain financial management experience; the second is to have certain mathematics, statistics, data management, and data mining technology and be able to carry out from a global perspective. Dig and analyze. The training of professional and technical personnel is a long-term and systematic work. In the future, in the process of applying big data technology to financial budgeting, it is necessary to introduce a group of experts in the field of data analysis to mine complex and uncertain objects and perform specific data mining on them. At the same time, train the staff of various departments of the Commissioner’s Office on data analysis and application so that they can accurately mine data through simple data tools or according to the requirements of supervision work. At the same time, due to the characteristics of high starting point and new technology for data analysis, this paper proposes a method of using technical personnel to analyze data to enhance cooperation with relevant universities and research institutes.

Auditing based on big data replaces the traditional auditing model. Local audit institutions should fully introduce big

data audit technology, combine big data technology with the existing digital audit platform, and through multi-angle, multifield, and in-depth analysis, find that the audited entity exists in financial, business, or fund management. The main measures are as follows: (1) Using big data technology, audit institutions should establish and improve dynamic analysis mechanisms. At present, some regions have realized the collection of quarterly data on budget execution, which improves the timeliness and timeliness of data and provides a good opportunity for local audit institutions to build a dynamic analysis mechanism based on big data. Audit institutions can carry out dynamic big data analysis on all aspects of the raising, distribution, and use of budget funds, so as to achieve the whole process of auditing of budget units and synchronous and continuous audit supervision and timely discover and resolve potential risks and risks of budget units. (2) Audit institutions should reform the data analysis model of budget execution audit. At present, the data model used by my country's audit institutions is relatively simple, and it is difficult to conduct a more comprehensive and accurate screening of audit doubts. Therefore, audit institutions should use big data technologies such as machine learning, natural language processing, and distributed stream processing to analyze the audit data of budget execution so that it can play a good role in improving the efficiency of audit work. For example, using machine learning methods, based on the core business indicators of the enterprise, build enterprise internal control evaluation indicators to build enterprise business risk evaluation and internal control risk evaluation models, so as to accurately identify the operational risks and internal control weaknesses of the budget execution unit. Use natural language processing technology to analyze text data of important meetings and internal programs of budget execution units. It overcomes the shortcomings of traditional audit that cannot identify and analyze a large amount of text data and changes the text and unstructured data that cannot be used effectively before. The use of distributed data processing technology can improve the automation of data collection, format conversion, data sorting, etc., thereby improving the standardization of budget execution audit work and thereby improving the overall budget execution audit work efficiency.

5.3. Carry Out Data Mining and Analysis Pilot Projects and Build a Budget Supervision System. The application of big data mining technology for budget supervision is still in the initial exploratory stage in my country, and it is relatively lacking in manpower, technology, and experience. When using big data for data mining and analysis, it is necessary to correctly select special projects with large data volume and high data quality for data mining and analysis, constantly summarize the phased results, and promote them within the scope of the commissioner's office so that big data technology has gradually become the basis for the Commissioner's Office to carry out budget supervision, and a budget supervision system of the Commissioner's Office based on big data technology has been built. Big data is the trend. With the continuous deepening of my country's fiscal budget supervision and fiscal informatization, the use of big data

technology to carry out fiscal reform and budget management will help promote the modernization of my country's government governance system and governance capacity and improve the scientificity and effectiveness of fiscal budget management. It plays an important role in deepening the reform of the fiscal and taxation system.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] W. U. Ya-Kun, H. X. Guo, and X. M. Wang, "Research overview on large data tech," *Journal of Liaoning University(Natural Sciences Edition)*, vol. 34, no. 14, pp. 321–321, 2015.
- [2] H. Mao, "The change and growth of the detective work under the background of "large data"," *Journal of Jiamusi Vocational Institute*, vol. 4, no. 9, pp. 132–135, 2015.
- [3] T. Bareith, T. Tatay, and J. Varga, "The impact of changes in financial supervision on the profitability of the hungarian banking sector," *Economies*, vol. 10, no. 7, p. 176, 2022.
- [4] A. G. Isaev, "Distribution of financial resources within the budget system of the Russian Federation and regions' economic growth," *Spatial Economics=Prostranstvennaya Ekonomika*, vol. 4, no. 48, pp. 061–074, 2016.
- [5] A. Sadhu and S. Chakraborty, "Non-traditional machining processes selection using data envelopment analysis (DEA)," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8770–8781, 2011.
- [6] J. Begenau, M. Farboodi, and L. Veldkamp, "Big data in finance and the growth of large firms," *Journal of Monetary Economics*, vol. 97, no. 8, pp. 71–87, 2018.
- [7] S. Moon and J. Jung, "An analysis of the youth policy budget of the local government - focusing on the Seoul 10-year(2006–2015) budget," *Journal of Community Welfare*, vol. 58, pp. 201–226, 2016.
- [8] P. E. Jarris, J. P. Leider, B. Resnick, K. Sellers, and J. L. Young, "Budgetary decision making during times of scarcity," *Journal of Public Health Management and Practice*, vol. 18, no. 4, pp. 390–392, 2012.
- [9] H. U. Ying, Q. W. Jiang, and X. J. Luo, "Research on the relationship among budget function, president characteristics and organizational performance of public hospitals," *Chinese Health Economics*, vol. 8, no. 5, pp. 132–136, 2019.
- [10] Y. Leng, Z. Chen, and Q. Zhang, "Distributed clustering and filling algorithm of incomplete large data," *Calculating machine Engineering*, vol. 54, no. 36, pp. 757–762, 2015.
- [11] P. Kilby and D. Scigliuzzo, "Tumbling after budget forecast," *International Financing Review*, vol. 2110, no. 2099, pp. 45–45, 2015.
- [12] S. N. Singh, S. Hillmer, and Z. Wang, "Efficient methods for sampling responses from large-scale qualitative data," *Marketing Science*, vol. 30, no. 3, pp. 532–549, 2011.

Retraction

Retracted: Application of GA in Furniture Modeling Style Design

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Yu and C. Liu, "Application of GA in Furniture Modeling Style Design," *Journal of Function Spaces*, vol. 2022, Article ID 5264185, 11 pages, 2022.

Research Article

Application of GA in Furniture Modeling Style Design

Heng Yu ¹ and Chunjing Liu ²

¹*Sangmyung University, Seoul 03016, Republic of Korea*

²*Liaocheng University, Liaocheng, Shandong 252000, China*

Correspondence should be addressed to Chunjing Liu; liuchunjing@lcu.edu.cn

Received 25 July 2022; Revised 20 September 2022; Accepted 22 September 2022; Published 3 October 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Heng Yu and Chunjing 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 rapid development of the national economy, the construction industry is unprecedentedly active and has made great achievements. However, most traditional structural design is based on the personal experience of designers. In order to improve the production efficiency of furniture, it is necessary to optimize it. Since the evaluation criteria mainly depend on people's subjective thoughts, it is difficult to describe the individual's adaptive function. In this paper, an interactive innovation evolution system based on evolutionary algorithm is proposed. Taking full advantage of the evolution function of genetic algorithm and the modeling advantages of ACIS platform, a prototype innovation evolution system is developed based on ACIS/hoops platform with the genetic algorithm based on tree structure as the innovation support, which can help designers complete the innovation modeling design. The results show that the total processing time of the sorting scheme obtained by this algorithm is reduced by 2475.2 s, the production efficiency is increased by 20.6%, and the job waiting time is reduced by 15079.2 s. The results show that the algorithm is feasible and effective in solving the furniture production scheduling problem. The furniture production scheduling scheme solved by genetic algorithm can provide certain reference value for production personnel to formulate production scheduling scheme and improve production efficiency.

1. Introduction

With the changing consumption concept of the society, factors such as product innovation, artistic appearance, agreeableness, and environmental protection are paid more and more attention and occupy a prominent position in the market competition [1]. This trend urges enterprises to raise the design of product innovation, appearance modeling, human engineering, and other aspects to a new level when they start to carry out new product development, which also urgently requires further breakthroughs in the research of industrial design, so as to improve the corporate image, product design level, and market competitiveness [2]. Product modeling design is the creative design of the product's shape, color, surface decoration, and material, so as to give the product a new shape and new quality. In the process of actual development, enterprises need to continuously improve according to the development of the times, and only in this way can they ensure the quality of their own business development [3].

At present, most digital model design software adopts these two modeling methods. In this software, designers control the modeling changes of products through characteristic parameters [4]. The optimization of product modeling design is based on the perfect matching of each local detail. There are many design schemes for each local detail, and the final matching effect is endless. The parametric modeling method forces designers to spend energy and time on the parameter adjustment of product local modeling, instead of grasping and optimizing the overall modeling effect of the product [5]. However, when designing products, it is very difficult for designers to produce a reasonable scheme due to the limitations of knowledge, design experience, and design knowledge [6]. In addition, design is a group collaborative work project. The complexity of modern products makes it impossible for a single designer to be competent for complex design tasks. Therefore, the research of the new generation of computer-aided design must provide designers with design tools and support frameworks in a distributed environment to support enterprises to produce

high-quality, high reliability, low-cost, creative, and competitive products [7]. The synthetic part of the design has the following characteristics: creativity, multiple solutions, approximation, incompleteness, and empirical synthesis. At present, the optimization design problem of product modeling scheme is mostly performed by humans, it is difficult to form an exact system product scheme, and it is more complicated when faced with a large scheme set, so it is necessary to find a suitable mathematical algorithm for intelligent optimization [8]. Computer-aided design technology has been developed for decades, and it is becoming more and more mature. It has been widely used in various fields such as machinery, automobiles, aviation, and architecture and has become an indispensable auxiliary tool for modern engineering design [9]. CAD technology promotes the renewal and transformation of traditional industries and disciplines, realizes design automation, and enhances the competitiveness of enterprises and their products in the market. This trend urges enterprises to raise the design of product-oriented innovation, appearance modeling, ergonomics, etc., to a new level when starting new product development, which also urgently requires further breakthroughs in the research of industrial design, so as to improve the corporate image product design level and market competitiveness [10].

Under the above background, this paper studies the optimization method of parameters in parametric product modeling design and builds a ladder solution method to make the human-computer interaction in the whole design process more organized, so as to make up for the current parameterization and characterization. The method proposed in this paper is used in the detailed design stage of product modeling. At this stage, the method proposed in this paper helps designers to adjust and optimize the parameters of product modeling.

2. Related Work

In recent years, personalized products are popular in the market, and consumers have higher and higher requirements for the appearance of products, which makes the products have to be constantly updated with the changes of users' needs. In addition, the design characteristics of sofa products and the size and shape of internal parts largely depend on the appearance of the products, and the design mainly depends on the shape of the products.

Fabisiak studied the generation of building plans using GA. The novel building plans can meet multiple fuzzy constraints and target management. They also showed how evolution can generate new buildings by learning from famous architectural styles [11]. Colim et al. use computer simulation technology to generate artworks, such as table lamps and sculptures. They successfully organized the International Conference on Generative Art in Milan, Italy, on the theoretical research and application of evolutionary computing in architectural design, industrial design, art design, and music creation. An in-depth study was done [12]. Umentani et al. developed a garment style-aided design system by using genetic programming technology, which encodes a series of related sizes of styles into chromosomes, and the

system evolves styles according to users' choices, and others proposed interactive GA; that is, the fitness function is obtained through interaction with users, thus solving the evaluation problem of style fitness [13]. Soleimani and Kannan apply GA to computer-aided design and discuss the new progress of GA in computer-aided design system [14]. Jia et al. have done a lot of exploratory work in the description model and calculation model of innovation principle and innovation process and creatively solved the problems of knowledge expression and realization such as pattern composition, color, and description through intelligent technologies such as synthesis and analogy generation design [15]. Yao et al. improved the GA to optimize the truss structure and combined the multiplier method and the pseudoparallel GA to improve the premature phenomenon of the simple GA, and it was also quite effective for the optimization problems with complex constraints [16]. Sayed proposed a new hybrid algorithm, relative difference quotient-GA, and gave several examples in this paper. The calculated results show that the hybrid algorithm can significantly improve the computational efficiency and the ability to search the global optimal solution [17]. Smorkalov and Vorobeichik proposed to use GA in the optimization design of steel structures, discussed in detail the principle and implementation steps of GA, and put forward useful improvement suggestions for the advantages and disadvantages of GA, providing guidance for the application of GA in practical engineering [18]. Lotta et al. developed a clothing style-aided design system by using genetic programming technology, which encodes a series of relevant dimensions of styles into chromosomes, and the system evolves styles according to the user's choice. The fitness function is obtained to solve the evaluation problem of style fitness [19]. Sahu et al. optimize the furniture production line on the basis of GA, improve the running quality and running cost of the furniture production line, continuously improve the operating efficiency of furniture manufacturers, and promote the development of China's economy [20].

This paper closely combines the innovative research of conceptual design with computer technology and develops an environment to support the innovative design of sofa products by using computational models and computer tools, using the high information storage capacity and visualization means of computers. In this paper, GA is introduced into computer-aided design, which makes the innovative design of sofa products more intelligent and makes the generated images more creative. We are combining technologies such as machine learning, GA, and artificial neural network to develop an environment with independent intellectual property rights that supports innovative product design.

3. Methodology

3.1. Design Theory of Furniture Modeling Style Based on GA. The operating efficiency of furniture production line directly determines the production efficiency and production cost of furniture. In order to improve the production efficiency of furniture, it is necessary to optimize it. Based on this, this

paper will first introduce the operation status of furniture production line. Secondly, the balance of furniture production line is analyzed. Finally, the optimization measures of furniture production line based on genetic algorithm are analyzed. Finally, the furniture production line is optimized effectively. Since the object of this furniture production line optimization is mainly time and designated production elements, it is necessary to calculate the corresponding production beat according to the production output. Ensure that each production process is carried out independently, and the operation elements can only correspond to one workstation. In the process of allocating workstation time, the corresponding constraints must be met. The production time of each workstation cannot be greater than the constraints. On this basis, the furniture modeling style design based on genetic algorithm is determined. Modeling gene is a product modeling element determined according to the modeling characteristics of product semantic description, and it is the basic design element to express modeling style. The coding mode of product modeling genes is determined according to the modeling characteristics of modeling genes. For specific product modeling design, it is only necessary to distinguish the modeling characteristics of product modeling elements, and it is not necessary to accurately express their modeling characteristics. Therefore, fuzzy semantic quantification method is generally used to express the modeling feature information of product modeling elements. The modeling evolutionary design model of product semantic constraints is shown in Figure 1.

It includes two main parts: (1) the transformation between modeling design space and evolutionary design space and (2) shape evolution design. Semantic quantitative description is applied to realize the transformation from modeling design space to evolutionary design space, including the determination of modeling design elements, the coding of modeling genes, and the semantic quantitative description of target product modeling. In the process of actual product modeling design, in order to avoid the extreme value caused by the big difference with the actual design target, the modeling gene is set to be $[0.1, 0.9]$ and encoded in the form of real numbers. In the modeling design of product semantic constraints, a modeling gene string corresponds to a possible modeling design scheme, that is, a solution. The length of the gene string is the same as the number of product modeling elements, and one genetic gene represents one modeling element, and it corresponds one-to-one. Genetic genes use numerical values $[0.1, 0.9]$ to represent the level of modeling elements, so that a genetic gene type can specifically represent the modeling or structure type of a product modeling element. Taking the board desk as an example, its modeling genes and judgment criteria are shown in Table 1.

In Table 1, the judgment criteria of morphological elements are 0 and 1, respectively, indicating the straightness and curvature of the line. The judgment standard of color is 0 and 1, which, respectively, represent the cold and warm, simple, and gorgeous of 128 colors. The judgment standard of the connection is chosen. “independent 0” means that there is no common part between two parts, “crossing 0.5”

means that one part is embedded in another part, and “containing 1” means that one part contains another part.

The furniture production line studied in this paper mainly produces sofas. In the actual production process, the sofas are composed of left three positions and right three positions and pedals, and two pillows are installed on the left three positions. The production process of the whole sofa requires a total of 24 processes. In the process of actual research, time measurement is carried out for each process, and a total of 6 times are measured for a process, so as to ensure the accuracy of time detection. The average value is calculated according to the measurement results of 6 times, which is taken as the actual working time. The assembly drawing of sofa is shown in Figure 2.

In the actual process of sofa making, 24 processes can be divided into 10 parts. First, take the wooden frame, and second, nail white gauze, install springs, nail nets, and install elastic belts Third, spray water, paste three position cotton, and install foot steps. Fourth, set adhesive cloth. Fifth, nail the cloth. Sixth, nail the feet, and install the hardware frame. Seventh, install the headrest, and install the coat. Eighth, put the glue, put the bag; ninth, inspection; and tenth, packaging. The above steps are the general process of sofa making. In the process of studying the optimization of furniture production line, we need to take the above steps as the main object and use GA to study it.

The interactive GA in which the designer participates can solve this problem. The designer's personal concept is added to the process of optimization, and the evaluation and selection based on the fitness function are replaced by the designer's choice. In this way, the advantage of the search breadth of the GA is used, and the final solution of the design scheme can be matched with the designer's original design concept. At the same time, the designer's participation also brings some limitations to the GA: (1) the limitation of population size. Because of the limited perception ability of human beings, there are not too many candidate solutions (individuals) participating in the search and solution, and the population size cannot be too large. (2) Designers, as human beings, cannot bear too much work, so the whole search process cannot be like the original GA, where the selected parent crosses and mutates to produce offspring, and then, it takes hundreds of cycles for reproduction to converge. (3) The product modeling scheme (individual) must be expressed in perceptual physical form to ensure the rationality of the designer's choice, but not in abstract coding (chromosome).

3.2. Furniture Design Model Based on GA. The algorithm design in the optimization process of furniture production line mainly includes the following contents. First is coding. This process needs to be carried out using the sequence of job elements. The job elements are allocated according to the actual situation of the workstation, and the corresponding serial numbers are arranged into chromosomes. Second is coding translation. Only the sequence of job elements can be displayed in the above chromosomes. Therefore, in the process of translation and coding, chromosomes need to be allocated to corresponding workstations. The third is

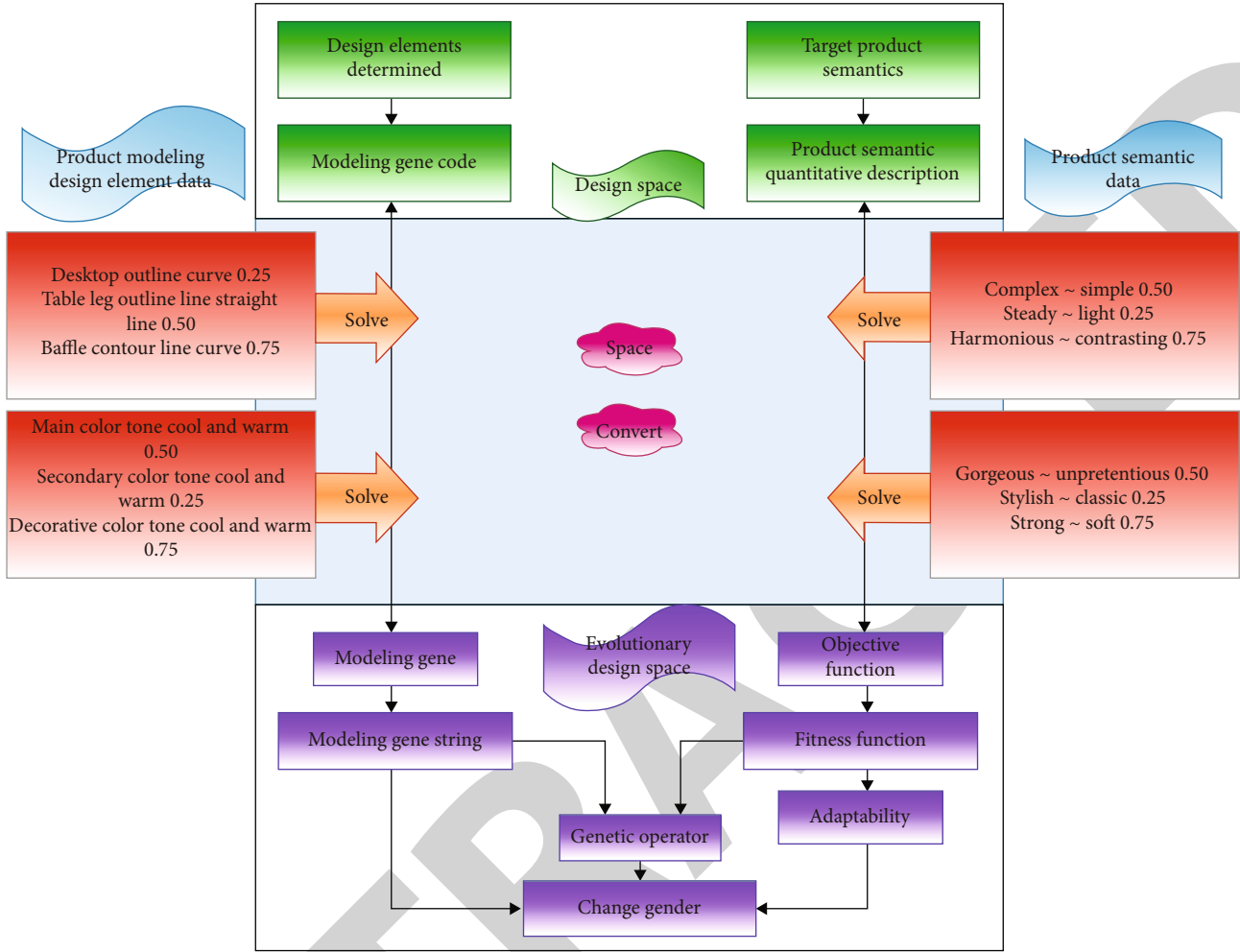


FIGURE 1: Modeling evolutionary design model with product semantic constraints.

TABLE 1: Modeling genes and judgment criteria of panel desks.

Shape feature category	Category number	Modeling gene	Judgment standard (value)	
			0	1
Morphological elements	1	Table top outline shape	Straight	Song
	2	Desktop bottom outline shape	Straight	Song
	3	Outline shape on both sides of the desktop	Straight	Song
	10	Desktop main color	Cold	Warm
	11	Desktop secondary color	Cold	Warm
	12	Desktop decoration color	Rustic	Gorgeous
Link relationship	19	Link form of table top and table legs	Cross 0.5	Contains 1
	20	Link form of table legs and baffles	Cross 0.5	Contains 1
	21	Link form between bezel and desktop	Cross 0.5	Contains 1

the selection operator. The fourth is the crossover operator. In this process, we need to use the crossover probability. The standard GA adopts fixed length binary coding. The advantages of this method are fine gene expression and long problem coding, which is conducive to solving combinatorial optimization problems. However, this method is not flexible enough, and it needs to map from coding domain

to problem domain. For the problems that the coding domain is consistent with the problem domain and the coding length changes greatly, the representation method of tree structure is more flexible. There are two ways to initialize the population: the first method is to manually enter the expression by the designer or user. The system provides designers with a floating panel for manual parameter input. This

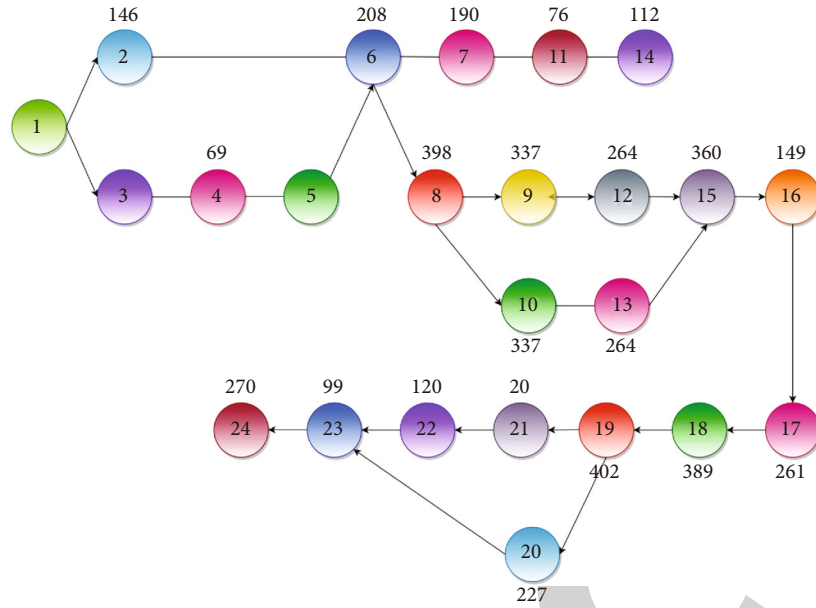


FIGURE 2: Sofa assembly.

method is suitable for designers and users with a certain mathematical foundation and requires a general understanding of the properties of the manipulated functions; the second is the random combination method. Each individual is actually a chromosomally characteristic entity. As the main carrier of genetic material, chromosome is a collection of multiple genes, and its internal expression, namely, genotype, is a combination of genes, which determines the external expression of individual shape. Therefore, at the beginning, we need to realize the mapping from phenotype to genotype, that is, coding. GA adopts natural evolution model, such as selection, crossover, mutation, and migration, and shows the process of GA. That is, the population is initialized randomly, and the fitness function of each individual is calculated. The fitness function refers to the function introduced to measure each chromosome in the problem in order to reflect the adaptability of chromosomes. All offspring mutate according to a certain probability. Then, the fitness function of the offspring will be recalculated, and the offspring will be inserted into the population, and the parent will replace it to form a new generation of offspring. This process will be executed circularly until the optimization criteria are met, as shown in Figure 3.

Adaptive crossover and adaptive mutation need to design formulas of adaptive crossover probability and adaptive mutation probability, which are closely related to individual fitness, according to the characteristics of GA, so as to effectively improve GA. The algorithm has a good adjustment effect for the problem that requires a small adjustment of the probability of crossover and mutation in the problem, but for the problem that requires a large range of adjustment of the probability of crossover and mutation, the optimization effect is not ideal. For example, in the early stage of evolution, too small variation range of crossover and mutation probability will make the excellent individuals basically unchanged, leading to the stagnation of optimization and

falling into local optimal solution. In this paper, the adaptive GA is effectively improved. The formula is as follows:

$$P_c = \begin{cases} P_{co} \times \log_2 \frac{f_m - f'}{f_m - f_a} + 1 & f^i \geq f_{avg}, \\ P_{co} & f^i < f_{avg}, \end{cases} \quad (1)$$

$$P_m = \begin{cases} P_{mo} \times \log \frac{f_m - f'}{f_m - f_a} + 1 & f \geq f_{avg}, \\ P_{mo} & f < f_{avg}. \end{cases}$$

For optimization problems, the situation is often very complex, and there are many types of objective functions and constraints. Before optimization, it is necessary to establish a mathematical programming model for the problem to be optimized. For a solving function minimization problem, the mathematical programming model is as follows:

$$\begin{aligned} \min \quad & f(X) \\ \text{s.t.} \quad & X \in R \\ & R \subseteq U. \end{aligned} \quad (2)$$

If the function $h \in L2(Rn)$ is radial, there is a function $\phi \in L2(R)$. For $h(x) = \phi(\|x\|)$, the following formula holds, where $\|x\|$ represents the Euclidean norm of x and its Fourier transform is also radial. A general expression of radial basis function is

$$h(x) = \phi \left((x - c)^T E (x - c) \right), \quad (3)$$

where ϕ represents the radial basis function, C represents the central vector of the function, and E is a transformation matrix, usually Euclidean matrix. It is a measure of the distance between the input vector x and the center C in the sense of the definition of matrix E .

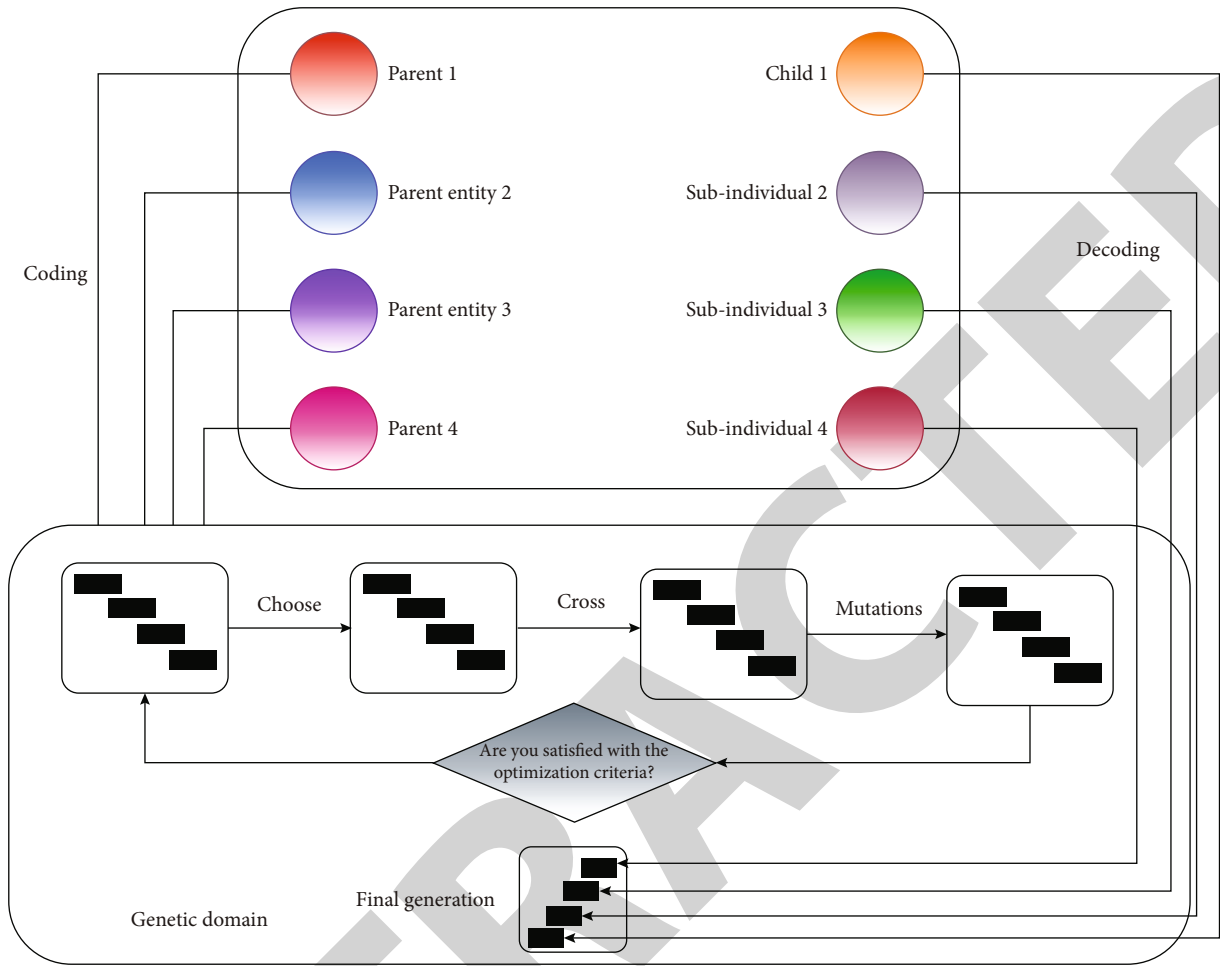


FIGURE 3: The process of GA.

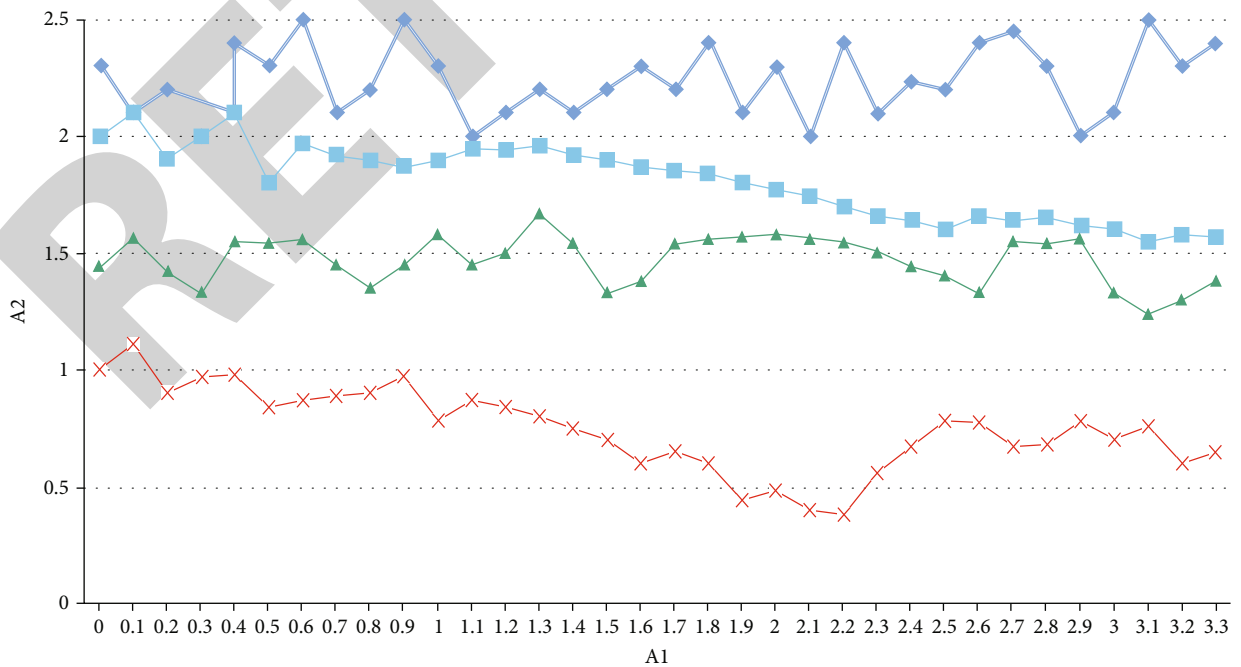


FIGURE 4: Test function T1.

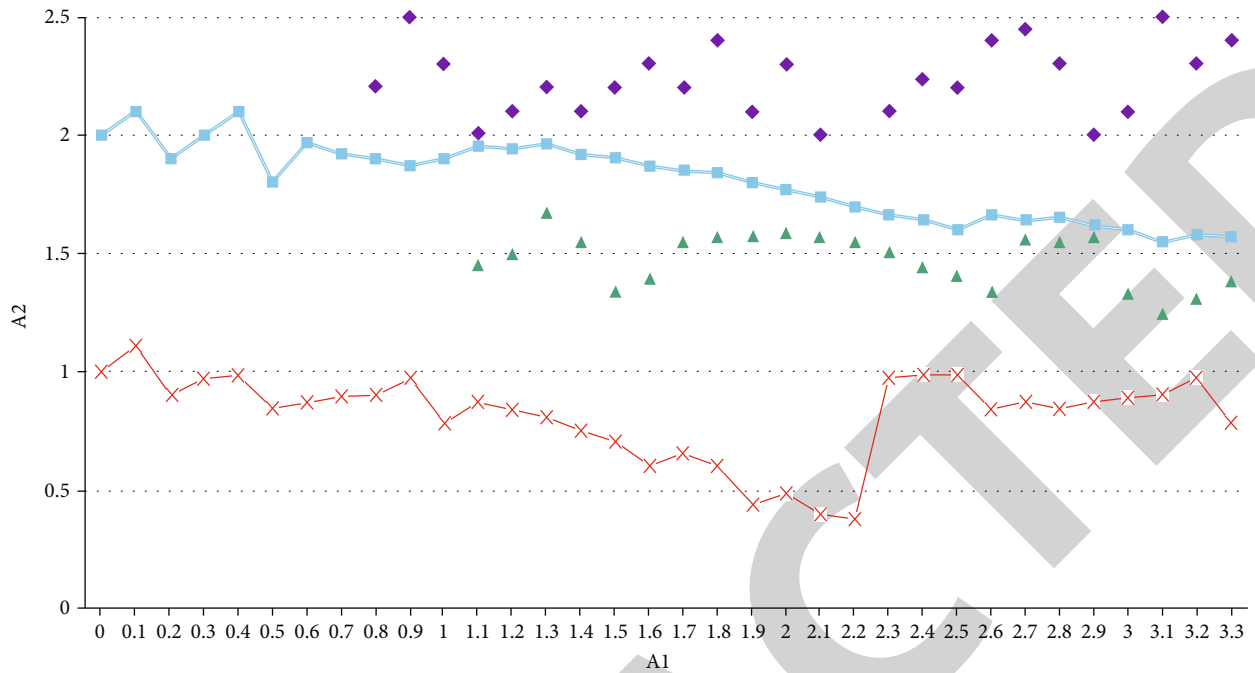


FIGURE 5: Test function T2.

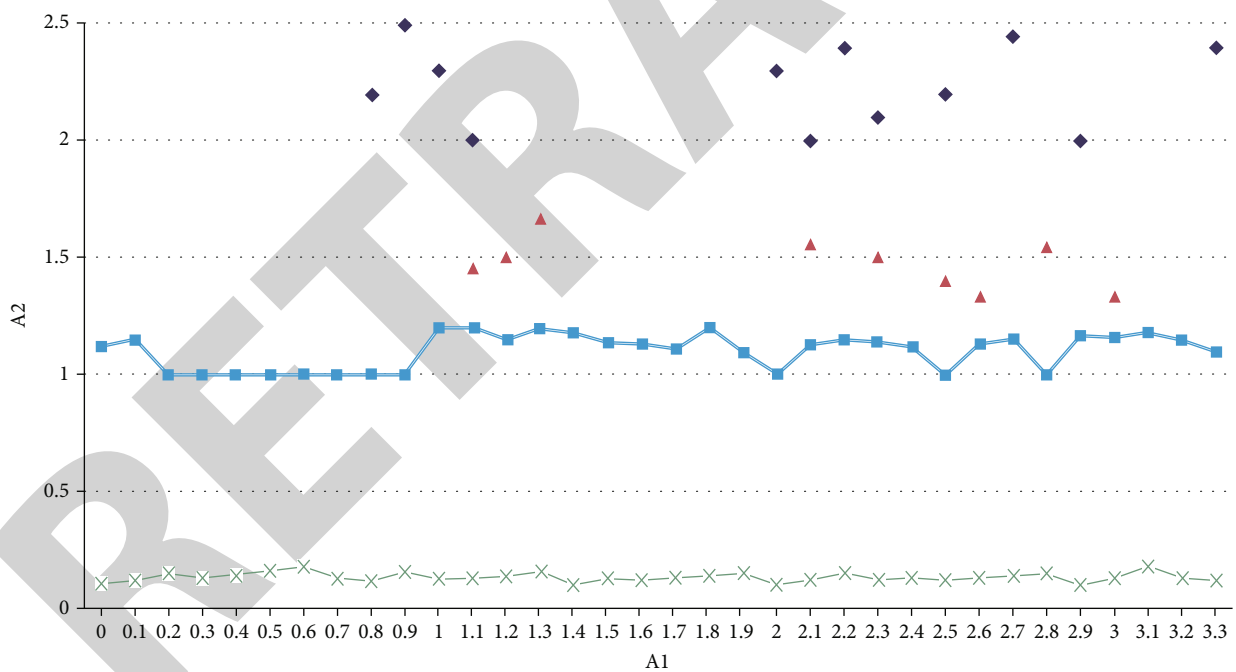


FIGURE 6: Test function T3.

If e represents a Euclidean matrix, in this case, $e = r^2 \mathbf{1}$. R is the radius of radial basis function; then, the above formula is simplified as

$$h(x) = \varphi \left[\frac{(x - c)^T (x - c)}{r^2} \right], \tag{4}$$

$$h(x) = \varphi \left[\frac{\|x - c\|^2}{r^2} \right].$$

The standard is a three-layer structure, multi-input and multioutput feedforward network composed of input layer, nonlinear hidden layer radial base layer, and linear output layer. The number of neurons in the input layer is the same as the dimension of the input vector. The number is the same as the output vector dimension.

The function of the hidden layer is to perform nonlinear transformation on the input vector, and its activation function is defined as a radial basis function with symmetrical

TABLE 2: Internal forces of conventional design method columns.

Column number	Bending moment (KN m)	Shear force (KN)	Axial force (KN)	Axial pressure ratio
1	50.23	16.55	163.45	0.190
2	72.54	32.53	560.46	0.135
3	12.45	26.45	465.45	0.129
4	63.87	13.78	1323.748	0.1654
5	56.52	24.18	1658.46	0.096
6	45.53	28.73	586.75	0.377
7	83.55	46.53	178.74	0.319

TABLE 3: Internal forces after column optimization.

Column number	Bending moment (KN m)	Shear force (KN)	Axial force (KN)	Axial pressure ratio
1	44.23	14.90	535.41	0.291
2	45.45	34.12	1095.52	0.132
3	78.42	38.56	658.13	0.093
4	59.53	28.33	371.52	0.452
5	67.95	22.11	331.05	0.196
6	48.75	22.41	253.044	0.047
7	86.49	41.52	456.154	0.175

properties. The Gaussian function is used as an example to illustrate the structure.

Gaussian function:

$$\phi_1(x) \exp \left\{ -\frac{\|x - c_t\|^2}{\sigma_t^2} \right\}. \quad (5)$$

Because of the special properties of radial basis function, it has the selective response ability to a certain range of input variables (i.e., the receptive field of hidden unit), resulting in the local tuning ability. The output activation function of the layer is a linear function, and the output of the hidden layer node is currently weighted. Its j th output node has the following form:

$$y_j = \sum_{i=1}^k w_{ij} \phi_i(x). \quad (6)$$

To a minimum, where $\xi_s(F)$ is the standard error term, which can be expressed as

$$\xi_s(F) = \frac{1}{2} \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \frac{1}{2} \sum_{i=1}^N [y_i F(x_i)]^2 \quad (7)$$

No matter how complex the mathematical function is, it is formed by some mathematical operators, operands, and mathematical functions through composition. The initial population can be generated by random selection in the set of effective operators and operands.

4. Result Analysis and Discussion

In order to compare the performance differences among the improved genetic algorithm, the basic genetic algorithm, and the immune genetic algorithm in this paper, we use the test function to evaluate them. These test functions have different typical characteristics: nonconvex, continuous/discrete noninferior optimal target domain, multipeak, and biased search space. These characteristics make it difficult for the algorithm to converge to the noninferior optimal solution domain and maintain the diversity of noninferior optimal solutions.

In this paper, three evolutionary algorithms are used to run each optimization test function 60 times. Among them, the same values are selected for the same parameters in the algorithm, such as the number of iterations, 300; population size, 80; cross probability, 0.75; and variation probability, 0.02. The function values are averaged, and the simulation results are shown in Figures 4–6.

In the figure, * indicates the improved GA in this paper, + indicates the immune GA, o indicates that the curve near the bottom of the basic GA is the noninferior optimal target domain, and the other is the comparison curve. According to the evaluation criteria of the algorithm, the improved GA in this paper is superior to the basic GA and the immune GA. The design results obtained by the conventional design method and the design results optimized by the improved GA are listed in Tables 2 and 3, respectively.

The overall cost change curve is shown in Figure 7.

The change curve of the cost of board and sofa fabric is shown in Figure 8.

Through the data before and after optimization, it can be seen that the total cost before optimization is 11778.51 yuan, and the total cost after optimization is 8706.20 yuan, saving 3072.31 yuan. Compared with the total cost before optimization, the total cost is reduced by 26.08%, and the optimization effect is remarkable. Among them, the plate cost changes greatly in the early stage of optimization, the reduction range is large and tends to be stable in the late stage of optimization, the cloth increases in the early stage of optimization, and the change tends to be flat in the late stage of optimization. The reduction of plate consumption will lead to the increase of steel consumption, which is in line with the objective reality. It can be proved that the design of this program and the design of improved GA can realize the optimization of the frame structure. The initial parameters given by the user are used as the approximate design intention of the bottle body shape. On the basis of these parameters, the program changes the parameters in the original solution space according to a certain probability factor and obtains the initial solution group. The generated 600 bottle shape design schemes are in the form of 3D model files for users to interactively select to calculate their weight coefficients, and design practice has achieved good results. The determination and ordering of the ladder levels are determined by the interactive choices of the tested users, which provide a specific design direction for the designer's design. After the optimization, the operation benefit of the furniture production line has been effectively improved. In terms of

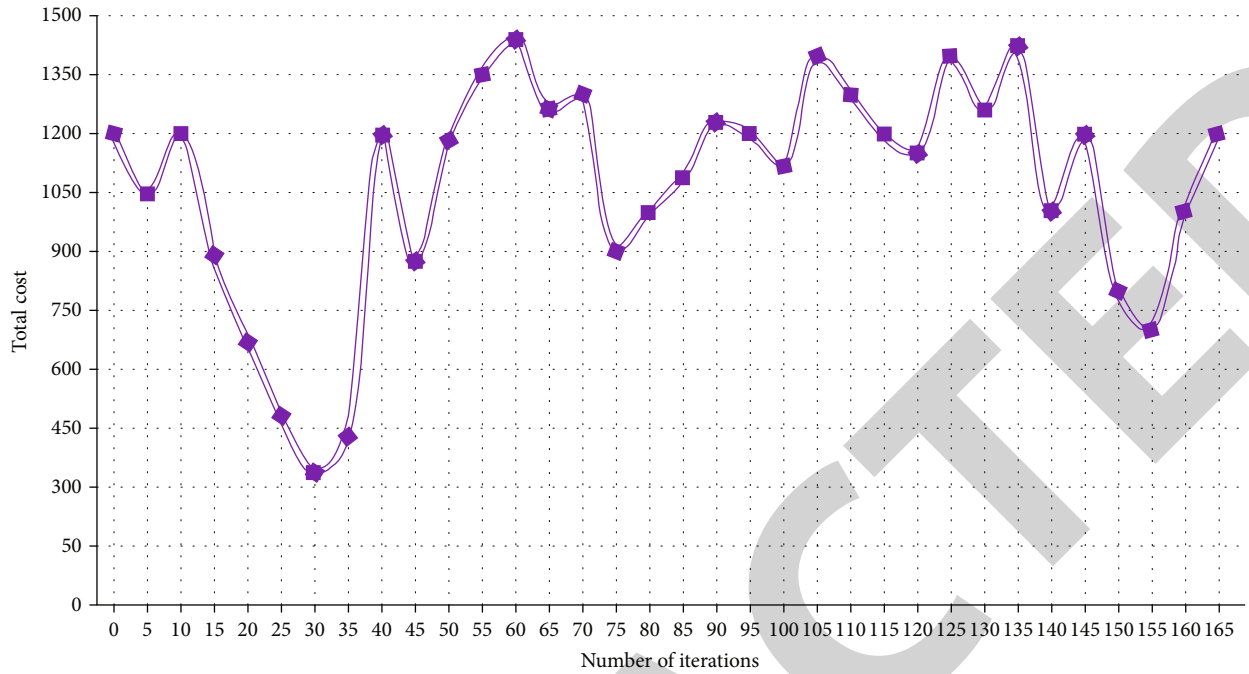


FIGURE 7: Total cost change curve.

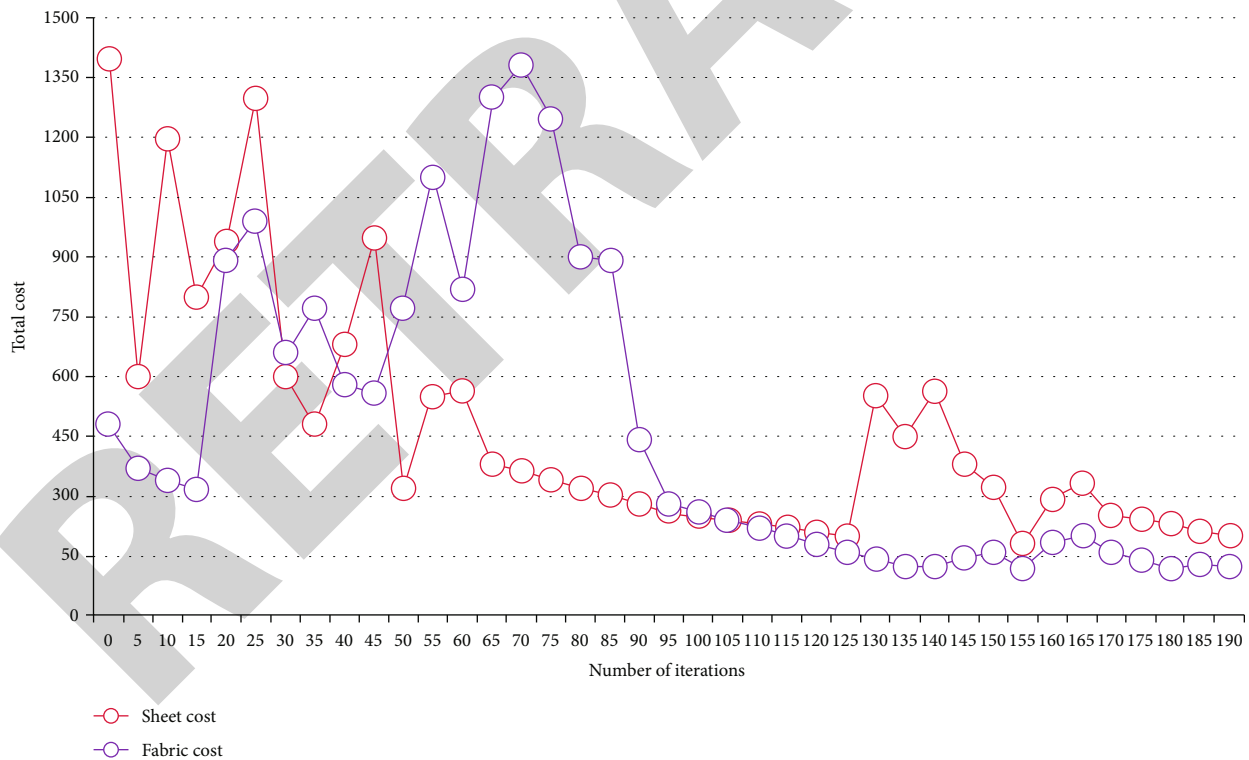


FIGURE 8: Subitem cost change curves.

economy, the benefit of the optimization of the furniture production line is $29.4\% \times 14 \times 3500 \times 12 = 172,930$ yuan (- the number of benefits improved \times the number of workers \times the monthly salary per capita \times the number of working months per year). It can be seen from this that the optimiza-

tion of furniture production line by GA has achieved remarkable research results, both in economic benefits and in labor costs, which has made a very obvious improvement, promoted the efficient operation of furniture production line, continuously improved the economic benefits of enterprises, and promoted the stable development of China's

economy. As people pay more and more attention to the furniture production line, how to improve the operation efficiency of the furniture production line has become a key issue concerned by relevant personnel. This paper studies the optimization measures of furniture production line based on genetic algorithm and finds that the research can greatly improve the production quality of furniture production line. Promote the development of genetic algorithm in furniture production line optimization in the future.

5. Conclusions

To sum up, as people's attention to furniture production lines gradually increases, and how to improve the operation efficiency of furniture production lines has become a key issue for relevant personnel. On the basis of the existing product scheme modeling, according to the semantic constraints, the rapid and intelligent drive between product semantics and product modeling scheme is realized, an optimized design scheme is generated, and the agile and intelligent product modeling design is realized. At the same time, taking the minimization of processing flow time as the objective function, a mathematical model of the scheduling problem of furniture assembly line production is built, and the GA is designed, and the solution program is programmed. In the workshop, the production sequencing experiment of six kinds of furniture products in the solid wood machining section is carried out as a comparative experiment of GA simulation experiment, and the processing hours of each process are measured to provide simulation data for the simulation experiment. The sorting results obtained from the simulation experiment are compared with the sorting scheme obtained from the production experiment. The results show that the total processing time of the sorting scheme obtained by the algorithm is reduced by 2475.2 s, the production efficiency is increased by 20.6%, and the job waiting time is reduced by 15079.2 s. It shows that the algorithm is feasible and effective in solving the furniture production job sorting problem. The furniture production sequencing scheme solved by GA can provide a certain reference value for the production personnel to formulate the production sequencing scheme and improve the production efficiency.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] S. Wang, "Application of product life cycle management method in furniture modular design," *Mathematical Problems in Engineering*, vol. 2022, Article ID 7192152, 10 pages, 2022.
- [2] R. Sirohi, A. Singh, and A. Tarafdar, "Application of genetic algorithm in modelling and optimization of cellulase production," *Bioresource Technology*, vol. 270, no. 90, pp. 569–776, 2018.
- [3] C. Mo, S. Zhao, Y. Ruan et al., "Research on reservoir optimal operation based on long-term and mid-long-term nested models," *Water*, vol. 14, no. 4, p. 608, 2022.
- [4] Y. P. Tsang, C. H. Wu, and K. Y. Lin, "Unlocking the power of big data analytics in new product development: an intelligent product design framework in the furniture industry," *Journal of Manufacturing Systems*, vol. 201, no. 7, pp. 65–556, 2021.
- [5] L. Oblak, A. Pirc Barčič, K. Klarić, M. Kitek Kuzman, and P. Grošelj, "Evaluation of factors in buying decision process of furniture consumers by applying AHP method," *Drvna Industrija*, vol. 68, no. 1, pp. 37–43, 2017.
- [6] J. Yao and R. Zhang, "Research on the combination in the function design of Ming-style furniture," *Journal of Beijing Institute of Technology (Social Sciences Edition)*, vol. 2, no. 5, p. 5662, 2015.
- [7] L. I. Ping, W. U. Yiqang, and Y. Zuo, "Research progress of the application of new materials in furniture," *Materials Review*, vol. 23, no. 780, pp. 59–378, 2015.
- [8] T. Shao, D. Li, Y. Rong, C. Zheng, and K. Zhou, "Dynamic furniture modeling through assembly instructions," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, pp. 1–15, 2016.
- [9] Y. Zuo and Z. Wang, "Subjective product evaluation system based on Kansei engineering and analytic hierarchy process," *Symmetry*, vol. 12, no. 8, p. 1340, 2020.
- [10] N. Hien, N. Phuong, T. V. Tran, and L. D. Thang, "The effect of country-of-origin image on purchase intention: the mediating role of brand image and brand evaluation," *Management Science Letters*, vol. 10, no. 6, pp. 1205–1212, 2020.
- [11] B. Fabisiak, "Characteristics of design process organization in selected furniture manufacturing companies," *Drvna Industrija*, vol. 58, no. 27, p. 85, 2016.
- [12] A. Colim, N. Sousa, P. Carneiro, N. Costa, P. Arezes, and A. Cardoso, "Ergonomic intervention on a packing workstation with robotic aid—case study at a furniture manufacturing industry," *Work*, vol. 66, no. 1, pp. 229–237, 2020.
- [13] N. Umentani, T. Igarashi, and N. J. Mitra, "Guided exploration of physically valid shapes for furniture design," *Communications of the ACM*, vol. 58, no. 9, pp. 116–124, 2015.
- [14] H. Soleimani and G. Kannan, "A hybrid particle swarm optimization and genetic algorithm for closed-loop supply chain network design in large-scale networks," *Applied Mathematical Modelling*, vol. 39, no. 14, pp. 3990–4012, 2015.
- [15] L. Jia, J. Chu, L. Ma, X. Qi, and A. Kumar, "Life cycle assessment of plywood manufacturing process in China," *International Journal of Environmental Research and Public Health*, vol. 16, no. 11, p. 2037, 2019.
- [16] J. Yao, D. M. Kaufman, Y. Gingold, and M. Agrawala, "Interactive design and stability analysis of decorative joinery for furniture," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–16, 2017.
- [17] B. T. Sayed, "Application of expert systems or decision-making systems in the field of education," *Information Technology in Industry*, vol. 9, no. 1, pp. 1396–1405, 2021.
- [18] I. A. Smorkalov and E. L. Vorobeichik, "Does long-term industrial pollution affect the fine and coarse root mass in forests? Preliminary investigation of two copper smelter contaminated areas," *Water, Air, & Soil Pollution*, vol. 233, no. 2, p. 55, 2022.

Retraction

Retracted: Analysis of the Impacts of Health Cost and Risk Preference on Farmers' Protective Behavior of Pesticide Application Based on the Autoregressive Threshold Model: A Case Study of Wuhu City in China

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Cai, Y. Gu, R. Li, and Z. Teng, "Analysis of the Impacts of Health Cost and Risk Preference on Farmers' Protective Behavior of Pesticide Application Based on the Autoregressive Threshold Model: A Case Study of Wuhu City in China," *Journal of Function Spaces*, vol. 2022, Article ID 9047754, 9 pages, 2022.

Research Article

Analysis of the Impacts of Health Cost and Risk Preference on Farmers' Protective Behavior of Pesticide Application Based on the Autoregressive Threshold Model: A Case Study of Wuhu City in China

Shukai Cai ¹, Yaxing Gu,¹ Runze Li,¹ and Zhonghong Teng ²

¹Department of Economic and Management, Anhui Polytechnic University, Wuhu, Anhui 241000, China

²Department of Physical Education, Anhui Polytechnic University, Wuhu, Anhui 241000, China

Correspondence should be addressed to Zhonghong Teng; tengzhonghong@ahpu.edu.cn

Received 27 June 2022; Revised 2 September 2022; Accepted 8 September 2022; Published 25 September 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Shukai Cai 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 is aimed at investigating the impacts of health cost and risk preference on farmers' protective behavior of pesticide in a case study of Wuhu city in China. Based on the field survey data from 523 farmers in the main grain-producing areas, the cost of illness (COI) method was employed to quantitatively measure the health cost (HC) of pesticide application, the Likert scale was used to measure the risk preference (RP) of pesticide applicators, and the autoregressive threshold model was used to test the impact of health cost and risk preference on farmers' protection behavior in the process of pesticide application. The key findings of this research case study reveal that when the health cost and risk preference are both lower than the critical value ($HC \leq 107.235$, $RP \leq 3$), the health cost does not affect improving the protection level of farmers in the process of pesticide application. However, when the risk preference exceeds the critical value ($RP > 3$), and the health cost exceeds a certain critical value ($HC > 107.235$), they have a significant positive effect on improving the protection level of farmers in the process of pesticide application. High health cost combines with higher risk preference ($HC > 107.235$, $RP > 3$) which can significantly improve the protection level of farmers in the process of pesticide application.

1. Introduction

China is one of the largest pesticide producer and consumer country in the world. Zivin [1] found that pesticide application can not only control the yield loss caused by diseases and insect pests but also cause negative effects on the health of the pesticide applicator and increase the risk of pesticide exposure if the protective measures are not standardized or is unsatisfactory. A recent report published by the World Health Organization (WHO) and United Nations Environmental Programme (UNEP) [2] estimated that there are 1 million human pesticide-poisonings each year in the world, with approximately 3,500 deaths. The most common health effects associated with pesticide exposure include headaches, skin and eye problems, salivation, hormone disruption, and

loss of consciousness (see for instance [3–9]). Luckily, the health effects of pesticides can be minimized among farmers and other pesticide operators by protection behavior such as the use of face masks, goggles, gloves, hats, protective clothing, and boots.

Previous research study has suggested that farmers' protective behavior of pesticide application is mostly affected by age, gender, education, household characteristics, policy characteristics, and governmental regulations (see for instance [9–13]). Recently, Wang et al. [14] have given significant attachment to the role of health cost of spraying pesticides and the farmer's risk preference. Akter et al. [8] investigates that there exists a large gap between the knowledge of potential pesticide risks, and pesticide application. Salzsar and Rand [15] found that more risk-averse farmers

use less pesticide. Farmers' perceptions of the risks in human health posed by pesticides, which can decrease their pesticide expenditure, and their risk attitude is the main factor of farmers' pesticide application behavior (see for instance [16–21]). Gong et al. [22] have found that risk aversion significantly increases pesticide use. Bagheri et al. [23] concluded that farmers' knowledge of pesticide use as well as their attitudes and perceptions concerning risks and safety play a crucial role in safe spraying operations in farms. However, the existing literature paid less attention to the influence of both health cost and risk preference on farmers' pesticide application behavior. Farmers who have paid higher health costs and have stronger risk awareness are more likely to realize the negative health effects caused by pesticide application and may use more personal protective equipment (PPE) to reduce pesticide exposure, thus reducing health costs. Existing studies have studied the impact of health cost and risk preference on farmers' protective behavior, respectively. But they have not yet seen the impact of health cost and risk preference on farmers' protective behavior within a unified analytical framework. In addition, numerous researchers have not taken note that the impact of health cost and risk preference on farmers' protective behavior may be nonlinear, in case they exceed a certain threshold limit, it will have a significant impact.

In this case study, we employed the panel threshold model to empirically examine the health costs and risk preferences of rice farmers. The nonlinear effects of health cost and risk preference on farmers' protective behavior were investigated with the help of econometric estimations. The results show that farmers pay high health costs in the process of pesticide application. However, the health costs of farmers have an impact on the level of self-protection in the process of pesticide application, only when the health cost exceeds a certain critical value. Also, it has a significant and positive effect on improving the level of prevention and protection of farmers in the process of pesticide application. In addition, the health cost enlarges the protection level difference in the process of pesticide application. However, if a high health cost is combined with the higher risk preference of farmers, the gap of protection level in the process of pesticide application can be significantly enlarged, which plays a synergistic role. This study will help to further explore the influence of health cost and risk preference on farmers' protective behavior. The findings of this paper is of great significance to further take comprehensive measures to reduce the health cost of pesticide application.

Finally, the paper structure is ordered as follows: the first section of this paper is the Introduction, we investigate the impacts of health cost and risk preference on farmers' protective behavior of pesticide; Section 2 presents data and empirical analysis, where qualitative survey-based questionnaire data were collected from 523 participants; Section 3 discusses the empirical model selection and its suitability in our study; finally, we compare our output with other related studies and examine the policy implications on the impacts of health cost and risk preference on farmers' protective behavior of pesticide in Wuhu city.

2. Data Sources and Empirical Analysis

2.1. Data Sources. The data sample used in this paper is obtained by the random survey questionnaire distributed among the students who were studying at Anhui Polytechnic University. Before conducting the survey, the authors followed the essential prerequisites in the preparation of questionnaire design. As first, we carefully consider the survey main content and designed the questionnaire with the help of relevant field experts and previous literature review. In November 2014, the authors conducted a presurvey which focuses on group interviews in Yijiang Town, Nanling County, Wuhu City, and Anhui Province. The relevant field farmers' information were obtained in face-to-face interview, and later on the questionnaire was modified for the reason of reducing the psychological pressure of the interviewees. The formal survey was launched in February 2015, with the main pesticide application season-ending, which helps farmers to memorize the year's pesticide use. Most of the respondents were farmers, and 98% of them were householders. The questionnaire included different variety of questions such as the criteria of farmers, age group, education level, pesticide spraying information, and health cost of pesticide application. A total of 600 questionnaires were distributed in the formal survey, where we received a total of 556 responses from the participants with a retrieval rate of 91 percent. Overall, after we carefully scrutinized all the questionnaires, we found that 523 of them were valid, and the effectiveness ratio of the survey was 94.065 percent. Moreover, after the completion of the questionnaire, a compensation of about 20 yuan was given to each participant which is equivalent to 1/5 of the local daily wage with the purpose of improving and maintaining the enthusiasm of the interviewees and the quality of the survey.

2.2. Personal Protective Measures and Hygiene. The pesticides frequently used by the surveyed sample farmers include chlorpyrifos, chlorpyrifos benzamide, thiazide, imidacloprid, thiamethoxam, benzamide, propylphos, abamectin, BT emulsion, trichlorfon, and rice blast. Many farmers stated that when spraying pesticides, they consider a various number of protective measures such as avoid spraying pesticides during wind, smoking, taking a bath after spraying pesticides to reduce the harmful impact of pesticide application on health; however, the survey shows that farmers do not use enough protective measures when using pesticides. None of the farmers in the survey used special protective measures such as cloths, masks, gloves, hats, and goggles. The expenditure on protective equipment were mostly not mentioned in the survey by the participants. This shows that farmers more likely do not considering the importance and the need to use protective equipment. A relatively small number of farmers take precautions such as wearing masks, but this number is far less than the expectations of the survey. On the other hand, the survey finds that the proportion of farmers who wear gloves in the sample is not satisfactory. Many farmers reflect that wearing gloves will affect the efficiency of pesticide application. Others stated that farmers do not use protective measures because

of discomfort, social interaction, limited number of equipment, the supply of equipment availability constraints, and cost constraints. The survey found that the protective measures taken by farmers in the process of preparing pesticide application are limited with an average value of 3.122 percent. More protective measures are adopted, including wearing long sleeve coat and trousers, washing hands, and bathing after spraying (see Table 1).

2.3. The Health Cost of Pesticide Application on Farmers. The health cost measurement methods of pesticide application mainly include the cost of illness method (COI), willingness to pay method (CVM), and prevention cost method (see for instance [24–27]). Although CVM can be used to measure the health cost comprehensively with lower cost, it also bears a lot of criticism, because the interviewees often ignore the constraints of real market conditions when expressing their willingness to pay (see for instance [28]). However, due to the integrity of data acquisition, the preventive cost method can only capture part of the health cost, which limits its application. Therefore, this paper mainly uses the disease cost method to calculate the health cost of pesticide application. Based on the observation of objective behavior and real market, the (COI) method has been widely used, which helps to measure the health cost of pesticide application relatively completely and objectively. In this paper, we use a number of econometric techniques to calculate the health cost of pesticide application, mainly including the medical cost of sensitive poisoning caused by pesticide exposure and the lost labor time. Specifically, it includes several expenditures such as: (1) the medical expenditure, transportation expenditure, and accompanying expenditure of family members in hospitals and private clinics; and (2) waiting time, treatment time, and the opportunity cost of being unable to work caused by illness. This study does not calculate the cost of chronic diseases, pain and discomfort, family care costs, and intentional pesticide poisoning.

In this paper, the health cost of pesticide application on farmers measured by the (COI) method is shown in Table 2. The total health cost is 91.846 yuan per year per person, of which the direct monetary expenditure is 48.587 yuan per year per person, including medical expenses of 36.263 yuan per year per person and transportation expenses of 12.325 yuan per year per person, while the opportunity cost of time loss is 43.265 yuan per year per person.

3. Effects of Health Cost and Risk Preference on Pesticide Application Behavior

3.1. Model Specification and Selection. The previous analysis shows that there are significant differences in the self-protection of the pesticide application process with different health costs and risk preferences in some places. One question is, how much impact does it have on farmers' self-protection behavior in the process of pesticide application? Considering that the relationship between health cost, risk preference, and farmers' self-protection behavior in the process of pesticide application is nonlinear, there is a complex mechanism of health cost and risk preference on farmers'

self-protection behavior in the process of pesticide application. There is a certain threshold value of health cost and risk preference when the health cost and risk preference are lower than the threshold value, the impact on farmers' self-protection behavior in the process of pesticide application presents a relationship, when the threshold value is higher it presents another relationship.

Moreover, to analyze the internal relationship between the phenomena more accurately, we established a threshold autoregressive model to test the nonlinear relationship between health cost, risk preference, and farmers' self-protection behavior in the process of pesticide application.

The threshold model divides the model into two or more intervals (also known as a regime). According to the threshold value and different equations which express each interval with the help of the threshold model, it is helpful to capture the zero point or critical value where the interval may occur, which is different from the Chow test of subjective exogenous setting structural mutation points, the "threshold model" divides the interval according to the characteristics of the data themselves. In addition to the decent characteristics of the general econometric model, it can also capture the threshold effect in the economy. From one threshold model setting to a multi-threshold model setting, one threshold model is extended to a multithreshold model setting $y_{it} = \mu_i + \beta_1 x_{it} I(q_{it} \leq \gamma) + \beta_2 x_{it} I(q_{it} > \gamma) + e_{it}$, $I(\cdot)$ as an indicative function, the observation value q_{it} is divided into two intervals according to whether the threshold variable is greater than or less than the threshold value γ , the observation value is divided into two intervals, when the minimum sum of squares of residual errors $S_1(\gamma)$ is searched, the corresponding threshold is the optimal estimation value $\hat{\gamma} = \arg \min S_1(\gamma)$. The cross-sectional threshold model involves two hypothesis tests: (1) test whether the threshold effect exists and (2) test whether the estimated threshold values are equal to the true values. In the first test $H_0 : \beta_1 = \beta_2$, the alternative hypothesis is $H_1 : \beta_1 \neq \beta_2$, the statistics $F_1 = (S_0 - S_1(\hat{\gamma})) / (\hat{\sigma}^2)$ do not meet the standard distribution, and the bootstrap method is used to obtain the critical value of the approximate distribution. The second test $H_0 : \gamma = \hat{\gamma}$, statistics are $LR_1(\gamma) = (S_1 - S_1(\hat{\gamma})) / (\hat{\sigma}^2)$. For the case of multiple thresholds, the model is set as follows:

$$y_{it} = \mu_i + \beta_1 x_{it} I(q_{it} \leq \gamma_1) + \beta_2 x_{it} I(\gamma_1 < q_{it} \leq \gamma_2) + \beta_3 x_{it} I(q_{it} > \gamma_2) + e_{it}. \quad (1)$$

Search the minimum residual square sum of the second threshold $S_2^r(\gamma_2)$, corresponding to the second threshold value $\hat{\gamma}_2^r = \arg \min S_2^r(\gamma_2)$. Observe whether the two thresholds are significantly different through the following statistics: $F_2 = (S_1(\hat{\gamma}) - S_2^r(\hat{\gamma}_2^r)) / (\hat{\sigma}^2)$, if F_2 is significant, it indicates that there is a second threshold, and then continue to search for the third threshold, and so on, until the last hypothetical threshold is not significant. Threshold variables can be exogenous variables or explanatory variables in the model. The results obtained by this threshold regression method can fit the data more accurately and precisely than

TABLE 1: Personal protective measures and hygiene.

Protective measures and hygiene	Wear long-sleeved coat	Wear trousers	Wear hat	Wear gloves	Wear mask	Avoid spraying pesticide during the wind	Wash hands	Taking bath
Proportion	90.249	87.381	37.476	17.017	9.560	47.419	87.381	91.587

TABLE 2: Farmers' health costs of pesticide application.

Variable	Maximum	Minimum	Mean
Direct monetary expenditure (yuan)	620.236	0	48.587
Medical expenses (yuan)	450.762	0	36.263
Transportation expenses (yuan)	85.689	0	12.325
The lost time cost (yuan)	320.829	0	43.265
The total time lost (hours)	48.348	0	5.624
Time opportunity cost (yuan/hour, with nonagricultural income)	15.894	0	8.215
Time opportunity cost (yuan/hour, without nonagricultural income)	12.543	0	3.092
Total health cost (yuan/year)	705.093	0	91.846

Note: (1) When calculating the rest time, if the farmer is hospitalized, we calculate it by 10 hours a day. (2) Considering that the calculation of the opportunity cost of time is mainly based on the income of farmers working every day, while the working hours of farmers are generally 10 hours a day, we use the daily income/10 hours when calculating the opportunity cost per hour; we use the annual average income/365 days/10 hours to calculate the opportunity cost of farmers without working income.

the ordinary regression model, especially when there is a nonlinear relationship between the explanatory variable and the response variables. The equation of the influence of health cost on farmers' behavior of pesticide application and protection measures are set as follows:

$$\begin{aligned} \text{AMT}_{it} = & \theta_1 \text{AEMP}_{it} + \theta_2 \text{AEMP}_{it}^2 + \beta_1 \text{HC}_{it} (\text{IND}_{it} \leq \gamma_1) \\ & + \beta_2 \text{HC}_{it} (\gamma_1 < \text{IND}_{it} \leq \gamma_2) + \beta_3 \text{HC}_{it} (\text{IND}_{it} > \gamma_2) \end{aligned} \quad (2)$$

The number of protective measures taken by (AMT) was used to reflect the difference in self-protection measures of different farmers. (HC) is the health cost of pesticide application for farmers. Moreover, to avoid the endogeneity between variables, this paper uses the health cost of pesticide application in the previous year as the value to measure the health cost of farmers. AEMP is the risk preference of farmers. γ represents the threshold. It should be noted that farmers' risk awareness is a dummy variable that cannot be directly measured. Therefore, this paper uses the Likert scale to indirectly measure the level of farmers' risk awareness. Although this method is slightly rough, however it has been proved by numerous research studies which is simple and effective.

3.2. Estimated Results. In the first round of threshold regression, the LM value and bootstrap p value of health cost (HC) as threshold variable accounted for 115.635 and 0.000, respectively; LM value and bootstrap p value of risk preference (RP) as threshold variable accounted for 97.723 and 0.000, respectively. The results show that the health cost and risk preference are likely to be the threshold variables influencing the self-protection level of farmers in the process of pesticide application at the significance level of 5%. Therefore, the health cost (HC) with a larger LM value is selected

as the initial threshold grouping index. The results show that the p value of the heteroskedasticity test is 0.062, and the original hypothesis of homovariance could not be rejected. Therefore, the outcomes show that there is no heteroskedasticity, and the estimation results of the model are acceptable. Afterwards, we take "social capital HC" as the threshold variable and the likelihood ratio sequence statistic LRN (R) as the threshold function. The estimated threshold value is 107.235. 326 samples fall into the low health cost group ($\text{HC} \leq 107.235$), and 197 samples fall into the high health cost group. No samples fall on the confidence interval $[107.235, 107.235]$, therefore, we can divide the samples into two groups: the low health cost group 1 ($\text{HC} \leq 107.235$) and the high health cost group 2 ($\text{HC} > 107.235$).

After the first round of threshold regression, the second round of threshold regression was performed for the low health cost group ($\text{HC} \leq 107.235$) and the high health cost group ($\text{HC} > 107.235$). In the low health cost group, the LM and bootstrap p values of the two threshold variables were obtained as follows: health cost HC (18.653, 0.004) and risk preference HC (21.968, 0.000). Therefore, the risk preference RP with a lower bootstrap p value was selected as the threshold variable of the second grouping. The results show that the test p value of heteroskedasticity is 0.160, thus, we cannot reject the hypothesis of heteroskedasticity. The result indicates that there is no heteroskedasticity existing between the selected variables. Afterwards, we use "risk preference RP" as the threshold variable and the likelihood ratio sequence statistic LRN (R) as the threshold function, and the estimated threshold value γ is 3. With the risk preference threshold of 3, the low health cost group can be further divided into two groups: "low health cost low-risk preference group" ($\text{HC} \leq 107.235$, $\text{HC} \leq 3$) and "low health cost high-risk preference group" ($\text{HC} \leq 107.235$, $\text{HC} > 3$). In these two groups, there will be no threshold (see Table 3 for detailed results).

TABLE 3: Threshold test with “risk preference RP” as threshold variable in the low health cost group.

Dependent variable: Level of self-protection	HC ≤ 107.235			HC ≤ 107.235 RP ≤ 3 Low health cost - low-risk preference			HC ≤ 107.235 RP > 3 Low health cost - high-risk preference		
	Coefficient	Standard error	Mean value	Coefficient	Standard error	Mean value	Coefficient	Standard error	Mean value
Independent variable									
Health costs (HC)	-0.322	0.008	72.262	-0.056	0.006	76.436	-0.056	0.154	68.812
Risk preference (RP)	-0.124	0.042	3.272	-0.265	0.084	2.127	-0.362	0.115	3.865
Observation sample size		326			192			134	
R2		0.492			0.124			0.256	

Heteroskedasticity Test (*p* value): 0.160.

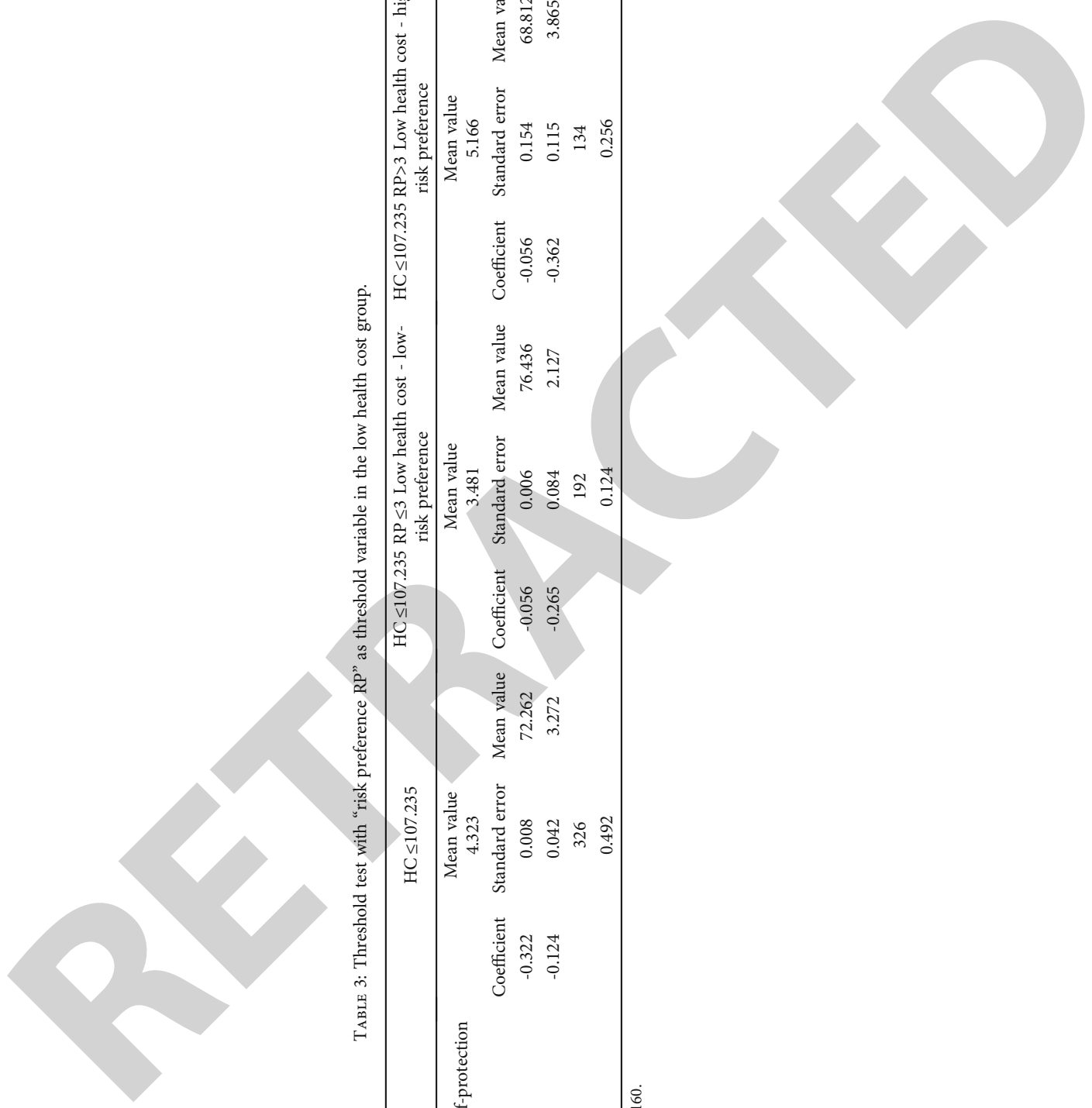


TABLE 4: Threshold test with “risk preference RP” as threshold variable in high health cost group.

	HC>107.235		HC>107.235 RP ≤3 High health cost - low-risk preference		HC>107.235 RP>3 High health cost - high-risk preference	
	Mean value	Standard error	Coefficient	Mean value	Coefficient	Mean value
Dependent variable: Level of self-protection	5.643			5.213		6.726
Independent variable						
Health costs (HC)	-0.096	0.008	-1.077	0.0695	-2.361	0.164
Risk preference (RP)	-0.514	0.065	-0.156	0.856	-1.365	0.325
Observation sample size	197		105		92	
R2	0.654		0.366		0.621	

Heteroskedasticity Test (p value): 0.236.

It can be seen from Table 3 that 326 samples fall into the low health cost group ($HC \leq 107.235$), and the average protection level of farmers during pesticide application is 4.323. According to the actual investigation, although many of these farmers know that pesticides are harmful to the human body, they have less knowledge and understanding about pesticide contact path, especially the impact of pesticide application on health. The findings reveal that these farmers have little knowledge about how to sensibly avoid and reduce pesticide vulnerability. In the process of pesticide application, they do not attach importance to the protection measures. They mainly wear long trousers and long sleeves. The proportion of farmers who wear hats, gloves, and masks is low, and the protective measures are simple.

By comparing the average values of protection level, health cost, and risk preference in the application process of the “low health cost low-risk preference group” and the “low health cost high-risk preference group”, we can find that the average health cost in the “low health cost low-risk preference group” is 76.436. The average value of health cost in the “low health cost high-risk preference group” is 68.812, and in the “low health cost low-risk preference group” it is 2.127, which is significantly lower than that of the “low health cost high-risk preference group”, which was 3.865. In contrast, in the average protection level of the two groups, it can be found that the average protection level of the “low health cost low-risk preference group” is 3.481, which is significantly lower than that of the “low health cost high-risk preference group,” which is 5.166. This phenomena shows a problem such as health cost has an integration and synergistic effect on individual risk preference, and health cost must rely on risk preference. Comparing the results of the two groups, when the health cost and risk preference are lower than the critical value, the health cost has no positive effect on improving the protection level of farmers in the process of pesticide application, while when the health cost is unchanged and the risk preference exceeds the critical value, the health cost plays a positive role in improving the protection level of farmers in the process of pesticide application.

The second round of threshold regression was carried out for the high health cost group ($HC > 107.235$), and the heteroskedasticity test p value was 0.227, and the homovariance hypothesis could not be rejected. Thus there was no heteroskedasticity between the selected variables. Moreover, two threshold regression analyses, LM and bootstrap p value health cost variable (48.270, 0.0000) risk preference (67.101, and p value 0.000), respectively, indicate that in the high health cost group, risk preference becomes the threshold variable for further grouping. Afterwards, we take “risk preference RP” as the threshold variable, and “likelihood ratio sequence statistic LRN (R)” as a threshold function, and the estimated threshold value γ is consistently 3. Therefore, the high health cost group can be further divided into two groups: the “high health cost low-risk preference” group 1 ($HC > 107.235, RP \leq 3$) and the “high health cost high-risk preference” group 2 ($HC > 107.235, RP > 3$).

According to Table 4, the protection level of farmers in the high health cost group is 5.643, which is significantly higher than 4.323 in the low health cost group. Comparing

the average values of protection level, health cost, and risk preference in the process of pesticide application between the “high health cost low-risk preference group” and the “high health cost high-risk preference group”, we can find that the average health cost in the “high health cost low-risk preference group” is 132.744, which is significantly higher than 109.626 in the “high health cost high-risk preference group”. However, the average risk preference in the “high health cost low-risk preference group” is 2.614, which is significantly lower than that in the “high health cost high-risk preference group” which is 3.986. The direct result is that the self-protection level of farmers in the “high health cost low-risk preference group” is 5.213, which is significantly lower than that in the “high health cost high-risk preference group” in the process of pesticide application. This shows two key issues, first, when the health cost exceeds a certain critical value ($HC > 107.235$), it has a significant positive effect on improving the protection level of farmers in the process of pesticide application, and health cost is an important factor to enlarge the gap of protection level in the process of pesticide application. Second, high health costs combined with higher risk preference can significantly improve the protection level gap in the process of pesticide application and play a synergistic effect.

4. Conclusion

Based on the field survey data from 523 farmers in the main grain-producing areas, this study found that the impact of health costs and risk preference on farmers’ protective measures is nonlinear when the health cost and risk preference are both lower than the critical value ($HC \leq 107.235, RP \leq 3$). On the other hand, the health cost does not affect improving the protection level of farmers in the process of pesticide application when the risk preference exceeds the critical value ($RP > 3$). The health cost plays a positive role in improving the protection level of farmers in the process of pesticide application when the health cost exceeds a certain critical value ($HC > 107.235$), it has a significant positive effect on improving the protection level of farmers in the process of pesticide application. Additionally, the increasing health cost combines with higher risk preference ($HC > 107.235, RP > 3$), which can significantly improve the protection level of farmers in the process of pesticide application.

There are several number of policy implications proposed in this paper. Firstly, it is essential to strengthen the publicity of health accidents in the process of pesticide application, because the health cost needs to reach a certain level or beyond a certain threshold value. The simple publicity of the negative effects of pesticides may not achieve the desired effect where more comprehensive measures should be taken. In addition, to improve the effect of publicity, the decision makers should focus on the publicity of increasing health cost accidents in the process of pesticide application. Secondly, this study shows that the combination of high health cost and high-risk preference will significantly change the behavior of farmers. Therefore, it is important to improve farmers’ risk attitude and enhance their awareness of safe production through publicity. It is very necessary to

make full use of new media such as TikTok and WeChat to publicize the health impact of pesticide exposure, improve the safety awareness of farmers, focus on how to use protective equipment, and especially publicize some typical adverse impact events caused by pesticide exposure. At the same time, regular training on pesticide application technology is provided for farmers to improve their mastery of pesticide application methods and pesticide application interval and understand the pesticide exposure risks caused by different pesticide application behaviors. Thirdly, more comprehensive measures should be taken to reduce the negative health effects of pesticide application. If the research provides more suitable and comfortable personal protective equipment, the personal protective equipment will be subsidized. Fourthly, due to the small scale and decentralized characteristics of agricultural production in China, the negative impact of pesticide application on publicity, government supervision, and policy implementation costs is significant. Our findings can gradually change farmers' decentralized self-control and self-control traditional ways by constantly improving the agricultural social service system, and improve the professionalism of pest control.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the research projects of the Humanities and Social Science Fund of Ministry of Education of China (18YJCZH163), Natural Science Foundation of Anhui Provincial (2108085MG245), and 2022 Academic Funding Project for Top Talents in Disciplines (Majors) of Universities in Anhui Province (gxbjZD2022022).

References

- [1] S. J. Zivin, "Insect population dynamics, pesticide use, and farmworker health," *American Journal of Agricultural Economics*, vol. 82, no. 3, pp. 527–540, 2000.
- [2] C. Wilson, "Exposure to pesticides, ill-health and averting behaviour: costs and determining the relationships," *International Journal of Social Economics*, vol. 32, no. 12, pp. 1020–1034, 2005.
- [3] B. M. Maumbe and S. M. Swinton, "Hidden health costs of pesticide use in Zimbabwe's smallholder cotton growers," *Social Science & Medicine*, vol. 57, no. 9, pp. 1559–1571, 2003.
- [4] D. Hu, Z. Chen, and Q. Sun, "Impact of contract production mode on Farmers' income and food safety: a case study of apple industry in Shandong Province," *China Rural Economy*, vol. 24, no. 11, pp. 17–24, 2006.
- [5] S. Isin and I. Yildirim, "Fruit-growers' perceptions on the harmful effects of pesticides and their reflection on practices: the case of Kemalpaşa, Turkey," *Crop Protection*, vol. 26, no. 7, pp. 917–922, 2007.
- [6] J. Huang, L. Qi, and R. Chen, "Technical information knowledge, risk preference and pesticide application by farmers," *Management World*, vol. 89, no. 5, pp. 71–76, 2008.
- [7] X. Fu and W. Song, "Analysis on influencing factors of farmers' purchase intention and purchase behavior of biopesticides – taking Sichuan Province as an example," *Agricultural Technology and Economy*, vol. 6, pp. 120–128, 2010.
- [8] M. Akter, L. Fan, M. M. Rahman, V. Geissen, and C. J. Ritsema, "Vegetable farmers' behaviour and knowledge related to pesticide use and related health problems: a case study from Bangladesh," *Journal of Cleaner Production*, vol. 200, no. 5, pp. 122–133, 2018.
- [9] W. Nunthasen, N. Ke, and W. Jaengsaengthong, "Farmers health effects and attitude towards agrochemicals use in rice production," *IOP Conference Series Earth and Environmental Science*, vol. 265, 2019.
- [10] J. Zhang, J. Wang, and X. Zhou, "Farm machine use and pesticide expenditure in maize production: health and environment implications," *International Journal of Environmental Research and Public Health*, vol. 16, no. 10, p. 1808, 2019.
- [11] Y. Wu, X. Xi, X. Tang et al., "Policy distortions, farm size, and the overuse of agricultural chemicals in China," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, no. 27, pp. 7010–7015, 2018.
- [12] L. Fan, H. Niu, X. Yang et al., "Factors affecting farmers' behaviour in pesticide use: insights from a field study in northern China," *Science of the Total Environment*, vol. 537, pp. 360–368, 2015.
- [13] L. Zhao, C. Wang, H. Gu, and C. Yue, "Market incentive, government regulation and the behavior of pesticide application of vegetable farmers in China," *Food Control*, vol. 85, no. 3, pp. 308–317, 2018.
- [14] J. Wang, M. Chu, Y. Deng, H. Lam, and J. Tang, "Determinants of pesticide application: an empirical analysis with theory of planned behaviour," *China Agricultural Economic Review*, vol. 10, no. 4, pp. 608–625, 2018.
- [15] C. Salzsar and J. Rand, "Pesticide use, production risk and shocks. The case of rice producers in Vietnam," *Journal of Environmental Management*, vol. 253, article 124241, 2020.
- [16] D. Pan, M. He, and F. Kong, "Risk attitude, risk perception, and farmers' pesticide application behavior in China: a moderation and mediation model," *Journal of Cleaner Production*, vol. 276, no. 10, 2020.
- [17] A. R. Kofi, "Farmers' risk preference and the adoption of risk management strategies in northern Ghana," *Journal of Environmental Planning and Management*, vol. 68, no. 12, pp. 1–20, 2018.
- [18] H. Hans-Peter, K. Abdul, L. Kathrin, W. Peter, K. Michael, and M. Hanns, "Cytotoxic and genotoxic effects of pesticide exposure in male coffee farmworkers of the Jarabacoa Region, Dominican Republic," *International Journal of Environmental Research and Public Health*, vol. 15, no. 8, p. 1641, 2018.
- [19] H. C. Laizer, M. N. Chacha, and P. A. Ndakidemi, "Farmers Knowledge, Perceptions and Practices in Managing Weeds and Insect Pests of Common Bean in Northern Tanzania," *Sustainability*, vol. 11, no. 15, pp. 1–11, 2019.

Retraction

Retracted: The Application of the Depth Model of Precise Matching between People and Posts Based on Ability Perception in Human Resource Management

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Cao, "The Application of the Depth Model of Precise Matching between People and Posts Based on Ability Perception in Human Resource Management," *Journal of Function Spaces*, vol. 2022, Article ID 5495437, 9 pages, 2022.

Research Article

The Application of the Depth Model of Precise Matching between People and Posts Based on Ability Perception in Human Resource Management

Hui Cao 

Department of Economics and Management, Taishan University, Shandong, Taian 271021, China

Correspondence should be addressed to Hui Cao; caohui123@tsu.edu.cn

Received 1 July 2022; Revised 2 September 2022; Accepted 6 September 2022; Published 22 September 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Hui 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.

To revitalize talents and give full play to the maximum utility of HR (human resources), it is not enough to accumulate talents alone. HR must be effectively allocated to realize the matching of people and posts. Competency is a personal characteristic of an organization that distinguishes its performance level in a specific job and organizational environment. In order to solve the problem that job seekers' job-seeking ability is difficult to match the job requirements, this paper combines neural network with traditional HRM (human resource management) algorithm based on ability perception and designs a depth model of accurate matching of people and posts in HR field, which can improve the quality of data training of traditional algorithm. The results show that compared with other algorithms, the $F1$ value of the proposed algorithm is obviously improved, and the proposed algorithm performs best, with the $F1$ value of 0.829. In this paper, the method of global network plus local network is used, which can effectively improve the hidden features of data and then improve the matching degree and recommendation accuracy of the algorithm.

1. Introduction

In today's knowledge age, science and technology represent wealth. The comprehensive quality of human resource management of economists in enterprises and institutions is directly related to the future development of the industry, so it is very important to improve the quality of personnel. In view of the negative impact caused by the fact that some enterprises and institutions do not pay enough attention to the economist, the manager or person in charge of the enterprises and institutions directly takes the role of the economist. Enterprises and institutions must rerecognize the importance of human resource management of economists and change their inherent ideas. Give full play to the main role of the economist's human resources to make more contributions to enterprises and institutions. This is a responsibility and a high requirement for human resource management of economists. The human resource management of economists should be aware of the competitiveness between enterprises and institutions, actively participate in

the planning, and mobilize the enthusiasm of employees while improving their own quality, so as to achieve the goal of common progress between employees and enterprises and institutions. In order to mobilize employees' work enthusiasm and enhance their sense of competition, the human resource management of economists can give old employees and new employees the opportunity to interact in the form of exchange forums. In addition, the human resource management of economists can also provide training for employees so that they can have a basic understanding of the culture, development strategy, and development status of enterprises and institutions. And further cover workplace etiquette, product knowledge, and other aspects. Promote employees to understand the current situation, advantages, and future planning of the industry more clearly, strengthen their sense of ownership, and encourage them to work harder. Perception model has a wide range of applicability, which is suitable for the HRM (human resource management) process of enterprises of various natures and can help enterprises significantly improve the scientificity and

rationality of HRM. The existing job recommendation system can provide accurate, fast, and personalized job recommendation services for job seekers by analyzing their historical search records, mining their job-seeking wishes, needs, and job-seeking tendencies [1, 2]. Through effective personnel-post matching, the quality of enterprise resource management can be comprehensively improved, which plays an important role in enterprise operation and work development. It is also very helpful for enterprises to achieve strategic goals.

In recent years, scholars have tried to apply the cutting-edge technology to the matching of people and posts. In early research abroad, Khan et al. used candidates and professional information to establish two-way recommendation between people and posts and realized the matching between people and posts [3]. Belsito et al. followed the idea of recommendation system and put forward a job recommendation system based on job seekers' preferences and interests [4]. Reachslin et al. believe that competence is the intrinsic characteristic of an individual, and there is a certain degree of causal relationship between this intrinsic characteristic and related performance in work and situation. The theory of this view of characteristics focuses on the study of people's competence from the perspective of discovering people's characteristics [5]. Baimenov et al. think that the company's operation is made up of orderly work, and different positions have different competency elements for employees, that is, the system theory of person-post matching [6]. Petros et al. pointed out that the matching process of people and posts is a dynamic matching process. The dynamic matching between people and posts is of great significance to help employees obtain career planning and personal development, give full play to their active role, and promote the development of the company [7]. Brueller et al. tried to establish the matching degree model between employees and specific positions by analytic hierarchy process [8]. Liang used AHP (analytic hierarchy process) and fuzzy comprehensive evaluation method to establish a measurement model of the matching degree between people and posts [9]. Mills et al. put forward the dynamic matching model between people and jobs; that is, the matching between people and jobs is a dynamic and active configuration process [10]; Hitka et al. conducted an empirical study on the competency models of top managers in enterprises, and the results showed that the competency models of top managers in the east and the west are similar. This study verified for the first time in China that competency evaluation can more comprehensively distinguish outstanding management cadres from ordinary management cadres [11]. Choi et al. studied various scales of HR (human resources) related to this, and this configuration type of person-post relationship mainly ensures the HR quality of various departments and positions in the enterprise through all links in the HRM process [12].

If we want to study the capability perception model, we need to fully understand not only the characteristics and needs of enterprises but also the development goals, missions, and business development of enterprises. In the process of enterprise development, the matching of people and

posts can ensure the smooth implementation of enterprise strategic objectives. In the period of production and operation, the strategic goal is an important guide, and all enterprise activities are carried out around the strategic goal [13, 14]. Excellent employees and ordinary employees will show significant differences in competency. Organizations can use competency indicators as one of the main bases for the recruitment, evaluation, and promotion of employees. Most of the studies have not adopted appropriate and reasonable empirical methods, and almost no one has done the corresponding research on the matching of people and posts after the competency model is built, let alone specific to a certain industry or field, and the research on the matching of people and posts with the competency of HRM employees in enterprises is even blank. In this paper, the ability perception model will be combined to construct the accurate matching model of job and post, and the rich semantic information contained in the long text of job and post features will be fully utilized to realize the accurate matching between job seekers and posts.

2. Research Method

2.1. Analysis of Enterprise Capability Perception Model. To distinguish people, the competency perception model should first classify and match people, positions, and posts. In order to realize the system matching among organizations, posts, and personnel, the effect of only using the general competency model is often not ideal, but we should combine the specific requirements of different organizations and posts to build different post competency models for different posts. At present, there are three main modes of developing competency model: based on the position, based on the overall value and core competence of the organization, and based on the key success factors of the industry. Among them, the post-based competency model is the most common development mode, which has strong applicability and operability, and is easy to be accepted and developed by organizations. Therefore, it has important guidance and practical significance to explore and develop the post competency model for optimizing human resource management. Quality is a necessary condition to produce high performance, and the capability perception model should be able to predict future performance and become a guarantee to support the sustainable development of enterprises in the future [15].

Post management by objectives requires the establishment of post management system, ability evaluation system, performance management evaluation system, and finally a scientific, reasonable, fair, and just HRM system which is competitive in the market and can effectively promote the realization of organizational strategic objectives. It is not difficult to find that employees' ability and quality influence employees' work performance, and to a certain extent, their ability and quality determine employees' work performance, by closely combining their positions and work behaviors.

Social role represents employees' understanding of social norms, and it is also an important embodiment of employees' behavior style, such as managers. Self-image is

a kind of self-display after employees have a clear understanding of their own abilities and roles, such as self-confidence, which often has a great impact on their performance. This study adopts the method of vectorizing each word after segmentation and then clustering the vectorized words. The results show that this method can effectively distinguish competency from other words so that they can be grouped into one category. The specific implementation process is shown in Figure 1.

Some competencies may be in high demand in all positions, and these competencies do not represent the position of HRM well. In this paper, the chi-square value is used as an index to reflect the importance of competence in the post [16]. The original assumption is that feature x_i and category c_j are independent of each other; that is, feature x_i is not representative of this category. The calculation formula is

$$\chi^2 = (t, c) = \sum_{i=1}^n \frac{(X_i - E)^2}{E}, \quad (1)$$

where X_i is the theoretical value and E is the actual value.

For the consistency judgment matrix, each column is the corresponding weight after normalization. For the inconsistency judgment matrix, each column is normalized to approximate its corresponding weight, and the arithmetic average of these n column vectors is calculated as the final weight. The specific formula is

$$W_i = \frac{1}{n} \sum_{j=1}^n \frac{a_{ij}}{\sum_{k=1}^n a_{ki}}. \quad (2)$$

It should be noted that the consistency of the judgment matrix should be checked in layer-by-layer sorting. Generally speaking, the judgment matrix is not required to strictly satisfy this property. This paper studies the difference between judgment consistency matrix and satisfaction consistency matrix. The limitation of determining the acceptance matrix by the consistency index is discussed. In order to improve the consistency of judgment matrix, a new method of correcting matrix by using the average characteristic of judgment is proposed.

In this study, structural hole index is used to measure the cohesion of a certain competency in workflow or team work [17]. The richer the structure, the more it shows that it plays a key role in the whole workflow or in the team. The calculation formula is

$$p_{ij} = \frac{a_{ij} + a_{ji}}{\sum_k (a_{ik} + a_{ki})}, \quad (3)$$

where p_{ij} is the proportion of the direct relation of i, j .

The evaluation ability level can adopt the behavior level description method; that is, when the model is built, the behaviors will be laid to different levels according to the difficulty and complexity of behaviors, and when judging, the employees will be anchored to the corresponding ability level according to the difficulty and complexity of behaviors

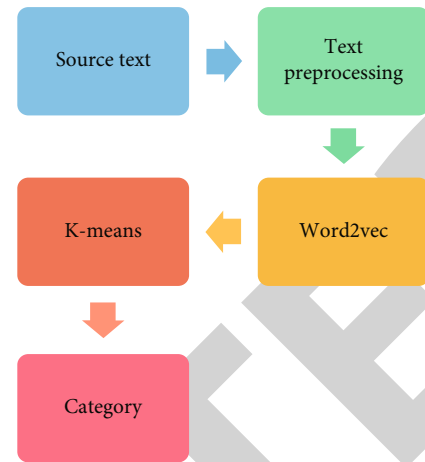


FIGURE 1: Specific process of competency feature extraction.

during the evaluation period, and the corresponding ability level scores will be given.

As far as HR quality is concerned, it is necessary to ensure the matching degree of employees' ability and quality with job requirements and company culture from the beginning of the recruitment process. After employees enter the company, they should strengthen job guidance and training to realize the dynamic matching and management between people and jobs, so as to have a competitive workforce. In the competency model constructed in this study, its hierarchical structure is shown in Figure 2.

The application of competency perception model can enable enterprises to get rich and useful information, and apply this information to recruitment and evaluation, so as to draw scientific and reliable conclusions. The construction of post competency model provides the basis and standard for people to enter posts and corresponding grades and is an effective way to realize the matching between people and posts [18, 19]. These abilities are difficult to observe and learn. The individual characteristics of factor mainly include a person's personality, motivation, values, and other characteristics, which are the deepest features and the least easy to be discovered and developed.

2.2. Establishment and Application of Depth Model of Accurate Matching between People and Posts. Person-post matching refers to the corresponding relationship between people and posts. Each incumbent has different abilities, qualities, and individual characteristics, and his or her individual needs and motivations are different. Only when the individual's needs are matched with the post compensation and organizational vision can the individual be motivated and work harder to achieve greater performance.

Personal abilities and qualities are compatible with job requirements, and personal requirements are matched with the salary level and organizational vision that the job can provide. In this form, the matching between people and posts is at its best. It is the most ideal matching form of "send," but it is difficult to realize the "send" form. In the process of post selection, if the post holder has work experience related to the post, then the company will save time and

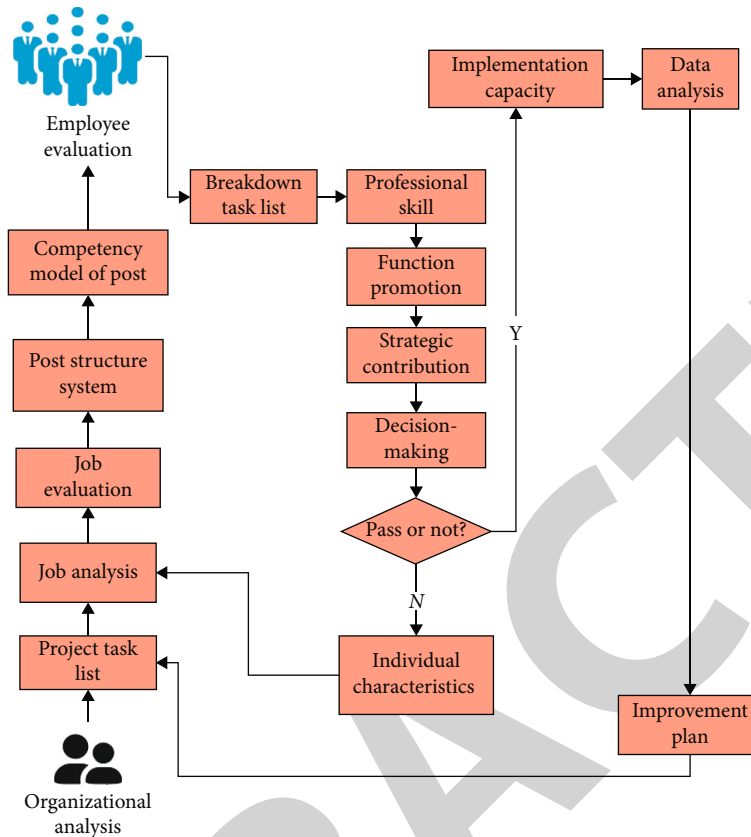


FIGURE 2: Post competency quality model.

cost for the post training of the post holder, and the post holder can adapt to the job quickly.

There are three necessary processes to realize the matching of people and posts, as follows:

- (1) The first step of job matching is to understand job requirements through job analysis, so as to realize job knowledge. For a job, it is to fully collect and observe all kinds of information such as the rights and responsibilities of the job and then analyze the work procedures and summarize the contents and rights and responsibilities
- (2) Know people: knowing people means judging people's post competence. Usually, many scientific methods, such as written test, interview, online psychological assessment, and group discussion, are used to identify people
- (3) Match with people and posts: the fundamental purpose of the whole match is to make people fit for their posts, make the best of their abilities, and finally reap the maximum organizational benefits

At present, there are two main aspects of job matching. The first is the matching between the professional skills of employees and the demands of jobs, and the second is the matching between employees' work attitude and salary return. Timely training can help enterprises get comprehen-

sive protection, and its training mainly includes four stages: demand analysis, training design, implementation, and result evaluation; enterprises should start to carry out employee training from multiple angles to ensure the perfection of training facilities and the rationality of training funds. In order to achieve the best match between people and posts, it is necessary to carry out post analysis and personnel quality evaluation according to the talent construction plan to recruit and select talents so that the match between people and posts can approach to rationality. When using certain evaluation methods, corresponding evaluation methods should be proposed for each quality element. There may be multiple evaluation methods for a certain quality element, which requires in-depth analysis, comparison, and accurate selection of discarded methods. Do not simplify or complicate the evaluation style. When introducing a new style, it should be verified to determine its applicability. In the contemporary evaluation, test method, oral test method, evaluation center method, and other evaluation methods are commonly used.

At present, most of the competency model and competency model are established by psychological evaluation method, which is inherently scientific, but it will be slightly complicated in the actual operation process of the enterprise, so it will undoubtedly reduce the actual operability of the enterprise and increase the operating cost. The essence of HR scheduling model is to analyze HR data and calculate the matching degree of people and posts. Then, according

to the score of matching degree between people and posts, the personnel can be dispatched, which can be abstracted as a recommendation model in essence.

In this paper, the basic neural network is improved, the hybrid recurrent neural network model is used, and the global model and the local model are combined, and the data features after the hierarchical operation of the model are taken as the network output, so as to realize the data processing. The loss function is as follows:

$$L = - \sum_n (y \log p + (1 - y) \log (1 - p)). \quad (4)$$

There are also feature attributes described by long texts in the characteristics of posts, and these corresponding long texts describe the specific ability requirements of posts and the skills of job seekers, which are the key information for accurate matching between posts and job seekers. By calculating the cosine distance between documents, the similarity of corresponding feature attributes is obtained:

$$S_{i\otimes j} = \frac{V_i \times V_j}{\|V_i\| \times \|V_j\|}, \quad (5)$$

where V_i represents the document vector of feature attribute i and V_j represents the document vector of feature attribute j ; $S_{i\otimes j}$ represents the similarity of feature attribute i, j long text.

In the process of bilateral matching of people and posts in recruitment, the two matching subjects, namely, job seekers and posts, respectively, give the demand information of each other according to their own conditions. The HR department of the enterprise uses multi-index method to process the index information through the information provided by both parties and then uses bilateral matching decision algorithm to effectively match people and posts. Most of the existing decision-making methods of person post matching assume that both parties of person post matching are completely rational. This is not consistent with the actual management activities, so this paper proposes a two-sided matching decision-making method considering the different psychological behaviors of the two sides. On the premise of analyzing and sorting out a variety of different types of evaluation indicators, first of all, the satisfaction value of both parties on the basis of complete rationality is obtained. Then, it introduces the regret avoidance psychology and prospect expectation psychology of both parties. This paper introduces regret theory and prospect theory to the decision-making method of person post bilateral matching. Through the construction and solution of the decision-making model, a matching method closer to the actual human resource management activities is obtained. It also proves that the results of bilateral matching are different under the influence of different psychological behaviors. The formula for calculating the satisfaction of the possessor to the post is as follows:

$$a_{kij}^1 = \begin{cases} 1, & k = j \\ \sum_{i=1}^T w_{kit}^1 a_{kijt}^1, & k \neq j \end{cases} \quad i = 1, 2, \dots, l_k. \quad (6)$$

When $k = j$, $a_{kij}^1 = 1$ indicates that the possessor has the lowest satisfaction with the original position; that is, the position applied for by the job seeker is at least better than the position currently occupied; when $k \neq j$, the occupant's satisfaction with the post is accumulated and obtained through the evaluation satisfaction under each index.

Assuming that the person-post matching coefficient based on competency is g , the actual competency of HRM employees is b , and the competency of HRM employees required by enterprises is b_0 , the corresponding weights f_i ($i = 1, 2, \dots, n$) and $\sum f_i = 1$ are given according to the different importance of different jobs. So,

$$b_0 = \sum_{i=1}^n f_i b_i, \quad (7)$$

where $\sum f_i = 1, i = 1, 2, \dots, n, n \leq 28$, n is the competency coefficient. Secondly, an expert evaluation team is formed by the enterprise to evaluate the competence of employees.

In the process of adjustment, enterprises should combine the specific situation. If there is a problem with employees' skills, managers should first consider vocational training for employees. If employees' working ability cannot be improved through training, managers need to consider post adjustment and whether to transfer them to posts with lower requirements.

Group the data to get a more comprehensive understanding of the characteristics of the HR model. At the same time, the data is saved in the data warehouse, which can support the subsequent data model training. The depth model of accurate matching between people and posts combined with ability perception is shown in Figure 3.

In the text similarity calculation layer, the attribute values with long text in the matching features of people and posts are vectorized by using Word2vec method, and then, the matching degree of corresponding features is evaluated by calculating the cosine distance between corresponding features. Embedding each word into the vector model, more strictly speaking, generates a vector space with a large number of text so that each word can find the corresponding vector, and the similarity between words can be calculated. Then, the data is abstracted, preprocessed, encoded by an encoder, and stored in a data warehouse. If the new evaluation results show that people and posts are matched, then the adjustment scheme can be determined; if not, the adjustment will be made according to the third step again until people and posts are matched to achieve dynamic balance.

3. Result Analysis

This paper extracts and analyzes the research resumes through network big data. According to the fields in the

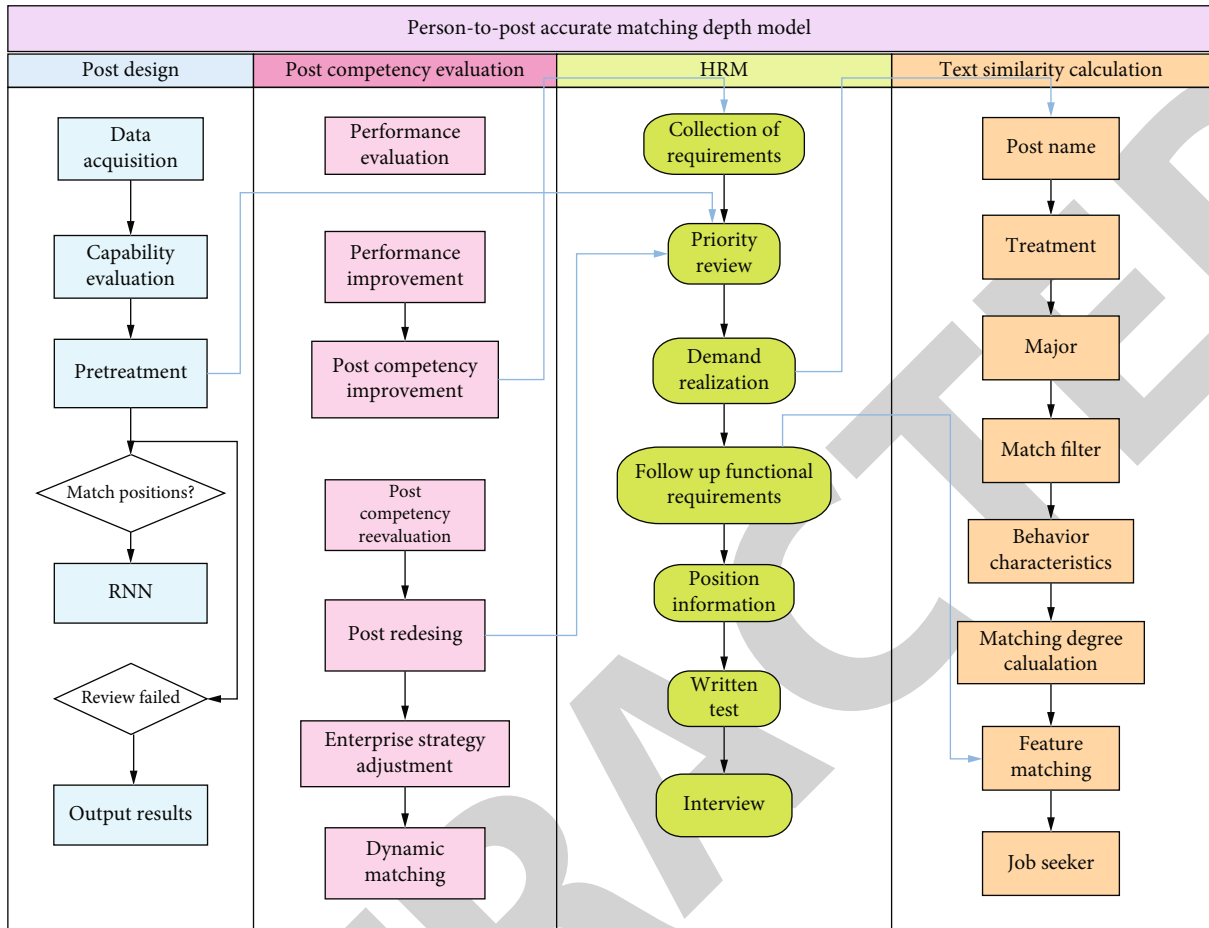


FIGURE 3: Depth model of precise matching between people and posts combined with ability perception.

obtained resume data and combined with the opinions of relevant experts, different variables such as working hours, enterprise scale, rank, highest degree, school grade with the highest degree, HRM major, student cadre experience, internship experience, and award-winning experience are extracted from the resume data and processed in a structured way.

Taking the average of the total scores of three experts as dependent variables and eight variables extracted from resumes as independent variables, multiple regression analysis was carried out to explore the factors in resumes that have significant influence on competency, and they were used as input variables of competency prediction model. The results are shown in Figure 4.

We can see that whether or not there is internship experience has a significant impact on competency at the level of 1%, and its coefficient is 1.32. Compared with those who have no internship experience, those who have relevant experience have a higher level of competence. This further proves the importance of practice for HRM work. High-level schools have certain advantages in terms of faculty, curriculum, and the combination of theory and practice. Students will get more resources and have more practical opportunities during their education, which will have a positive impact on their competence.

Combine the development strategy of the enterprise, fully consider the requirements of the enterprise in HRM posts, and design, adjust, and improve the current HRM posts. For job seekers, if they are matched to poor positions, there will be a big psychological gap, and then, they will lose interest and enthusiasm for work. As for the position, if you match a poor job seeker, it violates the idea of recruiting excellent talents, which is obviously unacceptable to the enterprise. All job seekers form a matching pair with the position that prevents it, and all matching pairs form a matching scheme. Job seekers who are not blocked by jobs and jobs that are not blocked by jobs are not matched.

Performance management can not only significantly improve employees' work enthusiasm and efficiency but also play an important role in promoting the smooth development of departmental work. Although HR planning includes three aspects, their common foundation is based on the establishment and analysis of the post ability perception model, so it can provide scientific, reasonable, and effective information for the HRM work of enterprises, help enterprises achieve accurate HR planning, and ensure the full use of HR.

The relative importance weights of many indicators in each level to the related indicators in the previous level have been obtained, and the total ranking is made from top to bottom. The results are shown in Figure 5.

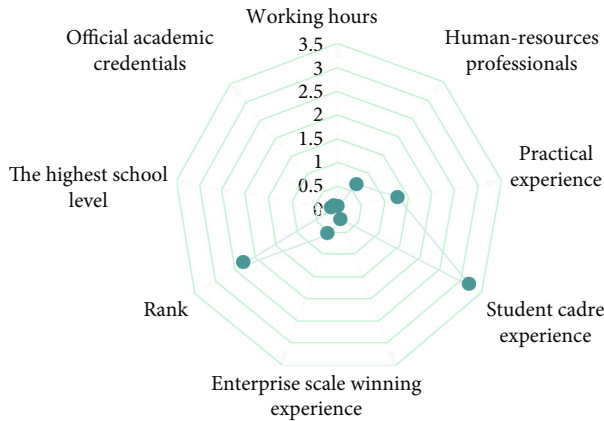


FIGURE 4: Regression result.

According to the calculation results, it can be seen that the strategic management and change management have relatively high weights, while HRM functional management and enterprise support management have relatively low weights in the latitude of enterprise HR supervisor competency evaluation. In addition to the necessary requirements for posts, enterprises also have an in-depth understanding of employees' mental health. With the all-round development of economy, the competition among enterprises also increases invisibly, which also intensifies employees' psychological pressure. Every employee has his own advantages and disadvantages, so managers need to put forward corresponding suggestions according to the actual situation of employees when they carry out HRM.

An enterprise's recruitment method based on capability perception model is an innovation based on the traditional recruitment method to meet the requirements of the development of the times. Conduct situational interviews and behavioral event interviews during the interview process. Add the weight of competency score to the comprehensive recruitment score, as shown in Figure 6.

Comprehensive practice shows that interviewers can evaluate the implicit characteristics of candidates more by means of scenario simulation and case analysis, so as to judge the matching degree of candidates and then make the decision on whether to hire or not according to the post ability perception model. Improve motivation, attitude, and behavior, so as to be better qualified for the current job or new job, and then, promote the improvement of organizational efficiency and the realization of organizational strategy and promote personal career development.

The most critical method of training demand analysis based on the competency perception model is to compare the difference between the employee's actual competency and the target post competency. The person position matching in the enterprise's human resource management is to select the appropriate personnel and match the determined positions. The person who is engaged in this position needs to have the ability to complete the job responsibilities well. These indicators should be measurable, observable, and instructive and have a key impact on the personal performance of employees and the success of the enterprise. At

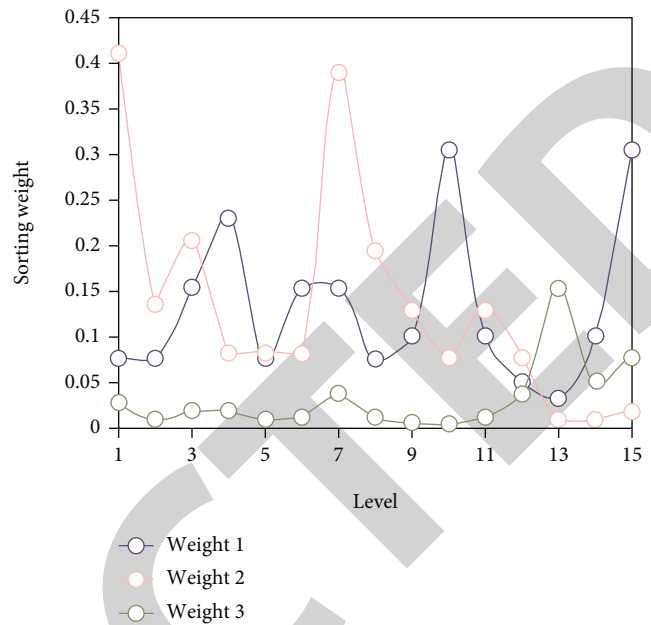


FIGURE 5: Comprehensive ranking result.

the same time, it is necessary to adjust various problems in the combination process to ensure the realization of the post objectives. Only by getting the correct post gap information can the training demand be accurately analyzed. The training content corresponds to each competency item, which constitutes the corresponding relationship between training content and competency item. Then classify it, and determine the training content and training methods according to it so that training courses can be developed in a targeted manner, and the main contents of training courses formulation can be guaranteed, including training course name, category, training method, training hours, training objectives, training contents, and assessment methods.

In this paper, the data classify and sort out the HR data of enterprises, and the data includes three types: personnel information, personnel evaluation matrix value, and personnel ability matrix value. Then, the feasibility of the proposed algorithm is compared. The proposed algorithm, Ref. [14] method, and Ref. [16] method are used for model training and experimental analysis. The evaluation indexes are accuracy, recall, and $F1$ value. The experimental results are shown in Table 1.

It can be seen that compared with the other two algorithms, the $F1$ value of the proposed algorithm is obviously improved, and the proposed algorithm performs best, with the $F1$ value of 0.829. In this paper, the method of global network plus local network is used, which can effectively improve the hidden features of data, then improve the training quality of data, and improve the matching degree and recommendation accuracy of the algorithm.

Person-post matching in HRM of an enterprise is to choose the right person to match the determined position and, at the same time, take corresponding measures to combine the two together. At the same time, it is necessary to adjust all kinds of problems in the process of combination, so as to ensure the realization of the post goal.

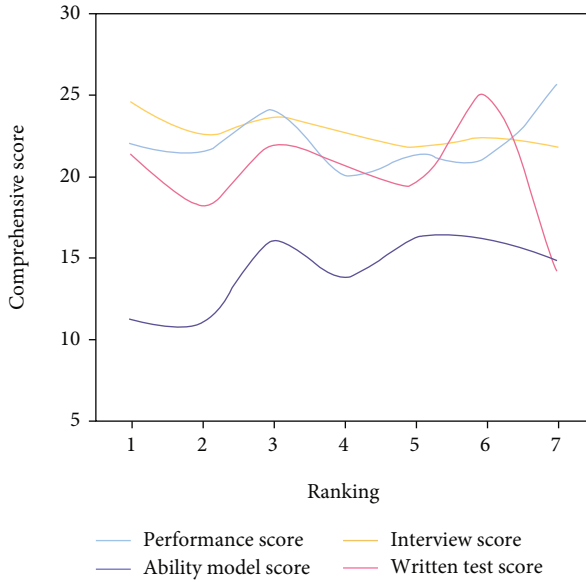


FIGURE 6: Comprehensive score of internal employment.

TABLE 1: Comparative experimental results.

Model	Accuracy	Recall	F1 value
Algorithm in this paper	0.816	0.833	0.829
Ref. [14]	0.698	0.603	0.667
Ref. [16]	0.771	0.729	0.753

The core of job matching is that the matching between people and jobs can reach the most reasonable state, that is, people can play the most effective role in jobs, and jobs can also give people the greatest satisfaction, so as to obtain the best performance. The post competency model can help employees to further clarify their quality development goals, so as to carry out career development planning more effectively, point out the development path for employees, and promote employees to improve their sense of responsibility for their own career and organizational development.

4. Conclusion

Job matching is a relatively complicated process. To achieve this, it is necessary for enterprises to optimize the distribution of jobs in combination with their own development in order to improve the job matching degree. In this paper, the depth model of accurate matching between people and posts based on ability perception is put forward, which takes into account the needs of job seekers and posts in reality, as well as the characteristics of occupiers in matching. By using genetic algorithm to solve the multiobjective optimization model, the best matching scheme under different requirements can be obtained. Compared with the other two algorithms, the F1 value of the proposed algorithm is obviously improved, and the proposed algorithm performs best. It shows that this model can effectively imitate the behavior of experts and make competency evaluation based on resumes, thus realizing the automatic operation of

matching ability and position in resume screening, effectively improving the efficiency, and providing a reference for enterprises to accurately position high-quality talents. However, the study still has some limitations. The model also needs to consider the different needs of job seekers and positions in the real world and needs to be expanded and supplemented 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 author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] S. I. Wong and S. R. Giessner, "The thin line between empowering and laissez-faire leadership: an expectancy-match perspective," *Journal of Management*, vol. 44, no. 2, pp. 757–783, 2018.
- [2] J. T. Kulas, D. H. Robinson, J. A. Smith, and D. Z. Kellar, "Post-stratification weighting in organizational surveys: a cross-disciplinary tutorial," *Human Resource Management*, vol. 57, no. 2, pp. 419–436, 2018.
- [3] Z. Khan, V. Soundararajan, and A. Shoham, "Global post-merger agility, transactive memory systems and human resource management practices," *Human Resource Management Review*, vol. 30, no. 1, article 100697, 2020.
- [4] C. A. Belsito, C. R. Reutzel, and J. D. Collins, "Human resource executives and post-IPO firm growth," *Journal of Organizational Change Management*, vol. 31, no. 7, pp. 1401–1418, 2018.
- [5] R. Weech-Maldonado, J. L. Dreachslin, J. P. Epané, J. Gail, S. Gupta, and J. A. Wainio, "Hospital cultural competency as a systematic organizational intervention: key findings from the national center for healthcare leadership diversity demonstration project," *Health Care Management Review*, vol. 43, no. 1, pp. 30–41, 2018.
- [6] A. Baimenov and S. Liebert, "Governance in the post-Soviet era: challenges and opportunities," *Public Administration Review*, vol. 79, no. 2, pp. 281–285, 2019.
- [7] G. P. Petros, B. S. Martina, and K. De, "The effect of human resource management on performance in hospitals in Sub-Saharan Africa: a systematic literature review," *Human Resources for Health*, vol. 16, no. 1, p. 34, 2018.
- [8] N. N. Brueller, A. Carmeli, and G. D. Markman, "Linking merger and acquisition strategies to postmerger integration: a configurational perspective of human resource management," *Journal of Management*, vol. 44, no. 5, pp. 1793–1818, 2018.
- [9] Y. Liang, "The emerging labour market and transformation from state amateurs to professional athletes," *Communist and Post-Communist Studies*, vol. 52, no. 4, pp. 379–390, 2019.
- [10] K. E. Mills, D. M. Weary, and K. M. Keyserling, "Graduate student literature review: challenges and opportunities for human resource management on dairy farms," *Journal of Dairy Science*, vol. 104, no. 1, pp. 1192–1202, 2021.

Retraction

Retracted: Innovation and Research of Digital Artwork Design Based on Big Data and Integrated Media

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Zhai, "Innovation and Research of Digital Artwork Design Based on Big Data and Integrated Media," *Journal of Function Spaces*, vol. 2022, Article ID 6333688, 10 pages, 2022.

Research Article

Innovation and Research of Digital Artwork Design Based on Big Data and Integrated Media

Wenbin Zhai 

College of Culture and Arts, Zhengzhou Tourism College, Zhengzhou 450009, China

Correspondence should be addressed to Wenbin Zhai; zhaiwenbin@zztrc.edu.cn

Received 26 July 2022; Revised 3 September 2022; Accepted 7 September 2022; Published 19 September 2022

Academic Editor: Miaocho Chen

Copyright © 2022 Wenbin 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.

We are in the era of big data, and art, which has developed along with human history, has been given new vehicles and forms of creation with the use of integrated media technology. This study first introduces artwork and its definition and classification, then briefly describes the definition and development of digital art, and focuses on the development, contribution, and drawbacks of digital artwork NFT (nonfunctional token). This study divided the design model into four layers based on the neural network calculation of factorization machine to build, and the data of each layer was analyzed and processed to build the FNN algorithm model. It was found that the largest proportion of NFT artworks circulating in the trading market is in paintings, at 55%. The HSV color space was selected as the color perception model, and then the picture pixel positions were determined, and redundant pixels were eliminated. To avoid data overflow, the spatial memory occupied by the pixels of the painting was calculated. The results showed that the growth of pixel space complexity was smooth and converged to a straight line, indicating that the original pixel data extracted in this experiment was stable. The circle was selected as the basic guideline for the creation of four digital art paintings with an overall uniform but innovative style.

1. Introduction

Computer technology has achieved an explosive development after 2000, and along with the popularity of computers and smartphones, the number of Internet users has shown an exponential growth. Technological advances have also led to the rapid development of the Internet, and the amount of information available to humans today is now dozens or even hundreds of times greater than before [1]. The massive amount of information has brought convenience and also brought the problem of information overload, which is a great challenge for both producers and users of information. Specifically for art creators, the content they produce is easily overwhelmed by the huge amount of data and does not stand out from the rest. Users of information are faced with a huge amount of information, and it is difficult to filter out the truly useful information, making time and energy wasted [2]. In the era of rapid changes, there are new technological miracles every day, and most of these miracles are created by having innovations. A well-known American entrepreneurialist once said that mankind is experiencing a silent revolu-

tion, and this revolution is a profound revolution in the spirit of innovation. In today's globalized economy, innovation has become the first driving force of world development. And in the field of art creation, innovation has the same important status.

Neural networks are new and advanced big data processing technologies that have received much attention in recent years and have been used in every bit of human life and production [3]. Neural network algorithms are inspired by the activity of human consciousness, and scientists have invented artificial neural networks in the field of artificial intelligence, which is the most fundamental root of the modern technological building of mankind [4]. Human consciousness is the product of constant transmission of information transmitters between neurons in their own physiological structure, and only by constant transmission and circulation can they think continuously. And humans constantly adjust and update the connection weights between neurons in order to perform better and faster thinking activities. Human consciousness is nonlinear and generates new signals with increasing knowledge and

experience, and these new signals cause new stimuli to the neurosynapses, which in turn generate new thoughts [5]. Artificial neural networks operate on the same principles as the human mind activity, where data is passed, processed, handled, and distributed at various nodes. Like human consciousness, it adjusts to changes depending on the input signals and has the same adaptive and autonomous learning capabilities as humans. Neural network nodes in computer technology are as important as human neurons, and the number of nodes and the way they are connected determines what kind of nonlinear mapping capabilities they have [6]. Factor decomposer-based neural networks are the earliest and currently more developed artificial intelligence network techniques, which use error backpropagation to train the data. The biggest feature of digital art is that it can strip off the internal relationship of color and graphics to directly express logic. This internal relationship is built on operation and data, and operation is the essence of cognition. The difference between digital art and previous art forms is that the internal logic used is similar to the essence of things, rather than guiding people to read and understand the internal relationship through appearances. Therefore, the final result presented by the data is the content directly reaching the essence of painting. It does not need to have the texture, color, composition, and other representations formed on the surface of traditional painting but a new image represented by the logic composed of internal relations. In essence, digital painting is not a type of painting art, and it is only a type of digital art with painting form or graphic characteristics.

Digital artworks are unique in the era of big data, and virtual artworks only emerge when technology has developed to the extent it has today [7]. Digital artwork is a visual representation created using computer technology and related software, where the related software varies depending on the type of artwork. Digital art paintings would use drawing software, digital art statues would use modeling software, and digital art songs would use arranging and mixing software. The fusion media technology provides the right conditions for the creation and development of digital art, which can only be created, disseminated, sold, and collected so smoothly and easily if all the multiple media are used wisely [8]. However, just like traditional art, digital artists also encounter creative bottlenecks. The huge demand in the art trading market indicates a huge potential for development; so, this study will use algorithms to create a relevant model based on big data melting media technology. The model will then be used to bring innovation and inspiration to the creation of digital artworks. In summary, this study has strong relevance and great digital business value.

2. Overview of Digital Artwork

Digital art is essentially the same as traditional art, an artistic practice, except that the tool for its creation has become computer technology [9]. Art has been present in our culture since the dawn of mankind. Painting, sculpture, music, photography, and games have progressed along with mankind, and they have given him spiritual food about beauty.

Because of the advancement of global digitization, more and more artists are creating digitally and do not materialize them but exist only as pixels [10]. This virtual property does not make digital artworks more or less valuable but more convenient to store and transmit and more topical. It is possible to distinguish digital art from traditional art in terms of output methods, which are often used for digital output. It is also flexible and diverse. For example, a video collection of Super Bowl football stars' scoring moments can be a work of art, a symbol abstracted from a real photograph can be a work of art, and an emoji packet that is widely distributed on the Internet is also a digital work of art [11–13].

Digital art is gaining more and more attention, and after being affected by the epidemic, the number of various types of digital art exhibitions is increasing year by year because of its ability to be held on the Internet. Various types of well-known academic institutions have also established big data integrated media disciplines and digital art research centers. However, because of its systematic nature, it has three major congenital deficiencies at the beginning. The first is the problem of copyright protection. Compared to traditional artworks, digital collections only need to copy the station body to complete the copy. And it is more easily tampered with the creator's ideas and style can be easily borrowed and copied. Secondly, the current dissemination channels are limited, and both offline and online channels have different degrees of restrictions on the display of digital artworks. Traditional places such as galleries, museums, shopping malls, and parks prefer to exhibit physical artworks, as traditional artworks usually do not require special equipment such as digital screens. Although the number of online art exhibitions is increasing year by year, it is still insufficient in relation to the huge number of artworks. The final point is that the lack of trading volume and the inability to circulate commercially greatly limit its development. Traditional art can be auctioned at Christie's and Sotheby's and can also be circulated in various art galleries, antique streets, and e-commerce platforms [14]. Because of its easily copied nature, for a long time, there was no trading channel for digital art at all.

Digital art is usually not the digitized original text data and the original audio/video data itself, but when these original materials are designed into a larger digital work through computer technology and information art and become a part of this large work, we can call it digital art. The emergence of nonfungible token (NFT), however, is well positioned to overcome all three of these shortcomings. NFT emerged as a new wave in the art world, becoming known in 2020 and causing a surge in the trading of digital art by the end of 2021. The emergence of NFT has led to sudden overnight riches for artists who were not famous and has led many already famous artists to devote themselves to its study and creation. The landmark event of NFT was the sale of *The First 5000 Days* by American programmer Michael Winkelmann at Christie's for \$70 million [15]. To this day, it is not only artists various creators who sell NFT but also traditional game maker Konami has created its Castlevania-themed NFT for sale, eventually making over \$160,000 in profit [16]. Figure 1 shows the iconic NFT content.

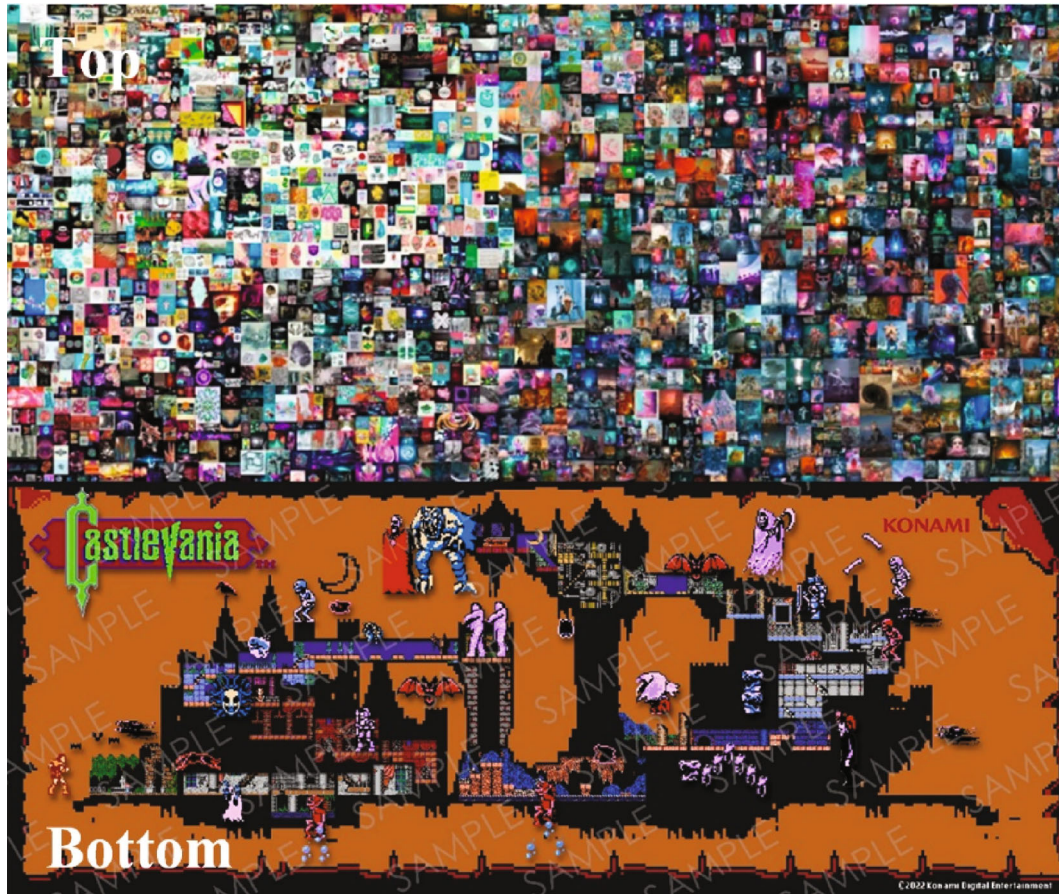


FIGURE 1: Iconic NFTs: The First 5000 Days by Beeple (top) and Konami's Castlevania-related works (bottom).

NFT is unique in terms of copyright, creation, and trading, and these characteristics set it apart from digital artwork in the traditional sense. In terms of copyright, everyone in the Internet can download, copy, share, and collect NFT artwork, but its owner holds the only genuine artwork. Each NFT item has a unique encryption code; so, it can be queried and traced by everyone in the blockchain, which eliminates the possibility of it being forged and pirated. In terms of creation, NFT is more free, and modern art can intercept the classic artwork for the artist's own expression and interpretation to form a new NFT artwork. The most outstanding feature is that in terms of transactions, the same as copyright transactions are also completed on the blockchain, which eliminates the possibility of fraudulent transactions. And as with securities investments, similar to stocks where one share can be purchased, NFT can be purchased in pieces [17]. These features, especially the trading measures, have led to an unprecedented boom in the NFT digital art trading market. Such a market has first attracted traditional visual content creators in advertising, film and video, and games. Numerous content creators have transformed into NFT art creators, allowing them to break away from the endless demands with their A-side clients and apply their accumulated technical and design talents to the digital art field. Second is the entry of traditional majors in the consumer field, with large companies such as Coca-Cola, Nike, and

Adidas relying on their own products and marketing resources in the layout of the NFT art field [18]. The booming market has also given rise to specialized trading sites, and currently, the world's first and largest NFT trading site is OpenSea, which accepts individuals and art groups to post and buy and sell NFT digital artworks on it.

Behind the boom in the market, NFT has also revealed many problems. The first and foremost is that the NFT art market is controlled in the hands of a few people, and its share basically follows the rule of two or eight. About 20% of NFT digital artwork holders occupy 80% of the value of the market, which makes it difficult for ordinary people to profit from this market, which is not conducive to the long-term development of the market [19]. Secondly, because of the lowering of the creation threshold, the identification of artwork value is in a vague and confusing state, and there is no authoritative organization or institution that can identify the reasonable value of a digital artwork [20]. This adds a great deal of uncertainty to the market, making it less stable.

3. Neural Network Algorithm Based on Factorization Machine

The currently used CTR designs are linear models, such as logistic regression, Bayesian and plain Bayesian, and FTRL. Linear models use a large number of sparse one-hot codes

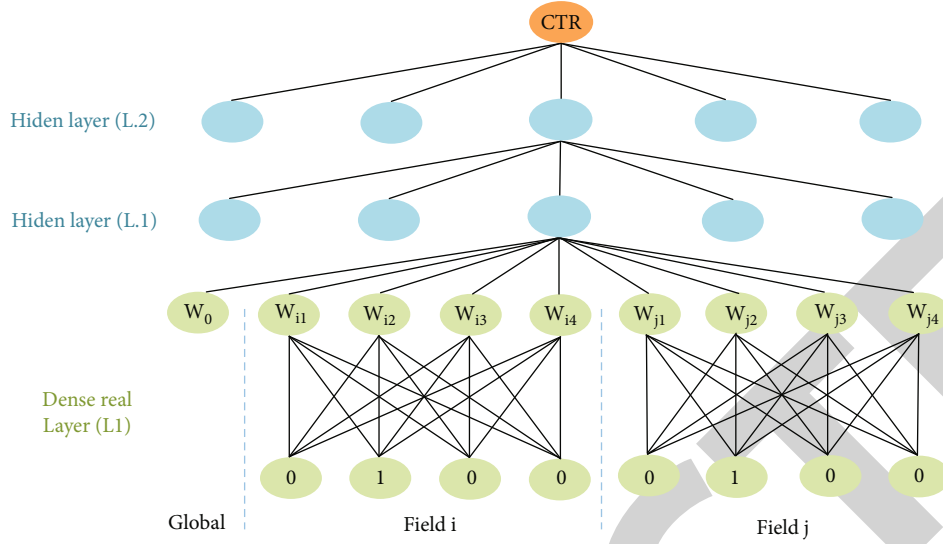


FIGURE 2: Four-layer FNN algorithm model structure diagram.

for computation, which are simple and effective but have poor performance and cannot design complex relationships between features. The nonlinear model is an algorithmic model that maps binarized features to a low-dimensional space on the basis of the linear model and designs the relationship between features by inner product, which eventually leads to improved performance [21]. The gradient boosting tree algorithm is its representative, by constructing a tree and thus achieving automatic learning. However, both linear and nonlinear models are severely limited in expressiveness and have poor generalization ability.

In the linear model, the common way to extract high-order features is based on manual and prior knowledge. When the feature dimension becomes large, the feature combination becomes extremely difficult, and the labor cost becomes more expensive. The factoring machine is a neural network system embedded with embedded layer. Its main function is to construct hidden variables between features, which measure the second-order combined features based on all features. Neural network algorithms can be learned and designed from the original features before achieving a high-dimensional feature representation through unsupervised training. The Factorization-machine supported Neural Network (FNN) algorithm selected for this study can perform cyclic supervised learning using FM. Thus, sparse features are effectively reduced, and continuous dense features are obtained. On top of the embedding layer, this experiment constructs a multilayer FNN for digital artwork design modeling.

With one-hot encoding, the FNN algorithm inputs category features that are fielded. It has multiple cells, each representing a specific field with only one positive value, and the others are negative. The features, after encoding, are represented in the following Figure 2.

The top layer above will output a real number, and the bottom layer is the FM layer. The probability of its click is expressed by the following equation.

$$\hat{y} = \text{sigmoid}(W_3 l_2 + b_3), \quad (1)$$

where l_2 is the input feature of the layer in which it is located, which can be calculated by Eq. (2).

$$l_2 = \tanh(W_2 l_1 + b_2). \quad (2)$$

\tanh is the activation function selected for this study, and l_1 in the function is the original feature input value, calculated as

$$l_1 = \tanh(W_1 z + b_1), z = (w_0, z_1, z_2, \dots, z_i, \dots, z_n). \quad (3)$$

The number of domains is represented by n in the above equation, and z_i denotes the vector parameter in the i -th domain, calculated as

$$z_i = W_0^i \times x[\text{start}_i : \text{end}_i] = (w_i, v_i^1, v_i^2, \dots, v_i^k). \quad (4)$$

Initialization of the vector FM of the first layer yields

$$y_{\text{FM}} = \text{sigmoid}\left(w_0 + \sum_{i=1}^N w_i x_i + \sum_{i=1}^N \sum_{j=i+1}^N x_i x_j\right). \quad (5)$$

With the above computational steps, the FNN algorithm can improve the model learning ability and get more potential data relationships. Continuing to add the weights of the implied layers to the algorithm model for calculation, the choice of using SGD for the prediction of FM weights can effectively reduce the computational effort. Firstly, all four layers are initialized using the crossloss function.

$$L(y, \hat{y}) = -y \log \hat{y} - (1 - y) \log (1 - \hat{y}). \quad (6)$$

Since the input elements have only one positive value, then back propagation can be used to update the weight values and make the FNN algorithm model more effective in learning.

Then, the regularization process is carried out, and there are two mainstream regularization methods, namely, L1 regularization and L2 regularization. L1 regularization is done by adding an absolute value to the loss function, and the common formula is

$$L(\theta) = L + \lambda \|\theta_i\|_1. \quad (7)$$

Similar to the L1 regular, the L2 regular also adds a parameter to the loss function but switches the Eulerian parametrization, and its expression is

$$L(\theta) = L + \lambda \|\theta_j\|_2. \quad (8)$$

A comparison of the two shows that the L1 norm prefers smaller scale algorithmic models because the sparsity of the solutions computed is stronger. Therefore, the L1 regularization does not give too much weight to the outlier features, and the result will lose some accuracy. Also, in order not to overfit, the L2 regularization is used in this study, and the regularization process is as follows.

$$\Omega(w) = 2\|W_0\| + \sum l = 3(\|W_l\|^2 + \|b_l\|^2). \quad (9)$$

Calculate the error generated by each iteration of the FNN neural network as

$$\delta_j = \frac{\partial L}{\partial z_j} = \frac{\partial L}{\partial a_j} \times \frac{\partial a_j}{\partial z_j} = \nabla a_j \times \sigma(z_j). \quad (10)$$

Vectorization is then performed to obtain the following:

$$\delta^L = \nabla a^L \odot \sigma(z^L), \quad (11)$$

where \odot is the product of the corresponding weight values. Recursion of the error for each layer again leads to the calculation of

$$\delta_j = \frac{\partial C}{\partial z_j} = \sum_k \frac{\partial C}{\partial z_k} \times \frac{\partial z_k}{\partial a_j}, \quad (12)$$

$$\sum_k \delta_k w_{kj} \sigma(z_j) = \sum_k \delta_k \times \frac{\partial(w_{ki}a_i + b_i)}{\partial a_j}. \quad (13)$$

The gradient of the algorithmic model weights of FNN is

$$\nabla w_{jk} = \frac{\partial C}{\partial w_{jk}} = \delta_j \frac{\partial(w_{ki}a_i + b_i)}{\partial w_{jk}}. \quad (14)$$

The bias gradients for each layer in the model structure are

$$\nabla b_j = \frac{\partial C}{\partial b_j} = \delta_j \frac{\partial(w_{ji}a_i + b_i)}{\partial w_{ji}}. \quad (15)$$

Within this gradient range, and taking into account the error of the obtained results, the final digital art design

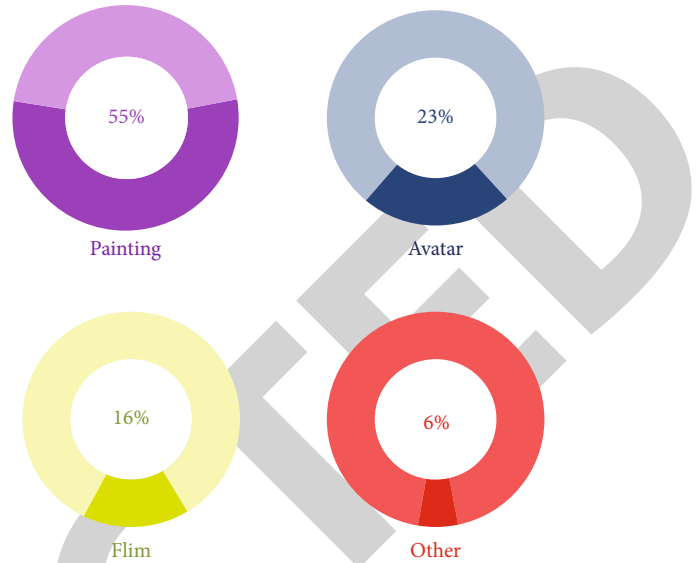


FIGURE 3: Proportion of different types of NFT artworks circulating in the trading market.

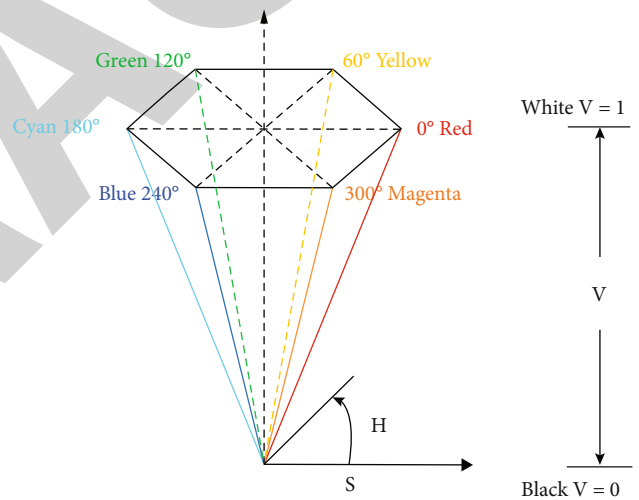


FIGURE 4: Hue, saturation, and value color space model structure diagram.

model based on the neural network algorithm of the factorization machine is obtained.

4. Examples of Digital Artwork Design and Its Analysis

4.1. Percentage of NFT Types. Before using the FNN algorithm model for digital artwork, also known as NFT goods for creation, this study conducted statistics and analysis of the collected big data. It was found that the proportion of different types of artwork in NFT is roughly the same as in traditional artwork, but there are some differences, and the results are shown in Figure 3.

The statistical proportion chart is clear that the largest proportion of NFT artworks circulating in the trading market is in paintings, at 55%; in second place are avatars that

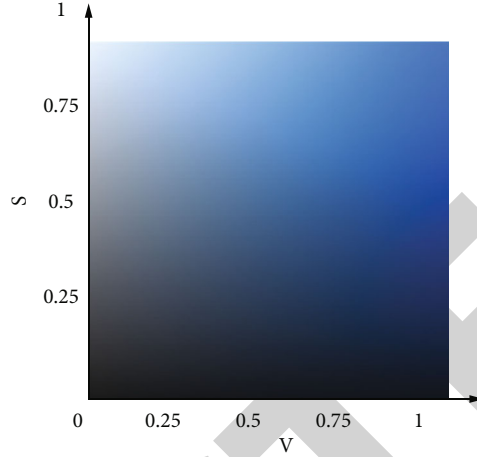
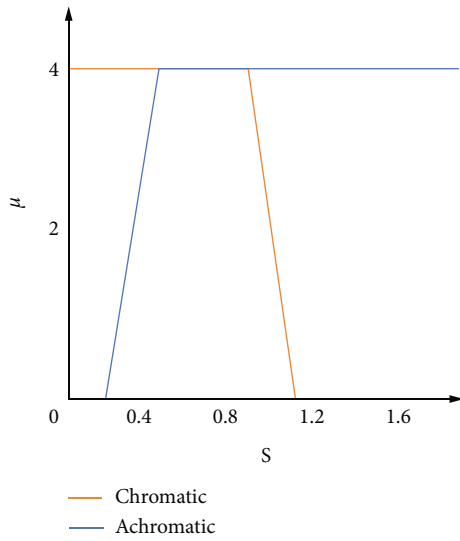


FIGURE 5: Variation of color with S and V at $H = 0$ (left) and S -component fuzzy set partition (right).

do not exist as traditional artworks, accounting for 23% of the total, because the rise of design software in the Internet has given rise to the creation of avatar art. This is followed by film and television works, which account for 16% of the total. The reason for the lower percentage is that the equipment needed to shoot a film and the technical threshold required of the creator are much higher than that of painting. By analyzing the above data, and considering the arithmetic cost of the algorithm and the creation cost, this study decided to use the model established in the previous section for the design of digital artworks of paintings.

When creating digital art, another difference from traditional art painting is that digital artwork needs to be presented on different display devices. Therefore, it is necessary to select a uniform color standard, so that the viewer can see the artwork on different display devices with a certain degree of uniformity. This experiment uses HSV (hue, saturation, value) color space, which is the most widely used model of human visual color perception. It characterizes color by its luminance, hue, and saturation, and its structural model is shown in the following Figure 4.

It is obvious from the figure that the HSV color space is structured as an equilateral hexagonal cone, with black at the apex of the equilateral hexagonal cone because the luminance of this point is 0. White is located at the center of the bottom of the equilateral hexagonal cone because the luminance of this point is 1, the point with the highest luminance. As indicated by the central axis on the left side of the figure, V represents the brightness of the color, and the colors on the central axis represent all the colors of this gradient from black to gray to white. h is the hue of the color, and the six prongs of the equilateral hexagonal cone represent the six main colors of this model, where 0° is red, 60° is yellow, 120° is green, 180° is cyan, 240° is blue, and 300° is magenta. These prisms are arranged equidistantly around the central axis, from 0° to 360° containing all the colors visible to human vision. s refers to the saturation of the color, a parameter used to characterize the purity of the color. The

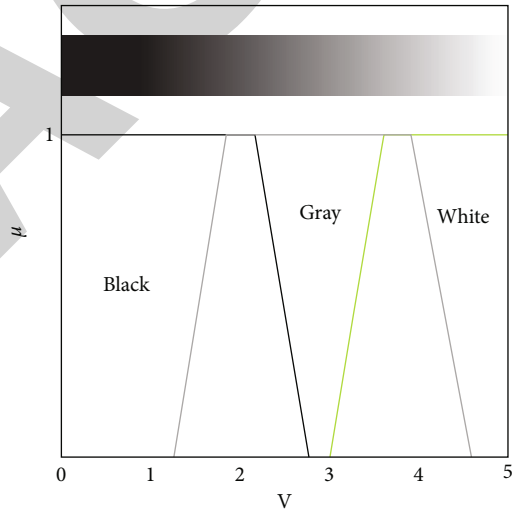


FIGURE 6: Plot of ceramic material color as a function of V and V -component fuzzy set partition.

dots in the diagram represent colors, and the closer the dot is to the central axis, the less saturated the color is. h , s , and v represent luminance, hue, and saturation, respectively, and are three separate color parameters, one of which can be adjusted at will without affecting the other two. Therefore, the NFT digital artwork can be designed in a wider range of choices, without considering the mutual influence of various color parameters, effectively improving the efficiency of the model calculation.

Figure 5 shows the color change (left) of S and V and the result of S -component fuzzy set division (right) when $H = 0$. The display of color on different screens follows the general rule of color, if the color is less saturated, it means that the purity of this color is lower. When saturation is low, it is more grayish in human visual perception. As shown in the left image above, when the saturation is 0 and the luminance

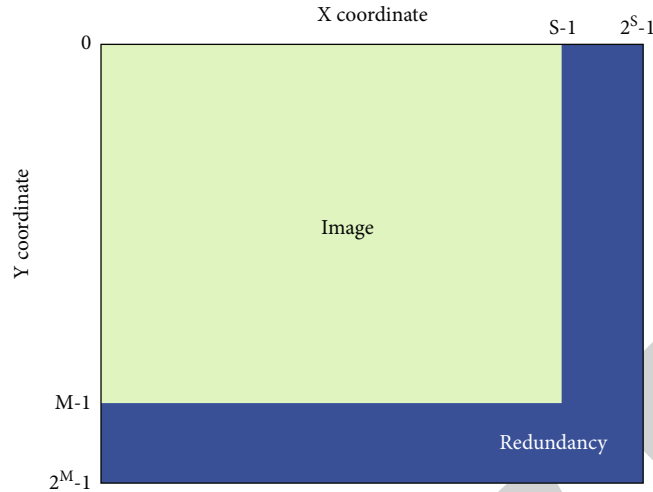


FIGURE 7: Effective color pixel and redundant color pixel map processed by the algorithm model.

is close to 0, the color near the vertical axis is perceived as black, regardless of the hue. When the saturation is still 0 and the hue is also close to 0, the color near the horizontal axis is perceived as black to gray to white as the brightness increases. Only when all three are in a suitable interval will humans perceive this color as it was traditionally perceived. Therefore, in the image difference prediction model, a trapezoidal blurring of saturation will be performed using the fuzzy method, and then the right image in the above figure is obtained. Again, because of the corrected subjectivity in the human eye's perception of color and noncolor, there is an overlap between the two regions.

As shown in the black and white gradient diagram above, when both saturation and color values are 0, the normalized interval of luminance is 0 to 1, showing a gradual process from black to gray and then from gray to white. This display leads to the inference that the most intuitive difference in noncolor for human perception is the difference in luminance. Therefore, trapezoidal blurring of noncolors using the fuzzy method again yields the right panel in the above figure. Similar to the trapezoidal blurring results in Figure 6, there are overlapping areas of black and gray and gray and white because of the human eye's own reasons.

Then, the pixel positions of the painting artwork screen to be created are determined, the target color pixels and redundant color pixels are analyzed using image processing, and the results are shown in Figure 7.

The green part in the above figure is the valid pixel area, which is the block of pixels that can be used by the FNN algorithm to design the model, while the blue part is the redundant pixel area, which needs to be removed during the design. The redundant pixels include both pixels that overlap with the valid pixels and pixels that contain useless information, both of which interfere with the computation of the model and slow down the computation. Both types of pixels interfere with the computation of the model and slow down the computation. In addition, redundant pixels deviate from the original color scheme of the design, and

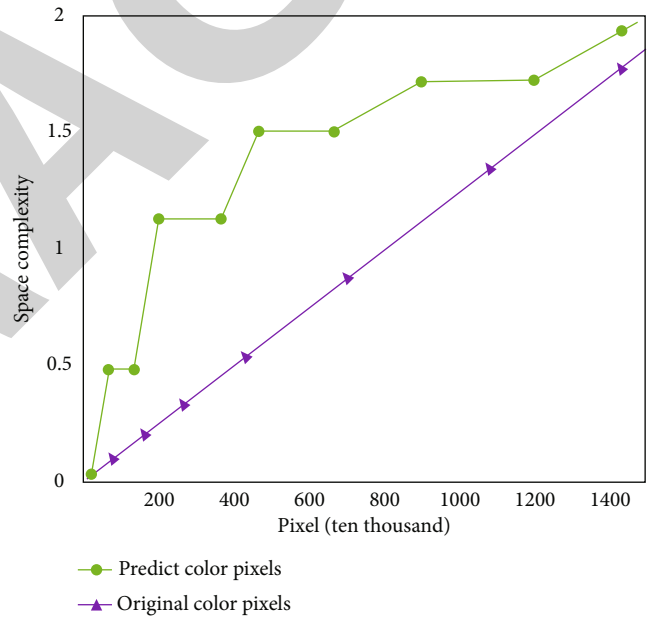


FIGURE 8: The spatial complexity of the original color matching pixels of digital artwork and the resulting pixels after design.

their presence gives the designed digital artwork a dirty look and must be removed.

In order to avoid data overflow, the spatial memory occupied by the extracted color-matching original pixels and the pixels obtained after performing the design are also calculated and considered. There is a special processing algorithm for overflow color. Like the chroma key, you need to know the color of the matting target, which is the basis for removing the overflow color. The YUV space of the original image is transformed, and then the original signal is processed according to the chromaticity and saturation information of the target color (that is, the color of the overflow color), the color similar to the overflow color is removed from the original signal, and then the image processed by



FIGURE 9: Four digital art paintings drawn with a circle as the artistic creation concept using big data fusion media technology.

the overflow color is superimposed with the new background to achieve the purpose of removing the overflow color. The growth of the original pixel space complexity is smooth and tends to be a straight line, indicating that the original pixel data extracted in this experiment is stable. In contrast, the pixel data of the color to be designed shows an exponential growth, indicating that the data obtained from the design contains larger information, and that the design results are more accurate with higher spatial complexity. The detailed and intuitive data structure is shown in the following Figure 8.

After determining the color rendering mode as HSV and the grayscale processing, the final idea of the digital artwork was confirmed. The circle is the perfect embodiment of geometry, and it has different artistic meanings throughout the world. Stars, planets, and satellites can be abstracted as circles; so, in many civilizations, the circle represents the sun or the moon and has a rich poetic character. Moreover, the circle has no beginning and no end and has the meaning

of eternal cycle, which can bring people endless imagination. Therefore, this design uses the constructed FNN algorithm model, incorporates the big data technology innovation into the digital art painting, and chooses the circle as the symbol to convey the imagery to create four digital art paintings.

Figure 9 shows four digital art paintings using big data fusion media technology and taking circles as the artistic creation concept. The above four digital art paintings are drawn according to the design concept by feeding the instructions to the algorithm model. The overall background is black, conveying the mood of mystery and the unknown and reminding people of the deep universe. In the first digital painting, the circle creates a viewpoint from the cave, and the bird-like creatures in the circle create the image of a scary giant. In the second painting, the silhouette is replaced by a cluster of buildings, and a mirage of urban architecture appears in the lower half of the circle, adding a sense of mystery to the unified style. The rows of straight lines in the upper part of the circle can be seen as clouds, which form

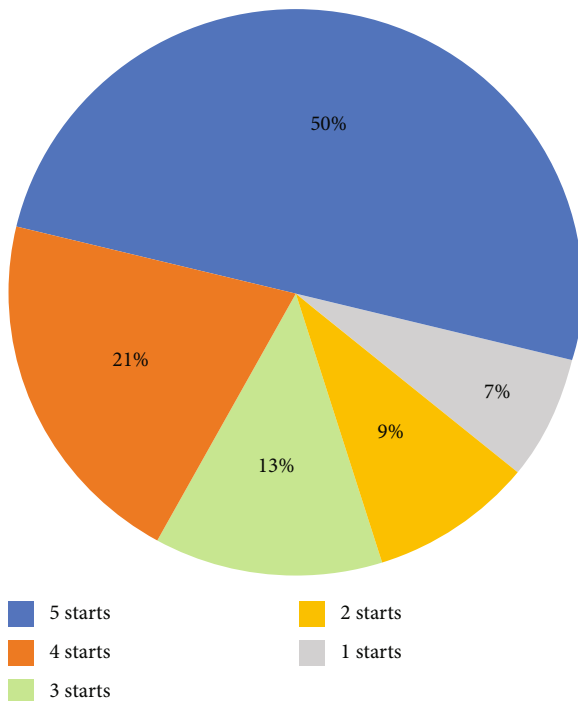


FIGURE 10: Graph of website users' ratings of four digital art paintings.

a great sense of oppression to the buildings on the screen and make the whole painting more artistic tension. The third painting abstracts the circle into a stack of books, and the overall wooden color style gives the painting a warm life. The last digital art painting is very different from the first three, in which the round body is made into a cave with an arched door, except for the gray ground near the cave, which is endlessly black. The archway is surrounded by blue and orange lights, and the lower part of the picture is reflected by water, giving the whole painting a visual, and the light and water expand the three-dimensional and spatial sense of the picture. After the creation, the study also uploaded the four digital art paintings to the NFT trading site OpenSea for sale.

After a period of selling on the website and average scoring by OpenSea, this study counted the popularity of these four digital art paintings, and the results are shown in Figure 10.

About half of the website users gave a positive rating of 5 out of 5 stars, and about 21% more gave a high rating of 4 stars, indicating that the paintings created using the data algorithm in this paper received a high average and good popularity. However, about 7% of users still expressed that they could not understand the artistic ideas conveyed by the paintings, which indicates that there is still room for improvement in this algorithmic model.

5. Conclusion

The proportion of different types of art in NFT is roughly the same as that in traditional art. Considering the operation cost and creation cost of the algorithm, this study decided to

use the model established in the previous paper to design and draw digital artwork. The most widely used visual color perception model HSV color space is selected. It characterizes color by brightness, hue, and saturation. Determine the pixel position of the artwork to be created and use image processing to analyze the target color pixels and redundant color pixels, including pixels overlapping with the effective pixel information and pixels containing useless information, both of which will interfere with the calculation of the model and slow down the calculation speed. The pixel data of the color to be designed increases exponentially, which indicates that the data obtained from the design contains more information, the design result is more accurate, and the spatial complexity is higher. Finally, the circle was selected as the basic criterion for creation, and the algorithm model constructed by big data integration technology was used for artistic creation. Four digital art paintings with unified overall style were obtained, but each one focused on performance. The results show that the spatial complexity of the original pixel increases smoothly and tends to be straight, which indicates that the original pixel data extracted in this experiment is stable. However, art is a huge system, which contains the endless pursuit of beauty. Therefore, there are still a lot of unknown topics that need to be studied and explored by relevant personnel.

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 is no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] M. K. Saggi and S. Jain, "A survey towards an integration of big data analytics to big insights for value-creation," *Information Processing & Management*, vol. 54, no. 5, pp. 758–790, 2018.
- [2] H. Wu and G. Li, "Visual communication design elements of Internet of Things based on cloud computing applied in graffiti art schema," *Soft Computing*, vol. 24, no. 11, pp. 8077–8086, 2020.
- [3] D. Pesce, P. Neirotti, and E. Paolucci, "When culture meets digital platforms: value creation and stakeholders' alignment in big data use," *Current Issues in Tourism*, vol. 22, no. 15, pp. 1883–1903, 2019.
- [4] N. Côte-Real, P. Ruivo, and T. Oliveira, "Leveraging internet of things and big data analytics initiatives in European and American firms: is data quality a way to extract business value?," *Information & Management*, vol. 57, no. 1, article 103141, 2020.
- [5] A. Gigante, L. Navarini, D. Margiotta, B. Barbano, A. Afeltra, and E. Rosato, "In systemic sclerosis, microvascular and hands digital arteries damage correlates with serum levels of endostatin," *Microcirculation*, vol. 25, no. 4, article e12449, 2018.

Retraction

Retracted: Analysis of Farmer Relocation Selection Behavior Based on Bayesian Network

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Li and Y. Guo, "Analysis of Farmer Relocation Selection Behavior Based on Bayesian Network," *Journal of Function Spaces*, vol. 2022, Article ID 9155510, 8 pages, 2022.

Research Article

Analysis of Farmer Relocation Selection Behavior Based on Bayesian Network

Song Li ¹ and Yanbo Guo ²

¹*School of Public Policy and Management, Guangxi University, Guangxi 530004, China*

²*School of Marxism, Hunan University of Finance and Economics, Hunan 410205, China*

Correspondence should be addressed to Yanbo Guo; guoyanbo@hufe.edu.cn

Received 6 July 2022; Revised 27 August 2022; Accepted 31 August 2022; Published 14 September 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Song Li and Yanbo 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 study of farmers' migration choice behavior can reflect people's needs for the future living environment from the side and has practical and important guiding significance for adjusting and improving the focus of future development. In order to make a detailed analysis of the prediction of migrant workers' behavior in cities, this paper studies the internal mechanism of the change of commuters' commuting patterns under the background of urban suburbanization. Using Bayesian network, this paper establishes a model of commuting mode transfer of urban immigrants, taking personal and family attributes as control variables. The influence of factors such as migration attributes and changes in building environment perception before and after migration is analyzed. Based on the survey data, a Bayesian network of farmers' land use behavior is established by using Bayesian network method to investigate the choice of farmers' migration direction. In the research process of migration direction, the research structure shows that 60.1% of farmers want to stay in the countryside, rather than migrate through Bayesian network algorithm. The results of regression analysis by establishing a multivariate logistic model show that there is a strong willingness to choose small towns and large cities. However, the willingness to choose not to relocate to local villages is relatively weak.

1. Introduction

In recent years, with the development of inland economy, a shortage of migrant workers has appeared in many provinces and cities in China. Scholars have made various analyses on the causes of the phenomenon of "shortage of migrant workers." Due to the need of economic construction and development, rural cultivated land has been requisitioned by the government, and many farmers no longer engage in agricultural production in the primary industry, but turn to urban development with more employment opportunities. In recent years, the construction of new countryside has been getting better and better. Many farmers who go out to work have returned to the villages where they lived since childhood. What are the factors that affect their relocation? Farmers who come to cities and towns to work for a living are playing drums, fearing that they have no technical skills and can only serve as cheap labor in cities. The analysis

of China's investment in agriculture, rural areas, and farmers has certain significance, but it fails to fully understand the root of the problem of "shortage of migrant workers" [1]. If the services for migrant workers fail to keep up, the phenomenon of "labor shortage" will be difficult to eliminate, and the information service for migrant workers is even more important.

Through the present situation, we will fully understand the influence of the general characteristics of peasant groups and other factors on their willingness to move to cities and towns in the future. Based on the sample survey, this paper analyzes the current situation of migrant workers' information behavior through the field investigation of migrant workers. The emergency risk of mutual fund assistance of farmers' professional cooperatives is a brand-new research perspective. An empirical analysis of the risk of mutual fund emergencies is not only conducive to promoting the development of rural cooperative finance but also has important

research value and guiding significance for preventing the occurrence of mutual fund emergencies [2]. In the process of urbanization in China, farmers' willingness to move directly affects the process and direction of urbanization in China and is related to the speed, scale, and effect of urbanization. The formulation of various policies and measures in the process of urbanization must take into account farmers' own wishes and choices, especially farmers' willingness to choose their future living space [3], so that we can have a clearer understanding of farmers' preference for future living space and the motivation of their choice.

On the basis of the existing research literature, from the perspective of farmers' willingness to choose central villages, small towns, and big cities, this paper takes farmers' own demographic characteristics, economic and social characteristics, and resource endowment as explanatory variables and empirically analyzes the influence of these characteristic variables on their willingness to move [4] by adopting multiple models. The main research contents of farmers' information behavior measurement include information demand, information search channels, difficulties encountered in the process of obtaining information, motivation of obtaining information, information absorption and utilization, and information evaluation. The chi-square test is carried out to eliminate irrelevant variables, and then, SPSS Clementine 12 data mining software is used to compare the prediction effects of TAN, Markov, and Markov-FS three different Bayesian network models and select the best Bayesian network model for farmers' future relocation. The farmers' land decision-making model is established, and the related influencing factors are analyzed in detail [5]. Based on Bayesian network, this paper constructs the transfer model of the main commuting modes of the relocated people and analyzes the related attributes of relocation by probabilistic reasoning, as well as the influence of the built environment changes caused by relocation on the change of commuting modes.

The innovative contribution of this paper is to study the internal mechanism of the change of commuters' commuting patterns under the background of urban suburbanization. Using Bayesian network, this paper establishes a commuting mode transfer model for urban immigrants and takes personal and family attributes as control variables to analyze the effects of migration attributes and changes in building environment perception before and after migration. In this paper, Bayesian network model combines probability theory and graph theory and can systematically describe the complex correlation between random variables. The structure of the interdependence between variables is directly revealed in the language of graph theory. The Bayesian network of farmers' land use behavior is established by using Bayesian network method to investigate the choice of farmers' migration direction. In the research process of migration direction, the regression analysis results of establishing a multiple logistic model show that the willingness to choose small towns and big cities is strong, while the willingness to choose not to migrate to local villages is relatively weak. In the analysis of farmers' income, SPSS chi-square test is used to eliminate irrelevant variables, which greatly reduces the amount of calculation of data mining, and Bayesian network model is used to compare the predic-

tion effect with the actual data, which has a certain reference value for the study of migration selection.

This paper studies the internal mechanism of the change of commuter commuting mode under the background of urban suburbanization. The research is divided into five parts. The first part expounds the general characteristics of farmers and the influence background of other factors on their willingness to move to cities and towns in the future. The second part analyzes the land use behavior of farmers by analyzing the survey results or applying regression models from the perspective of economics and sociology. The third part expounds the related technologies of this paper, including introduction and formula of Bayesian network, Bayesian network inference algorithm, and overview of Bayesian network data integration application. The fourth part analyzes the data of farmers' migration. It includes the data structure model based on Bayesian network expression and farmer migration. Finally, the full text is summarized. In the analysis of farmers' income, SPSS chi-square test is used to eliminate irrelevant variables, which greatly reduces the amount of calculation of data mining, and Bayesian network model is used to compare the prediction effect with the actual data, which has a certain reference value for the study of migration selection.

2. Related Work

Due to the difference of land use patterns between the inside and outside the city, the mismatch between suburban residence and employment development, and the imperfect public transport infrastructure in the new peripheral areas, residents commuting after moving out of the city is becoming more and more motorized, and the traffic congestion in suburban corridors is becoming increasingly prominent. From the angle of economics and sociology, this paper studies the land use behavior of farmers by analyzing the survey results or applying regression models. From the microscopic perspective, the Bayesian network of farmers' land use is studied. Combined with the personal and family situation and the economic and social environment of the original place of residence, the value orientation of various relocation destinations is formed, and finally, the destination with the highest relocation value is selected as the final choice will.

Zhu and Fan think that the relationship between built environment and travel behavior is the theoretical basis to explain the change of commuting behavior of the relocated people [6]. Doguc and Ramirez-Marquez think that relocation is a cyclical event of life, and its related attributes will have a certain impact on commuters' travel [7]. Another scholar said that travel behavior is subject to peoples' subjective feelings about the built environment, and the change of subjective environment cognition will affect individual choices. Li et al. think that the relationship between data can only be described by correlation, not causality. For example, the height and weight of a normal person are related, but it cannot be speculated whether birth height affects weight or weight affects height [8]. Prospero et al. think it is feasible to infer causality from data and put forward probability and causal inference algorithm for the first time, which is now Bayesian network [9]. Scanagatta et al. use Bayesian network to build regulatory

networks of multiple genes [10]. Chouaib et al. think that although the appropriate tangent point can be found in this way, it increases the complexity of the model, and only a few discretization intervals can be obtained, thus making the data lose more information [11]. Tong et al. think that developing small towns is the only way for China's urbanization. We should attach great importance to the irreplaceable functions of small towns, and the strategic position of small towns will remain unchanged. It is believed that the widespread "urban diseases" in big cities at present cannot be used as a reason to stop urbanization [12]. Leng et al. think that farmers' choice of relocation has gone through the cognitive stage of the destination, the judgment stage of relocation value, and the formation stage of relocation choice will [13]. In this process, farmers first make a preliminary assessment of the natural and socio-economic environment of several target destinations based on their own knowledge and experience and then get the benefits and costs of relocation of several target destinations. Rizvi et al. think that the willingness to move is the result of comprehensive evaluation and comparison between residents' residence and target residence [14]. Belur et al. believe that the external behavior of human beings is based on the geographical environment, and the decision-making behavior made by people is the corresponding result of the perception and evaluation of the geographical environment [15]. Khan et al. believes that farmers' decision-making behavior of moving is the result of their perception and evaluation of the external environment such as rural areas, small towns, and big cities [16]. Hoogesterger and Rivara think that farmers run, economic crimes, death/escape/kidnapping of the head of the mutual aid society, natural disasters that cause the cooperative to fail to operate normally, abnormal computer network paralysis, mass petitions and demonstrations, and other cases that may cause great harm are the risks and unexpected risk events of mutual funds [17]. Nayak et al. think that whether to develop big cities, medium cities, small cities, or small towns should be mainly regulated by the market, and the government can only guide them [18]. Delcroix believes that the development direction of urbanization in China is the development of medium-sized cities in urban circle, and at the same time, the further development of medium-sized cities will form a new urban circle [19].

3. Related Technologies of This Paper

3.1. Brief Introduction and Formula of Bayesian Network. Bayesian network, also known as belief network, is an extension of Bayes method. A typical Bayesian network consists of directed acyclic graph (DAG) and conditional probability table (CPTs). Bayesian network is a combination of probability theory and graph theory. Bayesian network has attracted the attention of different research fields since it was put forward, and several classical networks have been generated. When using Bayesian network to solve problems, the set of variables whose values have been determined is called evidence D . The set of variables that need to be solved is called hypothesis X . Target intention recognition based on Bayesian network is to solve the hypothetical variables under the condition of given evidence (battlefield events that have occurred). During the initialization of Bayesian network, the target intention and

the confidence of battlefield events are given in advance. When a new battlefield event is detected, the influence of the event (evidence) on the target intention (hypothesis) can be updated by Bayesian backward propagation. Until the confidence of a hypothesis (node state) in the goal intention exceeds the preset threshold, the goal intention is determined to be true. Bayesian network reasoning is an event probability speculation based on the established network model, and the probability of the target event after an event occurs is calculated by conditional probability. Network edge deletion is as follows: on the basis of the previous network structure, check the conditional mutual information of each point pair in the edge set, and delete the points below the threshold. The expressions of mutual information $I(X, Y)$ and conditional mutual information $I(X, Y|C)$ of the point pair are, respectively,

$$I(X, Y) = \sum_{x,y} P(x, y) \log \frac{P(x, y)}{P(x)P(y)}, \quad (1)$$

$$I(X, Y|C) = \sum_{x,y,c} P(X, Y, C) \log \frac{P(x, y|c)}{P(x|c)P(y|c)}.$$

The probability change of node variables in the network will transfer information through conditional probability dependence and change the probability distribution of target variables. The specific calculation formula is

$$P(XU, e) = \lambda_e(X) \Theta_{x|u} \prod_i \pi_x(U_i) \prod_j \lambda_{y_j}(X), \quad (2)$$

$$P(U_i, e) = \sum_{xu \setminus \{u_i\}} \lambda_e(Y) \Theta_{x|u} \prod_{k \neq i} (U_k) \prod_j \lambda_{y_j}(X), \quad (3)$$

where Y is the child node set of X and $\pi_X(U)$ represents the information passed to U by the parent node set X . $\lambda_Y(X)$ is the information transmitted from X to Y ; $\lambda_e(X)$ is an indicator for evidence; if X is entered in the e variable, the indicator is 1; otherwise, it is 0; $P(XU)$.

Union tree algorithm includes the Hugin algorithm and Shafer-Shenoy algorithm. When the network structure DAG and the node parameter set CPTs are determined, the network can be applied to practical cases to prepare for subsequent probabilistic reasoning. At present, a variety of structure learning algorithms have been developed, including the famous PC, K2, TAN, and EM. Some of these algorithms can learn the structure of the reference network under certain conditions. Some algorithms give up the original intention of constructing benchmark network and turn to the development of Bayesian classifier, such as TAN algorithm, whose Bayesian classifier is called TANC.

3.2. Bayesian Network Reasoning Algorithm. Network reasoning is based on a given Bayesian network. According to the conditional probability formula in Bayesian probability, the values of other nodes are deduced from the known values of any one or more nodes in the network. Data discretization refers to the process of breaking up continuous data into discrete data. This is not only the research hotspot of Bayesian network but also the research hotspot of the whole machine

learning field. Data discretization refers to segmenting continuous data into discrete intervals. The principle of segmentation is based on equal distance, equal frequency, or optimization methods. Data discretization is mainly performed on continuous data. After processing, the data value range distribution will be changed from continuous attributes to discrete attributes. This attribute usually contains two or more value ranges. It has been proved that data discretization is a NP problem. At present, a variety of data discretization algorithms are proposed, which are divided into two categories according to supervised discretization algorithm and unsupervised discretization algorithm [20]. Unsupervised discretization algorithms do not consider class attribute information, mainly including equidistance, equifrequency, and discretization algorithms based on kernel density estimation.

The process of reasoning posterior probability problem, maximum posterior hypothesis problem, and maximum possible explanation problem with network structure and prior probability belongs to NP problem. The network with simple structure adopts accurate reasoning algorithm, including joint tree (cluster tree) reasoning algorithm, symbolic reasoning, and elimination reasoning. Approximate reasoning algorithms are used for complex networks, including random sampling and circular message delivery.

Network reasoning includes four reasoning algorithms:

- (1) Variable elimination algorithm: its principle comes from the research of dynamic programming with indefinite order, which simplifies the reasoning calculation process by using the decomposition of joint distribution. This algorithm decomposes the joint probability into a series of parameterized conditional probability products through chain product rules and conditional independence and then transforms the formula. By changing the order of summation and product operations, it selects the elimination order of nodes during summation, thus reducing the computational complexity. The advantages of this algorithm lie in its universality and simplicity, and it can solve multiconnectivity, etc.
- (2) Junction tree algorithm: this algorithm can not only solve the reasoning in single connected network but also complete the reasoning calculation in multiconnected network, especially when there are multiple query nodes in the network. In the reasoning process, the message will spread to each node of the junction tree in turn and finally make the junction tree meet the global consistency. At this point, the potential function of a cluster node is the joint distribution function of all variables contained in that node [21]
- (3) Monte Carlo algorithm: Monte Carlo algorithm, also known as random sampling algorithm, is a kind of approximate calculation method widely used in numerical integration and statistical physics. Monte Carlo algorithm can be divided into two categories, namely, importance sampling algorithm and Markov chain Monte Carlo algorithm. Their main difference is that

the samples produced by the former are independent of each other, while those produced by the latter are interrelated

- (4) Approximate reasoning algorithm based on variational method: the basic idea of variational method is to transform the probabilistic reasoning problem into a variational optimization problem through variational transformation. Commonly used algorithms are naive mean field algorithm and circular propagation algorithm

The concept of the Shafer-Shenoy algorithm is simple and easy to understand, but because the Hugin algorithm avoids redundant computation, we mainly use the Hugin algorithm to build joint tree in reasoning process. Union tree algorithm is one of the most commonly used accurate reasoning algorithms, and it is also the basis of all accurate reasoning algorithms engines in BNT, as shown in Figure 1.

- (1) Establish a Moral graph, connect the parent nodes of the same node pairwise, and change the directed edge to the undirected edge
- (2) Triangulate the Moral graph and connect the nonadjacent nodes in the Moral graph with the least number of edges greater than 3
- (3) Identify all regiments
- (4) Establish joint tree: ellipse represents nodes, and rectangle is a set of divided nodes

According to the sequence of Bayesian network establishment, data preprocessing, structure learning algorithm, grading function, classification test method, and network reasoning are five parts, which provide suggestions for the subsequent use of Bayesian network to build a model.

The theoretical basis is provided as shown in Figures 2 and 3.

The BNT of MATLAB has given 13 mature structure learning algorithms. Except that the PC algorithm may produce a loop in the structure learning and cannot build a network, other algorithms can reasonably and effectively build a network structure. At the same time, the improved algorithms TAN_EM and MWST_EM are also used to build the network. Joint reasoning is the most commonly used reasoning algorithm in the learning process, also known as cluster tree propagation algorithm. It is the basis of accurate reasoning, with high computational efficiency and more accurate results. According to the sequence of building Bayesian network structure, this paper mainly introduces the data improvement, structure learning algorithm, scoring function, classification test, and reasoning of the network.

3.3. Summary of Data Integration Application of Bayesian Network. In terms of application, Bayesian network has been used in many fields. For example, the Bayesian network used in the revision function of Google search development; Microsoft developed MSBN software based on Bayesian network. And the medical diagnosis system was developed by

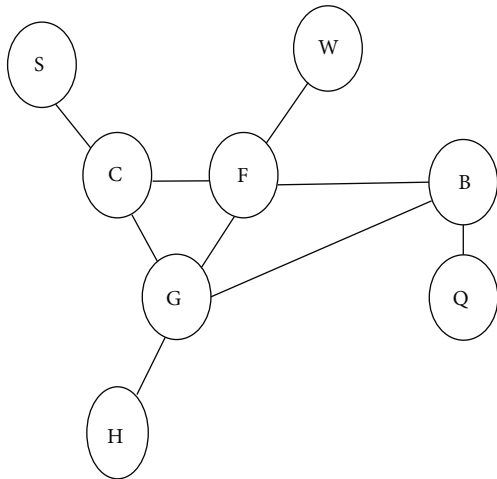


FIGURE 1: Triangulated Moral diagram.

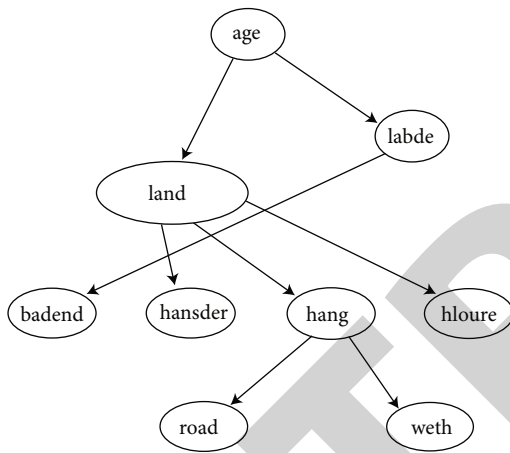


FIGURE 2: HC algorithm.

foreign countries based on Bayesian network. Besides medicine, Bayesian network is widely used in system diagnosis, pattern recognition, military simulation, and other fields.

However, there is still little research on the application of Bayesian network in farmers' migration choice and utilization decision. Although expert knowledge and experience can be used to establish the network structure, its modeling cycle is long, and it will consume a lot of financial resources and manpower. Therefore, it is not mature to apply Bayesian network model to farmers' land use decision. In addition to relying on the knowledge and experience of experts, we can also use the structure learning algorithm to build a network structure of experimental data and evaluate it. At present, R, Weka, SPSS, MATLAB, and other software have developed Bayesian network toolboxes, and the toolboxes developed will be different with different software. Modeling like Weka has low flexibility. Therefore, before establishing Bayesian network, we should choose the appropriate modeling software.

Bayesian network develops rapidly in LUCC field, but the total number of literatures is still small. According to statistics, during 1990-2010, but during 2000-2012, the number

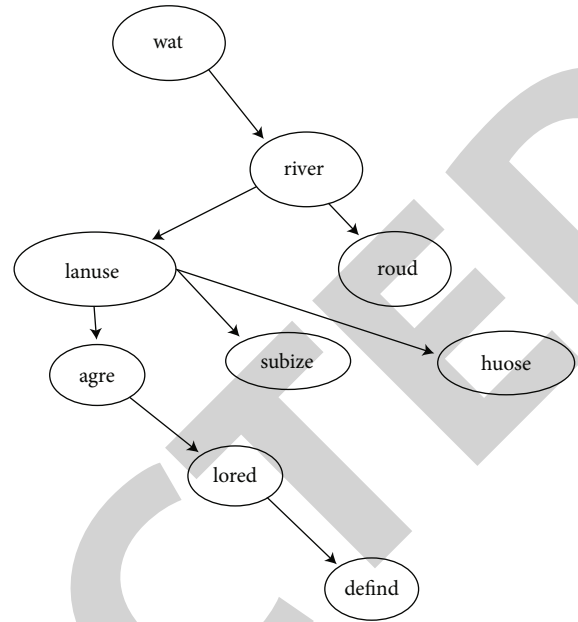


FIGURE 3: MW algorithm.

of foreign literature related to Bayesian network in ecological compensation and land use increased dramatically.

However, the application of Bayesian network in LUCC also has a lot of problems mentioned above, especially in data preprocessing and structure learning. However, there are also a few exemplary papers. For example, the Bayesian network incorporating the characteristics of land managers was established in Aalders, and even random distribution can be used to better predict the land use. It is concluded that the elderly farmers are likely to face no successors and the land use will be nonagricultural.

4. Analysis of Farmers' Migration Data

4.1. Expression and Analysis Based on Bayesian Network. With the increase of the number of variables involved in the network, the corresponding reasoning process becomes more and more lengthy and complicated. Therefore, the Bayesian network, which uses the belief propagation algorithm to improve the reasoning efficiency, is the product of the combination of graph theory and probability theory. The Bayesian formula (4) composed of structure and parameters is the foundation of the whole network.

$$P(A|B) = \frac{P(A|B) \cdot P(A)}{P(B)}. \tag{4}$$

Test the following determinants of logistics willingness to move: (1) small towns versus central villages, (2) large cities versus central villages, and (3) large cities versus small towns; so the following two logistic model formulas are established by taking choices 1 and 2 as reference classes:

$$\ln\left(\frac{p_2}{p_1}\right) = \alpha_2 + \sum_{k=1}^k \beta_{2k} x_k, \quad (5)$$

$$\ln\left(\frac{p_3}{p_1}\right) = \alpha_3 + \sum_{k=1}^k \beta_{3k} x_k. \quad (6)$$

The realization of scoring function based on information theory is based on coding theory and MDL principle. According to the principle of MDL, the structure learning of Bayesian network is to find the graph model with the shortest description length, which means that MDL scoring rules often find a simpler structure to balance the accuracy and complexity of the network. Generally, the penalty function of network complexity is usually expressed by the number of parameters.

4.2. Data Structure Model of Peasant Migration. Farmers' migration is an inevitable requirement for economic and social development to enter a new stage. Institutional changes such as identity change should be included in the scope of urban security to ensure the long-term livelihood of landless farmers. On the basis of previous studies, this paper attempts to use the data structure model of land lost farmers' migration decision and farmers' migration. This paper discusses the characteristics and mechanism of land lost farmers' migration, in order to provide reference for the migration model of land lost farmers. According to the migration direction p1, p2, and p3 of farmers, the probability of choosing rural areas, small towns, and large cities is expressed, respectively, by constructing an income chart and comparing the prediction accuracy of nodes. Compare the prediction accuracy of the models, as shown in Figure 4.

The frequency of data analysis with SPSS 19.0 software is shown in Table 1.

Among them, 60.1% of migrant workers chose to return to their hometowns, and 38.8% of migrant workers chose to stay in urban development.

Measuring the degree of influence of various factors on the transfer of two types of commuting modes, the inference analysis curve of the established network using the confidence propagation algorithm is composed of the sensitivity and specificity of the multiple critical values of a series of variables, the vertical axis TPR indicates the probability that the real value is positive and the prediction is positive, and the horizontal axis FFR indicates the probability that the actual value is negative and is predicted to be positive. When the ROC curve is closer to the coordinates in the upper left. The three categories of the class node "land use" are discussed separately and need to be evaluated by calculating the P, R, and I score F metrics. According to the confusion matrix obtained by the classification test, these indicators are calculated separately, as shown in Figure 5.

Network reasoning is often used to verify the practicability of the model. The network parameters are updated by parameter learning, and the results are random. After many times of learning, the results tend to mean distribution. Therefore, based on the survey data and network structure, the prior probability distribution of each node is set. In the network constructed by MWST+T+K2 algorithm, the classi-

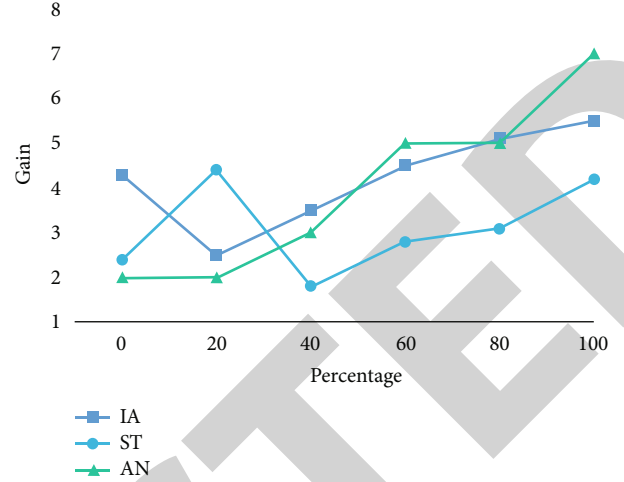


FIGURE 4: Income chart of three Bayesian network models.

fication level of MWST and MWST+T+K2 classifiers is the same under CV-5 test, and the classification accuracy is higher than that of NBC and TANC at the same time, and the best results are obtained.

In the research process, combined with the idea of software reuse in SoftMaker 8 project, we make full use of the historical risk data in the software project development process to build a risk historical database, realize the organic combination of historical knowledge and cases with current project risks, facilitate the improvement of Bayesian network reasoning data, and effectively improve the effect of risk management. MWST+T+K2 and GES networks have certain advantages in the four scores. The four scores of GES network are high, but the MWST+T+K2 network has more obvious advantages because of its long running time, high algorithm complexity, and inconsistent network structure with the research needs. At the same time, GES and MWST+T+K2 algorithms are used to process large data sets. Comparing their time consumption, it is found that GES takes a long time, which is not conducive to practical research. In addition, observing the MWST+T+K2 network, it is found that the relationship between the structural nodes is consistent with the objective facts, and it can be inferred that MWST+T+K2 network can be used as Bayesian classifier from the opposite causality in the network.

The similar node "land use" of MWST+T+K2 network has three states, namely, farmers' choose orchard planting, shed planting, and land idle, and its subnodes are labor and subsidy. Among them, the node "labor" has three values, indicating that the number of labor force of farmers is "less," "average," and "more." The four numerical values of "subsidy" are represented as "no subsidy," "small subsidy," "medium subsidy," and "large subsidy" in turn.

Figure 6 shows the land use reasoning data analysis chart. The research variable is the commuting mode after moving. When analyzing the sensitivity of the attributes of moving and the changes of built environment, the research takes the individual family attributes as the control variable and infers the influence of the changes of these two attributes on the changes of different commuting modes. For

TABLE 1: Analysis of whether farmers will stay in urban development in the future.

	Frequency	Percent	Valid percent	Cumulative percent
Not	1756.9	60.1	60.1	60.1
Valid Yes	1146.1	38.8	38.8	99.8
Total	2929.8	99.8	99.8	

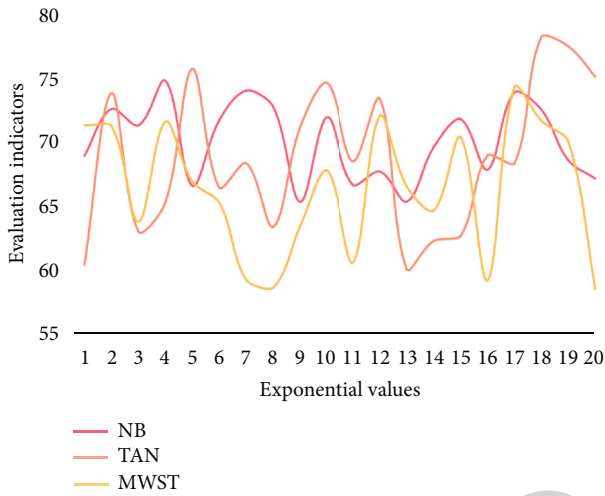


FIGURE 5: MW+T+K2 network evaluation index.

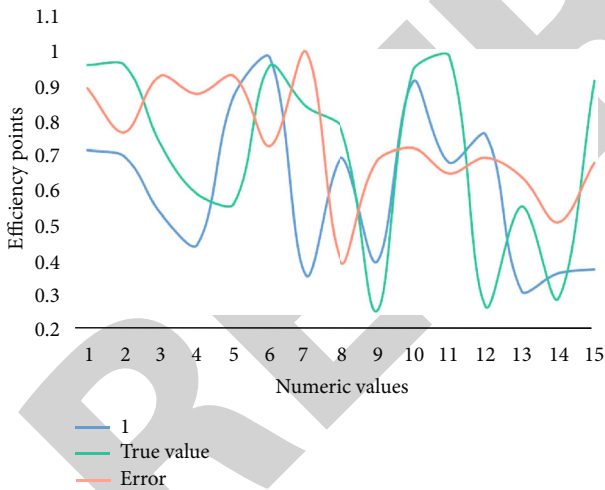


FIGURE 6: Land use reasoning data analysis diagram.

the transfer of non motorized migrants to public transport, in addition to family attributes such as family cars, car purchases and personal income, the change of commuter distance after relocation, of which transportation convenience is the main factor. Commuting distance is the main factor for nonmotorized commuting to bus commuting after relocation; that is, when commuting distance exceeds the service range of slow traffic, the possibility of commuters using bus after relocation will increase significantly Figure 7.

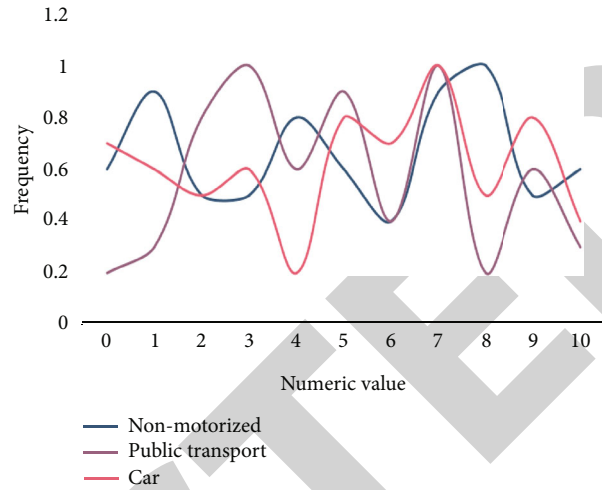


FIGURE 7: Farmers' migration travel mode.

Based on the research of MWST+T+K2 network, HC-K2 algorithm is put forward based on experience and experimental results, and a new Bayesian network-HC-K2 network for farmers' land use decision is established. At the same time, by comparing TANC, HC-K2, and MWST+T+K2, the classification accuracy of HC-K2 is higher than that of MWST+T+K2, which improves the classification problems of MWST+T+K2. According to the above data and model analysis, the choice of farmers' migration direction is reflected from the side.

5. Conclusions

This paper studies the internal mechanism of the change of commuters commuting style in the background of urban suburbanization, builds a model of urban migrants commuting style transfer by using Bayesian network, analyzes the influence of factors such as migration attributes and perceived changes of built environment before and after migration with personal family attributes as control variables, and establishes a Bayesian network of farmers' land use behavior by using Bayesian network method and based on survey data, so as to investigate and study farmers' choice of migration direction. In the research process of migration direction, the regression analysis results by establishing multiple logistic model show that the willingness to choose small towns and big cities is stronger, while the willingness to choose not to migrate to live in local villages is relatively weaker. The larger the construction area of rural housing is, the higher the willingness to move to smaller towns and not to move. Farmers who participate in the new rural cooperative medical insurance and whose houses have obtained the real estate license are more willing to move to larger cities. In the analysis of farmers' income, SPSS chi-square test is used to eliminate irrelevant variables, which greatly reduces the calculation amount of data mining, and Bayesian network model is used to compare the prediction effect with the actual data, which has certain references for the study of migration selection.

Retraction

Retracted: Design and Communication of City Brand Image Based on Big Data and Personalized Recommendation System

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Lin and G. Hu, "Design and Communication of City Brand Image Based on Big Data and Personalized Recommendation System," *Journal of Function Spaces*, vol. 2022, Article ID 9378800, 10 pages, 2022.

Research Article

Design and Communication of City Brand Image Based on Big Data and Personalized Recommendation System

Yan Lin  and Guofeng Hu 

School of Art and Design, Guangdong University of Science and Technology, Dongguan, Guangdong 523083, China

Correspondence should be addressed to Yan Lin; linyan1@gdust.edu.cn

Received 2 July 2022; Revised 23 August 2022; Accepted 31 August 2022; Published 14 September 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Yan Lin and Guofeng Hu. 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 define the positioning of the urban brand image, design the urban brand image, integrate and optimize the communication channels, improve the public participation awareness, and enhance the core competitiveness of the city. In this paper, a personalized recommendation search engine based on big data identifies keywords input by urban users. And give more accurate results based on some relevant information that can be extracted. This paper analyzes how to make better use of big data for tourism destination brand image management, and the existing shortcomings, and puts forward relevant suggestions. The industries related to cultural creative design and tourism elements constitute an intertwined cultural and tourism industry chain, and data technology plays an important role in the cultural and tourism industry chain. Through the development of tourism, tourists will produce comprehensive and diversified consumption in the city. Based on the analysis of big data, it can provide strong decision support for the government and industry managers, and realize the image design and communication of the urban brand identification system. Through the big data platform, establish the brand management strategy, improve the communication content of the city's brand image, and timely feedback the opinions and suggestions of tourists on the tourism destination, so as to adjust the communication strategy of the tourism image according to the feedback information of tourists. The results show that the big data personalized recommendation system can achieve ideal results in urban brand value and urban tourism related factors.

1. Introduction

With the accelerated development of urban civilization, a kind of cultural flavor permeates everywhere, and more cities are aware of the nationality and the positive spread price of urban traditional culture. More importantly, cities need to build new and unique urban cultural images on the basis of inheriting traditional culture. As a business model combining information technology with traditional business, big data is now in the ascendant, and is full of vitality. This allows data analysis to go beyond the traditional statistical sampling, and can analyze according to the correlation of different data sets, and make predictions for the whole or different individuals, thus influencing people's decision-making. In terms of machine intelligence, it also needs a lot of real data for training because it imitates human thinking [1]. Industries related to tourism elements form an inter-

twined tourism industry chain. Through the development of tourism, tourists will have comprehensive and diverse consumption in cities. Then the whole city and tourism-related industries form the agglomeration of consumption economy, thus promoting the development of urban economic benefits, employment opportunities, ecological environment, and urban planning. However, with the rapid development of the Internet, information has also increased geometrically. Traditional recommendation systems are prone to storage and calculation bottlenecks under massive data. By collecting, cleaning, and analyzing these data, we can find the characteristics of customer churn, tap the hidden value behind it, and give early warning to customers who may lose in the future, so as to reduce the customer churn rate. In the era of big data, in addition to the consumption information on the current website, consumers may also have a lot of shopping information on other

websites. However, with changes in consumer demand and increasingly fierce market competition, the appeal of traditionally positioned city images has declined. In order to improve economic benefits and tap the potential of economic growth, major cities are gradually looking for effective ways to develop urban brands, taking advantage of long-term cultural development and historical accumulation [2]. Each city has formed its own unique urban temperament and unique urban symbols. The energy of the city's social functioning is transformed into a symbolic form of stored value through the city's public utilities.

With the increase of the number of users, projects, and information, the recommendation system is also facing problems such as massive data processing. Especially when new consumers do not have any historical information on the current website, their preference information on other websites is particularly important. Personalized recommendation system can provide learners with personalized learning resources according to their learning objectives, learning foundation, hobbies, and status. The economic development of any city will experience the transfer of agricultural labor force to industrial labor force in the industrialization period. In the field of big data, people often say that big data is better than good algorithms. This shows the importance of user data, especially in the era of big data. Further promote the transformation of the focus of urban construction from the traditional urban development mode focusing on the pursuit of life, production, residence, and simple economic needs to the artistic, cultural, and ecological urban development mode focusing on people with cultural connotation and maintaining the balance of the ecological environment. Regional core competitiveness roughly includes industrial competitiveness, enterprise competitiveness, opening up, economic strength, science and technology, human resources, local governments, financial activities, natural environment and resources, infrastructure, living environment, etc. These indicator systems are very comprehensive and comparable [3]. In recent years, China's information technology has developed rapidly, "Internet +" has been widely used in various fields and has made great achievements in the commercial field. At this stage, whether it is personal knowledge or collective thought, whether it is social public opinion or legal norms, it is far behind the pace of algorithm development in the era of big data. However, compared with other groups, current urban brands are less satisfied with their active choice and travel experience, and the contradiction is more prominent. The tourist destinations they often choose are almost urban outbound tourism or first tier city tourism. From the perspective of information dissemination, the dissemination of urban brand image is actually the process of brand marketers disseminating urban brand information to the audience through the media [4]. Almost all traditional behaviors can be mapped to the network: online shopping, online social networks, online music, online reading and news, and various searches [5]. But at the same time, it needs to be further improved to make the big data personalized recommendation system more mature and perfect, which is worth thinking and studying by every big data practitioner.

The era of big data has come, and it has actually changed people's lives. The key to the era of big data is not only the exponential increase of data volume, but more importantly, various algorithms can greatly improve the ability of data collection and analysis. Tourism is a strategic pillar industry of the national economy, accounting for a significant proportion in the gross product of modern service industry. It is an important part of the service industry, and it can also promote the development of related service industries. Exploring the influence of the recommendation system on consumers' shopping behavior through field experiments can effectively avoid the endogenous problems and the lack of external effectiveness of traditional research methods [6]. With the continuous growth of the national economy, people's material life is extremely abundant, and tourism demand also shows a rapid growth trend. China's tourism is in a golden period of development, and tourism has gradually become the main force of tourism consumption, and it has become a market crowd actively sought by various tourist destinations. From the marketing point of view, running a city brand is like running an enterprise brand. Moreover, the assets of the city itself include the brand image of the city, which can create considerable benefits for the city [7]. Therefore, cloud computing and big data technology are used to integrate the computing resources and storage resources of multiple servers. At present, it is an effective solution to distribute the heavy computing and storage tasks to server clusters through the network, and to combine the scattered computing results. Faced with this problem, coding and disseminating the symbol system of the city image can make the symbol system of the city image dissemination orderly and more easily recognized and accepted by people. In addition, our city has entered the stage of accelerating the development of a modernized well-off society in an all-round way, and the economic structure and spatial structure of large, medium, and small cities have undergone tremendous changes [8]. Due to the application of artificial intelligence, cloud computing, big data, and other high-tech, the big data personalized recommendation system has become more and more mature and has achieved relatively ideal results. On the other hand, the rapid development of algorithms has in fact further highlighted the relative lag of algorithm ethics research. The research of personalized recommendation system has become a field of increasing attention from many scholars and researchers, and recommendation system has become another popular research direction supporting the development of cloud computing. Therefore, planning and disseminating the image of urban tourism, shaping, and enhancing the image of urban tourism can promote the development of tourism, so that the development of tourism can promote the development of the city and the adjustment of the industrial structure. Brand is the most valuable urban wealth. The most charming thing about brand value is that it can cause fundamental changes in people's consciousness, ideas, and thinking mode. Brand value also lies in its great potential to create wealth and continue to create new value. Now, Chinese cities at all levels have deeply felt that "brand is the commanding height of urban competitiveness" and have taken measures. While many

cities have achieved success in branding, some cities have also gone into misunderstandings. What are the misunderstandings in the process of urban branding? How to avoid going astray has become an important topic in the study of urban branding in China. This paper expounds this hot issue and puts forward corresponding countermeasures.

2. Related Work

With the increasingly fierce competition among cities in China, many cities have carried out the practice of city image dissemination, and it is no exception. In the study of city image communication, some scholars discussed the communication of city image from the angle of brand and marketing. Cui et al. put forward that the perception of tourism image should be divided into “gray area” and “halo area”, and discussed the tourism product integration mode of tourism image planning in “gray area” [9]. Niu et al. established: according to the nature of perceived objects, it is divided into two parts: hardware image and software image; it can be divided into five parts according to different sensory organs: visual image, auditory image, olfactory image, taste image, and tactile image [10]. Tai believes that the influencing factors can be divided into attraction factors, cultural factors, and related subsidiary factors. If these evaluation factors are consistent with tourists’ perceived image, tourists will be satisfied, and vice versa [11]. Ren put forward: CIS system has been used to study the design of city identity [12]. Zhang et al. put forward the concept of “local marketing”, he believes that: treat the city as an enterprise, regard resources as a product, shape the image of the city, and finally build a strong city brand [13]. Alalhesabi et al. discussed: Cities can be branded, and the main means of brand building is communication. Cities can promote themselves through advertising and other means of communication, increase awareness, and shape brand image [14]. Zigmund analyzed that the purpose of people’s travel is to relax and release the pressure of life through tourism activities [15]. Fan et al. established a model of the tourist experience content [16]. Shan et al. established: from the perspective of the formation process of tourism perception image, the tourism perception image is divided into 8 types: obvious induction, hidden induction, and autonomous native [17]. Tu believes that factors such as the perceived distance of tourists and the cultural events of the destination affect the perception of tourists’ tourism image [18]. Choi demonstrated the relationship between city image and city brand. He believed that “connecting image and association with city” is the final result of city branding [19]. However, it is not difficult to find that in the process of urban image dissemination, there are still few researches on the problems of the signifier and the signified of the urban image, as well as the communication strategy and the media strategy.

3. Strategies for Improving Personalized Recommendation System for Large Data

3.1. Optimize Bid Guidance Tool. In order to better analyze the development of a certain urbanization, this paper collects

indicators in representative sampling cities. The sampling city should be representative in the size of the city, economic development, geographical location, or other aspects related to the human settlements system of the country. Urban brand image design needs to be related to urban history, culture, style, natural environment, and other elements. Only when it fits well with the city’s status, characteristics, and economy can the city’s brand image be deeply rooted in the hearts of the people. For big data, controlling bidding is an important way to control traffic, which requires big data platform to analyze the reference basis of daily price change that is, ranking. In the 20th century, with the development of science and technology, especially computer science, people pay more and more attention to algorithms that are closely related to people. In addition, looking for the nearest neighbor users, the recommended results can only be the items selected by the nearest neighbor users, and some items that few people pay attention to cannot be found, so that the diversity of recommended results cannot be satisfied [20]. With consumer demand as the core, reorganize corporate behavior and market behavior, use various communication methods in a comprehensive and coordinated manner, send unified product information and service information with a unified goal and a unified image, and achieve two-way communication with consumers. The investigation on the tourism motivation of young and middle-aged and elderly people is shown in Table 1.

Page rankings can actually be calculated as degrees of a graph, but how they are calculated has been a problem for engineers until page raises the question of multiplying two-dimensional matrices and solves it by iterating. More than half of young people use social networks for at least three hours a day, based on their average daily use. As shown in Table 2.

After data processing, it is the selection of the algorithm. In the process of selecting the algorithm, consider the applicability of the algorithm and select the appropriate algorithm to build the model for the problem itself. When we decide which algorithm to use, the type and shape of data we have play a key role. Some algorithms can work with a small set of samples, while others require a large number of samples. Specific algorithms work on specific types of data. Average recommended consumption time comparison. As shown in Figure 1.

An important property of self is self-reference effect, which mainly means that the result of people’s coding information depends on how much self is implied in the information. Returns the average recommendation number comparison. As shown in Figure 2.

Therefore, the promotion of urban image can be defined as the process of carrying out both spiritual and material dissemination of the city. The method is to first excavate the core value of the city and summarize it through visual identification design. The reason for the division of city brand types is to enable city operators to precisely define the city brand image according to the city’s resource advantages and environmental characteristics. After the task, scheduling node receives the task; it will divide the task and assign the subtasks to the nodes with rich computing resources in the

TABLE 1: A survey on the tourism motivation of middle-aged, young and old people.

(a)					
Young and middle-aged					
Experience and growth insights	Study	Conformity and show off	Relax and relieve stress	Star chasing and curiosity	Share feelings
25%	6%	12%	14%	28%	15%

(b)					
Elderly					
Experience and growth insights	Study	Conformity and show off	Relax and relieve stress	Star chasing and curiosity	Share feelings
41%	6%	4%	22%	12%	15%

TABLE 2: Statistics on media exposure time of young people.

Time	More than 8 hours	6-8 hours	3-5 hours	1-2 hours
Percentage	11%	13%	45%	31%

cluster. The dissemination of city image has important political significance. Political image dissemination is the planning and dissemination of images in terms of political status, political development level, and political construction. However, the vast majority of city brand research is unilaterally limited to the stage of city mutual imitation, unable to fully penetrate into the strategic implementation. It also failed to systematically analyze the shaping of city brand image recognition system as a whole. In addition, big data platforms must use real time recommended prices as the core reference for price adjustment to make as many consumers as possible and help new customers adapt to the new launch environment as soon as possible. The formula is as follows:

$$\forall c \in C, s_c = \arg \max_{s \in S} u(c, s). \quad (1)$$

Euclidean distance is defined as follows:

$$d([x_1, x_2, \dots, x_n][y_1, y_2, \dots, y_n]) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \quad (2)$$

It is defined as follows:

$$d(\vec{x} - \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T \sum^{-1} (\vec{x} - \vec{y})}. \quad (3)$$

$$J_\delta = 1 - J(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|}$$

The weight of each item's feature in each cluster is calculated using the following formula:

$$W(f_i) = \frac{N(f_i)}{\sum_j f_j},$$

$$N(f_i) = \frac{\sum_{j \in k} f_j}{\sum u_m}, \quad (4)$$

$$\bar{r}(f_i) = \frac{\sum_i r_i}{\sum_{j \in k} f_j}.$$

The formula for the sum of squared errors is as follows:

$$\sum (x_i - \bar{x})^2. \quad (5)$$

3.2. Accuracy Optimization of Personalized Recommendation System. In practice, the big data platform can improve the accuracy of recommendation by using crowd optimization assistants, and further understand customers' shopping demands. Of course, due to the limitation of computing technology, when faced with a huge amount of data, the algorithms at that time generally estimated the big data by reducing the complexity of the data and using a method similar to statistical sampling. In order to reveal the causal relationship between data sets and get the commonness between data sets. In the concrete implementation of information retrieval system, it is necessary to quickly find the keywords contained in the document. For the number of documents, the number of terms is relatively small. At the same time, timely feedback tourists' opinions and suggestions on tourist destinations, so as to adjust the communication strategy of tourism image according to the feedback information of tourists. After discovering the user's interest contained in the user's search terms, that is, long-term interest, further study how to make personalized recommendations according to the content of the user's search keywords. However, the decision tree also has corresponding shortcomings, such as its complicated classification rules, because the decision tree uses greedy algorithm to construct branches, and only one attribute can be selected to split each time, and then it goes down in turn to continue splitting. Because self-reference is easy to retrieve, consumers can recall self-reference more quickly. The traditional communication mode basically considers how to transmit the information about the image of the tourist destination to the target audience and potential audience from the perspective of the tourist destination. So as to create a good brand image in the audience's mind. As the main body of the city, if the city residents can actively participate in it, it will be beneficial to the shaping and spreading of the city brand and make the city brand image more vivid. Collaborative filtering recommendation technology mainly adopts strong item recommendation within the group and association difference recommendation, while content filtering recommendation technology will adopt common domain knowledge analogy recommendation. In this system, the strong item recommendation within the group and the associated difference

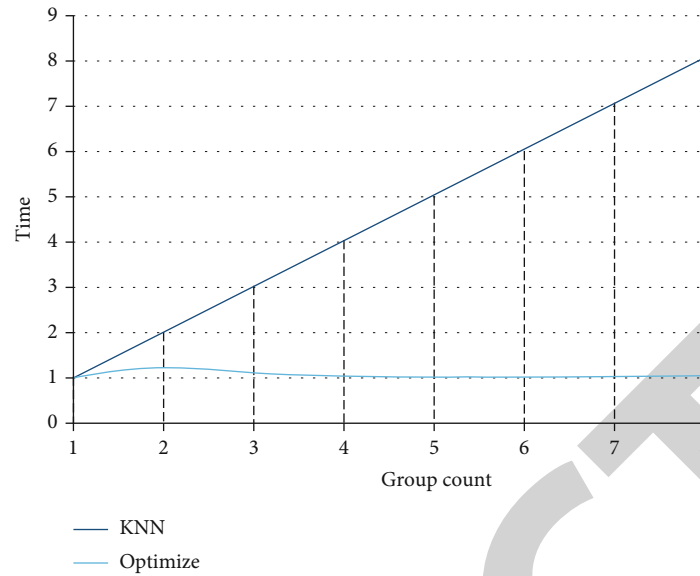


FIGURE 1: Average recommended consumption time comparison.

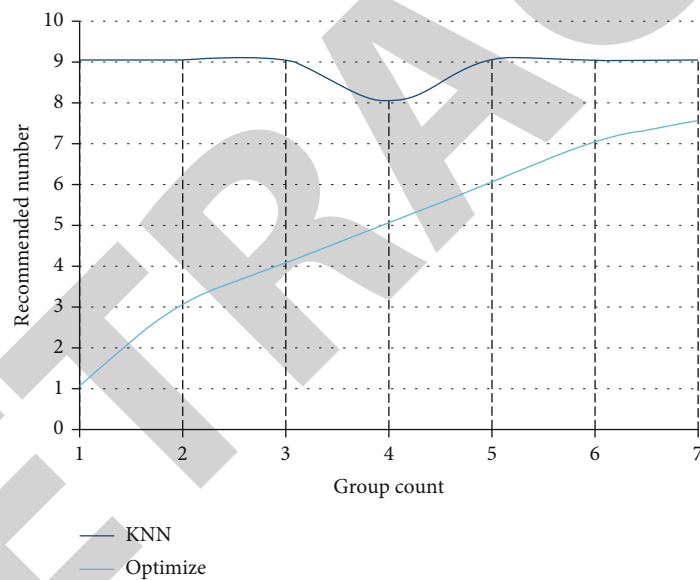


FIGURE 2: Comparison of the average number of recommendations returned.

recommendation algorithm are mainly used. The development of economic strategy not only requires rapid economic development, but also needs to promote rapid economic development through the dissemination of economic images to make people fully aware of the economic strength, economic foundation, and level of economic development. The dissemination of economic image is of great significance. Then it analyzes and sorts out the city image theory, and analyzes that the city image design only pays attention to material construction, ignoring the role of the brand on the city image. Crowd optimization tools can tag users as they collect user information, thereby forming different user tags, according to these different tags. The personalized recommendation system makes it easier to delineate different

shopping groups, thereby improving the accuracy of recommendation and improving the shopping experience of consumers.

4. City Brand Image Dissemination

4.1. *The Elements and Classification of Chinese City Brands.* The implementation of brand strategic planning begins with research. Because before determining a city brand, we must understand how the outside world views and evaluates the city, rather than being wishful thinking. The research starts with the resource advantages, future development, citizens' intentions, and the government's urban development planning of the region. The second is to use professional

institutions to conduct various forms of interviews and questionnaires to local and overseas audiences to understand the public, surrounding cities, and the international community's evaluation of cities. On this basis, an objective and scientific conclusion is drawn to provide decision-making basis for the next stage of work. Many elements have an impact on the city brand, but these elements have also become the key basis for dividing the city brand image. The computing node reads the data required by the mining task from the data storage node. And calculate according to the modified data mining algorithm, and finally send the mining results to the task scheduling node for merging, and the task scheduling node will store the final data mining results for use by the recommendation system. The economic image is based on hard power such as the level of economic development, economic foundation, and economic strength. It is the charm of the overall economic strength and drives rapid economic development. The city brand image behavior recognition system refers to the behavior of the city, and it is the activity mode of each city. The visual recognition system of city brand image is the performance outside the city, which will directly and tangibly show the spiritual appearance of the city. When big data starts to be used in big data, shopping bars like big rotation, guessing what you like, seeing and looking, and buying appear, which is an early form of big data recommendation system. Thus, most tourists agree that the impression before and after the tour is the same, indicating that the products and services provided basically meet the needs of tourists. As shown in Figure 3.

In the big data algorithm, besides dealing with more complicated data sets, another notable feature is that it no longer pursues the causal relationship and commonality between things. Instead, we should accurately mine the individual characteristics of different individuals, and deal with each analysis object in a personalized way to the greatest extent. On the cold start of the recommendation system, the optimization scheme is designed and implemented. The core is to use offline computing to solve the problem of instant computing and long tail effect in cold start recommendation. This requires us to formulate a media combination strategy through internal and external factor evaluation and public relations, so as to achieve a unified and most effective communication influence. And establish a long-term relationship with the target tourists, two-way interactive communication can be carried out, so as to effectively achieve the goal of communication and marketing. Therefore, user behavior data can be extracted from the log system of the search engine and the basic user information database of the search engine. The comparison of the curve results is shown in Figure 4.

However, the historical data of consumers in external companies can better represent consumers' shopping preferences if the current website makes use of such consumers outside the public. Therefore, the key to impress your audience is to find critical moments in which visitors will accept and resonate with them at some point throughout the travel process. The city's historical culture, natural landscape, historical celebrities, geographical location, humanistic customs, and architectural landscape are the

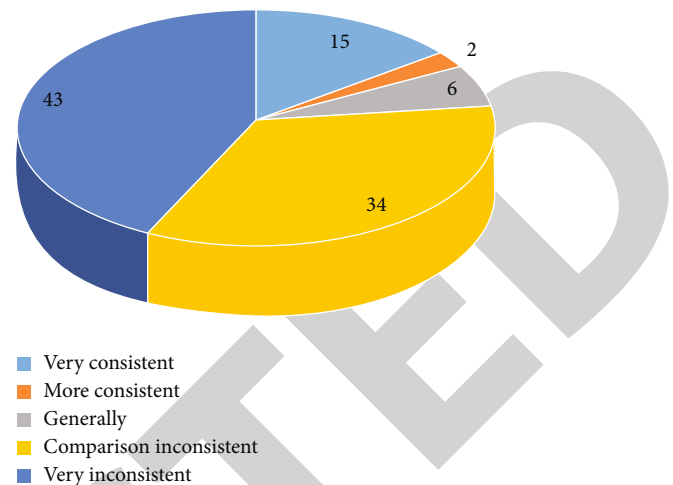


FIGURE 3: Analysis of impression consistency before and after tourism.

main components of city brand culture, such as mountain city, which takes its unique city culture as an important content of shaping city brand.

4.2. Big Data Communication Model of Chinese City Brand Image. Since the 21st century, the awareness of Chinese urban brand image communication has gradually increased, and the communication methods have become more and more diverse. After the learner logs into the system through the browser, the recommendation system makes recommendations based on the learner's data mining results. To build a local cultural brand, we should strengthen investigation and research, and explore the historical origin, development context, and basic trend of local culture in a multidimensional and in-depth manner. Strive to promote the organic integration of national style, traditional charm, characteristics of the times, and people's needs, so that the rich connotation of Fuzhou culture can adapt to contemporary culture and coordinate with modern society. The charm of the city image lies in the individualized cultural characteristics. Common things can only leave a cookie-cutter impression, while personalized things leave an impression that will last a lifetime. As a result, a dynamic subjective impression is formed in the minds of consumers or the public. Brand is a virtual vocabulary, it does not have an independent entity, does not occupy space but belongs to the intangible assets of a city, enterprise, and product. But the ultimate goal of a brand is to give the public a relatively simple and easy way to remember a product or business. Both students at school and young people at work have the same access to travel information. As shown in Figure 5.

At present, the personalized recommendation search engine for big data is mainly based on the keywords entered by users to identify, and based on some relevant information that can be extracted to give more accurate results. This works well for some shoppers with vague intentions. The development of algorithms corresponds to the development of productivity. Unlike the traditional relationship between productivity and means of production, the new goal of

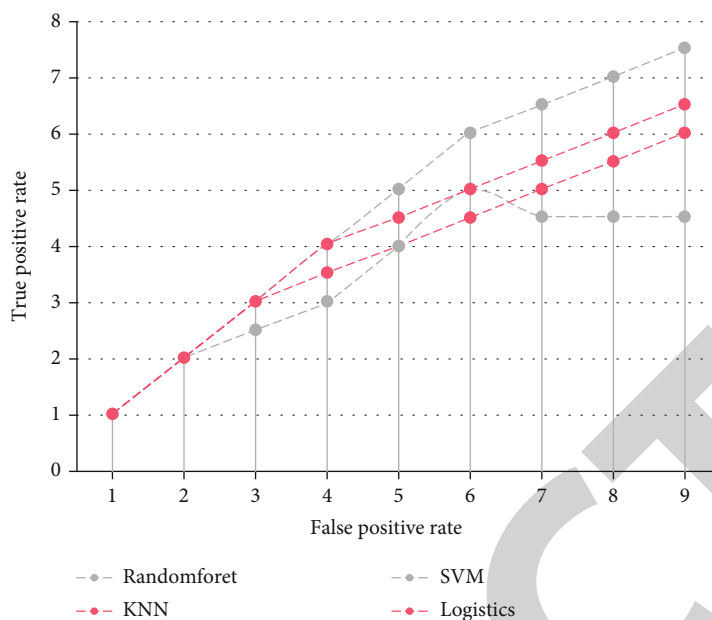


FIGURE 4: Comparison of Curve Results.

productivity development and algorithms, is to process larger and more complex means of production faster and better. The main function of the project index module is to calculate the index information of the project offline and store it on disk. The recommendation module is a module. The recommendation system provides instant recommendations for new users and directly faces users. Although the slogan “ecological leisure, hundred mile corridor” can have a certain impact, it cannot be used as a historical and cultural city to receive tourists. As shown in Figure 6.

Instead, it is necessary to first analyze the characteristics of user behavior in a specific search engine scenario, design a personalized recommendation model that can effectively mine the value of user data according to this characteristic, and then execute the corresponding recommendation algorithm to obtain a good recommendation effect. Let customers make retention suggestions for the first time in case of loss to reduce business losses. Workplace youths like national customs, most followed by natural scenery and urban gardens. Both show a general preference for festival activities, which are crowded by large numbers of people. As shown in Figure 7.

If the consumer comes from an external company with high correlation, because the goods or services provided by the current website are similar to those of the websites he visited before. Then the current website can provide goods close to the consumer’s previous preferences. Mainly, some local governments have put the construction, development, and management of urban brands on the agenda, and started systematic planning with a plan and layout. Good tourism image the construction of city tourism image is a very complicated problem. This is because its influencing factors involve many aspects. Generally speaking, the urban tourism image is mainly expressed through various image carriers, image hardware, and image software. Therefore, when we carry out the

construction of urban tourism image, we must start from various image carriers. Coordinate all factors, make overall planning for all aspects of the city, and at the same time, implement the principles of serving tourism and aesthetics. The city and its surrounding scenic spots and urban gardens are the main carriers of the city’s tourism image and the most important material basis for establishing its tourism image.

4.3. City Brand Image Design Strategy. Planning content of urban brand image design: thinking about urban positioning industry, transportation, tourism, and other aspects, based on the in-depth excavation of urban historical and cultural heritage; systematically and comprehensively provide the core demands of urban strategy. Urban positioning will organize interdisciplinary experts to conduct on-the-spot investigation, investigation, diagnosis, and evaluation, which is a necessary prerequisite for urban visual design system. City symbol is the core element of the city brand image system, the embodiment of the materialization of city culture and the concentrated embodiment of city values. It is unique and exclusive, and it is the core language of urban personalization.

4.3.1. Establish brand management strategy. The confusion of city brand image positioning will cause the core value of city brand to be unable to be effectively conveyed, and the competitiveness of city will be greatly reduced. Therefore, if the city brand image communication wants to convey the soul and temperament of the city in a unique way, it must carry out unique brand image positioning. Brand image positioning is the core link of city brand image communication. Positioning is conducive to the formation of core competitiveness, and once it is formed, it has strong extension ability and exclusiveness. With the intensification of competition among cities, distinctive and unique city

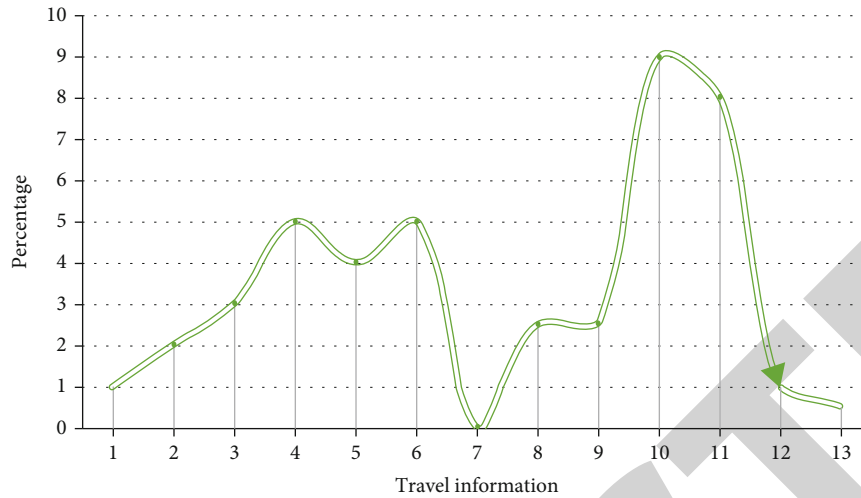


FIGURE 5: How to Get Travel Information.

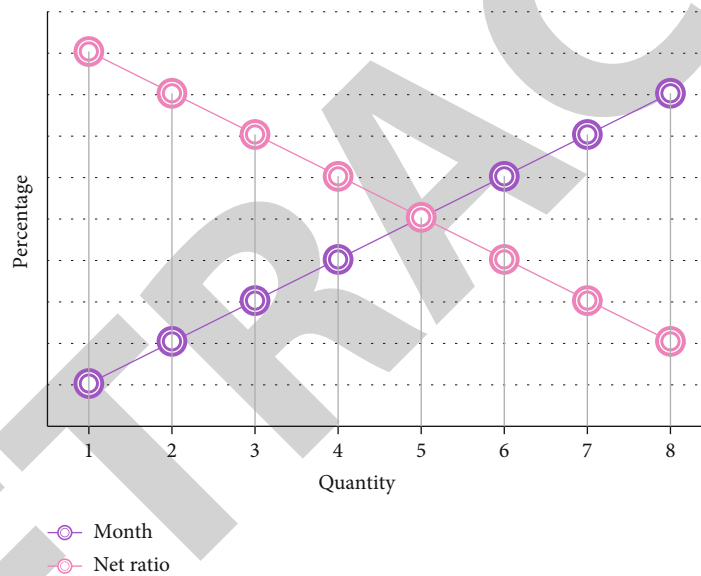


FIGURE 6: Tourist reception.

brand image plays a key role in improving the effect of city brand communication. The shaping and spreading of city brand image is a long-term and trivial work related to the development strategy of the city and the implementation of specific projects. The specific activities of city brand image communication should be consistent with the orientation and policy of city brand image, and corresponding adjustments should be made according to the specific conditions of specific activities. Driven by the “big data” technology, the media contact habits of the target audience can be clearly presented, and problems such as which programs the target audience chooses to watch, how long they are, and how much they like the programs can be solved, which facilitates the efficient integration of various communication means. In addition, with the continuous emergence of new media, the channels for the public to obtain information are constantly updated. Therefore, while the traditional means of commu-

nication are used in urban brand image communication, more attention should be paid to the application of new means of communication.

4.3.2. Upgrade city brand image communication content. As a new communication platform, social media can quickly transmit any kind of emergency to every corner of the earth, thus causing public crisis. Because of the different social groups and individual needs of the audience, the demand differentiation trend of the audience gradually becomes obvious, resulting in the diversity and differentiation of the audience’s demand for information content. This has led China’s media industry to gradually shift from the sender-oriented to the audience-oriented, and began to pay attention to the audience’s demand for different content. Due to the development of China’s commodity economy, the increasingly fierce competition in the media market, the

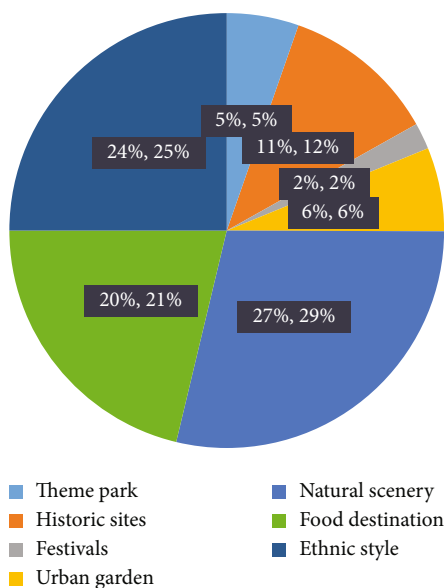


FIGURE 7: Preferences of tourist attractions.

continuous development of media technology and many other reasons, the audience has been given unprecedented initiative, and their self-awareness is also increasing. Taking the audience as the center and deciding the content style of communication has become the primary condition for the media to survive. The establishment of city image is an administrative concept established in combination with the current situation of government resources, the direction of urban development, social environment, and development trend. It is also to establish a good administrative image through the work recognized by the public and development goals, which coincides with the main development concept of the modern urban brand plan. Through the construction of city brand, the society can better understand the culture, development direction, and government work of a city. The combination of the construction of city brand and the establishment of government image is conducive to the realization of the public management function of the government, and the creative combination of the country and government concept with the city image. Promote the realization of the government's political, economic, cultural, and social service functions. City managers can pay attention to the trend of public opinion in real time with the help of data analysis charts, and should disclose information comprehensively and fairly when necessary, so as to minimize the possibility of rumors and negative emotions of the audience. When the crisis is in the outbreak period, the guidance of public opinion must strive for every possible time. Organizations or governments should release information at the first time, take the initiative in discourse as quickly as possible, and strengthen interaction with emerging media to reflect real public opinion.

5. Conclusions

This paper analyzes how cities in the big data era can better carry out brand image design. This paper analyzes a series of

brand image management work, such as the maintenance and evaluation of brand image. With the support of big data and related theories of urban tourism destination brand image management, this paper proposes that in the era of big data, the collection and analysis of data will help tourism destinations better manage their brand image. Cultural and creative industries and industries related to tourism elements constitute an intertwined cultural and tourism industry chain. Through the development of tourism, tourists will produce comprehensive and diversified consumption in the city. On the basis of inheriting traditional culture, cities should strive to create new and unique urban cultural images. Data technology plays an important role in urban planning and development as well as industrial development. Based on big data analysis, it can provide powerful decision support for government and industry managers.

However, the study has certain limitations. The research did not analyze from accurately targeting the national standard audience and mining their consumption demand. The research also needs to accurately position the city brand image and effectively integrate the brand communication means. Real time interactive public opinion guidance and intuitive and accurate effect evaluation.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Z. Liang and C. Zhikui, "Design and implementation of big data algorithm library teaching experiment platform," *Experimental Technology and Management*, vol. 37, no. 6, p. 6, 2020.
- [2] 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.
- [3] J. Miklós-Thal and C. Tucker, "Collusion by algorithm: does better demand prediction facilitate coordination between sellers?," *Management Science*, vol. 65, no. 4, pp. 1552–1561, 2019.
- [4] W. Huifang, S. Fang, Y. Chen et al., "Design of big data personalized recommendation system based on SSM framework," *Information and Computer*, vol. 34, no. 2, p. 3, 2022.
- [5] J. Jiang and H. H. Wang, "Application intelligent search and recommendation system based on speech recognition technology," *International Journal of Speech Technology*, vol. 24, no. 1, pp. 23–30, 2021.
- [6] K. Al Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riffi, and A. Sabri, "Hybrid recommender system for tourism based on big data and AI: a conceptual framework," *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 47–55, 2021.

Retraction

Retracted: Assisting Laboratory Management Based on Network Big Data Mining Technology

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xu, "Assisting Laboratory Management Based on Network Big Data Mining Technology," *Journal of Function Spaces*, vol. 2022, Article ID 2087287, 7 pages, 2022.

Research Article

Assisting Laboratory Management Based on Network Big Data Mining Technology

Yonghua Xu 

School of Computer Engineering, Jinling Institute of Technology, Jiangsu, Nanjing 211169, China

Correspondence should be addressed to Yonghua Xu; xyh@jit.edu.cn

Received 24 June 2022; Revised 20 August 2022; Accepted 25 August 2022; Published 13 September 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Yonghua Xu. 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 today's big data context, the composition of virtual networks is becoming increasingly complex, so it is neither easy nor difficult to strengthen the control of network security issues. This paper studies the important content of network big data mining-assisted laboratory management. On the basis of discussing the connotation of laboratory safety, combined with management practice, this paper puts forward some suggestions on laboratory safety management. This paper puts forward some views and thoughts on some problems existing in the management and experimental teaching of chemistry laboratory. By judging the stability of instruments and equipment, we know whether the calibration status of instruments and equipment has changed during this period, so as to ensure the sustainable use of instruments and equipment and the accuracy and traceability of laboratory test results. The use of laboratory management system can greatly improve the utilization rate of laboratory equipment, promote students' autonomous learning, and promote the standardization of laboratory construction. Improve the network monitoring method of big data mining technology and the application mechanism of big data network monitoring, in one type of laboratory. Teachers and students in the laboratory can reasonably arrange their own time according to their own interests to guide and operate experiments. Through continuous operation and practice, students' practical ability and experimental knowledge level should be improved as much as possible.

1. Introduction

With the progress of human beings and the development of science and technology, computers have become a necessary tool in people's life and work. Because of the weakness of traditional network security technology, information leakage has become a breeze [1]. At the same time, with the development of data mining technology itself, there are big data mining technologies represented by high-performance data mining, data stream mining, and complex data mining technologies. The traditional laboratory management mode has been unable to adapt to the current development trend. Therefore, modern management methods and technologies are used to carry out laboratory informatization construction. The experimental link has been relatively weak, and the national investment is seriously insufficient. Not only are the experimental instruments outdated and the number of sets insufficient, but also is the experimental environment poor, and the facilities such as lighting and ventilation are

imperfect. This situation has seriously affected students' learning enthusiasm and interest and restricted the cultivation and improvement of students' experimental research and hands-on ability. Traditionally, most of them are large-scale speeches handed down from street to street. If you want to get information, you need to go through private surveys and unannounced visits. Compared with online access, the cost is high and the efficiency is low [2]. Moreover, with the increase of the number of laboratory equipment, the number of laboratory equipment can no longer meet the normal teaching use, which seriously affects the training of students' experimental skills. It is precisely because of the crazy growth of data that "big data" is slowly being studied by all walks of life.

In some western developed countries, especially the United States, their scientific and technological level has always been in the leading position in the world. These countries have earlier researched on computer and network information technology, so they have developed very

rapidly. In recent years, although the Ministry of Education and each state have issued a number of documents to improve teaching quality, high-level teachers are encouraged to undertake basic course teaching tasks, and attention should be paid to the practical teaching of undergraduates [3]. Due to the hidden nature of the Internet, it brings us convenience and many uncertainties. If we can control those that endanger society in advance and seize the initiative, then we can better exert our advantages. With the improvement of the society's requirements for students' comprehensive quality, the laboratory teaching management system is gradually improving [4]. The traditional laboratory management model can no longer meet the current teaching methods, and the traditional management model needs to be broken. As an important part of laboratory management, strengthen laboratory safety management. It is an important task for laboratory management departments to build a normal operation and safety guarantee system for teaching and scientific research. The purpose of the reform of the laboratory management system is to improve the efficiency of the laboratory, realize the sharing of equipment and resources, and provide good material conditions and guarantee mechanisms for the reform of experimental teaching [5]. As a means of information processing, data mining technology has advanced functions such as analysis and prediction. Network security can also be guaranteed by improving the accuracy of network intrusion detection. As far as the current network security technology is concerned, although some progress has been made, there are still some problems such as weak protection capability and more security vulnerabilities, which are prone to network security accidents [6]. The network is the basis and barometer for network public opinion to guide work. Change traditional network public opinion-guided thinking with big data concept, apply big data mining technology to network monitoring well, and discover the information behind the network. However, the construction and management of information technology are not a simple process but require time and practice to reach a satisfactory level.

Laboratory is an important base for colleges and universities to cultivate innovative talents, conduct scientific research, and serve the society, and it is also an important symbol to reflect the level of teaching and scientific research and show the strength of running a school. Data mining technology is a means of processing information at present; on the one hand, it can realize the analysis and prediction of network data [7]. On the other hand, it can also effectively improve the efficiency and accuracy of network intrusion detection. It is of great theoretical significance and practical value to guide the network public opinion in the current environment and then help the government to make decisions and maintain the network social security. From the content point of view, information management mainly includes two parts: scientific management and resource management. Although the emphases of the two parts are different, they are closely related to each other. The laboratory management is weak, the utilization rate of experimental teaching resources is not high, and the benefits are low [8]. Undoubtedly, the above situation is incompatible with the

requirement of cultivating high-quality undergraduates. Traditional laboratory management is only for the management of laboratories, experimental equipment, and experimental personnel. And with the continuous improvement of the society's requirements for students' comprehensive quality, students' independent innovation ability has become the standard of today's new talents, and the drawbacks of traditional laboratory management methods have gradually emerged [9]. Therefore, how to scientifically control and manage the safety of the laboratory is an important issue faced by our laboratory managers.

The transformation of teachers' ideas is the key and premise of implementing innovative education. We should get rid of the shackles of traditional teaching ideas and strengthen our own ideas of advancement, openness, application, and democracy. We should give full play to the leading role of teachers and the main role of students and mobilize students' enthusiasm and initiative in learning. Secondly, we should create the best teaching procedure and the best teaching situation according to the teaching objectives and students' personality differences. Thirdly, teachers' awareness of knowledge authority should be changed, and the traditional indoctrination education should be changed into heuristic, discussion, and inquiry, so that students can study in a relaxed and happy mood. Encourage students to explore different ways to solve problems boldly, so as to further stimulate students' internal learning motivation. We should integrate all kinds of teaching resources. Because each student's basic knowledge, learning experience, self-study ability, and other factors are different, and their academic performance is different. Therefore, all kinds of teaching information resources with different degrees of difficulty should be set up in the teaching information resource database to meet the learning needs of students at different levels. At the same time, we should further build and timely update the multimedia teaching information resource system, including multimedia material library, multimedia textbook library, and multimedia courseware library. Actively optimize and combine various teaching media to realize the optimization of teaching process.

2. Related Work

No matter from foreign or domestic research, there have been some related academic papers on data mining technology and its application in network monitoring in the field of big data, but on the whole, there are still few related studies, lacking systematic and in-depth research. Luo et al. think in one article: big data has triggered a reexamination of scientific research methodology and is triggering a revolution in scientific research thinking and methods [10]. Yang et al. put forward in a paper: data mining helps network monitoring and guidance to choose the path, including network association analysis, network level division, network clustering, and network tendency analysis, and analyzed the practical value of network monitoring and guidance in the view of data mining [11]. Lou establishes scientific data management system applied in entry-exit inspection and

quarantine, taking the important nodes in the laboratory data management process as the research object, comparing the characteristics and advantages of traditional data management methods and laboratory scientific data management systems, and establishing a data management system suitable for inspection and quarantine laboratories through research and introduction [12]. Xu and Liu pointed out in their article: the progress and development of big data technology research have brought great opportunities for the development and application of data mining technology, and data mining technology will enter a new development period, giving the future development direction of data mining technology in the era of big data [13]. Huang established the laboratory of the Guizhou Geology and Mineral Center as an example; he expounded the big data and its enlightenment to the analysis and testing of geological and mineral resources. On the basis of introducing the connotation of big data, analyze the characteristics of big data [14]. Scholars such as Li and Xiao mentioned in their article that the patterns, trends, and correlations of big data can reveal social phenomena and predict the laws of social development. Network science and data science provide a new scientific methodology for social science research [15]. Wang proposed: combined with the impact of big data on social research, he analyzed the advantages of social research in the era of big data in data collection, data storage, data processing, and data presentation [16]. In the article, Rong and Gang designed a monitoring system deployed on the Internet, which monitors various information media such as web pages, forums, and microblogs on the Internet and automatically collects data on Internet pages [17]. Another trend is to analyze the environment, risk, research direction, and practical value of network monitoring in the big data environment from the macro level.

3. Exploration and Practice of Laboratory Safety Management

3.1. Standardize Infrastructure Construction and Eliminate Potential Laboratory Safety Hazards. The basic condition of laboratory is the first condition to ensure the safety of laboratory. Attach great importance to laboratory safety issues, start from the source, and incorporate safety construction standards into the planning and construction of laboratory infrastructure. The school occupies an important position in the teaching and scientific research work. But for a long time, experimental teaching has been attached to theoretical teaching. As the verification of curriculum theory teaching, the construction and management of laboratory are basically attached to the teaching and research section or research group. The birth of data mining technology comes from the irregularity, complexity, diversity, and other characteristics of the original data. Using data mining technology, valuable information in the original data can be extracted to make it play more roles [18]. Faced with the complexity of big data, some scholars try to use statistical methods and complex network methods to study how to reduce big data on demand [19]. With the rapid development of the Internet all over the world, people are full of curiosity and interest in

this emerging field, and the Internet has also started from a small research project. After more than 20 years of development, it has finally become the communication system that most people in the world rely on. By generating a function that maps data items to a real-valued forecast variable, regression analysis finds the dependence among variables and studies the forecast and trend characteristics of data series. The laboratory is provided with a card swiping machine. After entering the laboratory in the spare time, students will read their all-in-one cards on the card swiping machine, and the courses that the students have taken will be displayed in the system [20]. Active exploration and practice have also been carried out in the teaching of chemical theory and experimental courses, as well as students' scientific research training. The data structure of the power controller table is shown in Table 1. The data structure of the device information table is shown in Table 2. The data structure of the control association table is shown in Table 3.

A 24-hour monitoring of laboratories and hazardous substances shall be carried out to prevent outsiders from contacting and entering the laboratories, thus eliminating the loss and improper use of hazardous substances.

$$\begin{aligned} \text{Centroid} &= m_i = \frac{1}{n_i} \sum_{X \in C_i} X, \\ \text{Radius} &= \sqrt{\frac{\sum_{X \in C_i} (X - M_i)^2}{n_i}}. \end{aligned} \quad (1)$$

Sample deviation matrix A and sample covariance matrix S are

$$\begin{aligned} A &= \sum_{X \in C_i} (X - M_i), \\ S &= \frac{A}{n - 1}. \end{aligned} \quad (2)$$

When $q = 1$, the first-order Minh distance is

$$\begin{aligned} L_1(X_i, X_j) &= \lim_{d \rightarrow \infty} \sum_{k=1}^d |x_{ik} - x_{jk}|, \\ L_2(X_i, X_j) &= \left(\lim_{d \rightarrow \infty} \sum_{k=1}^d |x_{ik} - x_{jk}|^2 \right)^{\frac{1}{2}}, \\ d_{ij}M &= (X_i - X_j)' S^{-1} (X_i - X_j), \\ d_{ij}L &= \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{k=1}^p \frac{|x_{ik} - x_{jk}|}{(x_{ik} + x_{jk})}. \end{aligned} \quad (3)$$

3.2. Establish Rules and Regulations and Improve Safety Management Network. Only when the system is perfect, scientific and reasonable, and feasible can the safety construction of the laboratory have laws and rules to follow. The establishment of experimental center is an important

TABLE 1: Power control table.

Data item name	Data item type and length	Data item description
PC_PCID	Int	Primary key: controller number
PC_Mac address	Varchar, 20	Controller Mac address
PC_start date	Date time, 8	Put into use time
PC_manufacturers	Varchar, 60	Controller manufacturer

TABLE 2: Data structure of device information table.

Data item name	Data item type and length	Data item description
LE_LEID	Int	Primary key: controller number
LE_equipment name	Varchar, 20	Device name
LE_manufacturers	Varchar,50	Equipment manufacturer
LE_start date	Date time, 8	Equipment put into use time

TABLE 3: Data structure of the control association tables.

Data item name	Data item type and length	Data item description
CR_LE_LEID	Int	Foreign key: experimental device number
CR_PC_PCID	Int	Foreign key: power controller number
CR_state	Bit	Association status

attempt in the reform of laboratory system, and its brewing and establishment have been strongly supported by the school leaders. The neural network can also be regarded as a collection of multiple nerve units, so that the problem of large clusters of biological neurons connected by axons can be effectively solved. As the position of scientific research, the laboratory has many shortcomings, and it is bound to be impacted by big data. Big data mining technology is mainly applied in network security. The development of human society is a process of continuous updating. Big data plays an important role in the development of science and technology and the change of human concepts, and it exists everywhere. Or use the association rule method, through which the relationship between data items in the database can be described. If a certain item in one group of data and a certain item in another group of data appear at the same time, it can be judged that the two groups of data are related or interrelated. Under the same data set, with the increase of the number of nodes, the running time gradually decreases, which reflects the scalability of the algorithm. With the same number of nodes, the running time increases with the increase of data set. The main job of the laboratory administrator is to operate and manage all subfunctional modules. Therefore, I believe that laboratory managers must be cost-conscious and should implement institutionalized

simple management to achieve the goal of high efficiency and low cost. To make student lab report results. In addition, teachers can print and export data online for student scores or experimental reports through the score statistics function. It also implements the accountability system for laboratory safety work, with clear responsibilities and accountability. The laboratory director is the person in charge of the safety of the laboratory and is responsible for the school and college. Strictly implement the relevant safety management regulations of the University and college, and organize the formulation of laboratory safety management rules in combination with the actual situation of the unit. Regularly educate relevant personnel on prevention and safety laws and regulations, and urge them to consciously abide by various safety management rules and regulations. Regularly organize safety inspections, and keep safety records. Find hidden dangers and loopholes, and deal with them in time. If it is difficult for the office to rectify due to objective factors, temporary emergency measures must be taken.

4. Effective Application of Big Data Mining Technology in Network Security

4.1. Shortcomings Exposed in Traditional Network Security Technologies. At present, the network security technology used by people in daily life and work is mainly designed for a certain or a certain network security problem. Therefore, to some extent, these network security technologies can only solve some or some network security problems. However, these network security technologies cannot solve other related problems, let alone effectively protect the entire network system. For example, the access control and identity authentication technology used in people's network technology can only solve the problem of network user identity confirmation. However, the security of information transmission between users cannot be guaranteed. With the rapid development of network technology, the problem of network security has been paid more and more attention by the public. Nowadays, the data generated in medical, transportation, finance, education, and other fields is huge, and it belongs to the category of big data. The laboratory administrator is mainly responsible for inputting new experimental equipment information in the system; modifying, deleting, and querying the equipment information; etc. Once the equipment fails, it will report for repair. Only when a laboratory can meet these key factors can it have the best premise of providing users with high-quality services. Transaction response time mainly refers to the time it takes for the system to complete the operation after the user logs in to the system and performs the corresponding operation. The average transaction response time of the target system is shown in Figure 1.

In this way, users will be interested in the lab, and the lab will continue to be popular with users. In order to improve the management efficiency of each link, computer can be used to conveniently arrange the related teaching process of students, teachers, and laboratories. Under the new situation, lab safety work urgently needs us to actively change the

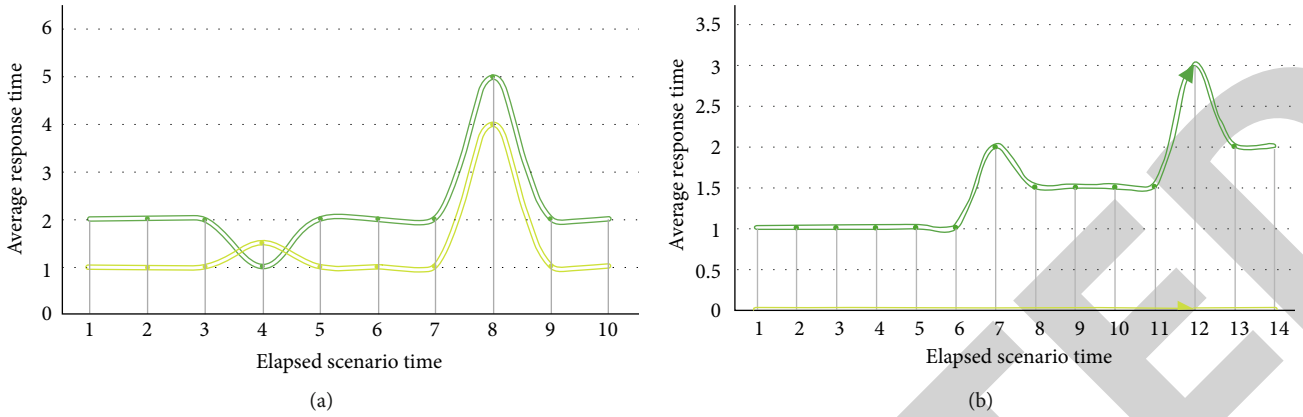


FIGURE 1: Average transaction response time.

mode of safety management, advocate safety services, and focus on prevention. Increase the breadth and depth of laboratory safety management, and use information technology, science, and technology and cultural means to change postevent management into preventive management. Trends provide a comparative analysis of the attention of multiple keywords, as shown in Figure 2.

According to the teaching needs, it is no longer restricted to “owning me,” but emphasizing “available for me,” so that the school can adjust and control from a macro perspective, so as to “turn the parts into whole.” The application value of data mining technology in the manufacturing industry is mainly reflected in the inspection of product quality, such as finding out the rules by researching and analyzing product data or by analyzing the production process to find out the main factors that affect the production efficiency and product quality. Since it is useful for public health programs, an attempt was made to further validate the model with the weekly ILI percentage for individual states. The CDC did not publish interstate data, and researchers used the percentage of ILI published by Utah to validate, as shown in Figure 3.

The laboratory incorporates the results generated by the above quality control methods and the laboratory’s self-quality requirements into the big data analysis to screen out the project data that is necessary for quality control. Under the conditions of traditional network technology, information storage also needs to rely on manual input, which is not only inefficient but also prone to input errors, resulting in a large manual workload.

4.2. Effective Application of Big Data Mining Technology in Network Security Management. To some extent, no matter what kind of network security problem, its infringement on network security is traceable, especially for network viruses. With the help of data mining technology, users’ data can be classified, collected, and evaluated through corresponding technical means, so as to achieve the purpose of dynamically scanning system data. In the process of applying big data mining technology to prevent network security problems, the application process is relatively complex. The amount of data involved is also relatively large, so it is necessary to clearly grasp the characteristics of each link and make a reasonable plan. Build multiple analysis modules

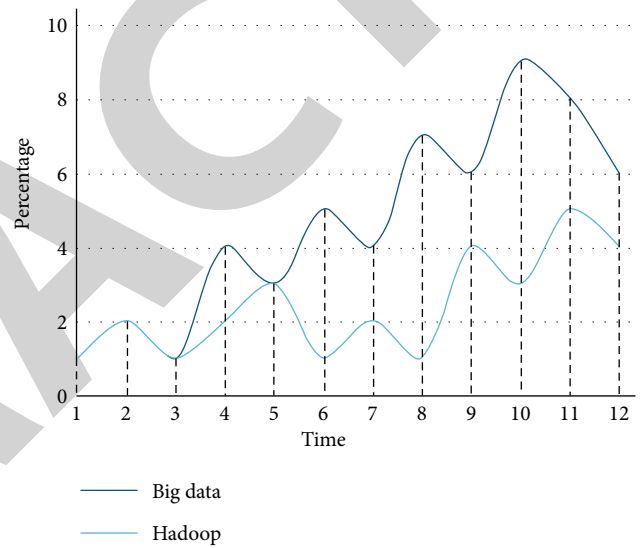


FIGURE 2: Search trend comparison score.

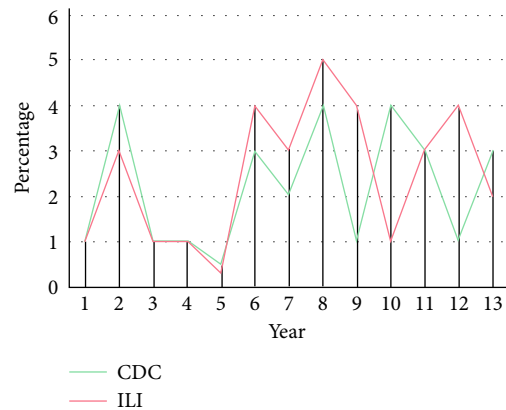


FIGURE 3: Model comparison diagram.

to ensure the security of network applications. Based on the systematic characteristics of large data technology, the specific working status of large data technology can be divided into four parts, namely, collecting information, preprocessing information, mining information, and pattern evaluation

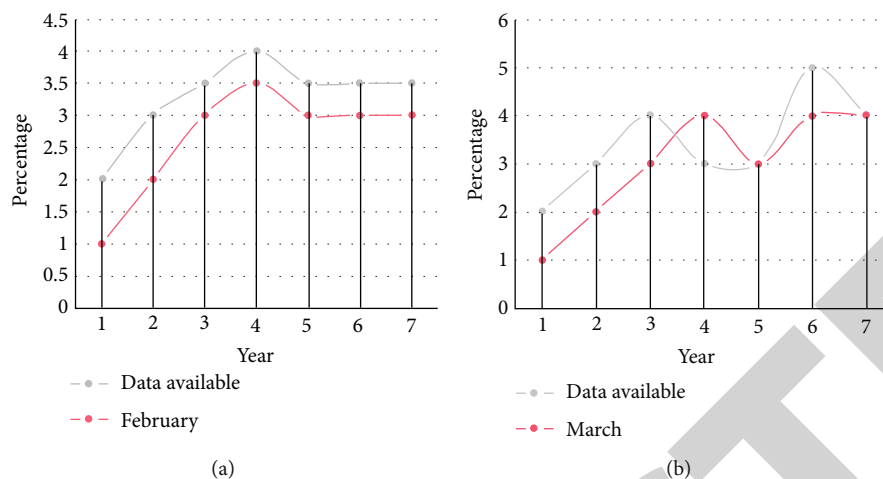


FIGURE 4: Comparison of data between two model areas.

information. The big data approach involves not only acquiring or storing data but also analyzing it to understand and discover its value. Among them, the director of the experimental center can arrange the experimental projects according to the school's teaching plan, manage the information of the experimental courses and the experimental projects, and can also view the optional records of the experimental courses. It not only serves postgraduate training and tests of various scientific research projects but also faces undergraduates [7], besides meeting the teaching needs of instrumental analysis experiments and comprehensive chemical experiments. If there are classes in the laboratory, it is judged whether the user's credit card swiping time is within the class time. If it is in the class time, according to the class scheduling information in the database, query the class and teacher of this class in the laboratory. Strengthening the supervision and control of laboratory environmental pollution has become the consensus of teachers, students, and leaders. The data is shown in Figure 4.

First of all, set goals and make implementation plans from the level of teaching reform. Now, the first phase has been completed, and the second phase construction will be adjusted on the basis of summing up the experience of the first phase construction. In this case, telecom enterprises must take effective measures to improve their technical level and service quality, so as to ensure customer satisfaction and loyalty. In practice, we should constantly strengthen the improvement of intrusion detection and improve the application scope of big data technology. Ensure the effectiveness of each work link, and ensure the security of the network environment.

5. Conclusions

With the rapid development of information technology and network technology, people rely more and more on science and technology. Network monitoring in the context of large data environment needs the convenience provided by large data mining technology. It is also necessary to improve the

network monitoring method of big data mining technology and improve the application mechanism of big data network monitoring. In a class I laboratory, teachers and students can reasonably arrange their own time. Conduct experimental guidance and operation according to their own interests, and improve students' practical ability and experimental knowledge level as much as possible through continuous operation and practice. However, there are still many specific problems to be solved. This paper puts forward some views and thoughts from the perspective of some problems existing in the management and experimental teaching of chemical laboratories. The use of laboratory management system can greatly improve the utilization rate of laboratory equipment, promote students' autonomous learning, and promote the standardization of laboratory construction. By judging the stability of instruments and equipment, we can find out whether the calibration state of instruments and equipment has changed during this period, which ensures the sustainable use of instruments and equipment and the accuracy and traceability of laboratory test results. Therefore, it is necessary to carry out more detailed and specific research work, constantly improve and innovate the data mining technology, and improve the intrusion detection efficiency and network security characteristics more effectively through application research. However, the research has certain limitations. It is also necessary to set parameters and deeply mine intrusion behavior paths. At the same time, the corresponding algorithm is applied to predict the intrusion behavior scientifically, so that abnormal intrusion detection can predict and detect unknown intrusion behavior in time and improve its protection effect. This needs further supplementary analysis in future research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Retraction

Retracted: Analysis of Consumer Behavior Data Based on Deep Neural Network Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Yuan, "Analysis of Consumer Behavior Data Based on Deep Neural Network Model," *Journal of Function Spaces*, vol. 2022, Article ID 4938278, 10 pages, 2022.

Research Article

Analysis of Consumer Behavior Data Based on Deep Neural Network Model

Shijiao Yuan 

Economy and Management Department, Hebi Polytechnic, Hebi 458030, China

Correspondence should be addressed to Shijiao Yuan; 20000454@hbzy.edu.cn

Received 28 June 2022; Revised 10 August 2022; Accepted 18 August 2022; Published 12 September 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Shijiao Yuan. 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 divides the research modes of consumer purchase behavior characteristics into three categories: experience-driven mode, theory-driven mode, and data-driven mode. An analysis algorithm based on customer consumption behavior is proposed, and the idea of combining customer consumption behavior factors such as satisfaction and loyalty is proposed. Through comparison, it is pointed out that the data-driven model is most suitable for analyzing the characteristics of online consumers' purchasing behavior. Using the decision support of knowledge base, different service schemes for customers with different evaluation degrees are realized. In order to improve the accuracy of sample classification and maximize the output function, genetic algorithm is used to optimize the samples. A deep neural network structure algorithm is proposed to classify customer transaction data samples. In this algorithm, the sheep nodes are not fixed, but the number of hidden layers and unit nodes of the neural network are dynamically determined according to the sample training. The research excavates various kinds of valuable information such as consumer preferences and consumption structure from the huge consumption data of consumers. It is not only helpful for enterprises to analyze consumers' consumption behavior and organize production but also helpful for enterprises to realize the concept of personalization.

1. Introduction

At present, the consumption concept of consumers has undergone great changes [1]. Nowadays, consumers no longer pay attention to the price of goods but pay more attention to the quality of goods, after-sales service of merchants, service attitude of sales departments, etc. [2]. Traditional commodity trading methods are now being affected by online consumption. The visual communication between merchants and customers has been interrupted. Customers can place orders without going to the store to compare goods, which greatly improves efficiency [3]. Western research on consumer behavior originated from "consumer analysis" [4]. However, most consumer behavior patterns are driven by experience or theory [5]. With the intensification of market competition, enterprises must focus on customer needs, strive to collect consumer information, mine customer consumption characteristics, formulate marketing strategies that meet market conditions, and improve market competitiveness [6].

It is mentioned in the science of consumer behavior that the consumption behavior of consumers has great volatility, and it is difficult to quantify with mathematical or logical rules [7]. The idea of data mining is to find hidden rules from irregularities [8]. This paper analyzes the consumption characteristics of users and finds out the consumption preferences of users in different places by tracking different consumption records. Because the offline stores cannot effectively conduct data statistics on consumers, they cannot well grasp the challenges they face [9]. At present, China's e-commerce is booming. With the implementation of the home broadband project, more and more consumers have access to the Internet and surf the Internet through the home network, which will increase the possibility of consumers' online shopping. In order to succeed in online shopping, consumers' support and participation are urgently needed. If you want to stimulate or motivate consumers to do online shopping, it will not work if you cannot clearly understand consumers' behavior [10].

Based on the idea of data mining, this paper considers two kinds of problems of consumers: consumption factors and consumption research objects.

2. Related Work

Shang et al. pointed out that online shopping itself has the advantages of convenience and quickness, which is favored by more and more consumers [11]. According to Boston Consulting Group's prediction, the current average level of 1,000 dollars in the United States exceeds [12]. Ke et al. pointed out that with the rapid development of the online shopping market, online shopping is no longer an optional supplement for consumers in addition to traditional shopping but has become an important shopping way for many consumers. The attention to online shopping behavior, especially the convenience provided by the Internet, and the variety of commodity types make a special purchase behavior—impulsive buying behavior likely to occur [13]. Zheng et al. pointed out that people would find ways [14]. Based on online word of mouth, Hong JL and others have a large number [15]. The fundamental reason for the uneven views of some consumers lies in the problem of consumers' decision-making. Different consumers have different consumption views, which leads to the difficulty of common consumption [16]. From this, it can be seen that online marketing is a way of direct targeting to convey specific marketing information to specific individuals, including one-to-one marketing through rich database content analysis and identification of online consumers' behavior patterns or their preferences [17]. Therefore, the network is an ideal medium with great potential for manufacturers to contact with potential customers and consumers and conduct relationship marketing. Based on the traditional model, Olan et al. and Sayeed et al.'s analysis and prediction of consumer behavior is only in the qualitative stage. Now, big data analysis can be used to track consumer shopping behavior and improve consumers' awareness of shopping platforms. loyalty to gain greater market share [18, 19].

Nik PG based on the key link, from the perspective of consumer behavior process problems [20]. This paper discusses the marketing strategy and marketing strategy combination that enterprises should adopt to carry out online marketing, aiming to provide guidance for enterprises to carry out online marketing.

3. Methodology

3.1. Customer Behavior Analysis Model. The consumption pattern evolves with the changes of productivity and production relations. The main reasons for the change of natural consumption patterns are the ecological environment, the degree of scientific development, and the situation of population and resources. The main reason for the change of consumption social model is the change of social needs based on consumption needs. In recent years, with the changes in the ecological environment, the degree of scientific development, and the situation of population and resources, China's consumption pattern has undergone unprecedented changes. This change reflects the change of social demand

dominated by consumption needs and the subtle influence of consumers on consumption patterns. The consumption pattern has undergone a revolutionary change. Now, customers do not have to go to the mall, just go to the homepage of the mall, and click the "Buy" button to buy the goods they want. The credit of shopping malls, even the credit problems of customers, the service attitude and quality of shopping malls, and so on, also appear. Building the model is done. Different algorithms are used for model training, and a unified evaluation standard is used to evaluate the effectiveness of the model, and then, the optimal model is selected to predict the product recommendation in the product subset to improve the accuracy of the recommendation. BP neural network has strong learning ability and nonlinear parallel processing and reasoning ability. A consumption behavior research model based on BP neural network. Before exploring the traditional prediction model, this section first designs the basic process of building the prediction model, which is applicable to the establishment of various models discussed in the paper, as shown in Figure 1.

Whether the customer is satisfied with the product includes three indicators: these data should be obtained through customer feedback information, that is, through questionnaire survey. Finally, the customer satisfaction is obtained by weighted summation:

$$M = \sum_{i=1}^m DC_i * ci, \quad i = (1, 2, \dots, m). \quad (1)$$

Among them, DC_i represents the weight of the i -th survey item, and ci represents the weight that the customer thinks the i -th survey item occupies in all survey items. Whether consumers are loyal to a product mainly includes the following indicators: the number of visits per week, the retention rate, and the number of purchases, where retention rate = number of visits/average number of visits per week. Based on the above, consumer loyalty can be expressed as

$$L = \sum_{i=1}^m ZC_i * ci | i = (1, 2, \dots, m). \quad (2)$$

According to the above definition, combined with the training and verification capabilities of the BP neural network, the customer behavior analysis model is obtained:

$$y_b^k = f \left(\sum_{i=0}^{N_1} \omega_{\theta} * x_i^k + \theta_k \right). \quad (3)$$

Among them, ω_{θ} is the weight corresponding to the analysis factor (satisfaction, loyalty, etc.) v_h^k in the output node x_h^k .

Typically, consumers are different in the likelihood of transacting with a business in a year versus a month. Then, the retention of the consumer is

$$B = \frac{N}{C * R}. \quad (4)$$

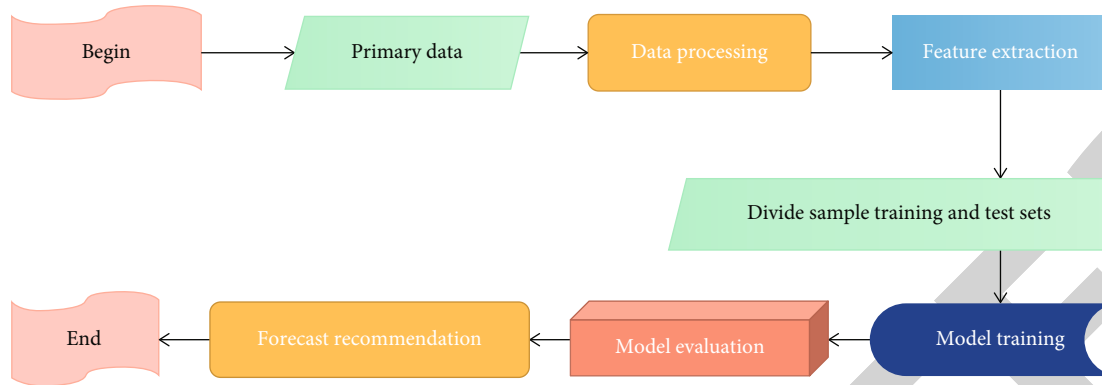


FIGURE 1: Flow chart of building prediction model.

According to the above definition, combined with the training and verification ability of BP neural network, the customer behavior analysis model is obtained: the output of the output layer node is

$$\begin{aligned}
 Z &= f \left(\sum_{i=0}^{N_2} \omega_{hj} * y_h^k + \gamma_j \right) \\
 &= f \left(\sum_{i=0}^{N_2} W_{hj} * F \left(\sum_{i=0}^{N_1} \omega_{ih} * x^k + \theta_k \right) + \gamma_j \right), \quad (5)
 \end{aligned}$$

where ω_{hj} is the weight of the output node Z corresponding to the hidden layer output node y_h^k . Z is the general evaluation of consumers. It reflects the error size function between the expected output and the calculated output of the network. The following is the output error of the i -th cell node: $E_k = 1/2 \sum_{k=1}^n (y_{ik} - T_{ik})^2$. The total error is

$$E = \frac{1}{2N} \sum_{k=1}^N E_k. \quad (6)$$

y_{ik} is the actual output value of the I node. For the neural network model, the hidden layer features are

$$h^{(1)} = \sigma \left(w^{(1)} h^{(i-1)} + b^{(1)} \right). \quad (7)$$

After training, the predicted value is close to the actual value, and the difference between them is defined as the loss function. Assuming that the training set is $\{(X^{(1)}, y^{(i)})\}_{i=1}^N$, n is the sample size, the overall loss function of the neural network model is

$$J(x, b) = \left[\frac{1}{N} \sum_{i=1}^N \|h_{wb}(x_i) - y(i)\|^2 \right] + \frac{\lambda}{2} \sum_t \sum_i \sum_j (w_{ij}^t)^2. \quad (8)$$

The first term is the mean square, which aims to control the error between the model output and the target, and the second term is the weight decay term, which prevents the model from overfitting through the weight decay magnitude.

Businesses get customer data from multiple channels, including consumption records, questionnaires, and feedback information, which contain a lot of important data, but also mixed with a lot of miscellaneous data that are not helpful for analysis. Therefore, it is necessary to use BP network model to analyze the information screening data after integration and cleaning. The whole process is shown in Figure 2.

Combined with the above figure, the whole process can be summarized into the following steps: (1) obtaining customer information, (2) integrate data from different sources through data migration and other ways and store them in another data warehouse, (3) using customer behavior analysis model for data analysis, and (4) after the analysis result is obtained, according to the degree of consumers and the knowledge base, the consumption strategy for this customer is obtained.

3.2. Consumer Behavior Data Processing Architecture. Recommend products to consumers with high accuracy and speed, focusing on designing effective analysis and prediction models. Before building a model, data processing and feature engineering are the basis for constructing a predictive model. Data processing refers to the analysis, calculation, sorting, and other processing of raw data. Feature engineering refers to the extraction of data features that are most suitable for the objectives of this research project on the basis of data processing. Consumer Internet behavior data is usually stored in the form of logs, and the logs related to consumer behavior analysis include consumer behavior logs, behavior event logs, and commodity category logs. Firstly, the interaction log is extracted from the user commodity interaction system to prepare the data related to the analysis and prediction of consumer behavior. Secondly, data preprocessing includes data cleaning, filling in missing values and removing outliers, removing duplicate data, ensuring the uniqueness of data, and dividing data sets according to time. The missing values of the paper are filled by the average. Thirdly, based on the overall description of the sample data in the form of charts and other forms, random sampling treatment is made according to the characteristics of unbalanced distribution of consumer behavior categories. Based on the original data, the data training set and test set are divided, shallow features are extracted

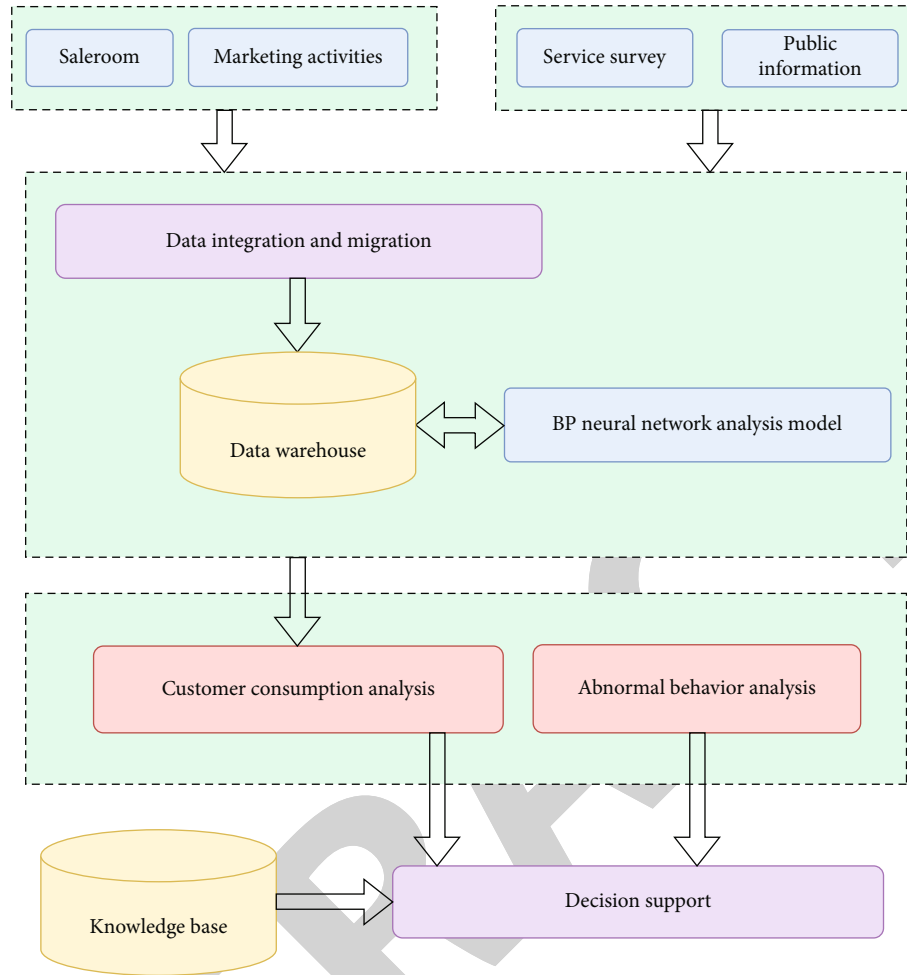


FIGURE 2: Customer behavior analysis model application model.

manually, the feature dimension is expanded, and then, the features are processed by methods such as normalization; finally, the prediction model is constructed and evaluated. In the construction of deep learning models, this paper makes a comparative study of prediction models. And identify the advantages and disadvantages of the model, so as to achieve recommendation prediction.

The flow chart of deep learning model construction is shown in Figure 3.

According to the characteristics of unbalanced category data, R DNN model and KM DNN model are introduced. Different models adopt early stop strategy. When the training times increase but the value of loss function no longer decreases, the training is stopped. Finally, the same AUC and F values are used as the method of effect evaluation. Data clustering is a Mini Batch KMeans clustering algorithm in sklearn-cluster module based on Python. Random sampling is based on Python's random module, and random generates random numbers to randomly extract negative samples from data. DNN is designed by using the deep learning library Keras based on Theano. Keras is modular and easy to expand, and it is guided by simple Chinese documents. It is easy for beginners of deep learning to get started, and it is also the foundation for researchers in the

field of deep learning to conduct in-depth research and excavation. In the user-product interaction log, there is very little data on consumer purchase behavior (marked as 1), and most of the data is the data that consumers have not purchased (marked as 0). There is a phenomenon of extreme imbalance between data categories. In classification problems, category imbalance often occurs, which is mainly reflected in the fact that there is less data in a certain category or several categories of samples. In real life, a small number of category samples are often the focus that deserves more attention. For example, the problem of advertising clicks. Users click on a small number of advertisements. Most users just browse without clicking. Focusing on advertisements with high click counts can help websites accurately put advertisements. The traditional classification model is built on the basis that the category samples are nearly balanced, and the classification model is more inclined. Therefore, when the categories are extremely unbalanced, the model may treat the data with less category samples as noise. Samples with few categories of data contain more important information. For example, in the identification of credit card fraud, if the fraud information is mistaken for normal information, the user will suffer heavy losses; in the medical diagnosis and treatment, if the patient is mistaken for a healthy

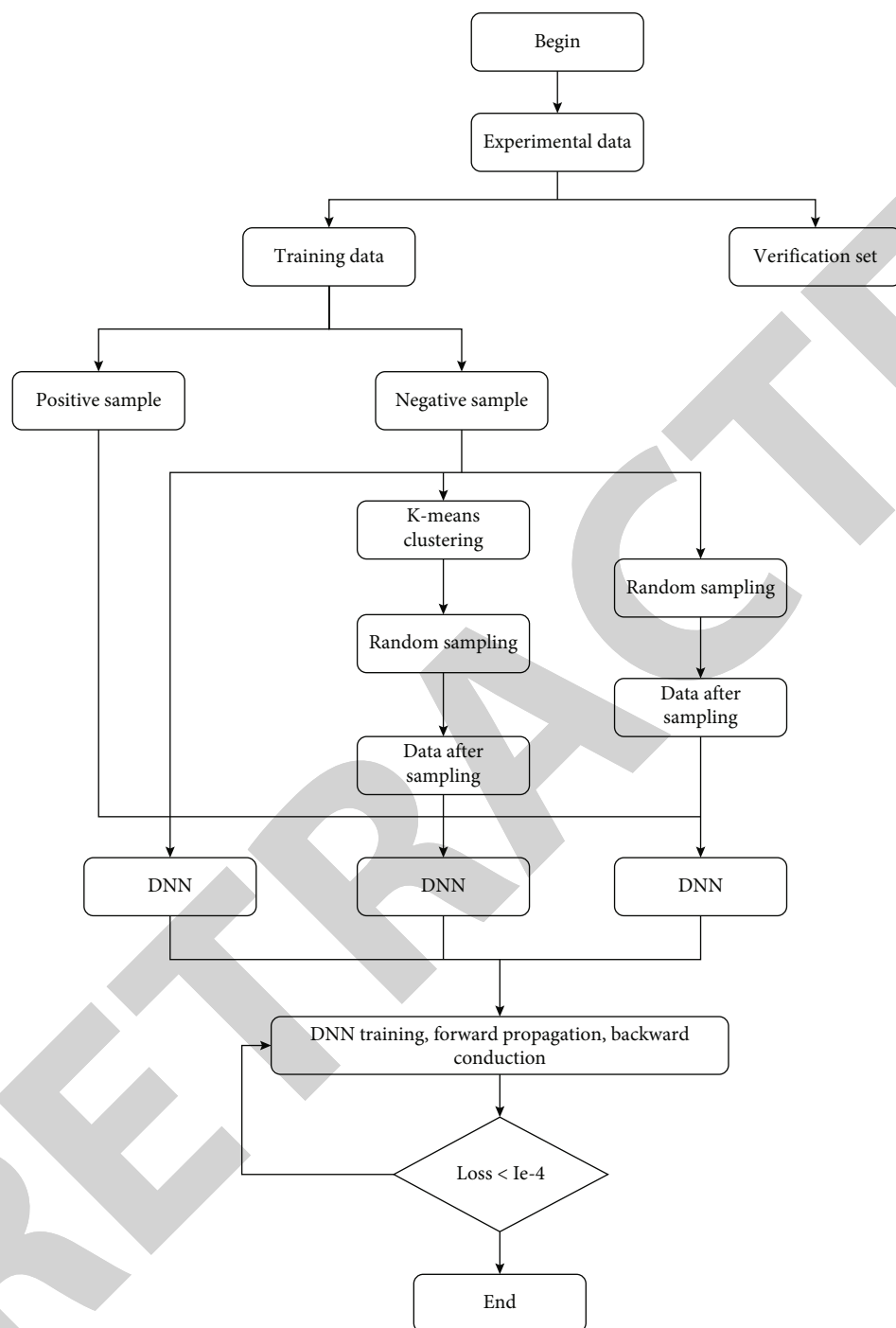


FIGURE 3: Flow chart of deep learning model construction.

person, delaying the diagnosis and treatment time may lead to serious losses and lead to aggravation or even life-threatening. Therefore, solving the problem of data class imbalance is our focus.

4. Result Analysis and Discussion

This paper selects a large retail enterprise as the experimental object, which has two ways of sales department and online sales, and has a complete examination paper investi-

gation mechanism. This paper selects its customers' consumption records and examination papers as sample data for analysis. It includes customers' personal information and historical consumption records, which has two ways: sales department and online sales, and has a complete test paper investigation mechanism. In this paper, the consumption records and examination papers of 50 customers are selected as sample data for analysis. This includes the customer's personal information and 10-month consumption records. Consumer data that is not of numerical type should

TABLE 1: Weights between the input layer unit node and the hidden layer unit node.

Hidden layer	Input layer									
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
H1	-4.1446	6.14784	-0.2854	5.3598	0.5466	-2.7655	3.454	-1.5866	-3.594	-3.4589896
H2	-0.5566	-4.56899	5.5956	3.5966	4.7486	2.45555	1.7578	-1.56456	-2.4988	2.49854
H3	-3.4846	0.498452	3.54152	-0.355	3.4884	3.4885	0.556	0.5456	4.15416	-1.98874
H4	3.4984	1.18545	-1.54485	-2.5845	0.54984	-3.4555	-0.569	2.9684	2.1685	-1.85858
H5	2.8451	-1.84556	-2.54865	-3.4554	-2.478	6.4552	-1.5866	-5.4854	-0.485556	1.9845

TABLE 2: Weights between hidden layer unit nodes and output layer unit nodes.

Output node	Input layer				
	H1	H2	H3	H4	H5
Y1	-9.4561	9.456156	-1.455696	-0.949856	1.94856
Y2	9.45664	-9.41568	0.4884523	1.258956	-1.55656

be converted into numerical type by defining data, that is, weight allocation. Lasagna used in this paper is based on Theano library, and Keras can use either TensorFlow or Theano. These libraries provide more flexible interfaces, which can be used to build neural networks and track the rapid development of deep learning research. All popular deep learning libraries also allow the use of high-performance graphics processing units (GPUs). For example, the conversion of consumer occupation \rightarrow weight: {laid-off or unemployed \rightarrow 0.3; General staff \rightarrow 0.5; Director \rightarrow 0.6; Department Manager \rightarrow 0.8 and so on}. According to the algorithm of the above dynamic neural network, the neural network can construct nine different network structures: $I = 10, J1 = 4, k = 2$; $I = 10, J1 = 5, k = 2$; $I = 10, J1 = 6, k = 2$; $I = 10, J1 = 12 = 4, k = 2$; $I = 10, J1 = 12 = 5, k = 2$; $I = 10, J1 = 12 = 6, k = 2$; $I = 10, J1 = 12 = 13 = 4, k = 2$; $I = 10, J1 = 12 = 13 = 5, k = 2$; and $I = 10, J1 = 12 = 13 = 6, k = 2$. The iterative termination conditions of the algorithm are as follows: the correct classification ratio of training samples is 0.05, and the learning rate is 0.53. After 21,000 iterations, the algorithm satisfies the iterative termination condition. The neural network structure is $I = 10, J1 = 5$, and $k = 2$. Tables 1 and 2 are the input layer unit node to the hidden layer unit node and the hidden layer unit, respectively.

In order to select the optimal individual X_i , let the population size be 10, the probability. According to the fitness, excellent chromosomes are retained. Classify and store purchased computers and non purchased computers according to the output value of the vector. When the output corresponding to the chromosome is to the salt element Y1 or only reaches the maximum function value 1, the information contained in the chromosome is the customer purchase behavior rule that should be extracted. Implicit nodes are the consumption factors such as satisfaction and loyalty. After processing and calculating the data in the above table, the trend chart of customer evaluation index is obtained, as shown in Figure 4.

The algorithm of dynamic neural network is proposed. According to the training results of neural network, the dynamic network structure is constructed. At the same time, this paper uses genetic algorithm to optimize the sample input disk of the neural network, looking for excellent individuals to make the objective function η , achieve the maximum value, then realize the classification of customer transaction sample data, and extract the behavior rules that represent the characteristics of customer consumption. In the data of computer samples purchased by customers given in this paper, the correct classification ratio of training samples is obtained. By synthesizing the weights, iteration times, and extraction rules of nodes in each layer of neural network, it can be seen that the algorithm proposed in this paper has the characteristics of small computation and high accuracy.

The results of the second statistical survey on the development of my country's Internet show that at this stage, the number of Internet users and the number of Internet computers in China have reached 137 million, respectively. The rapid growth of the total number of Chinese netizens has been noticed by the world, but the 137 million netizens only account for 10.5% of China's total population of 1.31 billion, an increase compared to 8.5% in the same period last year. This shows that although the total number of netizens in China is large and growing rapidly, the popularity of the Internet is still very low at present, but the future development space is relatively large, as shown in Figure 5.

Besides neural network can be used for sample data classification, decision tree is also a common method for data classification. ID3 is an algorithm that selects attributes as the split nodes of decision tree based on information gain. However, this algorithm is only effective relatively, resulting in lower accuracy. The results of this survey show that among netizens, those aged 31-35 account for 10.4% and those over 35 and those aged 36-40 account for a relatively low proportion. Netizens aged 35 and below accounted for 82.5%, and netizens over the age of 35 accounted for 17.5%. The age structure of netizens was still younger. From the perspective of the penetration rate, the penetration rate of netizens between the ages of 18 and 24 is the highest, reaching 38.8%, which is 10.2 percentage points higher than that of the previous year. The penetration rate of Internet users between the ages of 25 and 30 ranks second with 25.0% as shown in Figure 6.

Second, the educational background of Internet users can be seen from the educational level distribution map of Internet users. Figure 7 shows that Internet customers

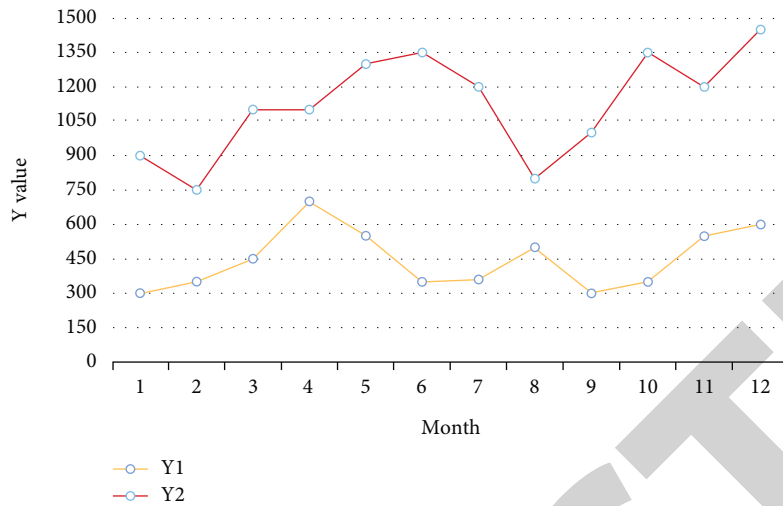


FIGURE 4: Change trend diagram of evaluation index Y.

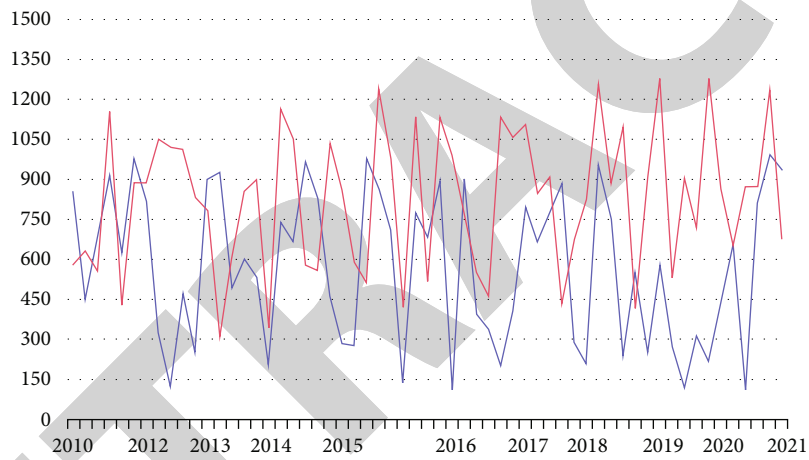


FIGURE 5: The total number of netizens in previous surveys.

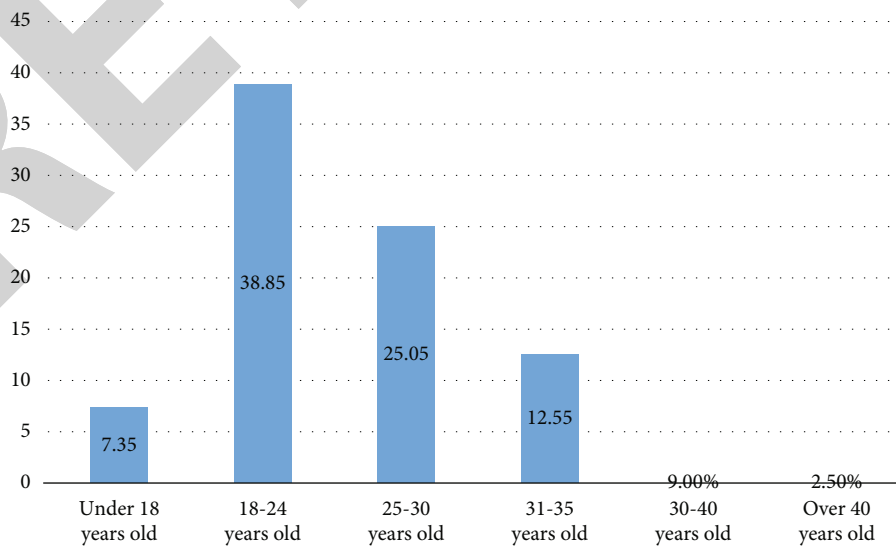


FIGURE 6: Internet users in different age groups.

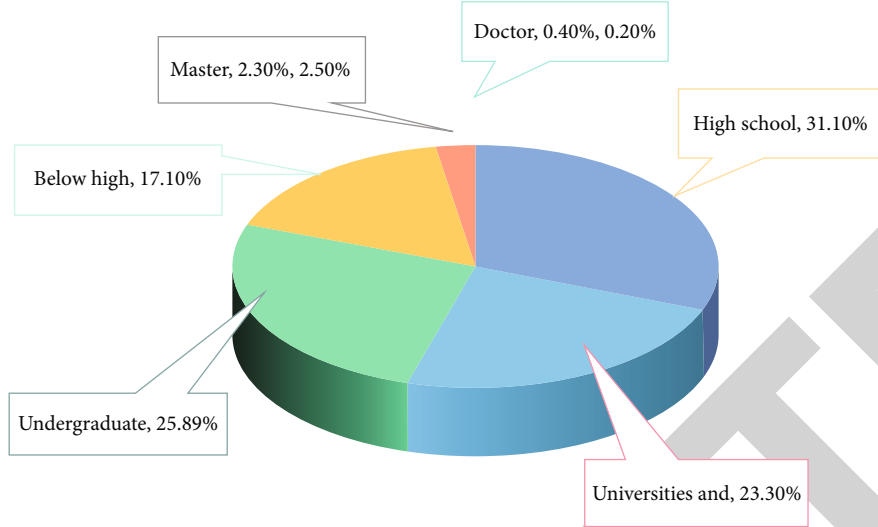


FIGURE 7: Distribution of education level of Internet users.

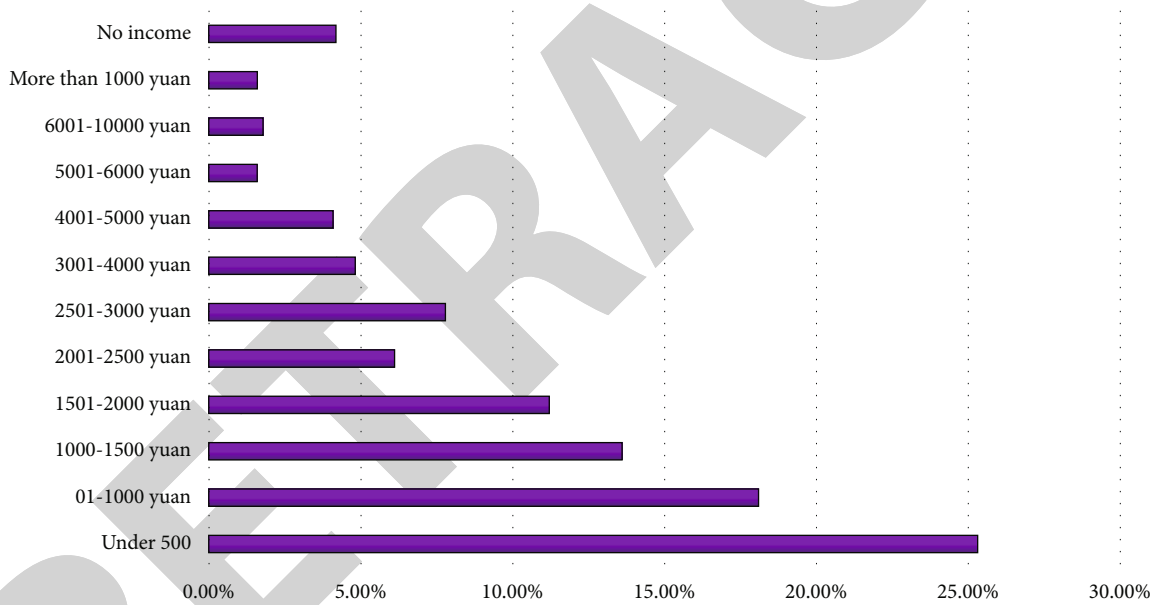


FIGURE 8: Monthly income distribution of Internet users.

generally have higher educational background. Among them, 17.1% have a high school education or below, 31.1% have a high school education, 23.3% have a college education, 25.8% have a bachelor’s degree, and 2.7% have a graduate degree. This shows that the use of the network is related to educational background. With the improvement of educational background, the use of the network is also improving, as shown in Figure 7.

Third, the monthly income of Internet users in this survey shows that the proportion of family Internet users whose monthly income is less than 500 yuan (including no income) is the highest, reaching 29.5%, followed by Internet users whose monthly income is 501~1000 yuan and 1001~2000

yuan (18.1% and 13.6%, respectively), 11.2% of Internet users whose monthly income is 1501~2000 yuan, and the proportion of Internet users whose monthly income is more than 2000 yuan is 27.6%. Low-income Internet users still occupy the majority, as shown in Figure 8.

The whole purchasing process of consumers is related to the products, prices, channels, promotion, and credit of enterprises, and any one of them will make the potential customers decide whether to buy or not. Therefore, it is necessary to closely link the purchasing process of consumers with the marketing strategy of network enterprises, so as to promote the occurrence of potential consumers’ purchasing behavior. Starting from the characteristics of online

consumers, this paper analyzes the purchase decision-making process of online consumers. Get the main factors that affect the purchase behavior of online consumers. Then, according to the research results, formulate effective marketing countermeasures for online enterprises. Reduce the blindness of enterprise network marketing and improve the possibility of success of online enterprises. In contrast, the algorithm proposed in this paper is suitable for the data set with large sample size, less affected by noise, and more suitable for mining the purchase information of customers with large sample size and extracting behavior rules. The monthly (or quarterly) consumption analysis of consumers can not only obtain the consumption level and ability of consumers but also understand the changes of customers in time. For example, if there is a downward trend in the Z value of consumers, you can find out which degree has changed and then take measures; for consumers with high potential, once the consumption level is found to increase, you can provide profitable services and so on.

5. Conclusions

This paper discusses the analysis and prediction of consumer behavior and explores the prediction of commodity recommendation based on consumer behavior. In order to improve the classification accuracy, genetic algorithm is used to optimize the samples, and customer purchase behavior rules are extracted by mining case data. Analyze the behavior characteristics of consumers, accurately identify and capture real online consumers (especially 20% VIP consumers), and carry out targeted online marketing activities according to the purchase behavior characteristics of these consumers. The research has far-reaching significance for online advertising and accurate recommendation, and its results also reflect consumers' consumption habits and consumption behavior rules to a certain extent. However, due to the small number of case samples, larger samples should be used to further verify the effectiveness of the algorithm 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 author declares no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] K. Y. Goh, C. S. Heng, and Z. Lin, "Social media brand community and consumer behavior: quantifying the relative impact of user- and marketer-generated content," *Information Systems Research*, vol. 24, no. 1, pp. 88–107, 2013.
- [2] A. Wulamu, Y. Zhang, Y. Xie et al., "A robust text classifier based on denoising deep neural network in the analysis of big data," *Scientific Programming*, vol. 2017, Article ID 3610378, 10 pages, 2017.
- [3] S. Coskun, L. Ozgur, O. Polat, and A. Gungor, "A model proposal for green supply chain network design based on consumer segmentation," *Journal of Cleaner Production*, vol. 110, pp. 149–157, 2018.
- [4] P. Brzustewicz and A. Singh, "Sustainable consumption in consumer behavior in the time of COVID-19: topic modeling on twitter data using LDA," *Energies*, vol. 14, no. 18, p. 5787, 2021.
- [5] C. Thakur, A. Diwekar, B. J. Reddy, and N. Gajjala, "A study of the online impulse buying behaviour during covid-19 pandemic," *International Journal of Research in Engineering, Science and Management*, vol. 3, no. 9, pp. 86–90, 2020.
- [6] C. Kathiravan, P. Mahalakshmi, and V. Palanisamy, "Online impulse buying behavior of consumer triggered by digital marketing," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2S6, pp. 648–653, 2019.
- [7] J. Chen and S. Lin, "An evaluation for consumer behavior with multi-characteristics based on programming genetic algorithm," *Boletín Técnico/Technical Bulletin*, vol. 55, no. 9, pp. 209–215, 2017.
- [8] Y. Li, M. Zhang, and W. Wang, "Online real-time analysis of data streams based on an incremental high-order deep learning model," *IEEE Access*, vol. 6, pp. 77615–77623, 2018.
- [9] Z. Kalinić, V. Marinković, L. Kalinić, and F. Liébana-Cabanillas, "Neural network modeling of consumer satisfaction in mobile commerce: an empirical analysis," *Expert Systems with Applications*, vol. 175, article 114803, 2021.
- [10] G. Çağil and M. B. Erdem, "An intelligent simulation model of online consumer behavior," *Journal of Intelligent Manufacturing*, vol. 23, no. 4, pp. 1015–1022, 2012.
- [11] S. S. C. Shang, Y. L. Wu, and Y. J. Sie, "Generating consumer resonance for purchase intention on social network sites," *Computers in Human Behavior*, vol. 69, no. APR., pp. 18–28, 2017.
- [12] Q. Zhou, Z. Xu, and N. Y. Yen, "User sentiment analysis based on social network information and its application in consumer reconstruction intention," *Computers in Human Behavior*, vol. 100, pp. 177–183, 2018.
- [13] K. Yu, Y. Liu, L. Qing, B. Wang, and Y. Cheng, "Positive and unlabeled learning for user behavior analysis based on mobile Internet traffic data," *IEEE Access*, vol. 6, pp. 37568–37580, 2018.
- [14] Y. Zheng, L. Shen, and Y. Zhou, "Analysis of students' consumption and learning behavior based on the big data of campus card," *Revista de la Facultad de Ingenieria*, vol. 32, no. 3, pp. 191–200, 2017.
- [15] J. L. Hong, "A study of consumer repurchase behaviors of smartphones using artificial neural network," *Information*, vol. 11, no. 9, p. 400, 2020.
- [16] A. Sharma, Y. K. Dwivedi, V. Arya, and M. Q. Siddiqui, "Does SMS advertising still have relevance to increase consumer purchase intention? A hybrid PLS-SEM-neural network modeling approach," *Computers in Human Behavior*, vol. 124, article 106919, 2021.
- [17] F. H. Sarah, C. L. Goi, F. Chieng, and K. M. R. Taufique, "Examining the influence of atmospheric cues on online impulse buying behavior across product categories: insights from an emerging e-market," *Journal of Internet Commerce*, vol. 20, no. 1, pp. 25–45, 2021.
- [18] H. Xu, K. Z. K. Zhang, and S. J. Zhao, "A dual systems model of online impulse buying," *Industrial Management & Data Systems*, vol. 120, no. 5, pp. 845–861, 2020.

Retraction

Retracted: Optimization of Digital Recommendation Service System for Tourist Attractions Based on Personalized Recommendation Algorithm

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Wang, Z. Qin, J. Tang, and W. Zhang, "Optimization of Digital Recommendation Service System for Tourist Attractions Based on Personalized Recommendation Algorithm," *Journal of Function Spaces*, vol. 2022, Article ID 1191419, 9 pages, 2022.

Research Article

Optimization of Digital Recommendation Service System for Tourist Attractions Based on Personalized Recommendation Algorithm

Yue Wang,¹ Zhaoxiang Qin ,¹ Jun Tang ,² and Wei Zhang³

¹College of Tourism, Inner Mongolia Normal University, Hohhot 010022, China

²College of Science, Inner Mongolia University of Science and Technology, Baotou 014010, China

³College of Government Management, Inner Mongolia Normal University, Hohhot 010022, China

Correspondence should be addressed to Zhaoxiang Qin; qinxz@imnu.edu.cn

Received 15 June 2022; Revised 17 August 2022; Accepted 22 August 2022; Published 12 September 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Yue 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.

With the deepening of tourists' demand for tourism services, the personalization of online tourists has gradually become an application of personalized recommendation technology. According to the application requirements of personalized scenic spot recommendation, this paper uses social networks and Bayesian networks to fully mine the matching degree between users and scenic spots for personalized recommendation. Add social network factors to the recommendation of tourist attractions, and fully tap the social network relationship between users. First, the users are clustered by the coupling bidirectional clustering algorithm. Then, DBSCAN (density-based noise application spatial clustering) algorithm is used to cluster scenic spots. Finally, two stable user sets and scenic spot sets are applied to the personalized recommendation algorithm to predict the user's next upcoming scenic spot. The algorithm is compared with some traditional algorithms in the dataset. The algorithm deals with the similarity of user attributes and user behavior and uses content-based algorithm to deal with the relationship between scenic spots; com datasets have better performance.

1. Introduction

With the development of the national economy, people's demand for tourism is becoming more and more strong. The rise of various tourism websites provides users with a platform to understand the destination scenic spots and arrange their itineraries. With the rapid popularity of smartphones and growing online tourism experience sharing platforms, tourists are increasingly inclined to decide what to visit in real time, rather than making detailed travel plans in advance. In the face of the uncertainty of user needs, how to combine the current actual state of users to make dynamic and personalized scenic spot recommendation for users has become an urgent problem to be solved in the tourism recommendation research. However, with the increasing amount of data, the problem of information overload on travel websites also becomes more and more serious. Therefore, providing a personalized travel recommendation

system is very critical for travel websites and users. The Internet continues to penetrate into people's lives; coupled with the smart devices and wireless communication technology, people can access the Internet more conveniently and quickly. As shown in Figure 1, according to the survey [1], my country increases every year according to a certain proportion [2]. The network has become increasingly diversified. Categories are usually applied to website design, such as the hao123 website, which helps users to quickly find the website they need [3]. The emergence of search engines makes it convenient for users to find the required information by searching for keywords, but sometimes users cannot accurately describe their needs, making it difficult for the system to distinguish the needs of users, and the recommendation system was born. From an objective understanding, search engines and recommendation systems have their own strengths in helping users filter information, which greatly facilitates users [4]. Search engine allows users to find

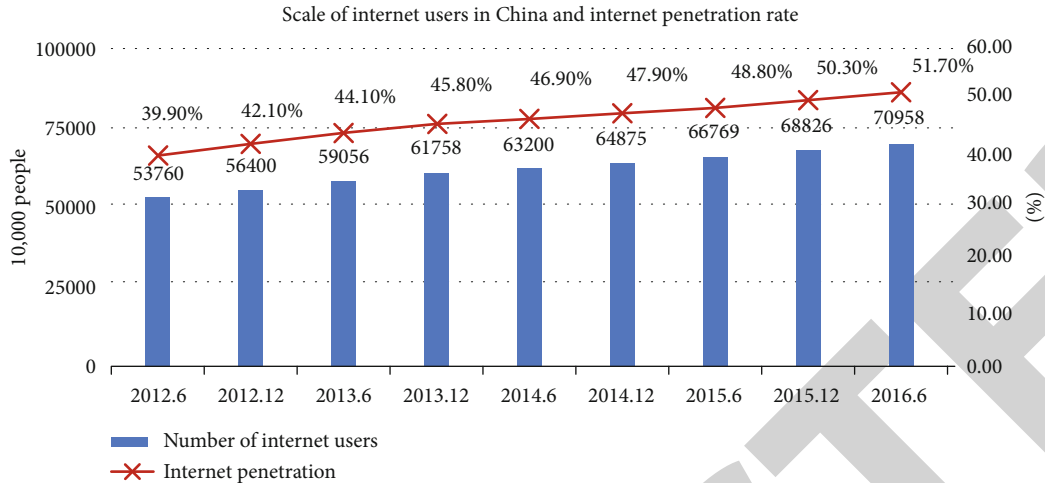


FIGURE 1: Number of the Internet users and Internet penetration rate.

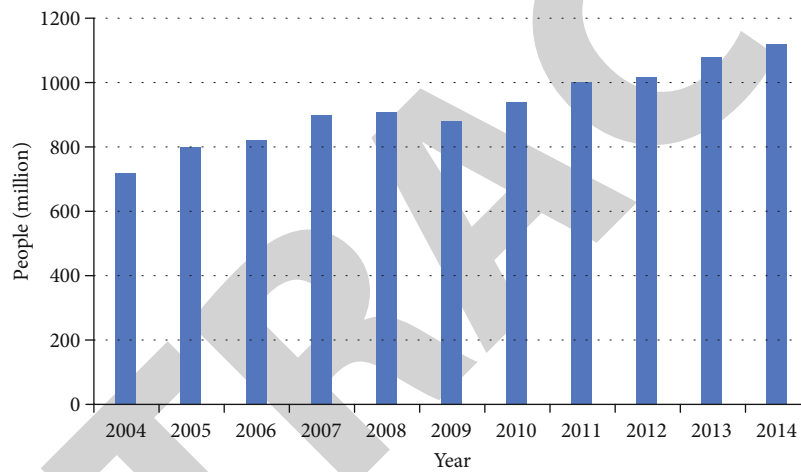


FIGURE 2: The number of international tourists from 2004 to 2014.

the information they need by searching keywords. However, search engines need users to actively provide accurate keywords to find information. Therefore, it cannot solve many other needs of users. For example, when users cannot find keywords that accurately describe their needs, search engines cannot do anything about it. Like search engines, recommendation system is also a tool to help users quickly find useful information. Unlike search engines, recommendation systems do not require users to provide clear requirements but model users' interests by analyzing users' historical behavior, so as to actively recommend information to users that can meet their interests and needs. Therefore, in a sense, recommendation system and search engine are two complementary tools for users. The search engine meets the active search needs of users when they have a clear purpose.

The task of recommendation system is to contact users and information. On the one hand, it helps users find valuable information for themselves; on the other hand, it allows information to be displayed in front of users who are interested in it, so as to achieve a win-win situation between information consumers and information producers. A rec-

ommender system is a software designed to help users quickly discover useful information, in which the user's historical behavior provides the necessary information for interest modeling, which is convenient for predicting the user's preferences so that the information that meets the requirements can be actively recommended to the user [5]. In recent years, due to the characteristics of recommendation systems that facilitate users to filter information, the application scope of recommendation systems has been expanding, from product recommendation, to news recommendation, music recommendation, etc. [6]. As shown in Figure 2, the analysis data of the tourism industry in 2015 shows that the number of tourists is increasing every year, which promotes this date to become a new application hotspot of the recommendation system. There are relatively complete websites recommended for tourist attractions in China, such as Tongcheng (<http://Travel.com/> and <http://Ctrip.com/>), which bring convenience to users' travel [7].

By analyzing the tourist attraction recommendation services of major websites at home and abroad, most of the recommendation methods focus on packaged services, lacking

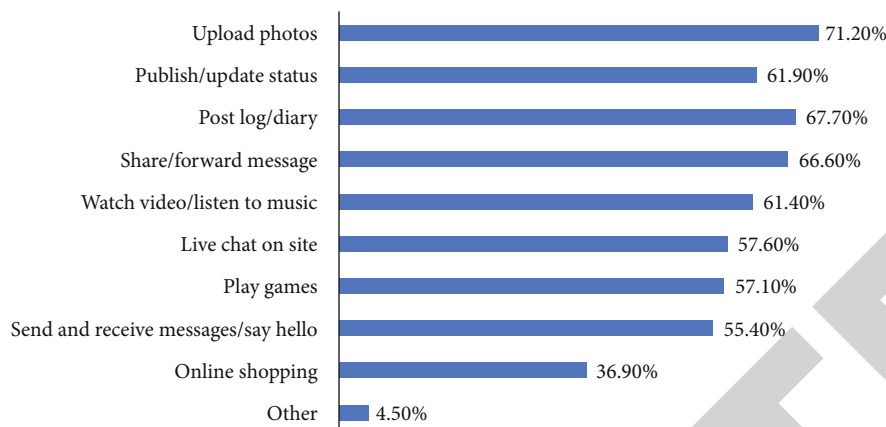


FIGURE 3: Netizens use the functions of social networking sites.

personalized, social, and real-time services [8]. According to a survey by the China Internet Network Information Center, a relatively high proportion of citizens use social networking sites to upload photos, post logs, and upload videos, as shown in Figure 3 [9].

2. Related Work

At present, should meet the preferences of users in many aspects, the development of personalized tourist attractions recommendation has become the general trend.

The recommended tourist attractions are more in line with the personalized needs of users [10]. Starting from the development history of foreign tourist attractions, Shaik et al. [11] made an intuitive in-depth analysis of the domestic and foreign tourism industry structure. Singh et al. [12] proposed the CityVoyager store recommendation system, which is a relatively novel and practical recommendation system, which mainly evaluates the user's preferences, thus laying the foundation for the recommendation of tourist attractions. Masaki [13] proposed to use location as a key factor to design an enhanced collaborative filtering system that can recommend personalized restaurant services to users, verifying the role of location-based services. Motia and Reddy [14] proposed the HGSM framework, which provides information for calculating the similarity of the user's geographic space by mining information from the user's geographic location history and the hierarchical attributes of the geographic space, and explores the relationship between users and users of intimacy and the association between users and geographic locations to address their interests and hobbies.

When tourists are in unfamiliar tourist locations, Ayswarya et al. [15] and others used Bayesian network and AHP to provide personalized recommendation services. First, they used the information related to tourist attractions on travel websites to build a tourist ontology tree to store tourist attractions. Shimizu et al. [16] summarized some systems and technologies related to travel recommendation before 2009 and proposed a recommendation system in the

mobile environment. Pradhan et al. [17] extracted the rating data of tourist attractions on tourism websites in order to construct a user-attraction matrix and then make final recommendations for tourist attractions.

3. Methods

3.1. Personalized Attractions Recommendation Algorithm Based on Social Networks. A complete tourism recommendation system should provide a complete set of personalized tourism recommendation programs, which is divided into two steps: scenic spot recommendation and path planning. For scenic spot recommendation, existing recommendation algorithms such as collaborative filtering and content-based and social network-based methods such as single data source are difficult to describe users' dynamic interests. The existing research generally regards it as TSP problem and NP hard problem, so most of them adopt heuristic algorithm. However, in tourism planning, only the length of the turnover path between scenic spots, the user time limit, and the opening time limit are usually considered. This chapter presents a social network (based on the social network with personalized tourist attraction recommendation, ptar). The purpose is to predict the next possible scenic spots for users and give real-time recommendations, which is different from the traditional recommendation algorithm. The algorithm processing steps are to first cluster users and attractions and replace the user and attractions matrix in the personalized rule-based algorithm proposed by Ma et al. [18] with the results obtained by clustering. Tourism planning is an applied science. At present, there is no perfect theoretical system of tourism planning. International discussions on tourism planning theory are often scattered. The main problems are as follows: (1) the theories and norms of urban planning and regional planning have replaced tourism planning. The standardization of scenic spot planning has replaced the tourism area planning, and there is little research on the theory and method of tourism planning itself. (2) There is little systematic thinking about the theory of tourism planning, and there is no systematic research on the ideological

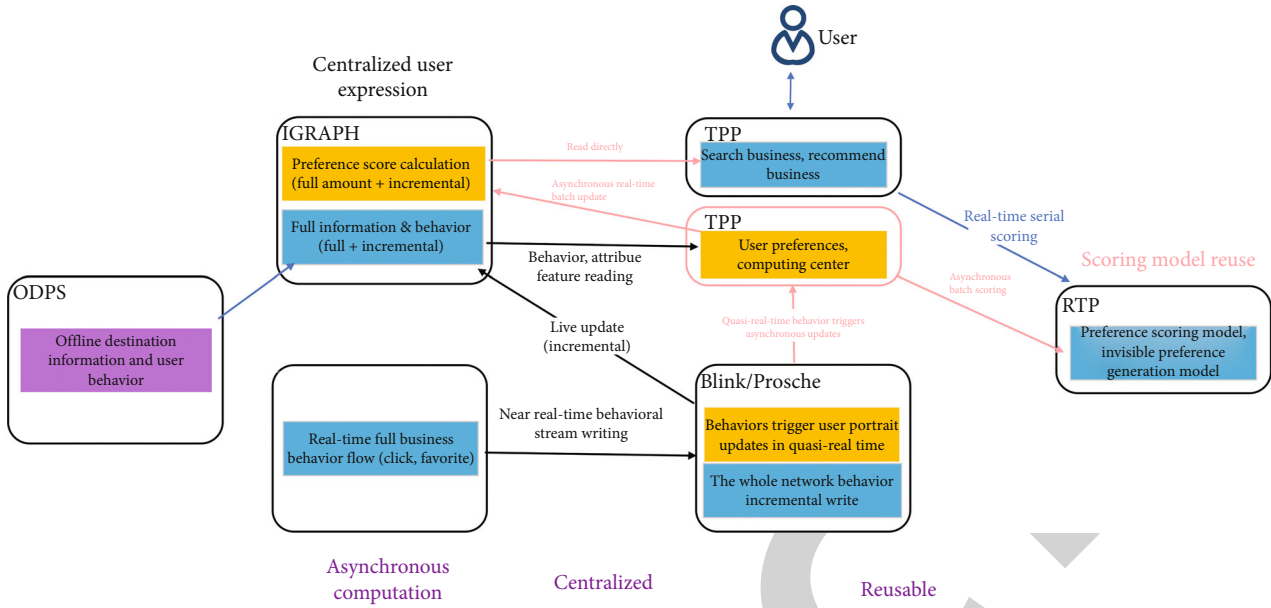


FIGURE 4: Recommendation algorithm for personalized tourist attractions based on social network.

methods and models of tourism planning. (3) The concept system of tourism planning is not standardized, and the understanding of tourism and its related concepts is even far from satisfactory.

For the above problems, this paper proposes a personalized travel recommendation system framework based on knowledge graph and user footprint. First, establish a knowledge map of the attractions, The knowledge map includes scenic introduction, comments, time, and other information, It provides data support for scenic spot recommendation and path planning; next, personalized scenic attraction recommendations through a deep interest evolution network. The model uses recurrent neural network and adds attention mechanism to describe the diversity and dynamic change of user interests. The model input mainly includes the scenic spot sequence in the user's footprint and the scenic spots to be recommended. Input information of scenic spots can be obtained through knowledge map, including scenic spot ID, scenic spot tag, comment tag, scenic spot introduction, and nearby scenic spots. Finally, the user's own travel days, the opening time of the scenic spot, the travel time of the scenic spot, and other factors are added to the fitness function. Appropriate heuristic algorithm is adopted for reasonable path planning.

As shown in Figure 4, the algorithm based on user data and social data is described as follows: First, users are processed, and users in certain locations are clustered together according to their preferences and geographic location information. Use the coupled bidirectional clustering algorithm to cluster users into a suitable category. Secondly, the scenic spots are processed. The frame processing is shown in Figure 5.

From previous research work, it can be concluded that clustering users is generally used to improve the recommendation accuracy by obtaining "appropriate" friend groups. The coupled two-way clustering algorithm (CTWC) is used

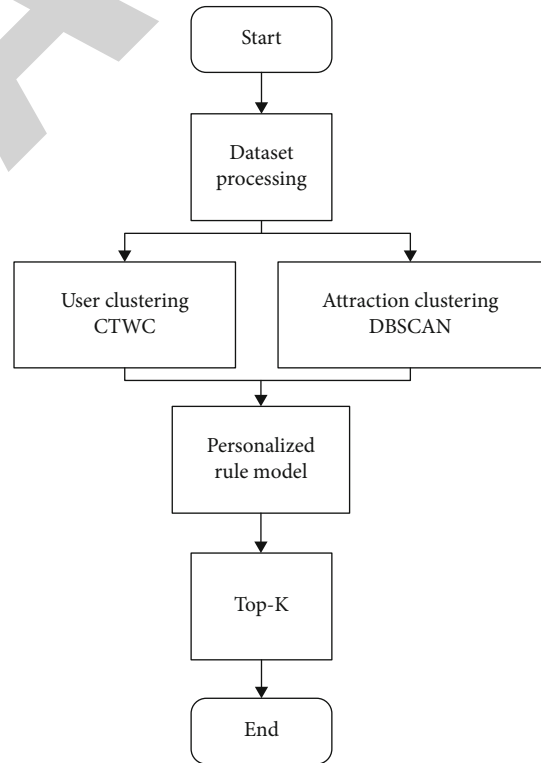


FIGURE 5: Algorithm processing frame diagram.

to cluster the dataset, which has been verified in the literature [19].

The scenic spot as P_c ; then, the scenic spot category is replaced. The quality of the clustering results is related to the dataset selected during the experiment, especially the prefiltering requirements for the scenic spots in the dataset.

TABLE 1: Clustering results of photo attractions with geographic information.

Clustering parameters		Attractions				Clustering results	
Eps (km)	MinPts	Great Wall	Tiananmen Square	Summer palace	Water cube	Number of categories	Noise rate (%)
0.5	10	300	350	2300	240	55	63.70
	20	150	312	2100	213	222	71.88
	30	117	286	1256	186	14	76.30
1	10	570	560	2765	350	70	46.29
	20	400	520	2460	321	23	58.60
	30	260	498	1486	246	15	63.02
3	10	724	700	6203	420	50	22.32
	20	680	668	5079	385	29	30.94
	30	648	623	4862	343	18	37.06

Therefore, the tourist attractions selected in this paper must have more than 3 users who have traveled to the scenic spots to avoid data collection.

The introduction of DBSCAN clustering algorithm is shown in Section 2.5. The date of scenic spot are shown in Table 1.

This paper introduces the friend relationship in the social network into the tourist research. A social network-based (PTAR) is proposed, which fully considers friends in social network relationships and friends with similar interests. According to the personalized rule algorithm, this paper defines the following objective function:

$$\begin{aligned} \min_{U,P} L(T, U^c, P^c) = & \frac{1}{2} \sum_{i=1}^m \sum_{k=1}^n I_{ik} (T_{ik} - U_i^c P_k^c) \\ & + \frac{\beta}{2} \sum_{i=1}^m \sum_{f \in F(i)} s(k, f) U_i^c - U_f^c F^2 \quad (1) \\ & + \frac{\lambda_1}{2} U^c F^2 + \frac{\lambda_2}{2} \|P^c\|_F^2. \end{aligned}$$

T represents the user-spots matrix, U_i^c represents the user i 's scenic spot set, and P_k^c represents the scenic spot k set. I_{ik} is an indicator function; if the user U_i^c gives the attraction P_k^c action, then $I_{ik} = 1$; otherwise, $I_{ik} = 0$. T_{ik} represents the number of photos, the user U_i^c gives to the scenic spot P_k^c , U_i^c and U_f^c represent the user vector, P_k^c represents the scenic spot vector, and $\|\cdot\|_F^2$ represents the Frobenius norm. α and β represent the added social network information, and $\alpha > 0$, and $\beta > 0$. $L(k, f)$ represents degree of correlation between the scenic spot P_k^c and the user U_f^c ; if the user acts on the scenic spot, it indicates that the correlation between the two is relatively high, and the user U_f^c has a high degree of correlation to the user U_i^c . The contribution is relatively small; $s(i, f)$ represents the relationship between the user U_i^c , $s(i, f)$ is a greater difference between users, that is U_i^c and U_f^c , the greater the difference between

users. The last two terms are a regularization term, in order to avoid overfitting of the objective function.

3.2. Experiment Evaluation Index of Evaluation Recommendation Algorithm. In this paper, quantitative description is used, and the prediction accuracy is selected to measure the ability of a recommendation algorithm. Different research directions use different evaluation indicators. The accuracy of the general forecast is represented as calculated in the following formulas:

$$MAE = \frac{1}{T} \sum_{u,i \in T} |R_{ui} - \bar{R}_{ul}|, \quad (2)$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{u,i \in T} (R_{ui} - \bar{R}_{ul})^2}. \quad (3)$$

In Formulas (2) and (3), R_{ui} is real rating u , R_{ui} is the item i obtained i through the recommendation algorithm, and T is test sets.

The prediction accuracy of Top-N recommendations is generally measured by precision and recall; let (u) represent the list of recommended actions on the training set, and $T(u)$ represent the list of actions on the test set. The formulas for calculating the recall rate (*Recall*) and the precision rate (*Precision*) of the recommended results are as follows:

$$\begin{aligned} \text{Recall} &= \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|}, \\ \text{Precision} &= \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|}. \end{aligned} \quad (4)$$

In the experiment, R represents the scenic spots actually visited by the user, and the recommendation result generated by the personalized scenic spot recommendation algorithm represented by T . It is recommended to use the TopK method

for the final result, that is, the length of the recommended result is a fixed value K , so for a given K value, the calculation formulas of precision and recall are as follows:

$$\text{Precision@K} = \text{Recall@K} = \frac{|R \cap T|}{K}. \quad (5)$$

3.3. Personalized Attraction Recommendation Algorithm Based on Bayesian Network Learning. It is a graph whose operations involve knowledge related to probability [20], often using probability as a way to predict the relationship between two. It is a typical Bayesian network diagram. From this, it can be seen that if the probability of X_3 is calculated, it needs to rely on X_1 and X_2 , which involves probability problems. The following is a brief introduction to Bayesian probability.

The Bayesian formula is a conditional probability formula, such as $P(A|B)$, which represents the probability of A when the condition of B occurs. The calculation formula is shown in Formula (6), where $P(A, B)$ represents the probability that A and B occur at the same time; the calculation formulas of $P(A|B)$ and $P(B|A)$ can be obtained as follows:

$$P(A|B) = \frac{P(A, B)}{P(B)}, \quad (6)$$

$$P(B|A) = \frac{P(A, B)}{P(A)}. \quad (7)$$

The formula for calculating the Bayesian posterior probability after processing is as follows:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}. \quad (8)$$

Bayes' theorem, often used in computing the likelihood of conditional probabilities, quantifies the relationship between the two. For example, let X be a data primitive with n attribute values (x_1, x_2, \dots, x_n); for the classification problem, Bayes' theorem provides a method consisting of $P(H)$, $P(X)$, and $P(X|H)$; similar to Equation (9), the formula is

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}. \quad (9)$$

Bayesian is each feature in a special set. And from the knowledge of probability, when A and B are independent of each other, $P(AB) = P(A)P(B)$. The formula for calculating is

$$P(X|H) = \prod_{k=1}^n P(x_k|H) = P(x_1|H)P(x_2|H) \cdots P(x_n|H). \quad (10)$$

User preference model is a very challenging problem in information system. At present, it mainly deals with the automatic discovery of user preferences and uses the model. As personalization and recommendation services become increasingly popular on the Internet and e-commerce, it is becoming more and more important to understand user preferences. Intelligent information system can analyze what users need and predict the products that users will choose in the future. On the basis of users' different preferences, the intelligent system can recommend products of interest to each user and provide personalized services. In the stage of predicting user preference, it is necessary for liking item, and probability the user liking the item can be presented in a probabilistic manner. Finally, users are recommended according to the probability. The specific process of inferring scenic spots for users is as follows.

First, define user sets and sights sets, and define users as $U = \{u_1, u_2, \dots, u_n\}$, representing n users, where $u_i (i \in n)$ represents user feature vectors (basic user attributes), including user statistics information (user label (sex) and gender ($uSex$)), age ($uAge$)), position ($uLocation$), and sign-in time ($uTime$), that is, $\langle uId, uAge, uSex, uLocation, uTime \rangle$ indicate a user feature vector. Then, by calculating the probability that the scenic spot is recommended to the user, this probability is called the recommendation degree. Finally, according to this probability, it is determined whether the scenic spot is worth recommending to the user. The recommendation degree is obtained by improving Formula (11), that is, the calculation formula of the recommendation degree recommended by the scenic spot to the user is as follows:

$$P(a_i|u_j) = \frac{P(a_i)P(u_j|a_i)}{P(u_j)}. \quad (11)$$

Definition the similarity between attributes of user demographic, namely $Sim(u, k)$, and its is:

$$Sim_f(u, k) = \frac{N_{f_u} \cap N_{f_k}}{N_{f_u} \cup N_{f_k}}. \quad (12)$$

$$Sim_b(u, k) = \frac{\sum_{a \in C_{u,k}} N_{s_{u_a}} N_{s_{k_a}}}{\sqrt{\sum_{a \in C_u} N_{s_{u_a}}^2} \sqrt{\sum_{a \in C_k} N_{s_{k_a}}^2}}. \quad (13)$$

Definition $Si(u, k)$ is as follows:

$$Sim(u, k) = bSim_b(u, k) + (1 - b)Sim_f(u, k). \quad (14)$$

4. Simulation and Results

4.1. Empirical Analysis of Recommendation Algorithms. The effect of these parameters on the experiment is discussed below.

- (1) The parameters α and β affect the calculation under different sparsity. Of course, training sets with different degrees of sparsity have different effects on α and β on

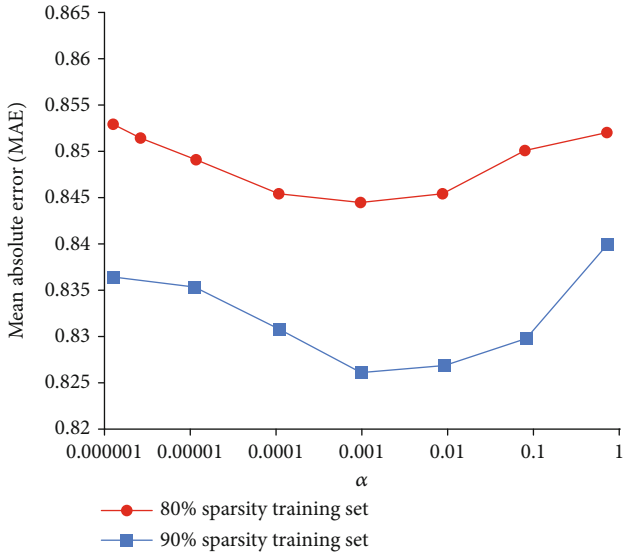


FIGURE 6: The effect of α on the mean absolute error under different sparsity.

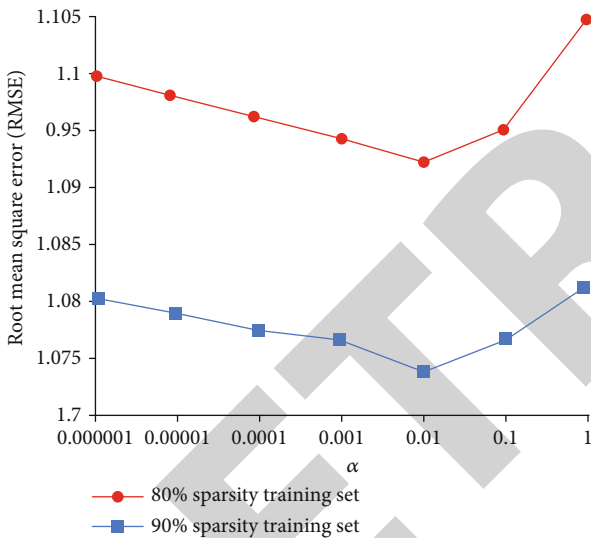


FIGURE 7: The effect of α on the root mean square error under different sparsity.

MAE and RMSE. Since α generally has the same trend of error as β , this paper only considers the parameter α alone (follow in Figures 6 and Figure 7)

From Figures 6 and 7, at the beginning, with the increase of α , MAE and RMSE have been decreasing (accuracy increases), but when α increases to a certain extent, MAE and RMSE increase again (accuracy decreases).

The initial step of the framework is users. The main step is to use the friend, and tag information regularization term finally merges the user's friend relationship and tag information to obtain the final recommendation. For α and β in this method is friend and label information should be included in the proposed method. Ultimately, one should find a suitable value to balance. In the experiment, the trend change of

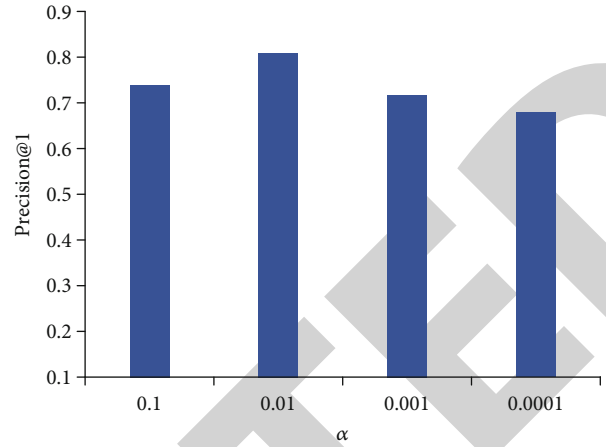


FIGURE 8: Influence of parameter α on Precision@1.

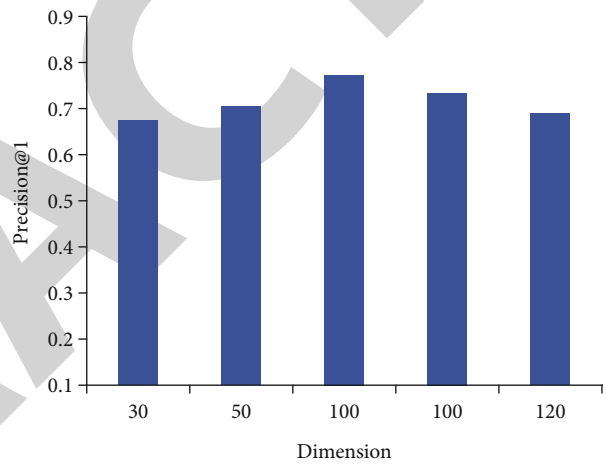


FIGURE 9: Influence of dimension of feature space on Precision@1.

the effect only affects the prediction accuracy of the experiment. Probability is an important index reflecting risk and uncertainty. Probability estimation has trend effect and will affect decision-making. The trend of the change of probability estimation and the trend of unilateral probability estimation reveal the influence of the trend effect of probability estimation on individual judgment, decision-making behavior, and irrational decision-making bias. Set the dimension of the space to 30, and the meaning of Precision@1 is accuracy of the recommended number of 1. The result is shown in Figure 8.

From Figures 8 and 9, as α increases, the accuracy gradually increases, and the recommendation accuracy also increases, but when the parameter α increases to 0.01, the accuracy decreases. Therefore, the value of α is finally set to 0.01 to predict the accuracy of the recommendation results.

In a dimensionality, in this experiment, the values of the parameters α and β were set to 0.01 according to the above experiments. Select the dimension = {30, 50, 80, 100, 120}, respectively, to verify the effectiveness of the method. In this way, in order for the algorithm to converge quickly, the learning rate is reduced by 10%. To verify the effectiveness

TABLE 2: Recommendation results for different contexts under the Gowalla dataset.

	Precision@10	Recall@10	NDCG@10	MSP@10
BGMF	0.07727	0.07928	0.07604	0.03736
MANC-U	0.08372	0.11279	0.07168	0.02538
MANC_P	0.08245	0.11357	0.01704	0.06144
MANC	0.08422	0.12135	0.07231	0.06373

TABLE 3: Recommendation results of different contexts under Yelp dataset.

	Precision@10	Recall@10	NDCG@10	MSP@10
BGMF	0.04252	0.06421	0.03306	0.03348
MANC-U	0.05375	0.07383	0.02271	0.03301
MANC_P	0.05363	0.07321	0.03474	0.03487
MANC	0.05459	0.07518	0.02302	0.03401

of the contextual attention module, we compare the impact of each contextual module on the performance of the model separately. As shown in Tables 2 and 3, the BGMF model is a model without user context and POI context, MANC-U is a model with a user context module on the BGMF model, and MANC-U is based on BGMF with interest point context model. It can be observed from table: (1) On both datasets, MANC outperforms MANC-P and MANC-U, mainly because MANC and POI neighborhood context. (2) MANC-U and MANC-P are higher than and BGMF in all four evaluation indicators. For example, in ranking top-10, MANC-U achieves 20.8% accuracy improvement over BGMF on Gowalla dataset, while MANC-P achieves 19.2% accuracy improvement over BGMF on Gowalla dataset.

Through the analysis of the parameters in the objective function, it is known that when α and β to 0.01, the efficiency τ is set to 0.5, the dimension is set to 80, and γ is set to 0.8; the objective function in the algorithm can reach the optimal solution. For rationality of each parameter setting in PTAR, we compares the performance of PTAR with the traditional collaborative filtering scenic spot recommendation algorithm, the collaborative filtering method based on the appropriate time of the scenic spot, and the tag-based scenic spot recommendation algorithm. The number of recommended accuracy rates is set to 1, 3, and 5 (Precision@1, Precision@3, and Precision@5), and Recall@5 represents the recall rate when the user recommends 5 scenic spots. The experimental results are shown in Figure 10. U-cf stands for user-based collaborative filtering recommendation algorithm. Ptlr (personalized travel location recommendation) represents a collaborative filtering method based on the appropriate time of the scenic spot, and a label based stands for algorithm. Among them, the user-based filtering recommendation algorithm (U-CF) calculates the user's preference for scenic spots. The formula is as follows:

$$p(u, i) = \sum_{v \in N_u} s(u, v) f(v, i). \quad (15)$$

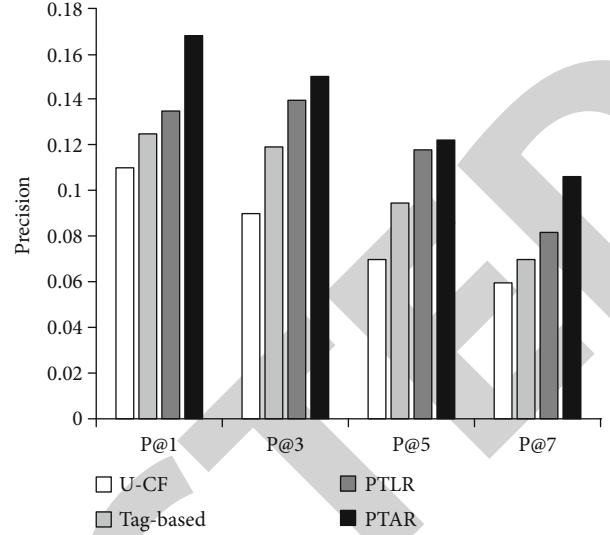


FIGURE 10: Comparison of algorithm accuracy.

In Figure 10, the horizontal axis P@1, P@3, and P@5 represent Precision@1, Precision@3, and Precision@5. This recommended accuracy. It can be clearly seen that the performance of the PTAR algorithm is obvious. It performs better than previous methods, and the recommendation accuracy rate is higher. From Precision@1, Precision@3, and Precision@5, it can be found that the recommendation accuracy rate of the tourist attraction about social network is very different because the number of recommendations. And the smaller the parts of recommendations, the higher the recommendation accuracy rate. So, setting a reasonable number of recommendations is crucial for algorithm research in future work.

5. Conclusion

The focus of this paper is on the research algorithm of tourist attractions. In order to meet the needs of tourists' personalized attractions to the greatest extent, this paper has done the following main research work:

- (1) Propose an algorithm based on social network: Firstly analyze the challenges faced by the recommendation method of it based on location social network; then use the algorithm for users and attractions, respectively, and use the coupled bidirectional clustering algorithm for users. The DBSCAN algorithm is used for scenic spots; finally, it is applied to the personalized rule algorithm to predict scenic spots and recommend tourist attractions to users. We need to verify the accuracy of the recommendation, and this algorithm is compared with some traditional algorithms, so this experimental results tell that this algorithm has a higher recommendation accuracy
- (2) We have a better prediction for tourist attractions that they are most likely to visit next, for the tourist attractions algorithm by proposed by this paper; this

Retraction

Retracted: Construction of Mathematical Model of Logistics Delivering Based on Intelligent Mobilization

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Liang and J. Guo, "Construction of Mathematical Model of Logistics Delivering Based on Intelligent Mobilization," *Journal of Function Spaces*, vol. 2022, Article ID 7386227, 11 pages, 2022.

Research Article

Construction of Mathematical Model of Logistics Delivering Based on Intelligent Mobilization

Haijun Liang  and Jing Guo 

Hebei Vocational University of Industry and Technology, Shijiazhuang 050081, China

Correspondence should be addressed to Jing Guo; guojing@hbcit.edu.cn

Received 27 June 2022; Revised 6 August 2022; Accepted 26 August 2022; Published 12 September 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Haijun Liang and Jing 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.

Distribution process is the core of logistics enterprise system. Efficient distribution process is the key for enterprises to improve logistics service level, gain competitive advantage, and win customers. In order to reduce the design error of the system and ensure the effective operation of the system, this paper establishes the system model and analyzes the properties of the model through the evaluation and analysis of various resources of the distribution system. The future development of intelligent logistics system is not only the key for logistics enterprises to win competition, but also a new measure to promote China's economic development. This is not only a typical NP hard problem, but also a major challenge for the rational and scientific development of the intelligent logistics industry. Based on the previous theoretical research results, this paper intends to explore the intelligent logistics distribution route selection scheme by using mathematical modeling methods such as particle swarm optimization, so as to provide new technologies and methods for logistics distribution operation and management decision-making. The results show that compared with the traditional genetic algorithm, the accuracy of the improved particle swarm optimization algorithm is improved by 10.36%. This method effectively improves the operation cycle and link efficiency and achieves the purpose of optimization. The improved particle swarm optimization algorithm proposed in this paper is not only more suitable and effective for enterprise decision makers to deal with subjective judgments in an imprecise environment. It is also based on the evaluation sequence of each indicator of the alternative obtained from the evaluation and the comprehensive evaluation of all indicators. By weighting and considering the results, the best solution for enterprise logistics and distribution is selected.

1. Introduction

As one of the basic industries and development arteries for the development of the national economy, it has received extensive attention from all walks of life. With the increasing application of automation technology and information technology in manufacturing and other entity enterprises, the competitiveness of manufacturing technology to improve products has become less and less, so enterprises focus on improving competitiveness in logistics, focusing on transportation. The functions of storage, loading and unloading, packaging, distribution, and information processing are organically combined, and it is hoped that professional and perfect supply chain services can enhance its core competitiveness and industrial-added value [1]. How to integrate the logistics delivering

system and improve the efficiency of logistics delivering has become a key problem to be considered in the study of logistics theory and the actual operation of logistics enterprises. The development of modern logistics information technology and control technology has created conditions for solving this problem. The integration of logistics system is to provide efficient and high-quality comprehensive services for logistics system. Coordinate and reorganize internal elements at different levels in physical and soft environments. Logistics integration service can not only generally reduce the operation cost of logistics system, but also provide efficient and high-quality comprehensive services to meet various flexible logistics needs. Logically speaking, logistics system integration services will create a broader range of large enterprises. In this emerging integrated service field, there are traditional operators, such

as express delivery, collect freight, and road transportation companies. There are also new entrants, such as state-owned railways, private railways, and postal companies. This integrated system provides unlimited development space for all these participants.

Combined with mathematical models, this paper realizes the optimization of logistics intelligent distribution path and realizes the modeling of the impact of intelligent distribution mode selection technology on logistics economic development. Finally, empirical analysis is carried out to draw validity conclusions. With the continuous development of the logistics industry, more and more attention has been paid to the research in this area. Vehicle logistics is a very important part of the logistics field. It is different from other logistics categories and has extremely high comprehensiveness and complexity. Jathe et al. propose a positioning technology based on WiFi round trip time (RTT), which will help to obtain robust positioning results. The proposed positioning technology is based on geometric method, using trilateral measurement and probability method based on the possibility of vehicle position, which strengthens the logistics planning of the whole vehicle [2]. At present, most of the research to solve this problem focuses on the following aspects: first, analyze the possible related factors and quantify these factors. When the general solution method is no longer applicable, a new method needs to be used for research. In order to achieve the goal of the enterprise, it is necessary to study the optimization problem of the enterprise logistics delivering network system. This paper makes full use of the idea of mathematical modeling to transform the complex practical problems such as vehicle stowage, commodity transportation, and transportation capacity constraints in the process of vehicle logistics operation. It has become a mathematical language, and a scientific and reasonable mathematical model has been constructed. Each model is interlocked, and the output data of each module will be directly used as the input data of the next module, which increases the operability of the model and basically realizes the intelligent logistics and distribution of the enterprise's own status quo. The location routing problem (LRP) is one of the most concerned problems in the enterprise logistics delivering system [3]. It is now widely accepted that the success of many businesses depends primarily on location-distribution network decisions. Because of the complexity of this problem and the integrity of the logistics delivering system [4], managers of many enterprises tend to make decisions based on the experience gained. The research goal of this paper is to consider the basic loading constraints of car carriers on the basis of batch orders and take the minimum cost and the minimum number of cars as the objective function to realize the optimal route and intelligent service vehicle distribution problem.

The innovative contribution of this paper is to analyze the location of logistics distribution center combined with machine learning method. The path planning problem of multiple distribution centers is studied. The structural framework of enterprise logistics system network considers the modeling of nodes and processes. The mathematical model of distribution center location path planning is established, and a new particle swarm optimization algorithm is proposed. It is applied to the optimization of distribution center location and road. Compared with

genetic algorithm, the performance of particle swarm optimization algorithm is better than genetic algorithm. The improved particle swarm optimization algorithm proposed in this paper is not only more suitable and effective to deal with the subjective judgments of enterprise decision makers in an imprecise environment, but it is also based on the evaluation sequence of each indicator of the alternative obtained from the evaluation and the comprehensive evaluation of all indicators. By weighting and considering the results, choose the best solution for enterprise logistics and distribution.

This paper studies the construction of the mathematical model of intelligent mobilization logistics delivering. The structure is as follows: The first chapter is the introduction part, which mainly expounds the research background and research significance of intelligent mobilization logistics delivering, and puts forward the research purpose, method, and innovation of this paper. The second chapter mainly summarizes the relevant literature, summarizes the advantages and disadvantages, and puts forward the research ideas of this paper. The third chapter is the method part, which focuses on the location-path method combining particle swarm algorithm and logistics delivering. The fourth chapter is the experimental analysis part. This part is experimentally verified on the dataset, and the performance of the model is analyzed. This part mainly reviews the main content and results of this study, summarizes the research conclusions, and points out further research.

2. Related Work

The whole vehicle path optimization problem aims at using the minimum number of cars and at the same time making full use of the loading capacity of the cars to achieve reasonable stowage and transportation, so as to achieve the goals of the lowest transportation cost or the shortest route. Scholars have long pointed out that facility positioning is closely related to transportation routes, and the positioning of factories, warehouses, and supply points is generally affected by transportation costs [5].

Delgoshaei et al. systematically studied the model and optimization algorithm of logistics distribution vehicle scheduling problem [6]. Ransikarbum et al. have done fruitful work on the problem of knowledge representation and intelligent modeling [7]. Yong et al. conducted researches from different perspectives, mainly focusing on the fields of mathematical modeling and optimization algorithms [8]. Konstantakopoulos et al. propose a method to classify multiple VRP variables related to freight transportation faced by most logistics and distribution companies in their daily operations and an algorithm to solve various problems. [9]. Qin et al. propose a comprehensive cold chain vehicle routing optimization model with the objective function of minimizing the cost per satisfied customer. For customer satisfaction, this paper uses the punctuality of delivery as the evaluation standard. [10]. Li et al. proposed a scanning algorithm to solve large- and medium-sized vehicles with load constraints and distance constraints, using the polar coordinate angle of each customer point to determine the direction of each route and then using an iterative calculation process to continuously optimize the total travel distance [11]. On this basis, Amaruchkul et al. innovated the

uniparental genetic algorithm, which uses the gene conversion position to recombine chromosomes [12]. Wang et al. proposed a mixed-vehicle multivehicle VRP model by introducing the fleet model and designed a hybrid algorithm combining tabu search heuristic and genetic algorithm [13]. Liu et al. applied the evolution reversal operator to enhance the local search ability of the genetic algorithm [14]. Kayvanfar et al. discussed a dynamic vehicle routing system based on online traffic information, where the system completes delivery tasks within a given time window according to customer requirements [15]. In order to use the dynamic time information, the system needs to establish the shortest path meter. The algorithm for the stochastic logistics problem is proposed. The algorithm is based on the decision-making process model, and the approximate optimal solution method is given. The algorithm has good practicability. Aiming at the increasing demand for flexibility in logistics delivering, a dynamic vehicle scheduling system is constructed that randomly arrives according to user orders during the planning period [16].

Since vehicle routing optimization problems often have many objectives and constraints, in order to simplify the solution process of such problems, some techniques and mathematical models can be applied to decompose or transform the problem into one or more existing research results based on the research results of previous scholars. This paper analyzes the results of the basic problems. Then the corresponding theory and algorithm are applied to solve each subproblem, and the optimal solution closest to the optimal solution of the original problem is obtained. The obtained optimal solution can also play a guiding role in practical problems.

3. Methodology

3.1. Using Mathematical Modeling to Complete the Optimization of Logistics Distribution Route. Mixed distribution means that the enterprise establishes a distribution system for small-scale distribution, while large-scale distribution can be outsourced and undertaken by a third-party distribution company [17]. Anderlueh et al. assigns some customers (“gray areas”) to echelons and the impact of three different urban layouts on the solution for free. The potential Pareto optimal solutions of two and three objectives are illustrated to effectively support the decision makers of sustainable urban logistics [18]. Virtual logistics refers to a logistics mode that uses computer network technology to carry out logistics operation and management and realize the sharing and optimal allocation of logistics resources among enterprises. That is, multiple member enterprises with complementary resources and technologies, in order to achieve the strategic objectives of resource sharing, risk sharing, complementary advantages, and other characteristics. Under the condition of maintaining its own independence, it has established a relatively stable partnership. Virtual logistics is the virtual integration of enterprise warehouses distributed all over the world into a large logistics support system by using increasingly perfect communication network technology and means. By completing rapid, accurate and stable material support tasks, we can meet the demand of multi frequency and small batch orders in the logistics market. The virtual logistics delivering mode is a new type of distribu-

tion concept that has appeared in recent years. It refers to the logistics delivering mode carried out by the dynamic alliance formed by the idea of virtual enterprises. The specific process is shown in Figure 1.

The positioning-routing problem of the logistics system generally considers two principles: customer service level and the total cost of the logistics system [19]. Customer service level is a relatively vague concept, which can be refined into some operational goals in actual systems, such as customer response time, product diversity, and product acquisition capabilities.

The optimization angle includes the following aspects: (1) maximum loading principle. In order to reduce the logistics cost as much as possible, it is necessary to minimize the situation that the car carrier is not fully loaded or empty. Therefore, during stowage, the maximum load capacity and space constraints of the car carrier should be calculated in detail according to the order details and customer needs so as to minimize the waste of space and ensure the maximum loading efficiency of the car carrier. (2) Principle of weight limit. The principle of weight limit refers to the fact that the maximum load should be considered when loading the car, but at the same time, the car cannot be overweight. If the total weight of the car carrier exceeds the carrying capacity of the road, on the one hand, it violates the “Road Traffic Safety Law,” and on the other hand, it will also affect the service life of the car carrier. (3) Loading and unloading principle. After the route scheduling plan is formulated, the logistics stowage should make a stowage plan according to the car route, so as to shorten the entire distribution cycle and improve the logistics delivering efficiency.

In the process of cargo transportation, it is extremely important to choose a reasonable delivery route. It can not only speed up distribution and improve service quality, but also effectively reduce distribution costs. Increase economic benefits. This paper constructs a planning model of delivery routes, which transforms the delivery problem into a theoretical optimal solution problem and a travel promotion problem in operations research. It is solved by programming, and the area is divided by ray rotation method according to the group method in the transportation route optimization strategy. The maximum bearing capacity of the deliveryman is 50 kg, and the volume of the goods is not more than 1 cubic meter. Use integer programming to plan the route in each area, so as to get the optimal route. This model can reasonably arrange the delivery route for logistics enterprises and improve the delivery efficiency. Saving delivery costs has a strong theoretical guiding role, so it has great practical value. In the process of cargo transportation, it is extremely important to choose a reasonable delivery route. It can not only speed up distribution and improve service quality, but also effectively reduce distribution costs and increase economic benefits.

Modeling perspective. (1) Type of input data: it may be deterministic or random. Many literatures started to study deterministic because it is relatively easy, but now, more research focuses on random data, especially the randomness of customer needs. (2) Objective function: The common objective function for the positioning-routing problem is to minimize the total cost, which is the sum of the positioning cost and the transportation cost. There are also some literatures that

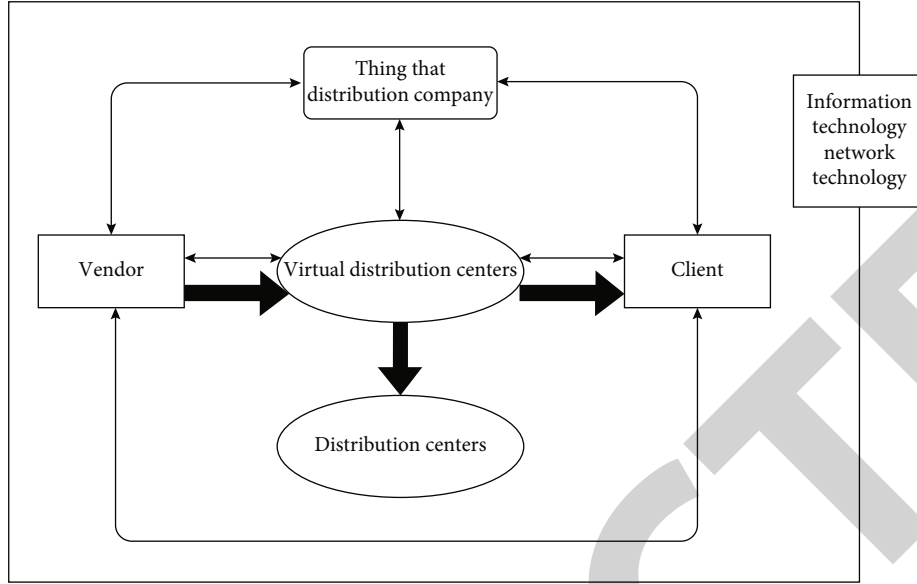


FIGURE 1: Virtual logistics delivering mode.

use different cost measurement methods, such as multiobjective methods and objective-dominant methods. (3) Solution space: The solution space can be discrete, network, or continuous. Considering the actual situation, most of these problems are discrete. But there is also a lot of work considering the positioning problem around transportation and the positioning problem of the traveling salesman. (4) Solving methods: There are mainly precise methods and heuristic methods. More and more literature uses heuristic methods, but precise methods are also very successful for some specific problems [19].

3.2. Iterative Update of Logistics Delivering Route Based on Particle Swarm Optimization and Genetic Algorithm. In order to further solve the problems of particle swarm optimization algorithm in the process of logistics distribution path, we can analyze here according to the characteristics of multiobjective constraint optimization in the logistics distribution path notebook. Then, the inertia weight of particle swarm optimization algorithm is improved, and a nonlinear variable particle swarm optimization algorithm is proposed. The vehicle number in logistics distribution can obtain the vehicles assigned by the distribution center to customers through the operational analysis of particles. Then, analyze the sequence of vehicles' driving paths in the logistics path. Then, determine the value size of the particles. Find the value analysis of the customer point where the logistics distribution vehicle J completes the distribution. The order analysis can be carried out according to the order from small to large value range, and then, the formal path order problem of distribution can be determined. Particle swarm optimization is to simulate the predation behavior of birds. The principle of the algorithm is as follows: POS first initialize a group of random particles, and each particle is a bird in the search space, representing a solution in the n dimensional space, where the position of the i th particle is $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$, and the velocity vector of each particle is $V_i = (v_{i1}, v_{i2}, \dots, v_{in})$. The individual extremum is the optimal solution found by the particle itself, that is, $pBest$, and

the global extremum is the optimal solution currently found by the entire population, that is, $gBest$. But at the same time, because of its obvious shortcomings, such as dealing with discrete optimization problems, it is easy to fall into local optimum and so on. Figure 2 is a flowchart of the particle swarm algorithm.

In each iteration, the particle updates its position and flight speed by tracking the individual extremum and the global extremum to search for the optimal particle in the solution space. The state update equation is as follows:

$$\begin{aligned} V_{i,t+1}P &= wV_i + c_1 * \text{rand}() * (P_{it} - X_i) + c_2 * \text{rand}() * (P_{gt} - X_i), \\ X_{i,t+1} &= X_i + V_{i,t+1}. \end{aligned} \quad (1)$$

Among them, the above formula represents the speed update of the particle, and the following formula represents the position update of the particle, which is the inertia weight, which represents the dependence of the particle on the current self-information; c_1, c_2 is the acceleration coefficient, c_1 express reliance on own experience, and c_2 represents the particle's dependence on the group experience. Compared with the genetic algorithm, the particle swarm algorithm has simpler rules [19]. The following is a flowchart of the particle swarm algorithm: the problem of localization and path change is considered, that is, to minimize the total cost and average waiting time or total loop length, and shows how to adopt heuristic ideas to deal with these goals. Their general idea is as follows: Let $G_-(V, E)$ be an undirected transport network with a set of V nodes and E edges, $|V| = n, |E| = m$. G represents the set of all points in the transportation network. For any node $v \in V$, two positive numbers w_v^-, w_v^+ are given, and the value of $w_v^- \leq w_v^+, w_v^-, w_v^+$ is the upper and lower bounds on the weighted w_v of node $v \in V$. And weight w_v can take any value on interval $[w_v^-, w_v^+]$. Let S be the Cartesian product over the interval $[w_v^-, w_v^+]$, and then, any $W \in S$ is called a weighted set

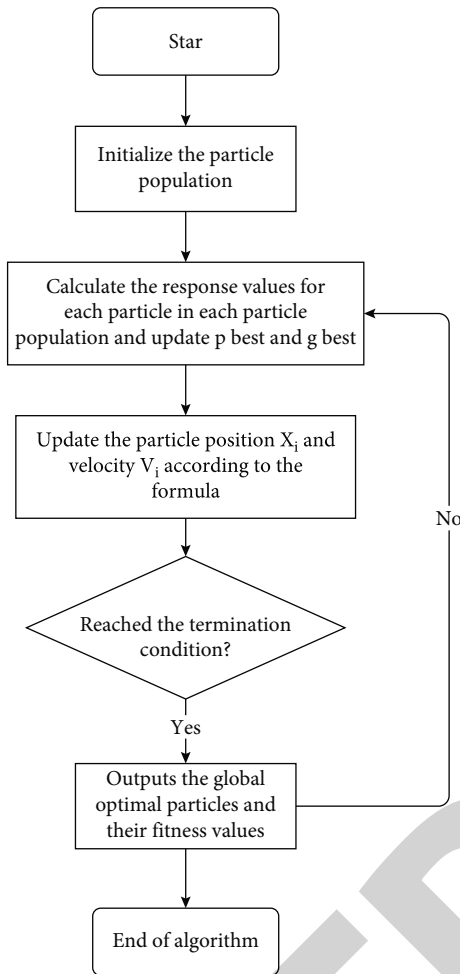


FIGURE 2: Flowchart of particle swarm optimization.

and denotes that a feasible weight is assigned to the network nodes. For any point a, b in network G , let $d(a, b)$ denote the shortest path between a, b , assuming that the matrix of shortest paths between nodes on G can be given (takes $O(n^3)$ time). Assuming that p identical facilities in network G need to be located, x_i represents the anchor point of the i facility, set $X = \{x_1, \dots, x_p\} \in G^p$ is called the positioning variable, and G_p takes the set of all p pairs in network G . For locating variable X and node $v \in V$, define the distance between nodes $v \in V$ and X as

$$d(v, X) = \min_{x \in X} d(v, x). \quad (2)$$

For positioning variable X and weighted set $W = \{w_v, v \in V\} \in S$, define

$$F(W, X) = \max_{v \in V} w_v d(v, X). \quad (3)$$

where $F(W, X)$ is the maximum weighted distance between location variable X and weighted set $W = \{w_v, v \in V\} \in S$. For any two location variables, $X, Y \in G^p$, define $REGR(X, Y) = \max_{W \in S} (F(W, Y) - F(W, X))$.

In common algorithms, an implicit selection mechanism is embedded in the process of determining the optimal location for each particle. Therefore, the idea of genetic algorithm can be used, and how to introduce the selection mechanism to form a new hybrid algorithm is the focus of current research. At the beginning of the iteration, each particle initializes its speed and position in space in a random way. Then, in the iteration process, the particle tracks the two extreme values to determine its position and velocity in the solution space. An extreme value is the optimal position of a single particle itself in the iterative process (that is, the spatial solution corresponding to the optimal fitness value). This is called the individual extremum of particles. The other extreme value is the optimal position found by all particles in the population during the iteration process, which becomes the global extreme value. If particles only track an extreme value, the algorithm is called local particle swarm optimization or global particle swarm optimization [20]. In evolutionary algorithms, the selection mechanism is usually used for the survival of the fittest areas, that is, selecting good areas and eliminating poor areas, which is conducive to more reasonable allocation of limited resources. The simulation results show that the algorithm has certain advantages in some test functions. The selection mechanism of hybrid particle swarm optimization and genetic algorithm is very similar. The hybrid particle swarm optimization algorithm uses the fitness value of each individual particle's current position, and assigns these particles a suitable order according to the fitness value of each particle, and then assigns good positions and velocities to the group with poor fitness values. Individuals, where the second half of the whole ranking with low fitness value is defined as the individual with poor fitness value, maintain the best position of each individual [21]. Therefore, the group can concentrate on the space with relatively good fitness value, and at the same time, it is also affected by the individual's own memory of retaining the best previous position. Through the research on the related algorithms of the logistics delivering path problem, the particle swarm algorithm is selected as the tool to solve the logistics delivering path problem in this paper. The limitations of particle swarm optimization are analyzed, and the ideas are summarized according to the existing improvement methods of particle swarm optimization (Figure 3).

4. Result Analysis and Discussion

A complete logistics network system, considering all its costs, including the cost of all transportation and positioning facilities per unit time, is used to compare the cost of enterprise operations under various network system structures, and guide enterprises to choose the appropriate logistics delivering network system.

When LRP considers multiple objectives such as customer service and cost, the objectives are generally divided into two: The first objective is to deliver goods on time according to the requirements of customers to improve the quality of logistics services, which is called the objective of completing delivery tasks on time. That is, when customers place an order, they often specify the arrival time of the goods and how to arrange distribution to meet the customer's time requirements [22]. This is one of the main goals of the enterprise logistics system;

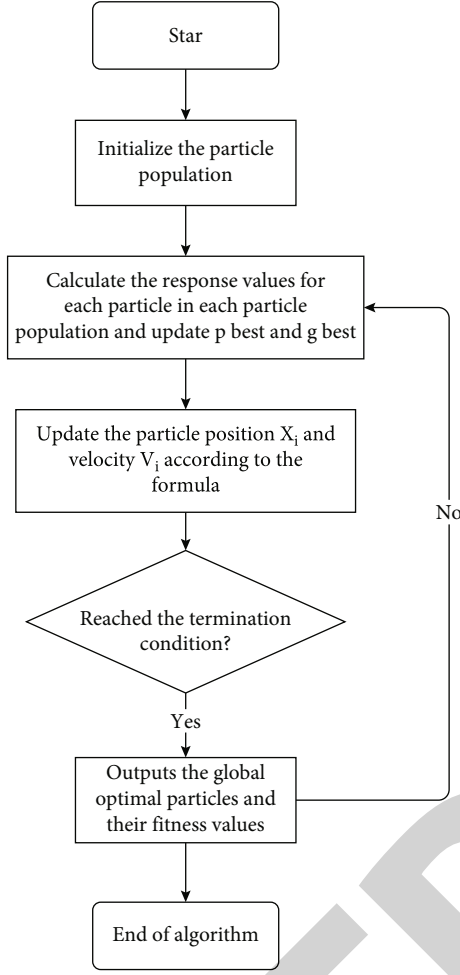


FIGURE 3: Flowchart of the improved particle swarm algorithm.

the second goal is to rationally optimize the logistics positioning location and distribution route to achieve the lowest total operating cost, which is called the lowest cost [23]. This is a requirement for enterprises to reduce the total cost of the logistics delivering system, including the cost of positioning facilities, the total operating cost, and the total cost of transportation routes. For the multiobjective enterprise logistics delivering system positioning-path problem, the relevant models established are as follows:

Variable parameters are as follows:

- (i) $P = \{1, 2, \dots, n\}$ means n feasible location facility
- (ii) $C = \{i | i = 1, 2, \dots, m\}$ refers to m customer in need of service
- (iii) $S = P \cup C$ refers to the sum of all facility positioning points and demand customer points (including distribution centers and warehouses)
- (iv) $V = \{v_k | k = 1, 2, \dots, s\}$ means s means of transport
- (v) uc_{ijk} refers to the average unit distance transportation cost of the k vehicle from customer i to customer j

- (vi) TC_k means the average unit cost of acquiring the k means of transport
- (vii) d_{ij} refers to the distance from customer i to customer j
- (viii) w_1, w_2 , respectively, refer to the weight coefficients of the two objectives
- (ix) q_j refers to the average number of customer $j \in C$ needs
- (x) Q_k refers to the capacity of the k -th transport vehicle
- (xi) F_t refers to the average unit time penalty cost of delaying the delivery of the goods as required by the customer
- (xii) t_{ijk} refers to the time it takes for the k -th vehicle to travel from customer i to customer j
- (xiii) T_K refers to the departure time of the k th car
- (xiv) T_i, T_j , respectively, refer to the time from customer i to customer j
- (xv) ET_{ij}, LT_{ij} , respectively, refer to the lower limit and upper limit of the time period from customer i to customer j
- (xvi) T_{ijf} refers to the time it takes to deliver the goods on time from customer i to customer j as required, and it can be the arrival time period specified by a customer in advance $ET_{ij} \leq T_{ij} \leq LT_{ij}, T_{ij} = T_k + t_{ijk}$
- (xvii) T_{ijf} refers to the delay from customer i to customer j when the goods are not delivered as required

The multiobjective LRP model to build the model is

$$f_1 = \min \left\{ \left[w_1 \sum_{i,j \in C, k \in V} (LT_{ij} - T_{ij})^2 + w_2 \sum_{i,j \in C, k \in V} (T_{ij} - ET_{ij})^2 \right] + \sum_{i,j \in C} F_t T_{ijf} \right\}, \quad (4)$$

$$f_2 = \sum_{i=1}^s TC_k \sum_{i=1}^m \sum_{j=1}^m x_{ijk} q_j + \sum_{i=1}^n G_i Z_i + \sum_{k=1}^s C_k \sum_{i=1}^m \sum_{j=1}^m x_{ijk} uc_{ijk} q_j d_{ij}. \quad (5)$$

Locate a transportation route arrangement problem location—routing problem; LR refers to a given number of potential facilities. Among these potential facilities, a group of facility locations should be determined. At the same time, a set of transportation routes from each facility point to each customer point is determined based on meeting the goal of the problem (usually the minimum total cost). The location of customer points and customer demand is known or estimable. The location of one or more facilities of the goods is known, and

TABLE 1: Simulation results.

Algorithm	Number of vehicles	Number of particles	Number of particles	Mean algebra	Experimental optimal average distance
Optimized particle swarm optimization	7	50	57	522.3	594.325
		100	59	433.2	594.246
		120	62	360.1	594.233
Traditional particle swarm algorithm	7	50	40	554.3	594.665
		100	45	459.6	594.364
		120	49	370.6	594.365

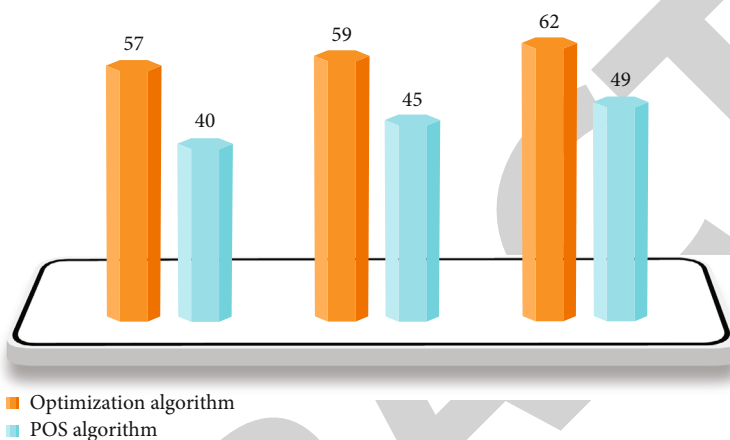


FIGURE 4: Comparison of the times of reaching the optimal path.

the goal of the problem is to establish those potential facilities to minimize the total cost. People have made a long-term and in-depth study on its model and algorithm. At present, there are few relevant papers in China, and the existing papers usually adopt the two-stage method to solve this problem. That is, the original problem is divided into two small problems: LA and VRP. First, solve the LA, and then, solve the VRP problem on this basis. In fact, for LRP models including time windows, there are still some limitations in solving practical problems in some specific cases. For example, the requirement for just-in-time production is very high, and manufacturing enterprises that implement zero-inventory production need a more real-time positioning-path problem method, that is, to realize the path selection and location decision under uncertain factors, which is the dynamic positioning-path problem. For example, a logistics delivering system has multiple production locations, multiple distribution centers, and multiple vehicles, and the decision is to optimize the entire logistics system through facility positioning and route selection decisions. The demand is not generated randomly and disorderly as in the traditional research method, but occurs according to a demand queue with a certain probability. Although there may be some unexpected events that lead to large changes in demand, usually the change in customer demand should be a deviation from the production plan within a certain limit. According to the previous introduction, the following compares the optimization algorithm with the traditional particle swarm algorithm to prove the superiority of the optimization algorithm. In this paper, the number of particles is simulated, and the advantages and disadvantages

of the optimization algorithm and the traditional algorithm are compared. Compared with traditional methods, the optimized particle swarm optimization algorithm has better results in the selection of the optimal distance than before, which proves that the optimized algorithm is more conducive to finding the optimal distance. The optimal solution improves the practicability of the algorithm in finding the optimal allocation scheme, and the optimal solution is more optimized. The simulation results are analyzed, and the results are shown in Table 1:

It can be seen from Figure 4 that whether the number of particles is 50, the number of particles is 100 or the number of particles is 120, the number of times the particle swarm algorithm has reached, and the optimal path is more than that before the optimization.

It can be seen from Figure 5 that whether the number of particles is 50, the number of particles is 100 or the number of particles is 120, and the average algebra of the optimized particle swarm algorithm is reduced, which is convenient for the realization of the algorithm, simple operation, and time saving.

It can be seen from Figure 6 that whether the number of particles is 50, the number of particles is 100, or the number of particles is 120, and the optimized particle swarm algorithm obtains better results in the optimal distance selection than before the optimization, which proves that the optimized algorithm is more conducive to finding the optimal distance. The optimal solution improves the practicability of the algorithm in finding the optimal distribution plan, and the optimal solution found is more optimized. It can be seen from the

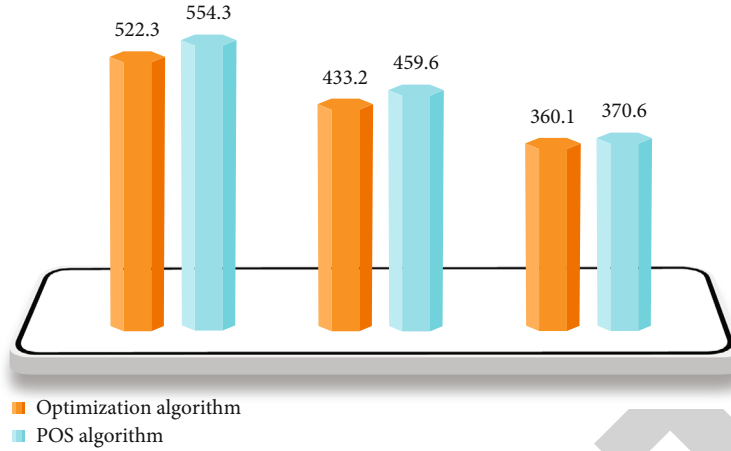


FIGURE 5: Average algebra comparison chart.

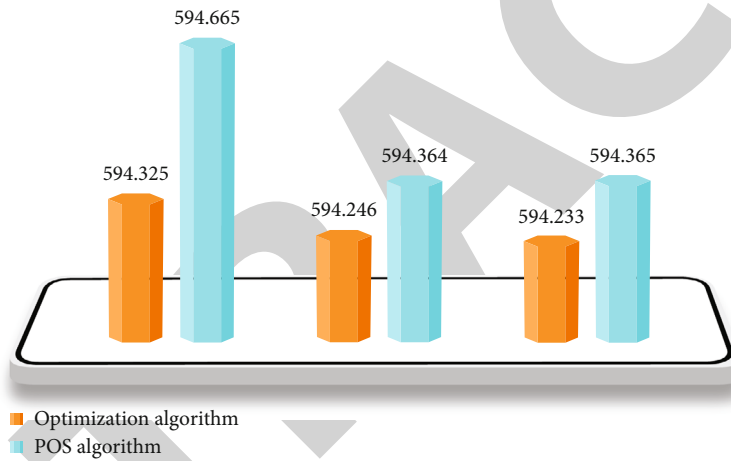


FIGURE 6: Comparison of optimal distances.

simulation results that the optimized particle swarm algorithm is effective and achievable in logistics delivering, and it is significantly better than the traditional particle swarm algorithm in reaching the optimal number of paths, average algebra, and optimal average distance.

The application scenario of this paper is the vehicle distribution scheduling optimization problem of multifreight center, multivehicle type, and multicommodity model. The calculation example in this section is intended to use 6 paths and 10 positioning data to solve the vehicle logistics solution Table 2.

The case uses Matlab2015b to program the calculation example, and performs the operation experiment on an Intel i5 2.30 GHz computer, and the preset number of iterations is 500. After ten running experiments, the best experimental results are as follows: the total transportation cost is 16547.298 yuan, the algorithm running time is 256.4985 s, and the required car transporter is 7.

The algorithm iteration diagram is shown in Figure 7. By observing and querying the Matlab workspace, it can be seen that the optimal solution of the experiment is obtained in the 65th generation of GA-PSO, and the population gradually

TABLE 2: The location and demand of each customer point.

Position	Abscissa/km	y-axis/km	Order
1	380	70	$2^*A + 3^*B + 2^*F$
2	350	100	$2^*C + 3^*D$
3	250	120	$2^*C + 4^*E$
4	225	139	$3^*C + 2^*F$
5	275	219	5^*A
6	255	175	$2^*D + 2^*F$
7	325	275	$3^*A + 2^*D$
8	425	155	$2^*B + 3^*F$
9	455	290	$4^*A + 2^*C$
10	475	155	$2^*B + 3^*C$

converges to the optimal solution in the 80th generation, that is, about 20 s after the algorithm runs. The algorithm can converge to the global optimal solution.

In order to verify the effectiveness of the algorithm and compare and analyze the performance of the algorithm, this

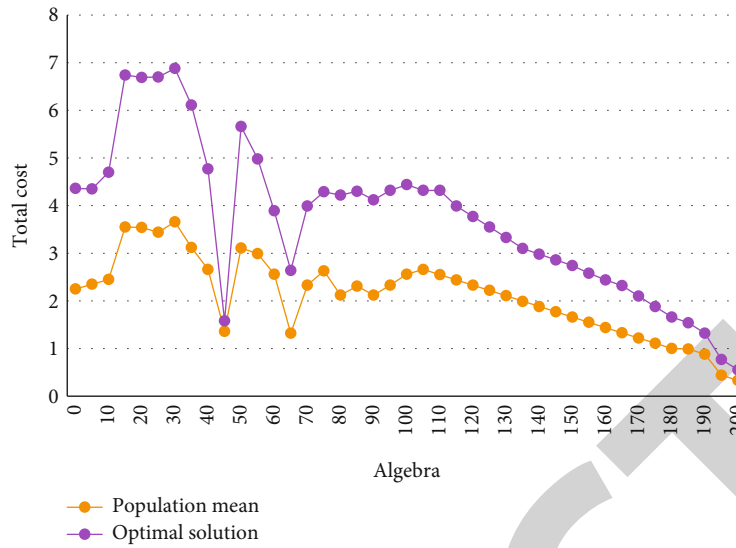


FIGURE 7: Iterative diagram of the hybrid algorithm.

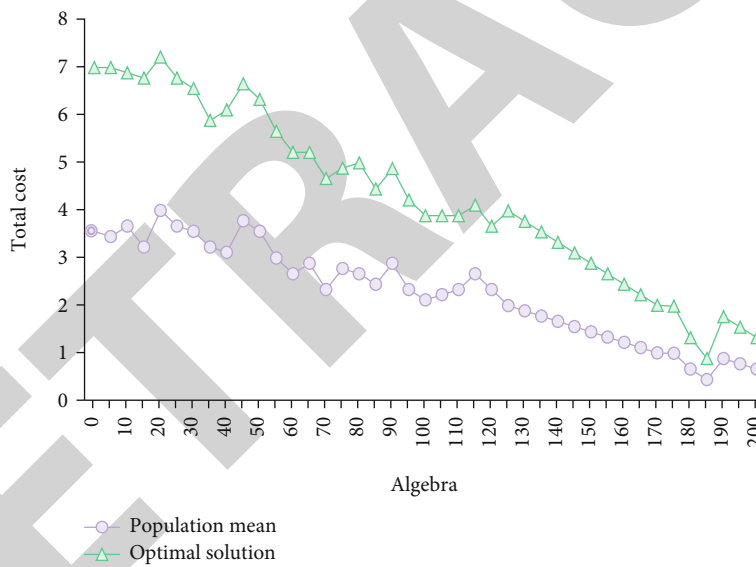


FIGURE 8: Iterative diagram of genetic algorithm.

paper uses the genetic algorithm to perform programming simulation experiments on the case. The randomization technique is used to guide the efficient search of a coded parameter space. Among them, selection, crossover, and mutation constitute the genetic operation of genetic algorithm. The core content of genetic algorithm is composed of five elements: parameter coding, initial population setting, fitness function design, genetic operation design, and control parameter setting. The coding method and population initialization in the genetic algorithm are the same as the GA-PSO hybrid algorithm. The convergence diagram of the genetic algorithm is shown in Figure 8:

The two algorithms have been simulated for ten times, respectively. Table 3 lists the comparison of the experimental results of the algorithms, and gives the maximum and average value of the ten operation costs of each algorithm,

the convergence algebra when the cost is the lowest and the results of each algorithm.

By comparing the simulation results of the two algorithms, it can be seen that the particle swarm hybrid algorithm has faster convergence and better optimization results. This shows that in a large-scale practical application, the conventional genetic algorithm takes too long to search for optimization, and once it falls into a local optimum, it cannot easily jump out, and cannot meet the timeliness of actual operation. The improved particle swarm algorithm has relatively better performance and better global search ability, which can meet the actual needs of logistics delivering scheduling.

The multivehicle logistics delivering model based on the shortest path designed in this paper can be solved by the optimized particle swarm algorithm; that is, the optimized particle

TABLE 3: Algorithm results comparison.

Algorithm	Total number of cars	Optimal cost	Convergent algebra	Highest cost/RMB	Average cost/RMB	CPU(s)
Particle swarm algorithm	7	16543.52	65	16853.43	16654.32	233.6579
Genetic algorithm	7	17889.12	70	17999.23	17165.43	325.5643

swarm algorithm can solve the logistics delivering problem. Compared with the traditional particle swarm algorithm, the particle swarm optimization algorithm designed in this paper introducing fuzzy classification and self-adaptive mutation mechanism has its own advantages and is an effective optimization algorithm, which can be used in practical operation. The two models based on the shortest vehicle travel path with the fewest vehicles and the two models based on customer satisfaction have different logistics delivering schemes; that is, for different optimal distribution targets, the resulting distribution schemes are also different.

5. Conclusions

In this paper, facility location and transportation combination optimization are studied as a whole. The influence of the two different factors on each other shall be taken into account. In the actual planning of logistics system, it is necessary to unify the location selection of logistics facilities and the decision-making of transportation routes. It is not advisable to ignore one of them and consider only one of them. The constituent factors of problem research and the main content of current research are given. This paper briefly describes the dynamic models of two common location routing problems in logistics distribution systems. From the perspective of analysis and comparison, firstly, the common location problems, routing problems, and the research status of location routing problems are described in detail. The path calculated by the data model improves the logistics distribution, improves the operation cycle and link efficiency, and achieves the purpose of optimization. Compared with the traditional genetic algorithm, the accuracy of the improved particle swarm optimization algorithm is improved by 10.36%. However, intelligent optimization itself is rich in content, and logistics system is a highly applicable field. Therefore, the research still has some limitations. The follow-up research directions mainly include as follows: when studying the logistics distribution path problem, this paper only studies the core optimization algorithm. In the future, GIS technology can be integrated into the core algorithm of this paper, so as to develop a more practical logistics distribution path system. Particle swarm optimization algorithm itself means a certain degree of parallelism, and a particle is independent of other particles in a cycle. Therefore, in essence, particle swarm optimization should be a distributed collaborative optimization 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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] M. A. Rahim and P. K. Banerjee, "A generalized model for the economic design of \bar{x} control charts for production systems with increasing failure rate and early replacement," *Naval Research Logistics*, vol. 40, no. 6, pp. 787–809, 2015.
- [2] N. Jathe, M. Lütjen, and M. Freitag, "Indoor positioning in car parks by using wi-fi round-trip-time to support finished vehicle logistics on port terminals," *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 857–862, 2019.
- [3] K. Hu and Z. Dong, "Design of spatial structure change model of coastal port logistics economy," *Journal of Coastal Research*, vol. 93, no. sp1, p. 1080, 2019.
- [4] X. Yu, "On-line ship route planning of cold-chain logistics distribution based on cloud computing," *Journal of Coastal Research*, vol. 93, no. sp1, p. 1132, 2019.
- [5] M. Z. Akkad and T. B. Nyai, "Multi-objective approach for optimization of city logistics considering energy efficiency," *Sustainability*, vol. 12, no. 18, p. 7366, 2020.
- [6] A. Delgoshaei, H. Norozi, A. Mirzazadeh, M. Farhadi, G. Hooshmand Pakdel, and A. Khoshniat Aram, "A new model for logistics and transportation of fashion goods in the presence of stochastic market demands considering restricted retailers capacity," *RAIRO-Operations Research*, vol. 55, no. 5, pp. S523–S547, 2021.
- [7] K. Ransikarbum and S. J. Mason, "Goal programming-based post-disaster decision making for integrated relief distribution and early-stage network restoration," *International Journal of Production Economics*, vol. 182, no. 10, pp. 324–341, 2016.
- [8] M. W. Yong and L. Y. Hong, "Improving bias estimation precision via a more accuracy radar bias model," *Mathematical Problems in Engineering*, vol. 2018, Article ID 3083258, 9 pages, 2018.
- [9] G. D. Konstantakopoulos, S. P. Gayialis, and E. P. Kechagias, "Vehicle routing problem and related algorithms for logistics distribution: a literature review and classification," *Operational Research*, vol. 22, no. 3, pp. 2033–2062, 2022.
- [10] G. Qin, F. Tao, and L. Li, "A vehicle routing optimization problem for cold chain logistics considering customer satisfaction and carbon emissions," *International Journal of Environmental Research and Public Health*, vol. 16, no. 4, p. 576, 2019.
- [11] H. Li, K. Xiong, and X. Xie, "Multiobjective contactless delivery on medical supplies under open-loop distribution," *Mathematical Problems in Engineering*, vol. 2021, Article ID 9986490, 7 pages, 2021.
- [12] K. Amaruchkul, "Planning migrant labor for green sugarcane harvest: a stochastic logistics model with dynamic yield

Retraction

Retracted: Application of Visual Guidance Design of Tourist City Combined with Deep Learning Computer Visual Symbols

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Cai and K. Liu, "Application of Visual Guidance Design of Tourist City Combined with Deep Learning Computer Visual Symbols," *Journal of Function Spaces*, vol. 2022, Article ID 8612543, 10 pages, 2022.

Research Article

Application of Visual Guidance Design of Tourist City Combined with Deep Learning Computer Visual Symbols

Jiaying Cai ¹ and Kai Liu ²

¹Liuzhou Institute of Technology, Liuzhou, Guangxi 545000, China

²Guangxi University of Science and Technology, Liuzhou, Guangxi 545000, China

Correspondence should be addressed to Kai Liu; 100000822@gxust.edu.cn

Received 11 June 2022; Revised 17 August 2022; Accepted 20 August 2022; Published 9 September 2022

Academic Editor: Miaocho Chen

Copyright © 2022 Jiaying Cai and Kai 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.

In recent years, computer vision has received rapid attention and development, and it also represents the most cutting-edge research direction of deep learning. With the rapid development of today's world economy, the arrival of the information age marked by the network has broken the boundaries of time and space in the sense of physics, and the social form in which the entire human being lives is more and more approaching a small "global village." The design and application of visual symbols need to achieve innovative combination and deconstruction of image symbols, indicator symbols, and symbolic symbols. When applying visual symbols to the design of urban tourism image, designers need to be clear about its role. The design and application of visual symbols of urban tourism image are of great significance, whether it is to condense and spread urban culture or beautify the environment. City brand image building can effectively protect the natural landscape and reduce the damage to the environment, and some city brand images can also become city landmarks. Therefore, designers should give full play to the advantages of local traditional culture, refine rich design materials, and skillfully integrate them into the design of visual symbols to further display the local humanistic characteristics, especially maximize the advantages and value of visual symbols, so as to make tourism. The city shows unique charm and unique humanistic spirit.

1. Introduction

With the increasing enrichment of people's spiritual life and the vigorous development of tourism, the competition between tourist cities has become increasingly fierce. This competition has gradually shifted from the competition of tourist attractions in the past to the competition of the city's overall tourism image to plan and develop urban tourism [1]. In recent years, during the golden weeks such as May Day, the national day, and the Spring Festival, people increasingly prefer to travel for the holidays. The major scenic spots have ushered in the peak of tourism, and the number of tourists has increased sharply, even exceeding the tourism carrying capacity of the scenic spots. On ordinary holidays and weekends, people will also choose the surrounding tourist attractions for sightseeing [2]. To a great extent, the development of China's tourism industry has driven the substantial growth of tourist attractions and the

surrounding service industries. China's tourism industry is showing a rapid growth trend with a bright future [3]. In the city, which is a space area centered around people or things, the space is relatively fixed, while the time is flowing. Through this "field" system of dynamic and static, human beings and the environment communicate with each other in material, energy, information, and spirit [4]. Due to the rapid development of tourism, there is an urgent need for an information system to display the image to the public. Therefore, the CI strategy is naturally introduced into the research of urban tourism to design the image of urban tourism. The meaning of the urban tourism identification system not only refers to CIS, but also it has given a new connotation [5]. In daily life, about 80% of information is obtained through vision. Good visual function enables people to capture a large amount of information in a short time, so that we can better understand the world around us. Then, the scenic spot guidance system is based on human visual

information processing, including recognition, discrimination, spatial perception and vision, and the integration with other senses to obtain information.

The urban brand image can break through the traditional urban development mode and release more energy for urban development [6]. With the development of economy, some cities gradually lose their characteristics, especially some tourism cities gradually fade the unique color of the city and only pay attention to economic benefits [7]. The theory of urban brand image has important strategic significance for urban development. It is necessary to popularize the importance of urban brand image, which will enable more and more tourist cities to obtain more benefits [8]. As an intelligent technology of computer vision data processing, deep learning technology can extract some data from objective things and through the construction of artificial neural network. After processing, it can be deeply excavated by computer, and the characteristics of data can be obtained through supervised learning or unsupervised learning, so as to achieve the advantages of faster processing speed, larger data volume, and higher accuracy of computer vision technology such as image recognition [9]. The decision-makers of domestic tourist cities have realized that the visual guidance system, as the synthesis and concentration of a city's image, culture, and characteristics, is of great significance to the sustainable development of the city, and at the same time, they have also increased the construction of the city's visual guidance system. However, due to the lack of in-depth understanding and effective guidance, its effects are also unsatisfactory, and some are even counterproductive, which will seriously hinder the healthy development of tourist cities [10]. The establishment of urban visual identity guidance system should integrate urban master planning with art design and effectively serve the interoperability design of urban master planning, urban space environment, and urban landscape design.

The research innovation lies in the research on some key technologies of deep learning in the field of computer vision symbols. Combining the old and new computer vision problems such as vision-oriented image retrieval and classification, some corresponding key technologies and solutions are proposed. Sort out the design principles of individuality and commonness, the unity of nationality and modernity, and the function and form of visual symbols of urban tourism image. Through semiotics, urban tourism image, visual image design, and other related theoretical monographs, it combines the contents of ecological symbols and humanistic symbols. This paper analyzes the image generalization of visual symbols of urban tourism image from the perspective of the integration of traditional culture and modern art and various interactive display modes. The design method of symbol implication and the future development trend are summarized.

2. Related Work

Wang H pointed out that nowadays, with the great changes in China's urban development, simple road network cannot meet the needs of fast-developing cities, and traffic facilities

such as urban trunk roads, expressways, viaducts, subways, tunnels crossing the river, and tunnels crossing the sea are constantly emerging. The development of this fast-speed urban traffic is accompanied by the expansion of urban area, and it is followed by the improvement of the overall appearance of the city [11]. Gao G et al. pointed out that the current research on the brand image of tourist destinations is relatively simple, and it is basically carried out on the basis of the theoretical framework of the enterprise [12]. Lu found that this topic combines the visual identity symbol system design with the tourism destination brand image and uses multidisciplinary knowledge to improve and supplement the research on the visual identity symbol design of tourism image [13]. This paper analyzes the method and path of visual identification symbol design in Taigu Sanduotang Museum. Xuchang proposed that this paper put forward new research ideas and system models around the technical bottlenecks and difficulties in the research of these computer vision fields. In addition, the idea of preference learning is introduced into the research of problems in the field of computer vision, which provides a new research perspective and direction [14].

Sarba et al. constructively put forward the concept of defining urban visual symbols in "Research on Visual Symbols of Urban Brand Image," which are mainly urban signs and urban symbols [15]. With the help of relevant theories of semiotics and visual perception, Tong builds a logo design system with visual characteristics, types, visual styles, and visual semantics as the main body and a symbol shaping system with visual characteristics, types, and setting principles as the main body [16]. Li et al. pointed out: "China's tourism industry is developing rapidly, gradually moving towards branding, and vigorously promoting the brand image of tourist attractions [17]". Amoros F pointed out: "The difference between region and environment creates different famous historical cities in China, and the charm of the city is reflected in the regional and cultural nature of its landscape design [18]". Under the guidance of urban planning, urban landscape design should continue the historical context, highlight traditional characteristics, highlight cultural charm, properly solve the contradiction between modernization and inheritance of tradition, and seek the combination of modern life and traditional culture. Chen et al. "pointed out:" driven by the accelerated development of urban construction, the urban public environment is becoming more diversified and complex. With the help of the guide sign system, the complex public environment space can become more efficient and orderly [1]. Cyr A found that only by combining tourism logo orientation system, regional cultural elements, and the surrounding environment can it be correctly positioned and accurately studied, so I hope to conduct in-depth research and analysis on topics related to such aspects [19]. Finally, most of the research literatures only discuss regional differences or regional cultures, but do not go deep into regional characteristic cultures. The elements of regional characteristic cultures are the essence and powder extracted from a wide range of regional cultures. One is how to better integrate with other technologies in different application fields. Computer vision can make

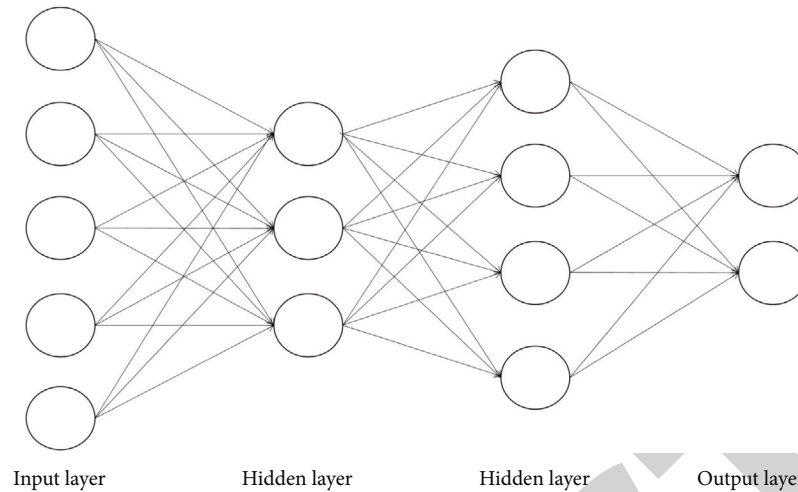


FIGURE 1: Artificial neural network model for deep learning.

extensive use of big data in solving some problems, which has gradually matured and can surpass human beings. But in some problems, it cannot achieve high accuracy. The second is how to reduce the development time and labor cost of computer vision algorithm. At present, computer vision algorithms require a lot of data and manual annotation and require a long research and development cycle to achieve the accuracy and time-consuming required by the application field. The third is how to speed up the design and development of new algorithms. With the emergence of new imaging hardware and artificial intelligence chips, the design and development of computer vision algorithms for different chips and data acquisition devices is also a challenge.

The research and analysis of the deep neural network framework in this paper have certain theoretical contributions to the development and application of deep learning, and at the same time, it has also played a positive role in promoting the research on many hot issues in the field of computer vision. □

3. Methodology

3.1. Deep Learning Applied to Computer Vision Analysis.

With the proposal of the concept of intelligent machine and the development and application of machine technology, computer vision technology has been rapidly studied and developed. The fundamental reason is that the perception of intelligent machines is mainly realized by visual technology, which helps intelligent machines have the ability to “see the world.” Usually, the collection, extraction, and processing of object features are realized by camera and processor. As an intelligent technology of computer vision data processing, deep learning technology can extract some data from objective things and construct artificial neural networks. After processing, it can be deeply mined by calculation, and the characteristics of data can be obtained through supervised learning or unsupervised learning. To achieve image recognition and other computer vision technology processing faster. As a branch of machine learning,

the idea of deep learning is to realize intelligent data processing by establishing an artificial neural network to simulate the human brain. The model is shown in Figure 1.

That is to say, the artificial neural network model established in advance is used to simulate the human brain for deep learning and analysis to realize the intelligence of the machine. Commonly used are the perception of images, sounds, videos, etc. Because of the multi hidden layer learning structure in the artificial neural network, it obtains data through input layer perception, multi hidden layer intelligent processing and output layer execution to complete a given task. Each node is equivalent to a processor and has a specific algorithm. The result is used as the input of the next layer node, and deep learning realizes signal transmission between neurons similar to human brain through this cascade of multiple hidden layers. And accumulate data experience under specific unsupervised training. From the perspective of the development goals of computer vision, one is to help people recognize and remember the external world, and the other is to help machines realize the perception of the surrounding environment. Among them, there are many researches on image recognition and processing, target state detection, scene analysis, and application, and the hot research fields are shown in Figure 2.

Take image recognition as an example, as shown in Figure 2. Urban visual guidance system is a comprehensive public information system that guides people to carry out activities in public places. Its primary function is to guide the direction, followed by strengthening the regional image. Locality or regionality is not the same as the application and expression of regional characteristic cultural elements. Regional characteristic cultural elements are the essence and core extracted from a wide range of traditional regional cultures. The uniqueness, identification, creativity, and other important contents of regional culture will eventually be transformed into visual symbols to express, spread, and develop cultural connotation. In order to make the communication and inheritance of current regional culture conform to the communication form of the times, we need to constantly excavate cultural symbols with local characteristics.

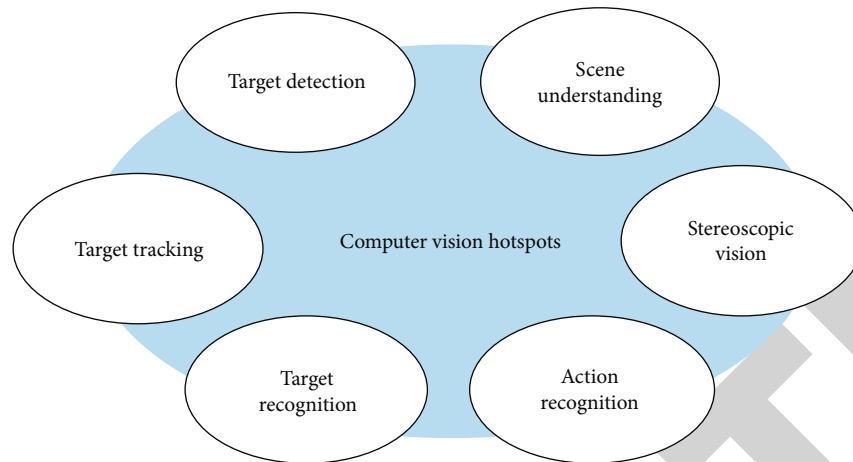


FIGURE 2: Computer vision application field.

Only by optimizing the extraction and redesign of visual elements can we better promote the development and innovation of traditional regional culture in the Internet era. Cultural symbols are the material carrier of national culture. Cultural visual symbols have strong graphic abstraction and are the visual signs of different national cultural differences. Find the differences of their respective cultures and present them in visual expression forms. The extraction and redesign of cultural symbols is a new concept, which is to mobilize one or more cultural elements or cultural symbols in the design process. Collect, compare, and refine, and then complete the design of concept or formal beauty through visual design methods such as deconstruction, reorganization, and addition. We should design visual symbols that can express the theme, on the one hand, in its external form, and, on the other hand, it is the expression connotation with special significance. The external form should conform to the visual form with aesthetic value, that is, it should be beautiful and generous, have the aesthetic feeling of artistic form, and also have the internal meaning corresponding to the shape. The traditional regional culture is summarized and refined, and the redesign methods of “decomposition” and “reconstruction” are used to dismantle and rearrange the traditional culture with local characteristics to make it conform to the aesthetic cognition under the current trend of the times. Give each graphic language element a new era of meaning and rich beauty, and let regional culture be protected and inherited through symbolic visual language. And with accurate information form, make people have correct association and behavior. Urban visual guidance system plays a great role in urban traffic dredging, urban planning, and dispersion of people flow. It also affects the style and appearance of the city and people’s quality of life to varying degrees and realizes the dialogue between people and urban space. Since the early 1990s, the urban visual guidance system has gradually formed its unique identity and integrity from the initial signs and indicators to the current guidance design. It integrates direction identification, public facilities identification, road search, warning service, urban culture, and artistic beauty. Street maps help people understand the overall situation of the surrounding environment and deter-

mine their location. So as to provide reference for selecting the next walking direction. Provide directional timetable of bus, train and flight information through driving route map, bus and subway route map and tourism route map. Guide lines that use different colors to guide specific destinations and various electronic information display and consultation facilities. The urban tourism orientation system is mainly composed of the following three parts: the orientation object of the urban tourism orientation, the orientation elements of the urban tourism orientation system, and the main information of the urban tourism orientation system, which together constitute the overall framework of the urban tourism orientation system, as shown in Figure 3.

3.2. Realization of Visual Guidance System Based on Tourist City. And with accurate information form, make people have correct association and behavior. Urban visual guidance system plays a great role in urban traffic dredging, urban planning, and dispersion of people flow. It also affects the style and appearance of the city and people’s quality of life to varying degrees and realizes the dialogue between people and urban space. Since the early 1990s, the urban visual guidance system has gradually formed its unique identity and integrity from the initial signs and indicators to the current guidance design. It is mainly to enable different users to efficiently get the required information and find the final destination in this more complicated modern spatial information environment. The guidance system and the identification system need and support each other, and in the same environment, they are integrated to a great extent. They are inseparable and belong to a large environmental system. The orientation information is conveyed through a logo guidance system based on graphic symbols, color perception, and text layout. The ultimate goal is to optimize the form of its information, guide people to generate accurate cognition, and make correct actions, which is a good action for the audience. The guide is an organic beautification behavior for the environment, which well achieves the dual satisfaction of vision and spirit. The development of the signage guidance system is a complex process. A successful signage guidance system integrates the elements of many different

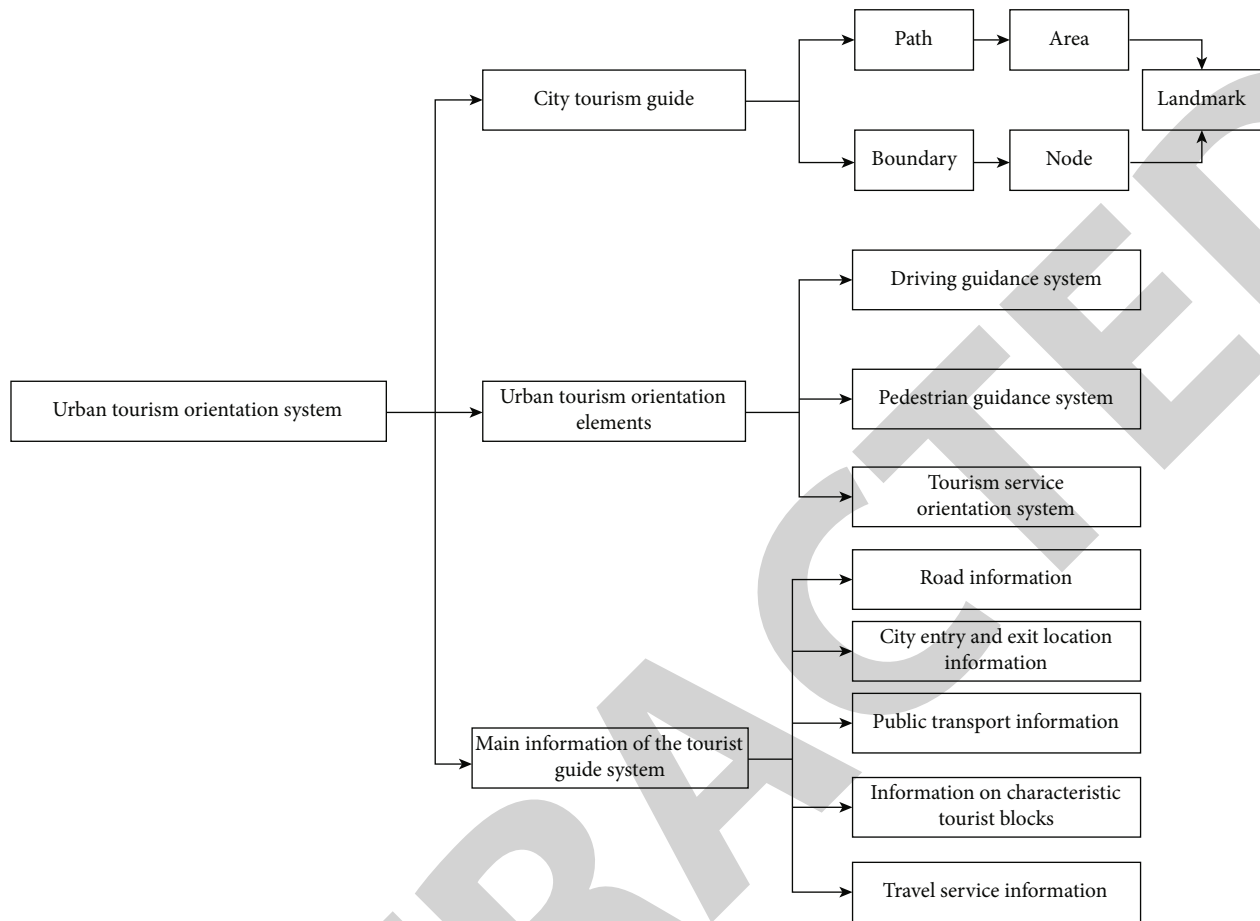


FIGURE 3: Component diagram of urban tourism guidance system.

disciplines and businesses, such as design, technology, architecture, physics, anthropology, spatial planning, and marketing.

The design goal of the tourism guidance system is to create a comprehensive, modern, unobtrusive, and easy-to-maintain work, which must have a high degree of recognition and information content. The design must actively show the versatility and characteristics of the scenic spot or region, so as to enhance the image of the region, prolong the stay time of tourists, and increase their return visit rate. The visual symbol of urban tourism image is the visual induction of standardization, systematization, visualization, and symbolization of all tourism resources in the city. Tourism resources have the function of further beautification, which can guide and help tourists to perceive the city more deeply, eliminate the anxiety caused by unfamiliar cities, and carry out continuous promotion of the city's tourism image. It is important to guide the public to improve the quality of life through tourism. Tourism can let people put down their troubles, and tourism will make people feel more about themselves and know themselves again in the world. Tourism is a special form of three-dimensional learning. Through tourism, there are gains, changes, and sublimations. Different cities have different tourism resources, cultural customs, history, and culture. The visual symbol of urban tourism is a symbol system refined based on the dif-

ferences in tourism regional characteristics, cultural history, and modern development. It visualizes and graphically processes the tourism characteristics of cities and facilitates the visual communication of urban tourism image information. For example, the visual symbol of PHUKET tourism is a deformed design on the word "Phuket." In the process of using visual communication to design urban tourism image, influenced by the modern design concept, it usually pays more attention to simplicity and abstraction. Pay more attention to the intuitive impact in the visual performance, and it is required to bring a strong feeling to the audience at the first time, which is contrary to the performance of traditional pattern elements. Applying traditional pattern elements to visual communication design helps to promote the humanization return of design language. Transition the relatively stiff design language to the direction of humanism and affinity, and promote the diversified development of visual communication design. The font style is full of the enthusiasm and vitality of Southeast Asian countries. The letter "U" is decorated with traditional Thai handicraft elements, and the left half of the "U" also symbolizes the waves rolled up in Phuket. The standard color of blue reflects the sky blue of Andaman Sea, where the sea water blends with the sky as shown in Figure 4.

These visual symbols of urban tourism show the connotation and culture of urban tourism, embody the character of



PHUKET

FIGURE 4: Tourist signs of Phuket.

the city, express the personality of the city, and have the uniqueness and identification of urban tourism, which can accurately provide information to tourists, attract tourists' attention, and induce tourism motivation, thereby effectively increasing the passenger flow of urban tourism.

Deep learning methods include many well-known deep neural network models, such as deep belief network (DBN), convolutional neural network (CNN), deep Boltzmann machine (DBM), and stack denoising self-encoder (SDAE). However, the most popular deep learning frameworks in recent years are convolutional neural networks and stacked denoising self-encoders. Obviously, the work of the decoder is an inverse process of the work of the encoder. However, the key factor of simultaneous training of the two networks is to minimize the error between the reconstructed input Z_i and the original input X_i . The mathematical formula used here to minimize the error is expressed as follows:

$$\min_{W, W', b, b'} \sum_{i=1}^k \|x_i - z_i\|_2^2 + a \left(\|W\|_F^2 + \|W'\|_F^2 \right), \quad (1)$$

where k represents the number of original input images. $y_i = f_\theta(W x_i + b)$ and $z_i = g_{\theta'}(W' y_i + b')$ denote the output of the encoder and decoder for each input image x_i , respectively. a represents the regression parameter, which is used to trade off the relationship between the error and the complexity of the two networks, which is measured here using the norm term $\|\cdot\|_F$.

Based on this, we collect this preference relationship between all positive and negative samples to build a preference relationship set, as follows:

$$PJ = \{v_i \succ u_j; i = 1..p, j = 1..n\}. \quad (2)$$

In order to further deduce the sorting function, we refer to the method used by Herbrich et al. In this way, we need to look for such a function $R : \mathfrak{R}^d \times \mathfrak{R}^d \rightarrow \mathfrak{R}$, for example:

$$\forall v_i, u_j \in \mathfrak{R}^d, R(v_i, u_j) > 0 \Leftrightarrow R(v_i, 0) > R(u_j, 0). \quad (3)$$

After that, this sorting function r can be simplified and defined as

$$\forall x \in \mathfrak{R}^d, r(x) = R(x, 0). \quad (4)$$

According to the set of preference relations generated by formula (4), we can use the following constraints to determine the function R :

$$\begin{aligned} R(v_i, u_j) &> 0, \\ R(u_j, v_i) &< 0, \forall (v_i, u_j) \in PJ. \end{aligned} \quad (5)$$

Therefore, we can construct a two classifier.

$$\begin{aligned} \min_{\omega, \xi} \frac{1}{2} \|\omega\|^2 + C \sum_{i,j} \xi_{ij}, \\ \text{s.t. } \langle \omega, v_i - u_j \rangle &\geq 1 - \xi_{ij}, \\ \xi_{ij} &\geq 0, \forall (v_i, u_j) \in PJ. \end{aligned} \quad (6)$$

In fact, ω here is the preference parameter we require. Finally, we can use this preference parameter to sort the candidate target image blocks. The sorting formula is as follows:

$$r(x) = \langle \omega, x \rangle = \sum_{i,j} \beta_{ij} \langle v_i - u_j, x \rangle. \quad (7)$$

4. Result Analysis and Discussion

In the performance comparison of target tracking algorithms, the first evaluation criterion is the tracking success rate indicator, which is used to measure the overlap rate between the predicted tracking target frame and the true value in each frame. Among them, the tracking effect in bolt and soccer video sequences is the best among all 15 target tracking algorithms, and the performance in coke, deer, and jogging2 video sequences is also second only to the best performance, but the performance in car4 video sequences is slightly worse, but it is also among the top four tracking algorithms. Network system is composed of several interconnected channels and their elements, which are arranged in an average and orderly way according to the networked spatial form. Usually, a spatial information system is composed of several interconnected roads and nodes formed by the interconnection between roads. This kind of system is mainly aimed at tourists who cannot determine their walking route and is often suitable for railway stations, squares, or other dense spaces with too many starting points or too many ending points. Tourists need to set their starting points to correspond to the action space required by the corresponding starting points to determine their walking route. In the denser areas of the city, the spatial setting of the tourist guidance system often uses the setting method of the network system, which can effectively point out the direction for tourists. The top 10 algorithms with the best performance among the algorithms are depicted and displayed. In the legend of the accuracy percentage map of each video sequence, we sort the output of all target tracking algorithms according to the standard that the distance between the center point of the predicted target tracking frame and the center point of the true value is not more than 20 pixels as shown in Figures 5 and 6.

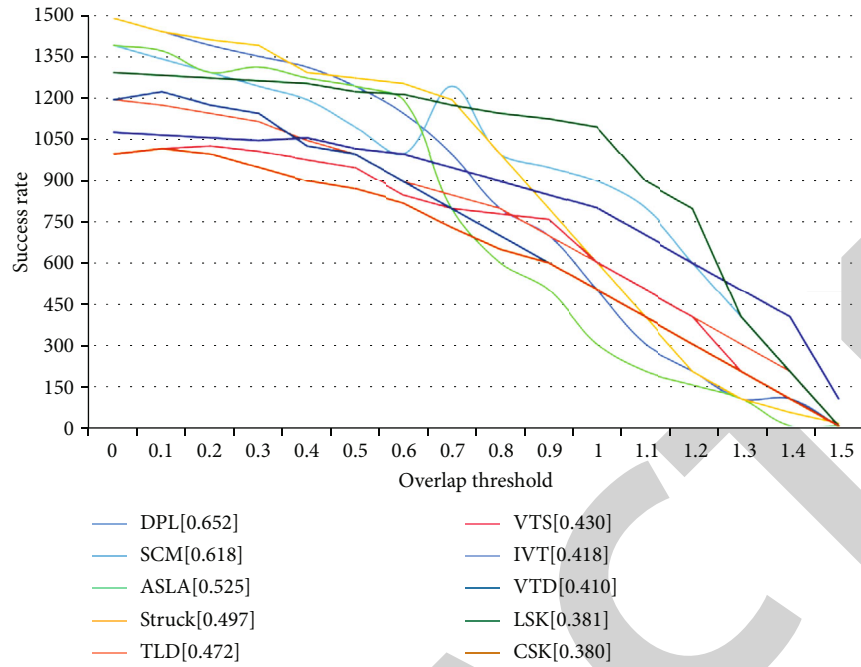


FIGURE 5: Success rate and accuracy plots across all video sequence frames.

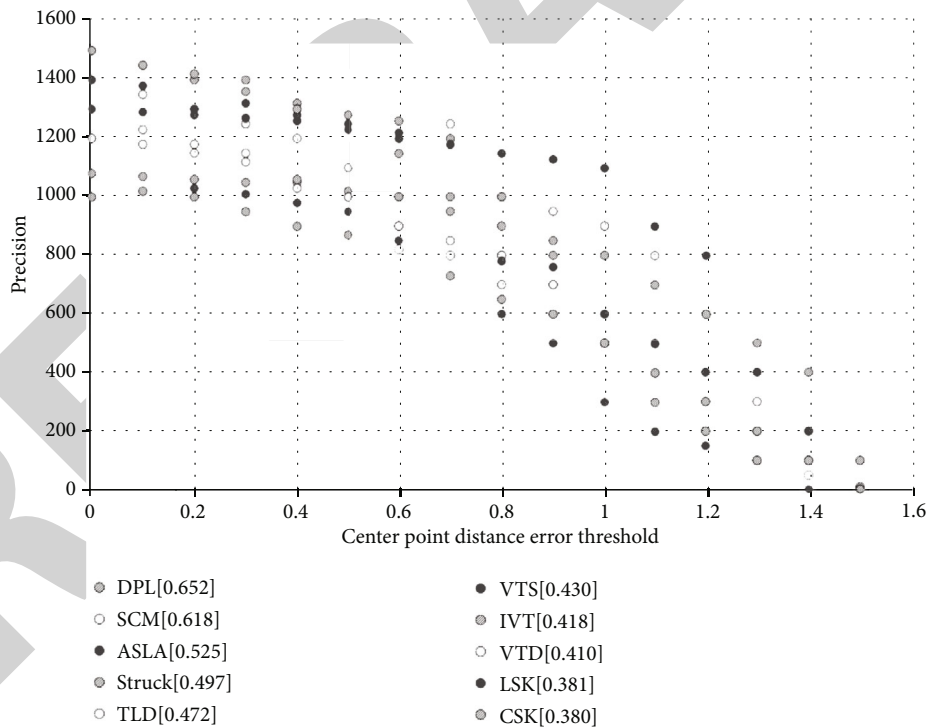


FIGURE 6: Overall average success rate of target tracking algorithm in all guidance sequences.

Tourism itself is a learning process to increase knowledge and broaden one’s horizons. Whether it is a cultural landscape or a natural landscape, tourists often get education and enlightenment from it and benefit a lot. While visiting the landscape, the function of picture explanation presented by the tourism logo guidance system enables tourists to get in-

depth information of scenic spots and strengthen their understanding of scenic spots. Design serves people, and audience feedback is an important criterion for evaluating the design level of a city’s visual orientation system. Since this case was put into use, it has received a large number of feedbacks from all walks of life through field visits, questionnaires, and other



FIGURE 7: Indoor guidance system design.

TABLE 1: The brand image of ancient “Fu Lu Shou”.

Source of color symbols	Color impression	Color expressions and functions
Building, brick	Blue, grey	Ordinary, moderate, stable
Yellow earth	Naturals	Beach, homeland, composure, stability
Fu Lu Shou Xi	Red, yellow	Red: enthusiasm, happiness, attention; yellow: noble, high reputation
Ethnic traditional clothing	Blue, blue	Sky, eternity, reason, elegance, moderation
Traditional furniture	Brown	Classical, earthy, gentle, harvest

methods. After summarizing and analyzing these feedbacks, we can draw the conclusion that the practical application effect of this case is good. The interactive electronic touch screen guidance system is more convenient and effective, which is a model of the combination of high-tech and cultural industries. However, if the interface design is too complex, this setting will be in vain, especially for the elderly. Therefore, how to make interaction design more humanized, more intuitive, and easy to use will become the primary problem beyond modeling. Form serves function, and style and culture will be the focus after solving the basic problems as shown in Figure 7.

The excellent environment of tourist attractions is one of the attractive factors of tourist attractions, and it is the foundation and guarantee for the sustainable development of tourist attractions. The environment of the scenic spot includes natural ecological environment and social cultural

ecological environment. There are many precious natural resources in the natural ecological environment of tourist destinations, such as rare tree species and rare species. There are many cultural heritages of human material civilization in the social, cultural, and ecological environment, such as landscape gardens and historical relics. These are nonrenewable resources.

In the design, the symbolic elements are summarized and refined in the visual graphic language, and the above information is integrated into a concise graphic symbol as much as possible by using abstract and concrete expression techniques. For example, the three stars of “happiness, wealth, and longevity” worshiped by the ancient people have lofty symbolic significance in the eyes of the common people. This kind of symbolic decoration technique is used in many places in Taigu Sanduotang. The “big mirror inlaid

with snail” is inlaid with the “Fu Lu Shou” three-star pattern. There is also this saying that when you look at the face of the lucky star as red, it is “the lucky star shines brightly.” If you look at Lu Xing’s face as green, it is “everything comes true.” If you look at the face of the birthday girl as yellow, it is “prolonging life.” Some blue-gray systems with lower purity represent the main tones of architecture and traditional clothing and can also trigger people’s nostalgic plots. From the perspective of harmony with the environment, khaki is also an important element. It can be said that a reasonable color positioning is the most intuitive reflection of the brand image, and it is also the first part of the visual experience. The details are shown in Table 1.

The tourism destination visual guidance system is a symbol interpretation system, which is completely a product of artificial design. Its main function is to guide and help tourists perceive the tourism image and activity function on the spot; conveniently, quickly, and clearly complete tourism activities; eliminate the tension caused by uncertainty when tourists enter the unfamiliar tourism destination environment; and help explain the tourism destination perceived environment, so as to achieve the clear and understandable features required by the tourism terrain image.

5. Conclusions

This paper studies and explores some key technologies of deep learning in the field of computer vision symbols. Combining the old and new computer vision problems such as vision-oriented image retrieval and classification, some corresponding key technologies and solutions are proposed. Try to sort out the design principles of individuality and commonness, the unity of nationality and modernity, and the function and form of visual symbols of urban tourism image. In view of this, this topic studies semiotics, urban tourism image, visual image design, and other related theoretical monographs. Various interactive display modes analyze the design methods and future development trends of the visual symbols of urban tourism image from three perspectives: image generalization, collection induction, and symbol implication. Guide and help tourists perceive the tourism image and activity functions on site; complete tourism activities conveniently, quickly, and clearly; eliminate the tension caused by uncertainty when tourists enter the unfamiliar tourism destination environment; and help explain the tourism destination’s perceived environment, so as to realize the clear and understandable features required by the tourism destination image.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] J. Chen, Z. Luo, Z. Zhang et al., “Polar transformation on image features for orientation-invariant representations,” *IEEE Transactions on Multimedia*, vol. 21, no. 2, pp. 300–313, 2018.
- [2] R. Deng, “The relationship between visual symbol elements and regional culture: a study of grey carving in the Pearl River Delta,” *World Scientific Research Journal*, vol. 6, no. 7, pp. 193–204, 2020.
- [3] J. Pongthanaisawan, W. Wangjiraniran, K. Chuenwong, and L. Pimonsree, “Scenario planning for low carbon tourism city: a case study of NAN,” *Energy Procedia*, vol. 152, pp. 715–724, 2018.
- [4] G. Li, L. Yu, and S. Fei, “A deep-learning real-time visual SLAM system based on multi-task feature extraction network and self-supervised feature points,” *Measurement*, vol. 168, no. 4-5, article 108403, 2021.
- [5] S. N. Suhaimi and T. A. Fauzi, “Visual semiotics: identity reflection in personal symbol creation,” *KUPAS SENI*, vol. 9, no. 1, pp. 93–103, 2021.
- [6] X. Lu, S. Yao, G. Fu, X. Lv, and Y. Mao, “Dynamic simulation test of a model of ecological system security for a coastal tourist city,” *Journal of Destination Marketing & Management*, vol. 13, pp. 73–82, 2019.
- [7] J. Gutiérrez, J. C. García-Palomares, G. Romanillos, and M. H. Salas-Olmedo, “The eruption of Airbnb in tourist cities: comparing spatial patterns of hotels and peer-to-peer accommodation in Barcelona,” *Tourism Management*, vol. 62, pp. 278–291, 2017.
- [8] P. I. A. Daniel, “Characterizing the location of tourist images in cities. Differences in user-generated images (Instagram), official tourist brochures and travel guides,” *Annals of Tourism Research*, vol. 73, pp. 103–115, 2018.
- [9] L. Zhou and J. Cho, “A new method of design based on genetic algorithm analysis of the application of traditional cultural symbols in visual communication design,” *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 13, pp. 3401–3408, 2019.
- [10] H. Chen, S. Wang, H. Guo et al., “Study of marine debris around a tourist city in East China: implication for waste management,” *Science of the Total Environment*, vol. 676, pp. 278–289, 2019.
- [11] S. P. Ramos and L. Mundet, “Tourism-phobia in Barcelona: dismantling discursive strategies and power games in the construction of a sustainable tourist city,” *Journal of Tourism and Cultural Change*, vol. 19, no. 1, pp. 113–131, 2021.
- [12] K. Waiyasuri and S. Chotpantararat, “Spatial evolution of coastal tourist city using the Dyna-CLUE model in Koh Chang of Thailand during 1990–2050,” *ISPRS International Journal of Geo-Information*, vol. 11, no. 1, p. 49, 2022.
- [13] Q. Lu, “Research on visual communication of cultural symbols based on central composite design methodology,” *Revista de la Facultad de Ingenieria*, vol. 32, no. 13, pp. 511–515, 2017.
- [14] L. Xuchang, “Design and application of Nanning tourism brand system based on visual orientation of regional characteristic cultural elements,” *Proceedings of the 5th International Conference on Arts, Design and Contemporary Education*, vol. 2, pp. 300–313, 2019.
- [15] Z. Sarba, S. Kmürücü, and M. E. Güler, “Visual sense of the cities and the impact on tourism: case of Izmir,” *International Journal of Social Sciences and Humanity Studies*, vol. 12, pp. 32–1054, 2017.

Retraction

Retracted: Innovation Capacity, Industrial Structure, and Regional Development of High Quality: Empirical Research Based on the County Level

Journal of Function Spaces

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 2023

Copyright © 2023 Journal of Function Spaces. 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. Zhao and H. Yao, "Innovation Capacity, Industrial Structure, and Regional Development of High Quality: Empirical Research Based on the County Level," *Journal of Function Spaces*, vol. 2022, Article ID 4446292, 13 pages, 2022.

Research Article

Innovation Capacity, Industrial Structure, and Regional Development of High Quality: Empirical Research Based on the County Level

Chuhao Zhao ¹ and Hui Yao ²

¹School of Public Policy and Management, Tsinghua University, Beijing 100084, China

²College of Humanities and Development, China Agricultural University, Beijing 100193, China

Correspondence should be addressed to Hui Yao; b20183120787@cau.edu.cn

Received 5 May 2022; Revised 26 May 2022; Accepted 28 May 2022; Published 8 September 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Chuhao Zhao and Hui Yao. 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 key to China's regional high-quality development lies in county-level coordination and promotion. Therefore, based on 2847 county-level administrative units in China, this paper uses empirical analysis methods such as dynamic spatial panel Dobbins model and intermediary effect model to study. Then we test the relationship between the industrial structure of county innovation capacity and regional development of high quality. The results showed that (1) the regional development of high quality at the county level in China was generally distributed in clusters, and the agglomeration had a further trend of strengthening. (2) The innovation ability and industrial structure also had a positive effect on the regional development of high quality at the county level, and the result was significant. (3) Innovation had both direct and indirect effects on regional development of high quality. Industrial structure and regional development of high quality had strong spatial and temporal dependence effects.

1. Introduction and the Presentation of Questions

As China has achieved a historic victory in achieving its goal of poverty alleviation, it has accomplished the huge task of eliminating absolute poverty for the largest population in the world. As China enters the postpoverty alleviation era, we need not only prevent vulnerable groups on the edge of poverty and areas with poor quality from returning to poverty, but also eliminate relative poverty. So we can narrow the income gap, achieve synchronous income increase, and effectively link poverty alleviation with the rural revitalization strategy. In 2021 the government work report was referred to “speed up the development of rural industry and strengthen county economies, broadening the employment channels” of farmers, and stressed that “strengthen rural basic public services and public infrastructure construction and promote urban and rural integration development” in the county. The characteristics of the county

economy in China are becoming the center, the town as the link. As the most basic unit of national economic operation, a series of measures, such as consolidating the achievements of poverty alleviation, promoting farmers' income increase, and promoting rural revitalization strategy, will be implemented at the county level [1]. The 14th Five-Year Plan points out that the theme of the 14th Five-Year Plan period is regional development of high quality. Not only has the economic development entered the stage of regional development of high quality, but also the society, ecology, culture, and national governance system have entered the stage of regional development of high quality. The county's regional development of high quality will be the key to achieve socialist modernization. At present, there are nearly 3,000 county-level administrative regions in China, which are extremely uneven in geographical distribution. Therefore, it is very important to find out the key factors affecting the regional distribution differences of regional development of high quality in county areas for comprehensively promoting

the rural revitalization strategy and regional collaborative development in the next step.

2. Literature Review

High-quality development is the latest development concept which is put forward by China as it enters the new era. Its core lies in paying more attention to the quality and efficiency of development, which has a certain enlightenment significance for the development concept of all countries in the world. Innovation is the fundamental power to promote the development of the economy to a higher stage and has been widely concerned by scholars for a long time. In 1934, Schumpeter emphasized the role of innovation for the first time in his *Theory of Economic Development*, which caused extensive attention to innovation in the academic world. Innovation is the main driving force supporting the construction of the modern economic system. "There is a strong coupling coordination relationship between technological innovation and the high quality of economic development [2], especially in the county level. Because scientific and technological innovation enterprises and first-class scientific research institutions are located in the urban circle. Therefore, innovative talents are first concentrated in large urban agglomerations. On the basis of the radiating and driving play of science and technology innovation, it is an emerging impact on the traditional process of transformation and upgrading of industrial structure. This will lead to the unbalanced spatial distribution of regional development of high-quality counties.

2.1. Innovation Capability and Regional Development of High Quality. Since Schumpeter first proposed the role of innovation, it has become a consensus that technological progress has a positive impact on economic development. Many scholars have empirically tested the positive effect of technological innovation on social, economic, and regional development [3–5]. Specifically, the role of innovation is reflected as follows:

Firstly, technological innovation diffusion drives industrial structure mutation and upgrading and market innovation. Technological innovation improves the overall performance of the market by changing the mode of enterprise benefit growth and promotes the transformation from capital and labor-driven to innovation-driven [6], from the imitator business model from 1 to N to the original model like from 0 to N [7]. Comprehensive innovation performance in different fields can be improved by comparing products, processes, and services [8].

Secondly, the transformation and application of technological innovation achievements will promote the formation of "industry-university-research-enterprise-city" business alliance [9]. Knowledge spillovers and diffusion among innovative enterprises not only interact with related industries and relevant innovation subjects but also form innovation connections with universities and research institutions, drive the integration and aggregation of scientific research resources, and provide the efficiency of industry-university-research collaborative innovation [10]. It will go through the intermediate product or technology transfer to the market demands ulti-

mately. So it will meet the realistic needs of economic society and technological innovation development [11].

Thirdly, the creation of an innovative system and innovative environment will force the improvement of government governance ability. Science and technology innovation needs innovative talents cultivation and other aspects of financial investment joint implementation. Asked the government to strengthen and support the attention of the scientific and technological innovation, an environment conducive to scientific research and development of innovation. Government research, for example, creates an environment that facilitates research and development. At the same time, it could reduce the uncertainty of the innovation system and also needs a series of science and technology policies as a guarantee, which is to improve national innovation ability and the comprehensive strength of the important means [12], through mastering key technologies to form leading achievements and promote the establishment of the modernization-oriented industrial system [13].

2.2. Innovation Capacity and Industrial Structure. China's industrial development system is characterized by the interactive evolution of economic growth and industrial structure change. And innovation-driven development plays a driving role in the evolution and upgrading of industrial structures. It can not only directly affect economic growth, but also affect total factor productivity through the factor allocation effect and then indirectly affect economic growth [14].

Firstly, human innovation improves labor productivity [15], because the level of labor education and the quality of labor force have effectively improved labor productivity and formed a dynamic mechanism of knowledge innovation. Accordingly, the production efficiency of the department is also improved.

Secondly, through improving the market demand structure of product innovation, innovative technology will give birth to new products. With the emergence of new products, consumers have more choices. As market participants exert pressure on product demand, market share forces market participants to introduce new technologies to promote the development and upgrading of new products. Therefore, the original industrial chain extends upward and downward and is improved [16].

Finally, R&D innovation changes the life cycle of enterprises. The departments that carry out technological innovation first will have crowding out effect on the old departments with backward technology and low efficiency, greatly shortening the life cycle of these enterprises and promoting the positive evolution of the overall industrial structure [17].

From the review of the above literature, at present, the focus of academic circles is on how scientific and technological innovation impacts the upgrading of industrial structure. But it has not been found that the three have been brought into a unified framework and their logical relationship and mechanism of action have been deeply discussed. Regional development of high quality, as an evolving overall concept of development, is the result of a highly modernized and evolving economic system, social system, and

TABLE 1: Evaluation index system of county regional development of high quality.

The first-level indicators	The secondary indicators	Basic indicators
The development of kinetic energy	Innovation power	The added value of cultural and creative industries accounted for the proportion of GDP
	The innovation of human	Number of universities and research institutes The number of 500 private enterprises
The development of the structure	Coordination of urban and rural areas	Engel coefficient of urban households
		Per capita disposable income to per capita GDP
	Ability to open	Per capita disposable income of urban and rural residents
		Number of commercial and trade enterprises
Development of effective	Green development	Number of logistics express outlets
		The density of road network
		Added value to accommodation and catering industry
	Results the shared	The surface water reaches or exceeds the proportion of the three types of water bodies
		Proportion of days with good air quality
		The reduction rate of water consumption per ten thousand yuan of GDP
		Electricity consumption per ten thousand yuan of GDP
Results the shared	The decrease rate of construction land per unit of GDP	
	Percentage of townships covered by public transportation	
	Proportion of congested roads	
		Second-hand housing price
		Coverage of medical and health institutions in townships and towns

institutional system [18], showing the characteristics of integrity, systematization, and structure. Knowledge economy, as the best practice to change the economic nature and human lifestyle [19], is the key to catch up with the first countries. With the transformation of the main social contradiction in China, the model of resource allocation will be transformed from government-led to market-led, which will determine the corresponding change of China's industrial system and industrial structure [20]. There is a close logical relationship between the transformation of old and new driving forces, upgrading of industrial structure, and transformation of development mode. Based on previous studies, this paper will expand as follows: First, from the perspective of research, we will take scientific and technological innovation ability as a direct influencing factor, and regional development of high quality will become its final result. Industrial structure upgrading will connect that two factors together just like a bridge. So this three factors will establish an effective connection and believe that scientific and technological innovation ultimately promotes regional development of high quality through the intermediary role of industrial structure. Second, from the perspective of research method, it investigates from the county level. In order to provide empirical evidence for theoretical hypothesis, dynamic panel space Dubin model which can overcome time lag effect, space lag effect, and endogeneity problem at the same time is adopted.

3. Theoretical Model

3.1. *Variable Selection.* Explained variable: comprehensive index of regional development of high quality ($rdhq_{i,t}$). As a

comprehensive indicator, high-quality development has a wide range of meanings, and replacing one or several indicators will inevitably lead to deviation. Therefore, it is concluded by establishing an indicator system. Regional development of high quality as a comprehensive indicator, based on existing studies [21, 22], constructed a county regional development of high-quality evaluation index system, including three first-level indicators of development momentum, development structure and development effect, six second-level indicators, and 19 basic index, as shown in Table 1. The comprehensive index of regional development of high quality of county was calculated according to the evaluation index system of regional development of high quality of county ($rdhq_{i,t}$). Firstly, the basic indexes in the index system were standardized to eliminate the dimensional influence. Secondly, the indexes were synthesized by factor analysis method in SPSS23.0.

Core explanatory variables: county innovation ability: existing studies [23] believe that R&D resources (capital and human expenditure) are the most important factor input affecting regional innovation performance. Therefore, the full-time R&D equivalent of scientific research personnel (ins_{it}) and the proportion of R&D expenditure in GDP (ie_{it}) are selected as the core indicators to measure the innovation capacity of counties [24, 25].

Industrial structure: according to existing studies [26], industrial structure upgrading refers to the process of industrial structure from low level to high level according to the law of economic development. It includes the improvement of the overall technical level, which is measured by the proportion of added value of the tertiary industry (is_{it}). The added value of the tertiary industry proportion can reflect

TABLE 2: Qualitative description of each variable.

Variable categories	Symbol	Meaning	Metrics and descriptions
Explained variable	rdhq	Regional development of high quality in counties	County high-quality development evaluation index system calculated
Explanatory variables	ins	Innovation main body	Researchers develop equivalents full-time
	ie	Innovation environment	Proportion of R&D expenditure in GDP
	Is	The industrial structure	Proportion of added value of tertiary industry
Control variables	ly	Per capita income	Per capita disposable income of urban residents
	ur	Level of urbanization	Permanent population urbanization rate
	ltp	Advances in technology	Amount of technology trading contract
	wl	Labor force level	Proportion of working-age population
	lop	Opening to the outside world	Foreign direct investment
	ps	Policy support	Proportion of added value of high-tech industry
	pg	GDP per capita	GDP per capita

the relation between the ratio between the industry, which is one of the commonly used indicators of the world and academia.

Control variables: in consideration of the influence of other factors on regional development of high quality of counties, control variables such as per capita income, urbanization level, and technological progress are selected, as shown in Table 2:

3.2. Model Specification. In order to test whether industrial structure can act as a mediating variable, the mediation effect model is adopted for empirical study, and the indirect effect of explanatory variable X on explained variable Y through intermediary variable M is referred to as the mediation effect [27], as shown in Figure 1:

$$Y = cX + e_1, \quad (1)$$

$$M = aX + e_2, \quad (2)$$

$$Y = c'X + bM + e_3. \quad (3)$$

The testing procedures of mediation effect in existing studies [28] include the following: First, it is necessary to test the regression coefficient C in Formula (1). If the regression coefficient C is found to be significant in the test results, it can be considered that there is a mediation effect in the model. However, whether the regression coefficient C is significant or not, subsequent testing steps should be carried out. Secondly, it is necessary to test the regression coefficients A and B in Formulas (2) and (3). If both are significant, indirect effects are considered. Finally, the coefficient C' in Formula (3) is tested to observe whether the coefficient is significant. If c' is not significant, it is considered that there is only a mediation effect but no direct effect, which is called complete mediation effect. If c' is significant, the direct effect is considered to exist at the same time, which is called partial mediation effect.

According to the above analysis, the innovation ability of science and technology, industrial structure, and regional development of high quality into test model, because of the

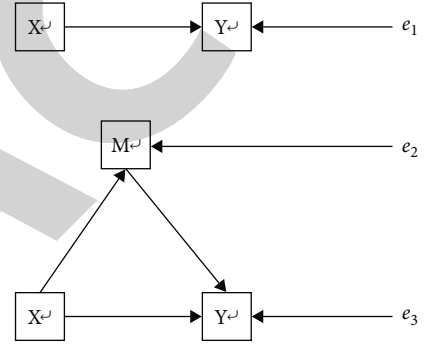


FIGURE 1: The working flow of mediating effect model.

difference of China's nearly three thousand county-level administrative divisions, is more obvious, and the county territory economic growth and development of high quality has obvious agglomeration in the spatial distribution of the trend, so we need to consider its inherent spatial spillover effects. In spatial metrology, the spatial autocorrelation means that the second-order effect of data is generated by the similarity of variables in adjacent regions, including Moran's I , Geary's C , Getis, and Join Count. Moran's I is selected for measurement, which is usually expressed as

$$I = \frac{n}{\sum_{i=1}^n (y_i - \bar{y})^2} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}}. \quad (4)$$

Among Formula (4), I is the spatial autocorrelation index with a value range of $(-1, 1)$, and $I > 0$ is a positive correlation, reflecting the agglomeration effect. When $I < 0$, the correlation is negative, reflecting the dispersion effect. $I = 0$ means no correlation, and the spatial features show randomness. w_{ij} is the spatial weight matrix, which represents the proximity between region i and region j . The significance test of Moran's I is realized by constructing a statistic $Z = (I - E(I)) / (\text{Var}(I))^{-1/2}$ subject to normal distribution. When Z value is significantly positive, interregional similarity tends to cluster; when Z value is significantly negative, interregional

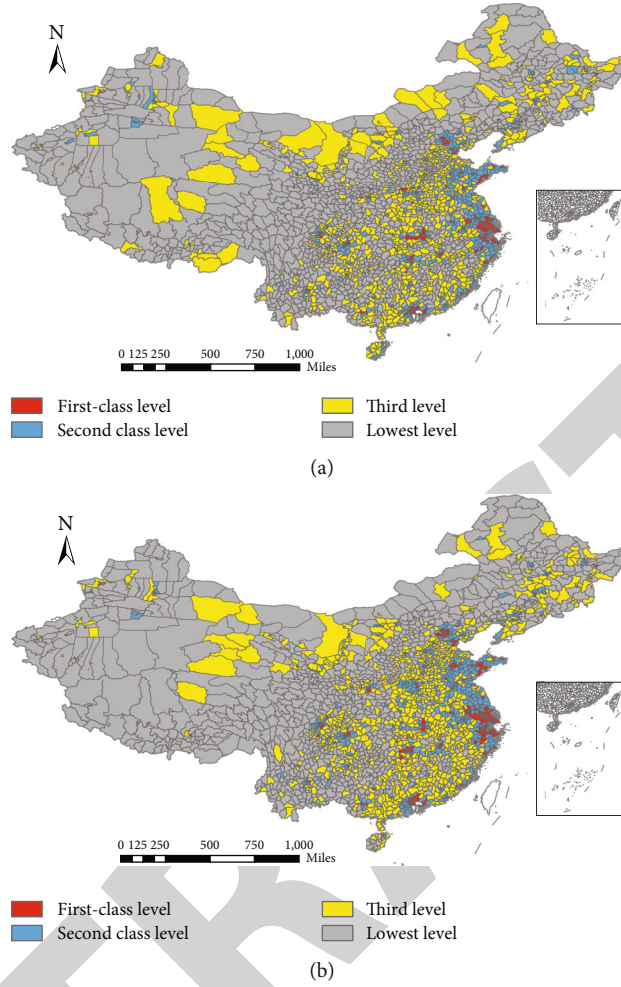


FIGURE 2: Spatial-temporal evolution of regional development of high quality at county level in 2019-2020. Note: The annual data of 2019 and 2020 are the average data of each month.

similarity tends to disperse; when Z value is 0, interregional similarity tends to randomly distribute [29].

County-level administrative region is the most basic administrative divisions in China, with a large number and wide distribution. Therefore, spatial attributes need to be considered in model testing, which is presented by spatial weight matrix. Considering the spatial spillover effect, the spatial Dubin model was used to analyze the general panel regression by adding spatial attributes. The spatial Dubin model has a good reflection on the spatial correlation of explanatory variables themselves, explanatory variables, and error terms [30]. Considering that high-quality regional development has strong path dependence on time, that is, time lag effect, and endogenous problems caused by two-way causality. In addition, considering that regional development of high quality has a strong path dependence on time, namely time lag effect, And endogenous problems caused by bidirectional causality [31]. Therefore, the explained variable, namely the lagged first-period variable of high-quality comprehensive development index, will be introduced into our research model, which is the standard static space panel Dubin model. In a word, we deal with the possible endogeneity problems in the model from two aspects of method and variable. So we can construct

the dynamic space panel Dubin model, as shown below:

$$\begin{aligned}
 rdhq_{i,t} = & \beta_0 + \beta_1 rdhq_{i,t-1} + \rho_1 \sum_{i=1}^n w_{ij} rdhq_{jt} \\
 & + \beta_2 ins_{it} + \rho_2 \sum_{i=1}^n w_{ij} ins_{jt} + \beta_3 ie_{it} \\
 & + \rho_3 \sum_{i=1}^n w_{ij} ie_{jt} + \beta_4 is_{it} + \rho_4 \sum_{i=1}^n w_{ij} is_{jt} \\
 & + \delta \sum X_{it} + \lambda \sum_{i=1}^n w_{ij} X_{jt} + u_i + \varepsilon_{it},
 \end{aligned} \tag{5}$$

$$\begin{aligned}
 rdhq_{i,t} = & \alpha_0 + \alpha_1 rdhq_{i,t-1} + \pi_1 \sum_{i=1}^n w_{ij} rdhq_{jt} \\
 & + \alpha_2 ins_{it} + \pi_2 \sum_{i=1}^n w_{ij} ins_{jt} + \alpha_3 ie_{it} \\
 & + \pi_3 \sum_{i=1}^n w_{ij} ie_{jt} + \varphi X_{it} + \pi_4 \sum_{i=1}^n w_{ij} X_{jt} \\
 & + v_i + \xi_{it},
 \end{aligned} \tag{6}$$

TABLE 3: Moran's I results of county innovation capacity and regional development of high quality.

Variables	I	$E(I)$	$sd(I)$	z	p value*	Variables	I	$E(I)$	$sd(I)$	z	p value*
rdhq1	0.113	-0.000	0.001	97.577	$p \leq 0.001$	rdhq13	0.114	-0.000	0.001	98.938	$p \leq 0.001$
ins1	0.120	-0.000	0.001	103.365	$p \leq 0.001$	ins13	0.128	-0.000	0.001	110.871	$p \leq 0.001$
ie1	0.048	-0.000	0.001	41.823	$p \leq 0.001$	ie13	0.041	-0.000	0.001	40.574	$p \leq 0.001$
rdhq2	0.111	-0.000	0.001	96.134	$p \leq 0.001$	rdhq14	0.102	-0.000	0.001	88.283	$p \leq 0.001$
ins2	0.126	-0.000	0.001	108.452	$p \leq 0.001$	ins14	0.132	-0.000	0.001	114.010	$p \leq 0.001$
ie2	0.048	-0.000	0.001	41.756	$p \leq 0.001$	ie14	0.041	-0.000	0.001	40.382	$p \leq 0.001$
rdhq3	0.104	-0.000	0.001	90.047	$p \leq 0.001$	rdhq15	0.103	-0.000	0.001	89.215	$p \leq 0.001$
ins3	0.127	-0.000	0.001	109.662	$p \leq 0.001$	ins15	0.132	-0.000	0.001	114.044	$p \leq 0.001$
ie3	0.048	-0.000	0.001	41.992	$p \leq 0.001$	ie15	0.039	-0.000	0.001	39.563	$p \leq 0.001$
rdhq4	0.103	-0.000	0.001	88.742	$p \leq 0.001$	rdhq16	0.099	-0.000	0.001	85.534	$p \leq 0.001$
ins4	0.127	-0.000	0.001	109.665	$p \leq 0.001$	ins16	0.132	-0.000	0.001	114.394	$p \leq 0.001$
ie4	0.048	-0.000	0.001	41.535	$p \leq 0.001$	ie16	0.039	-0.000	0.001	39.765	$p \leq 0.001$
rdhq5	0.112	-0.000	0.001	96.758	$p \leq 0.001$	rdhq17	0.101	-0.000	0.001	87.370	$p \leq 0.001$
ins5	0.125	-0.000	0.001	108.313	$p \leq 0.001$	ins17	0.133	-0.000	0.001	114.844	$p \leq 0.001$
ie5	0.047	-0.000	0.001	41.166	$p \leq 0.001$	ie17	0.039	-0.000	0.001	39.267	$p \leq 0.001$
rdhq6	0.099	-0.000	0.001	86.024	$p \leq 0.001$	rdhq18	0.102	-0.000	0.001	88.377	$p \leq 0.001$
ins6	0.125	-0.000	0.001	108.301	$p \leq 0.001$	ins18	0.133	-0.000	0.001	114.801	$p \leq 0.001$
ie6	0.047	-0.000	0.001	41.060	$p \leq 0.001$	ie18	0.038	-0.000	0.001	38.529	$p \leq 0.001$
rdhq7	0.104	-0.000	0.001	90.203	$p \leq 0.001$	rdhq19	0.102	-0.000	0.001	88.720	$p \leq 0.001$
ins7	0.124	-0.000	0.001	107.250	$p \leq 0.001$	ins19	0.133	-0.000	0.001	115.094	$p \leq 0.001$
ie7	0.047	-0.000	0.001	40.983	$p \leq 0.001$	ie19	0.038	-0.000	0.001	38.918	$p \leq 0.001$
rdhq8	0.104	-0.000	0.001	89.669	$p \leq 0.001$	rdhq20	0.103	-0.000	0.001	89.453	$p \leq 0.001$
ins8	0.127	-0.000	0.001	110.136	$p \leq 0.001$	ins20	0.134	-0.000	0.001	115.349	$p \leq 0.001$
ie8	0.047	-0.000	0.001	40.817	$p \leq 0.001$	ie20	0.038	-0.000	0.001	38.715	$p \leq 0.001$
rdhq9	0.103	-0.000	0.001	88.971	$p \leq 0.001$	rdhq21	0.101	-0.000	0.001	87.391	$p \leq 0.001$
ins9	0.127	-0.000	0.001	109.899	$p \leq 0.001$	ins21	0.133	-0.000	0.001	115.248	$p \leq 0.001$
ie9	0.047	-0.000	0.001	40.574	$p \leq 0.001$	ie21	0.039	-0.000	0.001	39.775	$p \leq 0.001$
rdhq10	0.112	-0.000	0.001	96.579	$p \leq 0.001$	rdhq22	0.102	-0.000	0.001	88.342	$p \leq 0.001$
ins10	0.127	-0.000	0.001	110.154	$p \leq 0.001$	ins22	0.132	-0.000	0.001	114.399	$p \leq 0.001$
ie10	0.046	-0.000	0.001	40.043	$p \leq 0.001$	ie22	0.038	-0.000	0.001	38.791	$p \leq 0.001$
rdhq11	0.112	-0.000	0.001	96.686	$p \leq 0.001$	rdhq23	0.106	-0.000	0.001	91.694	$p \leq 0.001$
ins11	0.128	-0.000	0.001	110.184	$p \leq 0.001$	ins23	0.132	-0.000	0.001	114.368	$p \leq 0.001$
ie11	0.047	-0.000	0.001	41.040	$p \leq 0.001$	ie23	0.038	-0.000	0.001	39.070	$p \leq 0.001$
rdhq12	0.111	-0.000	0.001	96.022	$p \leq 0.001$	rdhq24	0.114	-0.000	0.001	98.566	$p \leq 0.001$
ins12	0.128	-0.000	0.001	110.670	$p \leq 0.001$	ins24	0.133	-0.000	0.001	115.296	$p \leq 0.001$
ie12	0.047	-0.000	0.001	41.184	$p \leq 0.001$	ie24	0.038	-0.000	0.001	38.792	$p \leq 0.001$

$$\begin{aligned}
 is_{it} = & \eta_0 + \eta_1 is_{i,t-1} + \theta_1 \sum_{i=1}^n w_{ij} is_{jt} + \eta_2 ins_{it} \\
 & + \theta_2 \sum_{i=1}^n w_{ij} ins_{jt} + \eta_3 ie_{it} + \theta_3 \sum_{i=1}^n w_{ij} ie_{jt} \quad (7) \\
 & + \kappa X_{it} + \theta_4 \sum_{i=1}^n w_{ij} X_{jt} + \chi_i + \tau_{it}.
 \end{aligned}$$

According to the above analysis, the comprehensive index of regional development of high quality ($rdhq_{i,t}$) is the explained variable, the county innovation ability (ins_{jt} , ie_{it}) is the core explanatory variable, and the industrial structure (is_{it}) is the intermediary variable. Formula (5) is the Formula (3) in the corresponding mediation effect model. Equations (6) and (7) correspond to Equations (1) and (2). In order to further control possible endogeneity

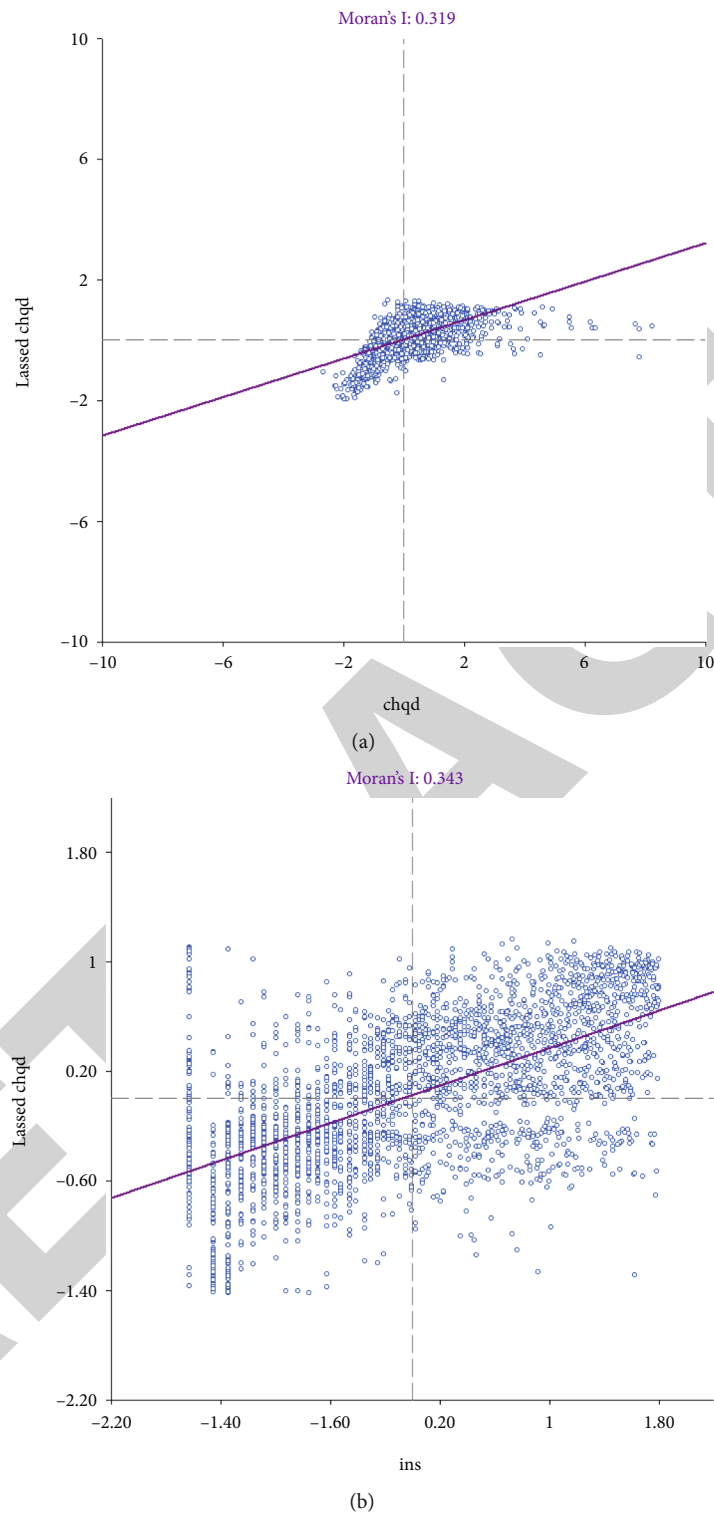


FIGURE 3: Moran's I scatter diagram of county innovation capacity and regional development of high quality. Note: Due to the limitation of the length chart, only the calculation results in December 2020 are given, and the calculation results in other periods are similar.

problems, all explanatory variables are empirically analyzed using data of one period behind. In addition, in model testing, we adopt natural logarithm for variable data measured by nonpercentage indexes to reduce the dispersion degree of sample data.

3.3. Data Sources and Estimation Methods. The data used in our analysis are divided into dynamic data and static data. The static data comes from China county-level statistical yearbook, and the dynamic data is obtained by crawler. In this way, the monthly data of 2847 county-level

TABLE 4: Estimated results of county innovation capacity, industrial results, and regional development of high quality.

Variable	Nonspace OLS	Nonspatial normal panel model (FE)	Nonspatial dynamic panel model (SYS-GMM)	Dubin model of static space (GSPA2SLS)	Dubin model of dynamic space panel (Han-Phillips GMM)
	Model 1	Model 2	Model 3	Model 4	Model 5
L.rdhq			0.907 (440.87)***		0.827 (329.71)***
ins	0.370 (98.21)***	-0.034 (7.15)***	0.032 (27.67)***	0.327 (89.46)***	0.076 (42.95)***
ie	1.391 (8.72)	-0.074 (0.36)	0.037 (2.22)**	0.966 (7.05)***	0.094 (5.84)***
is	0.412 (13.63)***	0.160 (4.44)***	0.101 (3.35)***	0.256 (8.94)***	0.205 (6.72)***
ly	15.233 (30.37)***	11.310 (4.66)***	1.744 (17.63)***	21.988 (39.92)	4.694 (30.37)***
ur	0.227 (41.96)***	-0.009 (0.18)	0.017 (12.89)***	0.191 (36.92)	0.026 (15.99)***
ltp	0.687 (25.96)***	-0.101 (2.60)***	0.067 (9.71)***	0.746 (29.38)***	0.182 (20.99)***
wl	0.280 (18.55)***	0.114 (3.44)***	0.035 (9.21)***	0.325 (21.25)***	0.040 (7.58)***
lop	1.195 (23.63)***	-0.518 (4.49)***	0.278 (21.68)***	1.320 (23.68)***	0.352 (19.35)***
ps	0.131 (13.22)***	0.085 (4.97)***	-0.018 (7.57)***	0.117 (11.84)***	0.051 (18.01)***
pg	6.993 (39.27)***	5.730 (11.76)***	0.678 (20.04)***	6.128 (36.19)***	1.040 (28.87)***
w.rdhq				1.599 (67.58)***	0.481 (67.75)***
w.ins				-0.455 (7.62)***	-0.793 (26.89)***
w.ie				0.054 (1.43)	0.321 (14.75)***
w.is				10.767 (17.98)*	-0.167 (181.73)***
_cons	-166.055 (83.35)***	-80.044 (12.04)***	-18.340 (25.16)***	-177.587 (90.58)***	-30.658 (35.32)***
R ²	0.7013	0.3755		0.7402	
N	66264	66264	63503	66264	
F (Wald) [p]	7503.14 [≤ 0.001]	34.50 [≤ 0.001]	2119.52 [≤ 0.001]	121.28 [≤ 0.001]	139.91 [≤ 0.001]
AR (1) [p]			-11.68 [≤ 0.001]		-11.36 [≤ 0.001]
AR (2) [p]			5.96 [≤ 0.001]		9.95 [≤ 0.001]
Sargan [p]			261.98 [≤ 0.001]		199.53 [≤ 0.001]

Note: the T value or Z value is shown in parentheses, and the values in brackets are concomitant probability. ***, **, and * represent the significance level of 1%, 5%, and 10%, respectively. The spatial dynamic panel model reported the Wald test, same as the nonspatial dynamic panel model (SYS-GMM). And the rest of the other models reported the F -test. Due to space limitation, the estimation results of spatial lag coefficient of each control variable are not given. The following table is the same.

administrative regions in China for a total of 24 months from January 2019 to December 2020 were obtained. Some missing data were filled by multiple interpolation method. Refer to the Han-Phillips generalized method of moments (GMM) estimation proposed by Han and Phillips [32]. At the same time to facilitate comparison, also report the traditional nonspatial OLS, than fixed effect panel, the nonspatial dynamic panel model of generalized moment estimation results, and the estimation results of generalized spatial panel autoregression two-stage least squares (GSPA2SLS) for static spatial panel model.

4. The Empirical Test

4.1. Time and Space Dynamic Evolution Characteristics of County Regional Development of High Quality. In order to directly reflect the spatiotemporal evolution characteristics of high-quality county development in China in 2019-2020, the above calculation results were plotted by Arcgis10.7 software and classified by natural breakpoint classification method. The results are shown in Figure 2, from

which the following characteristics can be seen: (1) the development of Chinese county high-quality distribution in general "Hu Huanyong Line" as the boundary presents the east-west less, especially focused on the coastal area; this is because the coast itself has better economic foundation, attaches great importance to science and technology innovation investment, and optimizes and upgrades the industrial structure and the strengthening of environmental protection and social welfare level, more conducive to the development of high quality; (2) in cluster distribution, the regional development of high-quality distribution of county areas is mainly the Beijing-Tianjin-Hebei region, Shandong Peninsula, Yangtze River Delta, Pearl River Delta, and Chengdu-Chongqing region as several obvious polar cores and the continuous and flake distribution as the center; (3) the agglomeration trend is further strengthened. According to the annual comparison, it is getting smaller regional development of high-quality areas (cities) and counties in the southwest, northwest, and northeast regions. And the number of regional development of high-quality areas (cities) and counties in the Pearl River Delta, Yangtze River Delta,

TABLE 5: Regression estimation results of county innovation capacity on regional development of high quality.

Variable	Nonspace OLS	Nonspatial normal panel model (FE)	Nonspatial dynamic panel model (SYS-GMM)	Dubin model of static space (GSPA2SLS)	Dubin model of dynamic space panel (Han-Phillips GMM)
	Model 6	Model 7	Model 8	Model 9	Model 10
L.rdhq			0.819 (50.00)***		0.809 (317.72)***
L.is					
ins	0.369 (98.38)***	-0.037 (7.66)***	0.068 (8.63)***	0.322 (89.54)***	0.082 (43.87)***
ie	1.442 (8.69)***	-0.051 (0.23)	-0.058 (0.92)	-0.988 (6.97)***	0.144 (8.98)***
ly	16.132 (32.10)***	14.682 (5.93)***	3.512 (7.60)***	23.126 (41.51)***	4.960 (32.07)***
ur	0.222 (40.78)***	-0.008 (0.16)	0.040 (6.62)***	0.192 (36.97)***	0.031 (19.24)***
ltp	0.684 (25.69)***	-0.112 (2.85)***	0.129 (5.76)***	0.763 (29.68)***	0.221 (25.46)***
wl	0.267 (17.66)***	0.115 (3.39)***	0.056 (4.81)***	0.311 (20.31)***	0.040 (7.38)***
lop	1.262 (24.64)***	-0.528 (4.57)***	0.391 (8.75)***	1.344 (23.89)***	0.447 (23.14)***
ps	0.166 (16.71)***	0.091 (4.90)***	0.009 (1.25)	0.150 (15.80)***	0.066 (23.14)***
pg	6.531 (38.85)***	5.311 (11.71)***	1.052 (6.80)***	5.477 (62.61)***	1.048 (30.01)***
w.rdhq				1.463 (65.78)***	0.510 (72.76)***
w.ins				-0.514 (8.98)***	-0.924 (29.41)***
w.ie				0.188 (5.82)***	0.611 (35.84)***
_cons	-152.641 (86.47)***	-76.034 (11.08)***	-28.334 (11.43)***	-162.415 (95.29)***	-26.642 (40.14)***
	66264		63503	66264	63503
F (Wald) [p]	936.70 [≤ 0.001]	33.77 [≤ 0.001]	135.48 [≤ 0.001]	114.42 [≤ 0.001]	169.91 [≤ 0.001]
AR (1) [p]			-23.71 [≤ 0.001]		-114.90 [≤ 0.001]
AR (2) [p]			4.32 [≤ 0.001]		10.26 [≤ 0.001]
Sargan [p]			223.58 [≤ 0.001]		197.83 [≤ 0.001]

Chengdu-Chongqing, and Central Plains Urban Agglomerations is obviously increasing.

4.2. Spatial Autocorrelation Analysis. According to Formula (4), the global Moran's I results and test values of China's county innovation capacity and regional development of high quality are calculated, as shown in Table 3:

Some conclusions can be drawn from Table 3, like that Moran's I in each period is significantly greater than zero, which means that there is an obvious positive spatial correlation between county innovation ability and regional development of high quality. By calculating local Moran's I and plotting the Moran's I scatter diagram of county innovation ability and regional development of high quality, the results are shown in Figure 3:

The results in Figure 3 show the scatter diagram between county innovation capacity and regional development of high quality; among them the first quadrant is H-H type (high-high) region, the second quadrant is L-H type (low-high) region, the third quadrant is L-L type (low-low) region, and the fourth quadrant is H-L type (high-low) region [33]. Among them, different counties have different correlation characteristics. On the whole, there is an obvious spatial autocorrelation between innovation capability and regional development of high quality at county level in China. Therefore, if such spatial correlation

is ignored in the empirical process, there will be regression bias [34].

4.3. The Impact of County Innovation Capability and Industrial Structure on Regional Development of High Quality. Before constructing the spatial econometric model, it is necessary to construct the spatial weight matrix to reflect the spatial correlation between regions. The appropriate spatial weight matrix has an important influence on the empirical results. Spatial weight matrix is commonly used in existing research, including geographic relating right weight matrix, inverse distance weighting matrix, economic weight matrix, and nested weight matrix [35], because the county innovation is related to the input and output aspects of systemic economic activity, not only need to consider geographical factors when considering the spatial correlation; therefore, according to existing studies [36], the nested matrix is used to construct the spatial econometric model, in which the geographical distance spatial weight matrix (W_{ij}^g) is constructed by the inverse ratio of the great circle distance method, as shown in

$$W_{ij}^g = \begin{cases} \frac{1}{d_{ij}}, & i \neq j, \\ 0, & i = j. \end{cases} \quad (8)$$

TABLE 6: Estimated results of mediation effect between innovation capability and industrial outcome at county level.

Variable	Nonspace OLS	Nonspatial normal panel model (FE)	Nonspatial dynamic panel model (SYS-GMM)	Dubin model of static space (GSPA2SLS)	Dubin model of dynamic space panel (Han-Phillips GMM)
	Model 11	Model 12	Model 13	Model 14	Model 15
L.is			0.898 (295.21)***		0.758 (321.52)***
ins	-0.258 (53.92)***	-0.234 (16.12)***	-0.039 (15.30)***	-0.263 (54.81)***	0.224 (44.12)***
ie	0.941 (18.20)***	2.015 (2.21)**	0.157 (3.60)***	1.044 (14.28)***	0.492 (11.33)***
ly	10.848 (18.50)***	316.418 (28.60)***	2.082 (6.04)***	7.158 (12.18)***	-3.807 (7.27)***
ur	0.007 (0.96)	0.299 (2.51)**	0.003 (0.73)	0.038 (5.35)***	0.089 (17.21)***
ltp	-0.041 (1.03)	-0.768 (4.75)***	-0.022 (1.07)	0.073 (1.88)*	0.120 (5.00)***
wl	-0.068 (2.83)***	-0.054 (0.39)	-0.016 (1.41)	0.011 (0.48)	0.451 (26.33)***
lop	2.452 (33.91)***	1.090 (2.07)**	0.308 (8.81)***	0.725 (10.01)***	0.179 (3.33)***
ps	0.748 (59.45)***	0.462 (6.73)***	0.125 (14.36)***	1.013 (84.13)***	0.398 (36.35)***
pg	-5.078 (27.19)***	-11.760 (7.43)***	-1.238 (8.41)***	-5.985 (34.69)***	-3.675 (35.65)***
w.ins				0.529 (8.32)***	-1.231 (11.45)***
w.ie				0.282 (6.18)	1.728 (23.21)***
w.is				1.328 (53.95)	1.427 (58.00)***
_cons	81.172 (39.92)***	-210.341 (8.76)***	15.529 (11.39)***	95.505 (50.60)***	-0.053 (0.04)
	66264	66264	63503	66264	63503
F (Wald) [p]	139.69 [≤0.001]	200.88 [≤0.001]	1584.19 [≤0.001]	324.84 [≤0.001]	162.99 [≤0.001]
AR (1) [p]			-35.95 [≤0.001]		-113.58 [≤0.001]
AR (2) [p]			10.50 [≤0.001]		11.15 [≤0.001]
Sargan [p]			168.33 [≤0.001]		624.45 [≤0.001]

Among Formula (8) d_{ij} is the great circle distance between district (city) county I and district (city) county J , which can be obtained by latitude and longitude calculation. The spatial weight matrix of economic distance (W_{ij}^e) is constructed by selecting the per capita GDP of each district (city) and county as the matrix element, as shown in

$$W_{ij}^e = W_{ij}^g \text{diag} \left(\frac{\bar{Y}_1}{\bar{Y}}, \frac{\bar{Y}_2}{\bar{Y}}, \dots, \frac{\bar{Y}_n}{\bar{Y}} \right); W_{ij}^{re} = \begin{cases} W_{ij}^e, & i \neq j, \\ 0, & i = j. \end{cases} \quad (9)$$

Among Formula (9) \bar{Y}_j is the average GDP per capita of region (city) and county (I) during the observation period, and \bar{Y} is the average GDP per capita during the total observation period. The advantage of using nested matrices is that both geographical distance and economic ties are considered. Dynamic panel spatial Dubin model is used to regression county innovation capacity, industrial structure, and regional development of high quality, and the results are shown in Table 4:

It can be seen from Table 4 that the time lag term (L.rdhq) and spatial lag term (w.rdhq) coefficients of regional development of high quality are significantly non-zero (at the significance level of 1%), indicating that the time and spatial correlation should be fully considered in the regression analysis; that is, it is reasonable to use the dynamic spatial panel model. By comparing the regression

coefficients of exogenous variables in all models, it can be seen that if the spatial correlation between districts (cities) and counties is ignored, the impact of each factor on regional development of high quality will be overestimated. Compared with model 1 and model 2, it was found that the regression coefficient of innovation environment (ie) did not pass the significance test due to traditional OLS estimation. The spatial lag term of each explanatory variable was added into model 4, and it was found that the spatial lag term of per capita income (ly), urbanization level (ur), and innovation environment (w.i) were not significant. The regression coefficient of the explanatory variable for the first-order lag term of model 5 is positive, indicating that the county has strong inertia. The high-quality core innovation subjects (INS), environmental variables (ie) and industrial structure (is) of regional development are all significant positive regression coefficients. It shows that the county's innovation ability, innovation ability and innovation ability play a positive role in the upgrading of the county's industrial structure and high-quality development.

From the perspective of spatial lag term coefficient, the spatial lag term coefficient of regional development of high quality (w.rdhq) and innovation subject (w.ins) will be overestimated in model 4 without the first-order lag term of explained variables, and the regression coefficient of innovation environment (w.ie) is not significant. Comprehensive view model 5 county regional development of high quality has significant time lag effect and spatial spillover effect; on

the one hand, the county regional development of high quality is a systematic, structural system of comprehensive, from volume expansion to structural optimization, with the inheritance of inertia and time stability of the economic operation [37]. On the other hand, regional development of high quality at the county level needs to focus on the integration of industry and city to promote the effective connection between urban functions and industrial development. In this process, cities and towns become more attractive to the population, which will promote the overall urbanization and industrialization process of the region [38].

4.4. Study on the Intermediary Effect of Industrial Structure. Based on the above analysis, the mediating effect of industrial structure on county innovation capacity and regional development of high quality was investigated, and regression analysis was conducted according to Formulas (1) and (6). The results are shown in Table 5:

As can be seen from Table 5, the regression coefficient of innovation subject (ins) and innovation environment (ie) of the core explanatory variables on the quality development of the explained variable (rdhq) is significantly not zero, suggesting that there is a mediation effect at this time. Combined with (2) and (7), the regression analysis is conducted, and the results are shown in Table 6:

It can be seen from Table 6 that the regression coefficient of the core explanatory variables innovation subject (ins) and innovation environment (ie) on the explained variable industrial structure (is) is significantly not zero. Combined with the results of Table 4, it is believed that the mediation effect of industrial structure exists, and there are both direct effect and indirect effect, namely, partial mediation effect. The estimated results of model 15 are consistent with the previous theoretical expectations. The change of the main social contradiction reflects the transformation from the total contradiction of supply and demand system to the structural contradiction. The basis for promoting structural changes in economic and social operation lies in the expansion of reproduction of the knowledge middle class, the promotion of advanced internal industrial structure through the improvement of employment capacity and the upgrading of consumption structure, and the realization of leap-forward growth and regional development of high quality led by knowledge-intensive industries [39].

5. Conclusion

After the victory in poverty alleviation, rural revitalization should be comprehensively promoted. The key battlefield of Rural Revitalization is the county, and the high-quality development of the county is the key to rural revitalization. This paper takes 2847 county-level administrative units of data as an example, based on the dynamic panel Dubin model, the mediation effect model, and the empirical analysis methods, such as the county innovation ability and the inspection between industrial structure and regional development of high quality. It is found that county innovation is the key factor affecting high-quality development, and industrial structure upgrading plays a mediating role, so it

is necessary to pay attention to county-level industrial structure adjustment to promote high-quality development. The main conclusions are as follows:

- (1) The regional development of high quality at county level in China generally shows a cluster distribution, and the agglomeration tends to be further enhanced. Innovation capability and industrial structure have a significant positive impact on the regional development of high quality at county level. "The seventh census" data to see the future population continue to eastern provinces and center has changed the trend of the urban agglomeration. And the 19th congress points out that the role of the county economy in the future will be more and more obvious. The innovation ability is the key to improve the county, which will promote the fundamentals of the industrial structure
- (2) Innovation has both direct and indirect effects on regional development of high quality. When the people is going into the knowledge economy society, the innovation to promote the development of high quality has a decisive role. On the one hand, the human society has entered a knowledge-based economy society, and the change of social development stage requires the economic structure to change from simple copying and imitation to innovation leading. A high level of scientific research and subvert technology breaks through promoting the commercialization and industrialization of scientific and technological achievements, promoting the establishment of knowledge-intensive and high-value-added industries; the endogenous power of knowledge to economic growth has become more and more obvious, and knowledge groups have become the main groups to create wealth and economic benefits, and their expansion of reproduction promotes the upgrading of consumption structure, such as the increase in the proportion of spending on science, education, culture, and health, which is conducive to the accumulation of human capital and the improvement of the quality of labor force, and promotes the advanced process within the industrial structure
- (3) Industrial structure upgrade has intermediary effect. It has strong space-time-dependent effects with regional development of high quality. On the one hand, the industrial structure has strong path-dependent characteristics and inheritance inertia in time dimension with regional development of high quality. And the current industrial structure and regional development of high-quality characteristics will inherit and retain the characteristics of the previous period to a large extent. There is also a significant spillover effect in space, which is not only driven by synergy under the model's typical demonstration, but also driven by competition under the "political tournament".

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the project of the National Social Science Foundation Youth Project: Research on Public Value Evaluation and Promotion Strategy of Provincial Digital Government in China (No. 20CGLO60).

References

- [1] X. Li, Y. Lu, and R. Huang, "Whether foreign direct investment can promote high-quality economic development under environmental regulation: evidence from the Yangtze River Economic Belt, China," *Environmental Science and Pollution Research*, vol. 28, no. 17, pp. 21674–21683, 2021.
- [2] H. Jian and H. Jinxin, "Evaluation of the coupling relationship between regional scientific and technological innovation and high-quality economic development in China," *Science and Technology Progress and Countermeasures*, vol. 36, no. 8, pp. 19–27, 2019.
- [3] Y. Xue, L. Hu, and Y. Feng, "Researches on influence of different technological innovation methods on development level of China's economic intensity based on dynamic panel generalized moments," *World Sci-Tech R & D*, vol. 37, no. 5, pp. 434–439, 2017.
- [4] N. Brandão Santana, D. A. Rebelatto, A. E. Périco, H. F. Moralles, and W. Leal Filho, "Technological innovation for sustainable development: an analysis of different types of impacts for countries in the BRICS and G7 groups," *International Journal of Sustainable Development and World Ecology*, vol. 22, no. 5, pp. 1–12, 2015.
- [5] Y. Su and X. L. An, "Application of threshold regression analysis to study the impact of regional technological innovation level on sustainable development," *Renewable & Sustainable Energy Reviews*, vol. 89, pp. 27–32, 2018.
- [6] X. Chen and L. I. Yunfeng, "Manufacturing technology innovation dynamic capability and regional development of high-quality: based on innovation leading perspective," *Science and Technology Progress and Countermeasures*, vol. 37, no. 6, pp. 92–101, 2020.
- [7] F. Zhulan, *How will China's innovation weaknesses be remedied*, vol. 25, no. 5, 2016People's forum, 2016.
- [8] U. E. Haner, "Innovation quality—a conceptual framework," *International Journal of Production Economics*, vol. 80, no. 1, pp. 31–37, 2002.
- [9] M. Chen and H. Wang, "Import technology sophistication and high-quality economic development: Evidence from city-level data of China," *Economic Research-Ekonomska Istraživanja*, vol. 35, no. 1, pp. 1106–1141, 2022.
- [10] T. Yigitcanlar, J. M. Corchado, R. Mehmood, R. Y. M. Li, K. Mossberger, and K. Desouza, "Responsible urban innovation with local government artificial intelligence (AI): a conceptual framework and research agenda," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 7, no. 1, p. 71, 2021.
- [11] Y. Guozhong and M. Jinhuang, "On the diffusion path of diversified technological innovation in national high-tech zones," *Science and Technology Management Research*, vol. 14, pp. 18–23, 2016.
- [12] M. Kafouros, 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.
- [13] J. C. Guan and R. C. M. Yam, "Effects of government financial incentives on firms' innovation performance in China: Evidences from Beijing in the 1990s," *Research Policy*, vol. 44, no. 1, pp. 273–282, 2015.
- [14] W. Haibin and Y. Huixin, "Innovation driving and modern industrial development system: an empirical analysis based on provincial panel data in China," *Economic quarterly*, vol. 15, no. 4, pp. 1351–1386, 2016.
- [15] Z. Zhidong and L. Changwen, "Technological innovation and industrial structure upgrading in the Yangtze River Economic Belt: the moderating role of marketization," *Science and technology progress and countermeasures*, vol. 37, no. 7, pp. 26–34, 2020.
- [16] T. Changqi and P. Yongzhang, "The spatial effect of technological innovation intensity on industrial structure upgrading under economic agglomeration," *Industrial Economics Research*, vol. 3, no. 88, pp. 91–103, 2017.
- [17] C. A. I. Yurong and W. Huiling, "Research on the influence mechanism of innovation input on industrial structure upgrading," *Exploration of Economic Issues*, vol. 1, pp. 138–146, 2018.
- [18] X. Xie, S. Zeng, Y. Peng, and C. Tam, "What affects the innovation performance of small and medium-sized enterprises in China?," *Innovation*, vol. 15, no. 3, pp. 271–286, 2013.
- [19] R. M. Unger, *The Knowledge Economy*, Verso, New York, 2022.
- [20] Y. Yang, J. Hong, G. Song, and S. Hong, "Technology policy, technology strategy and innovation performance: evidence from Chinese aircraft and spacecraft manufacturing," *Science and Public Policy*, vol. 44, no. 5, pp. scw089–scw630, 2017.
- [21] Z. Yi, H. Feng, and L. Junfeng, "Does agglomeration of producer services improve the quality of urban economic growth?," *Journal of Quantitative And Technical Economics*, vol. 5, pp. 83–100, 2019.
- [22] H. Yang and X. Zhu, "Research on green innovation performance of manufacturing industry and its improvement path in China," *Sustainability*, vol. 14, no. 13, p. 8000, 2022.
- [23] K. Y. Cheung and L. Ping, "Spillover effects of FDI on innovation in China: evidence from the provincial data," *China Economic Review*, vol. 15, no. 1, pp. 25–44, 2004.
- [24] L. Hedong, "Research on R&D efficiency and its influencing factors in China: an empirical analysis based on stochastic frontier function," *Studies in science of science*, vol. 29, no. 4, pp. 548–555, 2011.
- [25] W. Wu Xianhui, H., L. Qiang, and W. Yang, "An empirical study on the influencing factors of China's regional innovation system – a case study of Shenzhen," *Science & technology progress and countermeasures*, vol. 28, no. 7, pp. 26–31, 2011.

Retraction

Retracted: Modeling of Energy Saving and Comfort of Building Layout in Extreme Weather Urban Residential Area under the Background of Spatial Structure and Form Evolution: Taking Yichun as an Example

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Lu, G. Sun, E. Wang, and Z. He, "Modeling of Energy Saving and Comfort of Building Layout in Extreme Weather Urban Residential Area under the Background of Spatial Structure and Form Evolution: Taking Yichun as an Example," *Journal of Function Spaces*, vol. 2022, Article ID 6892035, 11 pages, 2022.

Research Article

Modeling of Energy Saving and Comfort of Building Layout in Extreme Weather Urban Residential Area under the Background of Spatial Structure and Form Evolution: Taking Yichun as an Example

Ming Lu ¹, Guofei Sun ¹, Endong Wang ², and Zichong He ¹

¹School of Architecture, Harbin Institute of Technology, Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology, Harbin 150001, China

²Yichun Natural Resources Bureau, Yichun 153000, China

Correspondence should be addressed to Guofei Sun; sunguofei@hit.edu.cn

Received 27 June 2022; Revised 18 August 2022; Accepted 22 August 2022; Published 5 September 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Ming 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.

After the founding of the People's Republic of China, the state formulated the first five-year plan for national economic development with the development of heavy industry as the core. The assistance of the Soviet Union completely opened the prelude to China's large-scale industrial construction. With the continuous increase of urban energy consumption in China and the huge pressure it brings to the environment, urban planning and design aiming at energy conservation and climate adaptability has attracted more and more attention. Residential building layout is an important part of urban planning and design, which is closely related to building environmental energy consumption and urban microclimate. Therefore, it is necessary to study the energy saving and climate adaptive design strategy of residential building layout. Yichun is a cold city, located in the northernmost part of the three northeastern provinces, with a population of about 1 million. Its climate is characterized by continuous low temperature in winter, and the extreme value of daily lowest air temperature in Yichun City since 1981 is -42.2°C . The highest temperature in summer is 35°C . The unique geographical location and climate have created Yichun's unique urban quality, and outdoor cultural activities are also unique. Yichun is a place with four distinct seasons, and each season has its own characteristics. The beauty of the four seasons contrasts greatly. The best travel time in Yichun is from June to September. In this study, the scientific nature of environment numerical simulation software is verified. Taking the surrounding residential open space as the research object, reasonable grid simulation size and initial boundary conditions are set. The quantitative relationship between design elements and thermal comfort level is established through software simulation. After repeated and a large number of hourly simulation verification, combined with the spatial distribution map and numerical distribution map of thermal comfort, the design elements and open space are established. On this basis, the optimization strategies of open space in residential areas in extreme weather areas are put forward, including the reasonable layout of activity areas, careful selection of open positions, attention to the proportion of open space, attention to plane enclosure and corner units, encouragement of secondary restrictions of activity space, reasonable selection of underlying surface materials, and reasonable layout of water bodies. This paper mainly studies the design of outdoor communication space in Yichun residential area in winter. The average temperature in winter is $-14^{\circ}\text{C} \sim -2^{\circ}\text{C}$. Through the study of design methods and countermeasures, an operable Yichun residential area design method is found.

1. Introduction

According to the statistics of the International Cold Region Association, more than 600 million people in the world have

lived in cold regions [1]. However, China has a vast land area under the background of severe cold climate. The winter is long and cold. Every year from November to the next April, the city will face the attack of severe cold, ice, and cold

wind [2]. At the same time, because winter is often affected by cold air from Siberia, compared with other countries at the same latitude in the world, the winter temperature in cold cities in my country is much lower, and the average temperature in January is often below -18°C [3]. In the research of building layout energy-saving and comfort modeling in cold regions, through numerical analysis and modeling design, the quantitative characteristics of influencing factors such as solar energy, illumination, cold wind avoidance, energy, building materials, and heating are analyzed, and the influence and contribution of the above factors on building layout energy saving are analyzed. By making full use of sunlight, energy is saved, and comfort of communities is improved, which provides an accurate model and data basis for the design and evaluation of building layout energy-saving and comfort modeling in cold regions [4]. As building energy consumption accounts for a large proportion of the total social energy consumption, building energy conservation has become one of the mainstreams of the world energy conservation wave [5].

Due to air pollution, global greenhouse effect, urban heat island effect, and ecological environment deterioration caused by massive energy consumption, energy consumption has become a serious problem that human beings have to face up to [6]. The concept of sustainable development has made people increasingly aware of energy, environment, and ecological affairs. Issues related to energy use and energy supply are the main factors considered in architectural design and planning and design [7]. For a long time, the design standards and planning control indicators of open space in residential areas are limited to the hard indicators such as sunshine time, greening rate, and floor area ratio, and no effective method to improve the thermal comfort of open space in residential areas has been put forward [8]. From the perspective of fitness for people's physical and mental health, people's physical health is to strengthen their physical quality through exercise, which belongs to physical hardware; people's psychological health refers to people's psychological feelings. A bad environment will give people feelings of depression, boredom, rigidity, etc., and make people feel uncomfortable, so they lose interest in it or even stay away from it. Generally, through the treatment of the physical form, color, texture, and other aspects existing in the space environment, it will dye a certain cultural atmosphere, alleviate people's inner discomfort and insecurity, and enhance neighborhood communication [9]. Due to the special climatic conditions, the open space in cold regions faces the problem of low utilization rate. However, residents in cold regions also long for public activities and comfortable open spaces in residential areas. Therefore, the overall climate in cold regions cannot be changed. Under the objective facts of the environment, in a limited season or time period, select a certain type of representative open space in the cold area and use the existing operational design methods and technical conditions to improve the cold area. The thermal comfort of the open space in the residential area, so as to improve the utilization rate of the open space in the cold area, is the main purpose of this study [10]. Compared with the general building layout, buildings in cold environment

adopt small shape coefficient to reduce the area facing the cold, so as to reduce the loss caused by heat exchange. Good thermal insulation materials can greatly improve the quality of buildings against cold. Usually, buildings in severe cold areas have large windows on the south side, small windows on the north side, or even no windows. Reasonably arranged heating equipment will greatly improve the comfort of buildings in cold areas.

In this paper, Yichun, a cold city, is taken as the research object, and the microclimate environment of public space in typical residential areas in transition season and the use of interviewees are investigated and analyzed within the range of transition season determined in previous studies. Some planning suggestions are put forward to prolong the outdoor season of cold cities, so as to provide reference for the construction of livable cold cities in China.

2. Related Work

From 1961 to 2020, China's annual average precipitation showed an increasing trend, the annual average precipitation days showed a significant decreasing trend, and the annual cumulative number of rainstorm station days showed an increasing trend. In 2020, the cumulative number of rainstorm station days in China was the second since 1961. Extreme heavy rainfall events in China are increasing, extreme low temperature events are decreasing, and extreme high temperature events have increased significantly since the mid-1990s. Since the late 1990s, the average intensity fluctuation of typhoons landing in China has increased. The average number of sand dust days in northern China shows a significant decreasing trend, reaching the lowest value in recent years and rising slightly. China's climate risk index is on the rise. In 2020, China's climate risk index was the third highest since 1961. From the perspective of annual climate change, in the past 30 years, most parts of the country have shown a warming trend, except that most of the Sichuan Basin, a small part of the Yunnan Guizhou Plateau and a small part of the Northeast Qinghai Tibet Plateau have shown a cooling trend. The national average temperature tendency rate is 0.24810 a , and the average precipitation tendency rate is $9.207\text{ mm}/10\text{ a}$. In the whole country, the area showing a warm and wet trend is the most extensive. It includes most of the northeast, Inner Mongolia, Xinjiang, Qinghai Tibet Plateau, the middle and lower reaches of the Yangtze River, and the southern part of the North China Plain. Peng et al. [11] have established a certain research foundation for the research on the wind and heat environment of open space in cold residential areas, and the microclimate of urban open space has gradually attracted the attention of planners and designers and has successively carried out relevant theoretical and practical research. Communication is a unique way of life for people. Communication is ubiquitous. From the ancient times to the present, at home and abroad, as long as there are people, communication is easy to occur. Wang and Sunaga [12] pointed out that the design of outdoor communication space in urban residential areas is more seriously affected by topography and regional climate compared with southern cities in cold cities like

Yichun. Ascione et al. [13] put forward a series of planning countermeasures, such as creating semi-indoor space, paying attention to the demand of pedestrians in winter, improving the activity support of open space, and improving the monotonous landscape environment, in view of the planning and design of public space in cold regions. In the study on livability of urban environment in cold regions and the design countermeasures of urban squares in cold regions, Lah et al., starting from the seasonal characteristics of urban squares in cold regions, put forward countermeasures for climate protection and improving thermal comfort in terms of location selection, scale, spatial level, landscape type, and detail design of squares [14]. In reality, the quality of communication space does not affect residents' outdoor leisure activities, and the lack of design considerations has brought serious problems. For example, many facilities are scattered and seldom used. The outdoor environment of residential areas is not a simple patchwork and accumulation. Simões and Pierce [15] pointed out that the spatial pattern of a residential area refers to the objective expression of the regional characteristics in the urban style, which includes the location setting, functional layout, building arrangement, road system, and surrounding environment of the residential area. According to the five elements of urban design, such as path, area, edge, node, and landmark, Arán-Ais et al. [16] put forward the design countermeasures of outdoor public space in cold cities, such as adopting closed or semiclosed trail system and adopting asymmetric sidewalk design to improve safety and thermal comfort. Based on the software, Davies et al. [17] have made a lot of simulations on the microclimate environment of the block layer gorge; quantitatively analyzed the different plane, section, trend characteristics, and the influences of the underlying surface, green plants, and water bodies on the wind environment of the block; and put forward the corresponding technical countermeasures with the improvement of thermal comfort as the measurement standard. The relevant theories of outdoor thermal comfort are summarized. From the root of the formation mechanism of human thermal sensation, the various elements affecting the thermal environment are explored and summarized. At the same time, the outdoor thermal comfort evaluation software is launched by programming with VB software on the computer. Zhao et al. [18] put forward the concept of outdoor season based on this and believed that in the late spring to early autumn seasons of cold cities, people can achieve a comfortable state by increasing or decreasing clothes under natural conditions. Zhang et al. [19] put forward the concept of transition season on this basis. They believe that there is a transition season between winter and outdoor season in cold cities. Although the temperature is low, there are still a considerable number of citizens in sunny and windless weather conditions. Méndez-Abreu et al. [20] put forward a modeling method of building layout energy saving and comfort in cold areas based on greenhouse effect analysis and built a microclimate living environment under the building layout of cold areas. The indoor space greenhouse effect analysis method was adopted to improve the comfort experience of cold areas through energy saving and thermal insulation design, which

has a good guidance. Kapsalis et al. [21] studied the interaction between ecosystem service function principles and circular economy from the perspective of interorganizational systems. Kanteraki et al. [22] explored how modern architectural designs can be applied to construction and domestication while following traditional construction methods and compared with buildings built and domesticated under bioclimatic architecture.

From March 2021 to July 2022, the author conducted research on the paper, taking Yichun as a representative of cold cities to investigate the residential area. The research methods include observation of outdoor activities of residents in the residential area, field survey of the site, visit survey of residents, and field photos. Through this research work, a large number of data about outdoor communication of residents in residential areas have been collected as the basis of this paper.

3. Methodology

3.1. Constrained Parameter Model and Energy-Saving Objective Function Construction. Through the optimized energy-saving design of building residential areas, carbon dioxide emissions are reduced, costs are reduced, and benefits are improved. The energy-saving and comfort modeling of building layout in cold residential areas analyzed in this paper mainly includes the following aspects: (1) sunlight and sunlight collection [23]: in the layout of buildings in cold regions, the indoor light environment of buildings is an important condition to ensure the constant temperature in the buildings, and a good light environment can also protect human health and improve comfort. Therefore, in the building layout energy-saving and comfort modeling of cold regions, it is necessary to optimize the brightness distribution to avoid damp and cold. (2) Selection of building insulation materials [24]: the thermal insulation material can collect excess heat and release it in a timely and stable manner to avoid condensation, mildew, and peeling of the building. In the selection of building thermal insulation materials, the building thermal insulation materials designed in this paper select organic matter as the base of the main wall, which has excellent adhesion [25]. (3) Temperature problem: in the living environment, in cold regions, temperature is an important indicator that affects human comfort in buildings. Building layout in cold regions requires energy conservation and comfort modeling to achieve energy conservation and environmental protection in addition to ensuring the overall thermal balance of the human body. The selection of thermal insulation materials and the description of thermal insulation wall design for the energy-saving model of building layout in cold regions designed in this paper are shown in Figure 1.

Parametric design is an important method of serialized product design, which is one of the key technologies of the new generation CAD system. Many scholars have done a lot of exploration work in theory and practice and achieved a lot of results. At present, the main research methods are variable geometry method and knowledge-based reasoning method. Based on process construction method and

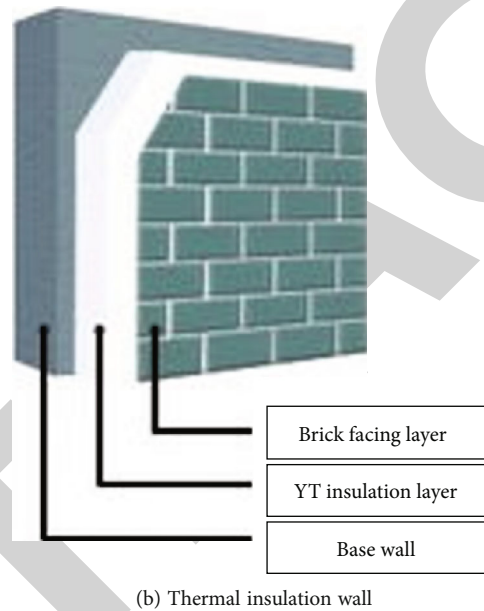
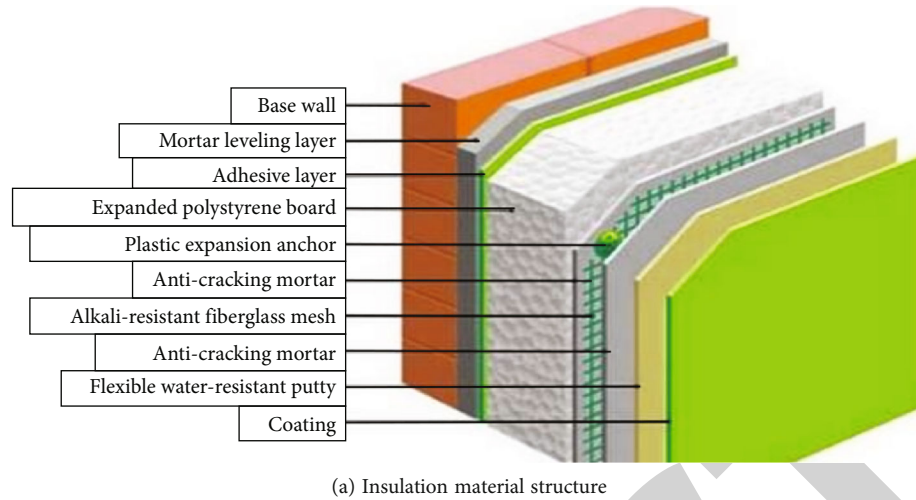


FIGURE 1: Selection of thermal insulation materials and design of thermal insulation walls in building energy-saving models in cold regions.

constraint propagation solution method, these methods focus on the representation and solution of constraints. Although the implementation of the constraint model is also described, it is not unified with the constraint representation model, and a constraint representation model in line with the programming theory is given. And these parametric design methods are general methods for general fields. Finding a parametric design method that can be applied to all industries, even if it is possible, requires long-term research. This paper uses object-oriented theory to analyze geometric constraint system and uses classes to represent primitives and constraints. The class hierarchy of the system is established, and an object-oriented parametric representation model of directed hypergraph is proposed. This model has the advantage of directed hypergraph, and it is easy to implement with object-oriented programming method. This method applies object-oriented technology throughout the whole process of system analysis, modeling, and implemen-

tation and has the characteristics of high efficiency and reliability. And the system is easy to expand and modify, especially suitable for parametric CAD system modeling and program implementation of serial products in a specific industry. The implementation scope of economic energy conservation is oriented to the whole society, and the implementation objects are industries. This paper mainly studies and discusses the methods of building energy-saving target system in the macro field from the perspective of economic energy conservation. It provides calculation methods and analysis tools for scientifically and reasonably decomposing the target of reducing energy consumption per unit of GDP to relevant departments.

In the modeling of energy saving and comfort of building layout in cold area, given the multilayer quantitative information level of building layout in cold area, such as illumination, temperature, humidity, green vegetation, and heating, they are, respectively, recorded as U_1, U_2, \dots, U_n .

Use $P(s_1), P(s_2), \dots, P(s_n)$ to describe the light intensity fluctuation data of buildings in cold regions. When the judgment conditions are $h_1 > 0, h_2 > 0, \mu_1$ and μ , the correlation of evaluation indicators such as light intensity and humidity of buildings in cold regions can be calculated by (a_1, a_2, \dots, a_n) and (t_1, t_2, \dots, t_3) ; the modeling parameters $d_1(t)$ and $d_2(t)$ for building layout energy saving and comfort in cold regions have accompanying characteristic stability, and the multidimensional search iterative equation of the information parameter is described as follows:

$$\begin{aligned} u^{(n+1)}(x, y) &= u^{(n)}(x, y) + \delta u_1^{(n)}(x, y), \\ u_1^{(n)}(x, y) &= M\Delta_s u^{(n)}(x, y) + N\Delta_t u^{(n)}(x, y; d). \end{aligned} \quad (1)$$

There are Ω matching pairs of light intensity in residential areas, and the oxygen release of green vegetation is $E \in \gamma(s)$, which forms a vector space in building zoning [26]. The heating quantity $AB = V$ of buildings in cold residential areas is a unilateral detection function. When $AB = V, AB =$, the heating edge set A converges in a finite field, and we get

$$D(C_1, C_2) = \begin{cases} \text{true if } D \text{ if } (C_1, C_2) > M\text{Int}(C_1, C_2), \\ \text{false otherwise,} \end{cases}$$

$$M\text{Int}(C_1, C_2) = \min(\text{Int}(C_1) + \tau(C_1), \text{Int}(C_2) + \tau(C_1)). \quad (2)$$

Based on multiparameter autoregressive analysis, the characteristic function of heat distribution of building layout in cold area is calculated, which satisfies

$$\varphi(\omega) = E[e^{j\omega X}] = \begin{cases} \exp\left\{j\mu\omega - |a\left[1 - j\beta \operatorname{sgn}(\omega) \tan\left(\frac{\pi a}{2}\right)\right]\right\}, a \neq 1, \\ \exp\left\{j\mu\omega - |a\left[1 + j\beta \operatorname{sgn}(\omega) \frac{2}{\pi} \ln|\omega|\right]\right\}, a = 1, \end{cases} \quad (3)$$

where $n = 1, 2, T$ represents the area number of residential buildings and $VI(x)$ is a k -dimensional lighting intensity state matrix. Through the above analysis, the energy-saving objective function is constructed, and the comfort modeling is realized by seeking the optimal solution of the objective function.

3.2. Comfort Simulation of Open Space in Cold Residential Area. In the above-mentioned parametric model analysis, through the optimal design of the buildings in the cold area, the comfortable experience and feeling of the residents in the severe cold area are increased, and the concept of energy saving, environmental protection, and green buildings is effectively advocated. Aiming at the disadvantages of the traditional model, this paper puts forward a design method and modeling scheme of building energy-saving layout in cold residential areas based on hierarchical integration of building divisions and feature extraction of energy-saving contribution parameters [27]. This paper analyzes the influencing factors of building energy conservation in cold areas, such as

solar energy, heating, thermal insulation effect of thermal insulation wall, and temperature and humidity [28]. In the existing research on the design of building energy-saving system, most of them focus on the analysis of the effect of a single factor [29]. There is no comparative analysis of the importance of various factors on residential energy consumption [30]. There is no research on the effect of energy-saving design based on residential system in combination with the existing actual situation in China, especially the meteorological conditions. It has little guiding significance for residential energy conservation in China. Therefore, in this study, the energy-saving characteristics of various factors in the residential system are different [31]. Therefore, by comparing the energy consumption of residential buildings with different characteristics, we hope to find the relevant factors affecting residential energy consumption and their importance and their overall impact, that is, the energy-saving effect [32]. According to the carbon dioxide emission of energy consumption as the test target parameter, the contribution weight function of building optimization and energy conservation in cold areas is constructed as follows:

$$S(x) = \sum_{i=1}^N v_i(x) \nabla^2 v_i(x), \quad (4)$$

where $n = 1, 2, \dots, n$, $VI(x)$ is a k -dimensional random variable, and the influence weight factor matrix of light, temperature, and humidity $Q_1 \geq Q_1 \geq 0, p > 0, R_1 \geq R_1 \geq 0, Z_1 \geq Z \Lambda > 0$, and $Z_3 > 0$. According to the above analysis, in the architectural design, the contribution weight coefficient matrix of thermal insulation materials to building energy conservation is obtained by using the thermal insulation materials of the platform, female wall, and wall surface of the main building as follows:

$$K = [K_1^T K_2^T K_3^T K_4^T K_5^T]^T, \quad (5)$$

$$P(\gamma_{w3} | x_{w3}, \theta, \beta) = \frac{1}{Z(\beta_i)} P(\gamma_{w3} | x_{w3}, \theta) (\gamma_{w3} | \beta_i),$$

where $Z(\beta_i) = \sum_{\gamma_{w3}} P(\gamma_{w3} | x_{w3}, \theta) (\gamma_{w3} | \beta_i)$ is the gray statistic value of building energy-saving and comfort modeling indicators for building layout in cold regions. The fitness evaluation formula for building layout energy-saving and comfort modeling in cold regions is

$$P(x_{w3}, \gamma_{w3} | \Theta) = \prod_{x_{ij} \in w_3} \prod_{k=1}^K a_k g(x_{ij}, \gamma_{ij} | \mu_k, \sigma_k^2). \quad (6)$$

In the above formula, θ is the lighting characteristics of sunlight, μk is the adsorption amount of carbon dioxide by building layout energy-saving and comfort modeling in cold regions in different time periods, and αk is the control coefficient of sunlight lighting in the planning and design of communities. Through the design, this paper optimizes the architectural layout of cold residential areas and improves

the energy saving and comfort of residential areas, as shown in Figure 2.

The test site is located on the east side of Shenzhen Road and on the north side of Xing'an Street, with a total construction area of 210,946 square meters and a plot ratio of 1.70. There are 7 six-storey residences and 5 18-storey residences, which were delivered on 2017-01-01. The overall plan of the community planning is shown in Figure 3.

The field test needs to measure the hourly temperature, humidity, and wind speed of each measuring point within 3 days, so the measuring instruments are mainly those that measure microclimate data. The purpose and measurement accuracy of the instruments used in the experiment are shown in Table 1.

This experiment adopts the method of flow observation and records the air temperature, relative humidity, and wind speed at the pedestrian height (1.5 m above the ground) of each measuring point every half hour at different positions of the open space, including the center of the open space, the entrance and exit positions, the shadow areas of buildings, near the water surface, and near the structures. Explore the influence of main building height, D/H ratio, open space scale, underlying materials, entrance and exit orientation, and other factors on the thermal comfort of open space.

4. Result Analysis

A total of 11 test points were arranged in the test site to explore the impact of building height, d/h ratio, open space scale, entrance and exit orientation, and other factors on the thermal comfort of open space. In addition, the observation data from the satellite ground station of Yichun Meteorological Bureau are obtained, which can be used as the comparative data of experimental tests. The weather station is located at 45.75° north latitude, 126.46° east longitude, 1264.6 meters above sea level, and about 10.1 kilometers east-south of the experimental plot. The location of the weather station is in line with the regulations of the International Meteorological Organization, and its observation data can represent the macrometeorological conditions of Yichun City. The temperature of the testing place is recorded every half hour by a multifunctional environmental tester. Here, the temperature test results of representative measuring points 2, 4, 5, 6, 7, 9, and 10 at different times are compared and analyzed by a line chart, and the test results are shown in Figure 4.

It can also be seen from the figure that the highest temperature in three days occurred at 11:30 on April 12th, measuring point 9, and the temperature value of measuring point 9 was almost the highest among the measuring points. It may be because measuring point 9 is located outside the shadow area of the building, and it is wrapped by a large area of evergreen coniferous trees, which has a certain advantage over other measuring points in resisting the cold wind. This may also be the reason why the temperature curve of measuring point 9 changes more gently than other curves. In contrast, the curve with the temperature test value at a lower level is that of test point 7 and test point 10. Compared with the temperature test values of test point 6 and test point 9 in

the nonshaded area, the temperature difference is about 1°C, and the higher the average temperature is, the closer the time is to noon, the greater the temperature difference is. The shadow of the visible building affects the absorption of solar radiation at the measuring point and causes the temperature value to be lower than the average level of the open space. The manual observation results of each measurement point at the test site are shown in Figure 5. The wind speed value of each measurement point does not have strong regularity like temperature and humidity but shows strong randomness.

The determination of initial boundary conditions is divided into two parts. In the first part, the meteorological data of local meteorological stations are used as the simulated comfortable boundary conditions. The meteorological data include air temperature, wind speed 10 meters above the ground, wind direction, relative humidity, and absolute humidity. The meteorological data comes from the data of Yichun Meteorological Data Platform on April 12, 2021. In the second part, the modified microclimate data of the actual open space is used as the secondary boundary condition. The second boundary condition is modified because the data of macrometeorological station cannot fully represent the simulated microclimate data in open space. After the initial simulation test and the wind direction frequency in the hot and cold alternate seasons, the initial boundary conditions of the simulation are determined. After the verification of the onsite measurement, the initial simulation conditions are consistent with the actual situation, so the initial air temperature is further determined to be 276.15 K (the average measured in three days), and the wind speed 10 meters above the ground is 4.05 m/s (the average speed of the meteorological station in three days), the wind direction is 315° (that is, the northwest wind direction, combined with the specific wind direction of the day and the prevailing wind direction in spring), the relative humidity is 40.5% (the average air relative humidity at 8:00 in the morning within three days), and the heat transfer coefficient of the building exterior wall is 1.7 W/m² K; the albedo is 0.3, the heat transfer coefficient of the building roof is 2.2 W/m² K, the albedo is 0.15 (gray hard pavement), and the indoor temperature is set to 20°C (the average room temperature in Yichun). After the simulation test, it should be verified and corrected according to the comparison between the simulation results and the measured results, so as to obtain the secondary boundary conditions used in the actual simulation. After the initial simulation test and the wind direction frequency in alternate hot and cold seasons, as shown in Table 2, the initial boundary conditions of simulation are determined.

The initial wind speed and wind direction are fixed; that is, the wind speed and wind direction entering the open space from the external environment are constant; while the actual measured wind speed has certain transient and contingency, the record of the wind speed takes the maximum wind speed within one minute, and the wind speed and wind direction are gradually changed. And there are instantaneous maximum wind speeds that are much higher than the average wind speed. Even so, from the practical significance of scientific research application and simulation, hourly simulation of the instantaneous wind environment

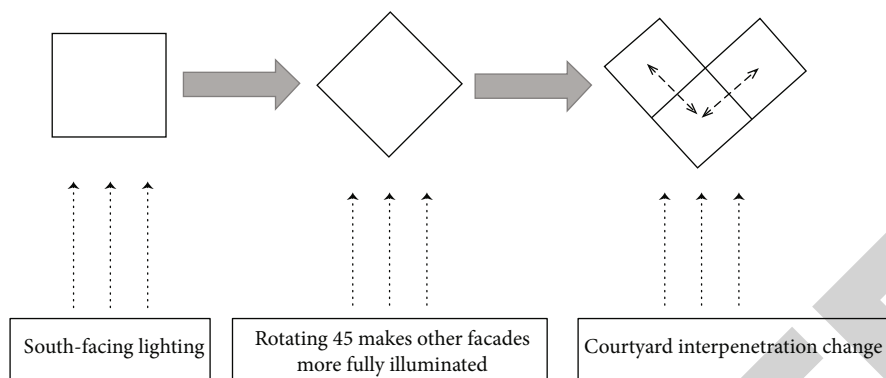


FIGURE 2: The building layout of the improved cold area.



FIGURE 3: Floor plan of Songyun City community at the test site.

TABLE 1: Instruments and accuracy used for on-site testing.

Measuring instrument	Instrument accuracy	Measurement parameters
American TSI9555-P multifunctional environmental tester	Temperature 0.1°C, wind speed reading ± 1%, humidity ± 3% RH	Humidity, temperature, wind speed
KANOMAX 6004 handheld digital anemometer	±(Indication value 5% + 0.1) m/s	Wind speed
Testosterone415 anemometer	Temperature 0.1°C, wind speed ± (0.05 m/s)	Wind speed, temperature
Laser rangefinder	0.05 m-60 mI ± 1.5 mm	Distance
WBGT2006 wet bulb black bulb temperature index meter	20-40°CI ± 0.5°C	Mean radiant temperature

in open space is of little significance for optimizing the thermal comfort of open space, and simulation of the average wind speed in open space is more beneficial for optimizing design.

5. Discussion

This paper will creatively expand the “current situation” mentioned in its analysis and put forward its own suggestions, design, and manufacturing measures to enhance the novel features of its analysis. And make its research results more widely

applicable to similar weather and geographical morphological characteristics around the world. This kind of question involves the construction industry, which has two aspects:

- (a) One dimension refers to the improvement of building materials, which belongs to bioclimatic building design, gypsum mixture and natural agricultural biomass residue (fiber) and/or mortar of gypsum and aerosol film, so as to reduce the loss of thermal bridge; improve the performance of heat insulation, sound insulation, and waterproof; and avoid the

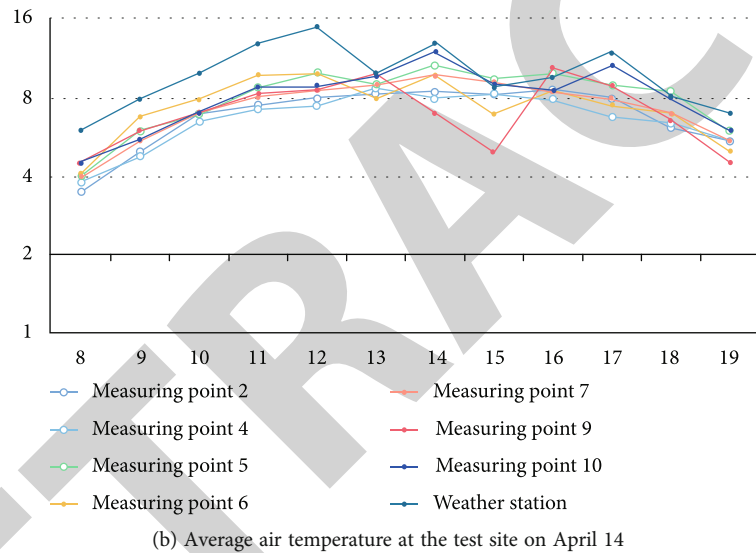
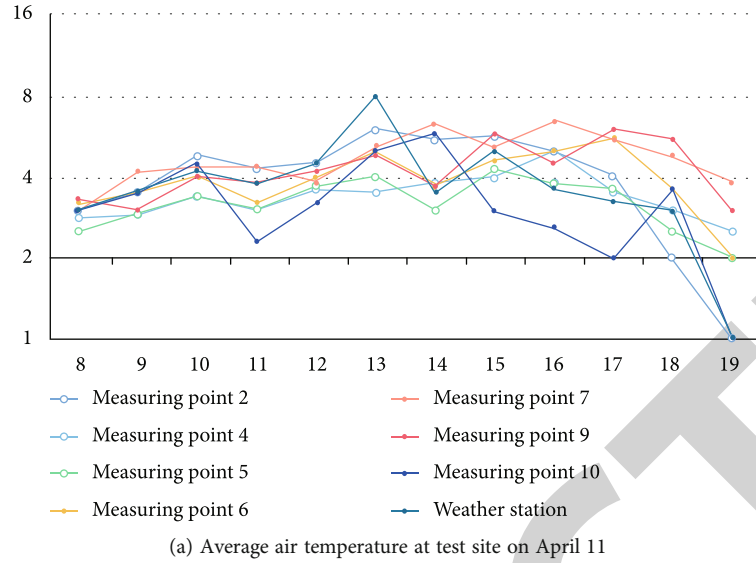


FIGURE 4: Average air temperature of each measuring point at the test site.

infiltration of water in its fractal structure. Other advantageous properties include higher resistance to temperature fluctuations and fracture, as well as better mechanical, hydrogen, and physical behavior (tensile and compressive stress) through solar radiation aging and humidity events

- (b) Another aspect is the improvement made in the principle of circular economy, that is, recycling and reusing, rather than disposal of building demolition degradation substances, for the second round of use (to avoid the direct retirement of debris substances for natural and water resources, so as to make the local environment bear the burden of environmental pollutants)

6. Regional Studies

The annual average temperature in Yichun is $-4-9^{\circ}\text{C}$. The average temperature during the day is 9°C . It is recom-

mended to wear suits, jackets, windbreakers, casual wear, jackets, suits, thin sweaters, and other warm clothes. At night, the average temperature is -4°C . It is recommended to wear cotton padded clothes, winter coats, leather jackets, tweed coats, woolen hats, gloves, down jackets, leather jackets, and other heavy warm clothes. The cities with the highest annual average temperature in Yichun are Tieli (9°C), Wuying (8°C), Wuyiling (7°C), and Jiayin (7°C). The cities with the lowest annual average temperature in Yichun are Wuyiling (-7°C), Yichun (-7°C), Wuying (-8°C), and Jiayin (-6°C). Yichun is located in the Tangwang River Basin in the hinterland of Xiaoxing'an Mountains in the northeast of Heilongjiang Province, China, from $46^{\circ} 28'$ to $49^{\circ} 21'$ north latitude and $127^{\circ} 42'$ to $130^{\circ} 14'$ east longitude. It is adjacent to Hegang and Tangyuan in the east, Qing'an and Suiling in the west, Yilan and Tonghe in the south, and Xunke across the river from Russia in the north. Yichun, with a cold temperate continental monsoon climate and cool summer, is a good place for summer vacation. In winter, the

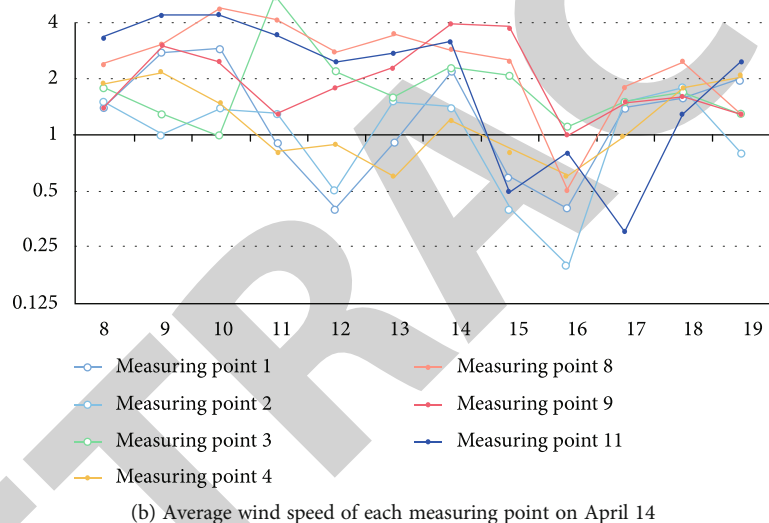
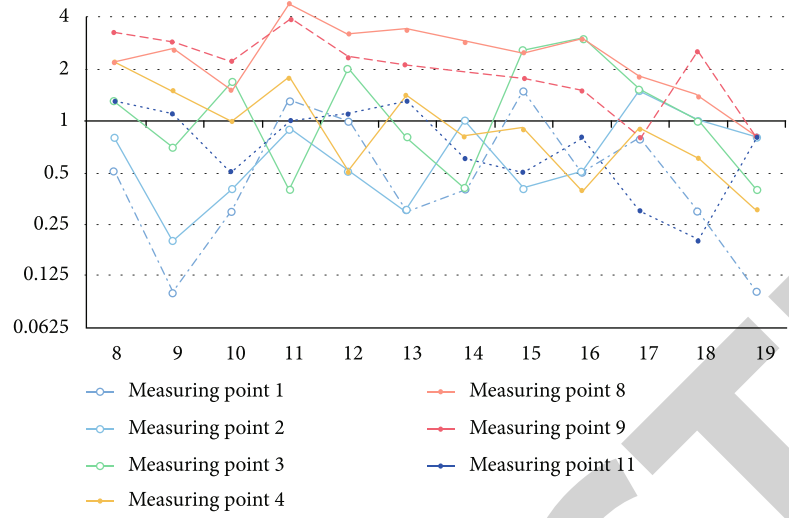


FIGURE 5: Average air temperature of each measuring point at the test site.

TABLE 2: Meteorological data of Yichun City from 7:00 to 8:00 on April 12, 2021.

	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00
Air temperature (°C)	3.2	5.1	6.8	6.5	8.9	6.8	9.5	8.5	9.3	8.1	7.5
Wind direction	338	310	286	336	360	286	350	256	214	362	265
Wind speed (m/s)	64	50	45	46	52	32	33	38	35	32	56

world wrapped in silver will definitely make you feel worthwhile. Yichun is rich in forest tourism resources, with a forest coverage rate of 82.2%. It has the largest and best preserved Korean pine virgin forest in Asia. Wuying Fenglin Nature Reserve and dailing Liangshui Nature Reserve have been approved by UNESCO to be included in the world network of man and biosphere reserves. Wuying National Forest Park has been rated as a national AAAA tourist area. In the forest sea of Xiaoxing'an Mountains, the biological community is rich and diverse, with more than 1000 kinds of wild animals and plants distributed, and Taoshan Forestry

Bureau has built the first open wild animal breeding and hunting ground in China. Landscape: Fenglin national Korean pine primeval forest nature reserve, Taoshan International Hunting Ground, Wuying National Forest Park, Tangwanghe Stone Forest, Jiayin Maolan valley scenic spot, dragon bone mountain dinosaur fossils, Meixi Huilong Bay, Anti-Japanese Alliance site, etc.

This section focuses on the microclimate simulation analysis of the residential open space. The development of urban high-rise buildings is in a blowout mode. The development and construction of high-rise buildings are closely

related to urban landscape and urban space. High-rise buildings internalize and privatize urban space. The three-dimensional and compressed urban space changes the sunshine, wind direction, and humidity of the surrounding environment, forming a unique microclimate in the interior of high-rise buildings. Firstly, the scientificity of the environment numerical simulation software is verified. The peripheral residential open space is taken as the research object, and the reasonable grid simulation size and initial boundary conditions are set. Then, the design factors such as d/h, east-west building length, opening position, underlying surface materials, and water distribution are quantitatively simulated. The relationship between the design factors and the proportion of open space is obtained. On this basis, the optimization strategy of active area, opening position, space proportion, plane enclosure, underlying surface, and water layout is put forward.

7. Conclusions

This paper takes the residential area of Yichun City as the research object, finds the problems existing in the communication space through investigation, and studies the outdoor communication space in the residential area suitable for the cold city of Yichun. The results of the investigation and analysis on the public open spaces of typical residential areas in Yichun City show that the outdoor activities of the respondents in the transition season are closely related to the microclimate environment. This paper gets some planning enlightenment from the survey results of climate comfort in transition season, which is used to improve the climate comfort of cities in cold regions in transition season and promote the livable construction of cities in cold regions. Through the optimization design of building layout in cold residential areas, the energy saving and comfort of living environment are improved, the satisfaction evaluation of building layout energy-saving and comfort modeling in cold residential areas is improved, the resource conservation and utilization are realized, and the living comfort experience is improved. The outdoor communication space in urban residential areas does not simply refer to the space for communication and exchange but builds a comfortable and pleasant outdoor communication platform for people through outdoor exercise, leisure, entertainment, recreation, consumption, and other activity spaces; provides people with more opportunities for outdoor diplomacy; and extends the time of outdoor communication. First of all, the research of this paper is based on the field investigation and research on more than a dozen different distribution and different scale settlements in Yichun, on the basis of interviewing and analyzing the behaviors, preferences, and needs of different groups of people, combined with the corresponding domestic and foreign theoretical knowledge. Through the research, the design theory of the outdoor communication space in the residential area of Yichun cold land is obtained. Through the research of design principles, design strategies, and design methods, the theory of outdoor communication space design in urban residential areas in cold areas of Yichun can be actively and beneficially applied to practice

reflecting certain operability, not just staying on the surface of words, which is beneficial to the overall improvement of the living quality of residential areas in Yichun.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] J. Rouleau, L. Gosselin, and P. Blanchet, "Robustness of energy consumption and comfort in high-performance residential building with respect to occupant behavior," *Energy*, vol. 188, p. 115978, 2019.
- [2] S. Chen, P. Cui, and H. Mei, "A sustainable design strategy based on building morphology to improve the microclimate of university campuses in cold regions of China using an optimization algorithm," *Mathematical Problems in Engineering*, vol. 2021, Article ID 2304796, 16 pages, 2021.
- [3] V. Dermardiros, A. K. Athienitis, and S. Bucking, "Energy performance, comfort, and lessons learned from an institutional building designed for net zero energy," *ASHRAE Transactions*, vol. 125, 2019.
- [4] A. M. S. Smith, E. K. Strand, C. M. Steele et al., "Production of vegetation spatial-structure maps by per-object analysis of juniper encroachment in multitemporal aerial photographs," *Canadian Journal of Remote Sensing*, vol. 34, supplement 2, pp. S268–S285, 2008.
- [5] P. Guéguen, M. Langlais, S. Garambois, C. Voisin, and I. Douste-Bacqué, "How sensitive are site effects and building response to extreme cold temperature? The case of the Grenoble's (France) City Hall building," *Bulletin of Earthquake Engineering*, vol. 15, no. 3, pp. 889–906, 2017.
- [6] J. Wang and J. Kang, "Study on quantitative relationship between residential building body design parameters and energy efficiency in cold areas," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 16, pp. 642–649, 2017.
- [7] F. Chi, R. Wang, and Y. Wang, "Integration of passive double-heating and double-cooling system into residential buildings (China) for energy saving," *Solar Energy*, vol. 225, no. 3, pp. 1026–1047, 2021.
- [8] T. J. Sanger and R. Rajakumar, "How a growing organismal perspective is adding new depth to integrative studies of morphological evolution," *Biological Reviews*, vol. 94, no. 1, pp. 184–198, 2019.
- [9] S. T. Moghadam, C. Delmastro, S. P. Corgnati, and P. Lombardi, "Urban energy planning procedure for sustainable development in the built environment: a review of available spatial approaches," *Journal of Cleaner Production*, vol. 165, pp. 811–827, 2017.
- [10] Z. Zhou, Q. Deng, W. Yang, and J. Zhou, "Effect of seasonal adaptation on outdoor thermal comfort in a hot-summer and cold-winter city," *Advances in Building Energy Research*, vol. 14, no. 2, pp. 202–217, 2020.

Retraction

Retracted: Application of Neural Networks in Financial Time Series Forecasting Models

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Li, "Application of Neural Networks in Financial Time Series Forecasting Models," *Journal of Function Spaces*, vol. 2022, Article ID 7817264, 9 pages, 2022.

Research Article

Application of Neural Networks in Financial Time Series Forecasting Models

Xinhui Li 

Department of Applied Economics, Cheongju University, Cheongju 25803, Republic of Korea

Correspondence should be addressed to Xinhui Li; lxinhui91@cju.ac.kr

Received 17 May 2022; Revised 26 July 2022; Accepted 29 July 2022; Published 30 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Xinhui 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.

At present, the economic development of the world's major economies is showing a positive and positive state. Driven by the development of related industries, the development of the financial field is also changing with each passing day. Various activities in the financial industry are in full swing, and the forecasts of related prospects are also full of uncertainties. Summarizing the laws of financial activities through technical means and making accurate predictions of future trends and trends is a hot research direction that relevant researchers pay attention to. Accurate financial forecasts can provide reference for financial activities and decision-making to a certain extent, promote the steady development of the market, and improve the conversion rate of financial profits. As an algorithm model that can simulate the biological visual system, the convolutional neural network can predict the numerical trend of the next period of time based on known data. Therefore, this paper integrates the support vector machine with the established model by establishing a convolutional neural network model and applies the prediction model to the prediction of financial time series data. The experimental results show that the model proposed in this paper can more accurately predict the trend of the stock index.

1. Introduction

With the development of the economic market, the financial system is becoming more and more mature. At present, the financial market occupies an important position in the national economic system, and the development of the national economy can also be reflected by the performance of the financial market. Because the financial market has unlimited business opportunities, it attracts many investors. Investors can analyze the movement trend of the financial market based on the historical information they have and then formulate investment plans based on the results of the analysis to obtain higher returns. The validity of historical financial data and whether the acquired historical data can be accurately analyzed determine the effectiveness of investment strategies. Therefore, obtaining effective financial data and analyzing and mining the data to obtain effective information to predict the movement trend of the financial market is a research hotspot in the academic and financial circles.

An important basis for formulating investment strategies is the analysis of historical financial data and the prediction of future data based on the analysis results. The nonlinearity, instability, high noise, and other characteristics of financial data are the important reasons why it is difficult to predict accurately [1]. Financial time series involves many fields, such as stock market, price index, and national income and output. Designing an appropriate financial time series forecasting model requires mining the hidden relationship between the overall economy and fiscal markets [2], but this task is extremely difficult. Most of the reasons why financial time series are difficult to predict are caused by different economic systems and business cycles, which usually show irregular changes and fluctuations, so it is difficult to predict irregular changes and fluctuations. The stock market plays an important role in the development of Japan's national economy, maintaining a healthy and stable economy from the perspectives of financing, fair pricing of securities, optimization of resource allocation, and macroeconomic management. It plays an important role above [2]. Stock

market prices are changing rapidly, and equity investment has the characteristics of high risk and high return.

On the premise of ensuring the safety of the principal, they can obtain higher than bank interest. Income is the main purpose of its securities investment. Therefore, people urgently need to analyze the stock data to find the changing laws, so as to guide them to make correct and effective investments and obtain the maximum net investment benefits. To improve the excess returns of equity investors, many researchers have analyzed financial time-series data and found legislative changes that pose serious challenges to securities market analysis techniques [3]. More and more researchers are getting involved in stock market analysis techniques. Financial time series faces many challenges in forecasting due to the large amount of data, nonlinearity, nonlinearity, high dimension, macro-environmental influence, high feature correlation, and high noise.

In recent years, due to the continuous in-depth research on artificial intelligence by scholars from all over the world, related algorithms and solutions have been continuously applied in the related fields of financial time series [4]. The establishment of complex system forecasting models can effectively process data information, which provides a new direction for forecasting financial time series. At the same time, the neural network can effectively reduce the misjudgment of stock market investment caused by human factors such as personal mood fluctuations [5, 6]. Therefore, neural networks are well suited for financial time series stock forecast modeling, and the technique can be used to predict stock trends over time.

The forecasting technology of financial time series involves many disciplines, so it is a relatively complex system engineering. On the one hand, there are few algorithm models for the general identification and analysis of financial time series, which can well handle the dynamic changing stock market in a complex environment; on the other hand, the existing technology still cannot handle financial time series stocks well. Big data volume information in the market. In addition, my country's stock market is not mature enough compared with developed countries, and market forecasting technology is still in the market cultivation stage and has not been widely used in actual economic life [7]. Therefore, the research on the modeling and forecasting of financial time series is of great importance significance.

In this paper, the convolution neural network model is established, the support vector machine is combined with the established model, and the prediction model is applied to the prediction of financial time series data. The proposed two prediction models are still the best for Nasdaq stock index, and BP neural network prediction model is better than support vector machine prediction model. It can be seen from the correlation coefficient table that cnn-svm hybrid prediction model has the best prediction effect on the next trend of all five stock indexes. CNN prediction model is a mixed prediction model to predict the next trend. The innovative contribution of this paper lies in the feasibility of the two models in stock index prediction. The hybrid prediction model of convolutional neural network and support vector machine provides the best prediction effect. Only

convolution neural network has better prediction effect than traditional BP neural network and support vector machine.

This paper is organized as follows. The first chapter introduces the socioeconomic background of the study of financial forecasting problems and analyzes the motivation of this paper. The second chapter introduces the research status of related fields at home and abroad and summarizes the research significance of this paper. The third chapter shows how to build a stock index prediction model based on convolutional neural network, explains the influence of various parameters of convolutional neural network on the prediction result, and determines the appropriate convolutional parameters. The fourth chapter tests and analyzes the proposed scheme and compares it with other methods to verify the effectiveness of the scheme. Chapter 5 summarizes the research content of this paper and looks forward to future research directions.

2. The Related Works

A variety of methods have been applied to financial forecasting, and the commonly used methods can be broadly classified into qualitative forecasting methods, quantitative forecasting methods, and data mining-based forecasting methods. The following is a brief introduction of these three aspects of forecasting methods and focuses on the current status of research on financial forecasting based on data mining.

Qualitative forecasting method is a more intuitive and subjective financial forecasting method, which is mainly based on the past and present information of the forecast object as well as personal experience and judgment ability by the experts concerned to make advance judgment and speculation on the future development trend and pattern of financial activities. Quantitative forecasting method is a kind of forecasting method based on statistical theory [8].

Data mining techniques are now widely used in the field of financial forecasting such as stock and futures [9]. Lu et al. [10] established an AR model for the Frankfurt stock market to predict and analyze the stock price, and the prediction results were more accurate. The moving average method is simple and fast, but the prediction accuracy is not high. Salehi et al. [11] predict the future price of the stock market based on historical data and uses the number of different poles and zeros of the ARMA model to predict the stock price of the next day. By comparing with the ARMA model, Luo et al. [12] found that the minimum mean square error was combined with the fractional integral autoregressive moving average model. The combined method is more accurate in predicting the stock price. Exponential smoothing is low cost, simple, and practical, and its prediction accuracy is high, even comparable to many more sophisticated and more statistically based methods. If the time series has obvious seasonal periodicity, its statistical characteristics also show regular seasonal changes. At this time, the seasonal coefficient method can be used for prediction, thereby improving the accuracy of the prediction results. Shahvaroughi et al. used BPNN for stock price prediction, and the experimental results showed that BPNN had a smaller prediction error than the comparison experiments in the paper

[13]. Kumar and Yadav [14] use the Shanghai Composite Index to calculate the volatility trend of stock prices. Since the changes in asset prices on trading days are not considered, especially on trading days with large changes in asset prices, the estimated value of the volatile GARCH model will be lower than the actual volatility. Syriopoulos et al. [15] proposed a fuzzy GARCH modeling method to predict stock market returns, which considers time-varying fluctuations and adjusts the adaptability of the model through gradual model construction. Cheng et al. used support vector machine to forecast the Korea Composite Stock Price Index (KOSPI) and showed better prediction performance than BP neural network and case-based inference methods [16]. Xiangfei Li et al. combined both support vector machine and empirical mode decomposition (EMD) methods to predict the error sequence of the initial forecast set and correct the original forecast value using the error forecast value. Experiments show that this method is better than the prediction by support vector machines alone [17]. Kumar et al. constructed a wavelet support vector machine to predict the volatility of two simulated stocks by replacing the Gaussian kernel function of the support vector machine with a wavelet kernel function, which showed better prediction performance than the Gaussian kernel [18]. Support vector machines have a great advantage in solving high-dimensional nonlinear data problems with robustness and generalization capabilities.

From the above research status at home and abroad, it can be concluded that domestic and foreign scholars mainly use LSTM neural network to predict stock closing price and return rate and obtain a high accuracy rate in short-term prediction accuracy, which is better than traditional time series prediction methods. Due to the many and changeable factors affecting the stock market, domestic and foreign scholars mainly focus on the comparison of model effects between different models and mainly improve the training and prediction accuracy of LSTM neural network from three aspects. These include the improvement of the LSTM neural network structure or the combination of different models, the innovation of dataset input, the improvement of the prediction accuracy of the improved model compared with the basic network prediction accuracy, and the solution of model overfitting [19]. In addition, the sequence dependence characteristics of financial time series data are an indispensable part of financial time series data forecasting. Although ARMA, GARCH, and other models can consider the sequence dependence characteristics of financial time series data, the determination of the sequence correlation length may be complicated. Or it is difficult to achieve automatic identification of sequence-dependent lengths based on the research problem [20]. Therefore, traditional econometric models have certain limitations for forecasting financial time series data with complex characteristics.

3. Convolutional Neural Network Based Stock Index Prediction Model

3.1. *Convolutional Neural Networks and Support Vector Machines.* Due to the huge and diverse characteristics of financial markets, financial time series usually do not have

stable statistical characteristics and are highly nonlinear. Financial time series is a kind of time series data, which has strong timeliness and strong dependence on the front and back of the data. It is impossible to adjust the order. They are generally two-dimensional data. Because time series have strong sequence lines, and there are generally dependencies, cycles, and other relationships before and after the data, it is possible to predict future data based on existing data through statistical knowledge. The high intensity of price volatility in financial markets results in nonstationary financial time series. Therefore, the salient features of financial time series are nonlinear and unstable characteristics. This also means that using linear forecasting models can only extract linear relationships in financial time series, but not nonlinear relationships. Therefore, complex nonlinear relational data can only be extracted and analyzed through machine learning models, deep learning models, or nonlinear data models. Combining the advantages of convolutional neural network and support vector machine, this paper innovatively proposes a new solution. The structure of the financial time series forecasting model designed according to the convolutional neural network in this paper is shown in Figure 1.

The training process of the algorithm model in the scheme proposed in this paper can be divided into the following two main steps:

- (1) Forward propagation process. A sample is randomly selected from the training sample set (X, Y_p) , where X is the network input and Y_p is the desired network output. The information of the input X is propagated from the input layer to the output layer through layers of computation, and the actual output of the input X is calculated according to the following equation:

$$O_p = F_n(\dots(F_2(F_1(XW_1)W_2) \dots)W_n). \quad (1)$$

- (2) Backward propagation process. This process is used to perform error back propagation. That is, it calculates the difference between the actual output O_p and the desired output Y_p :

$$E_p = \frac{1}{2} \sum_j (y_{pj} - o_{pj})^2. \quad (2)$$

SVMs have unique advantages in handling function approximation problems, performing complex tasks such as solving other pattern recognition problems such as function approximation. SVMs use structural risk minimization as a criterion, which allows training the resulting classifier to obtain a globally optimal solution.

Figure 2 depicts the basic principle of SVM. The two samples in the figure are represented by black circles and white circles, and the curve H represents the optimal

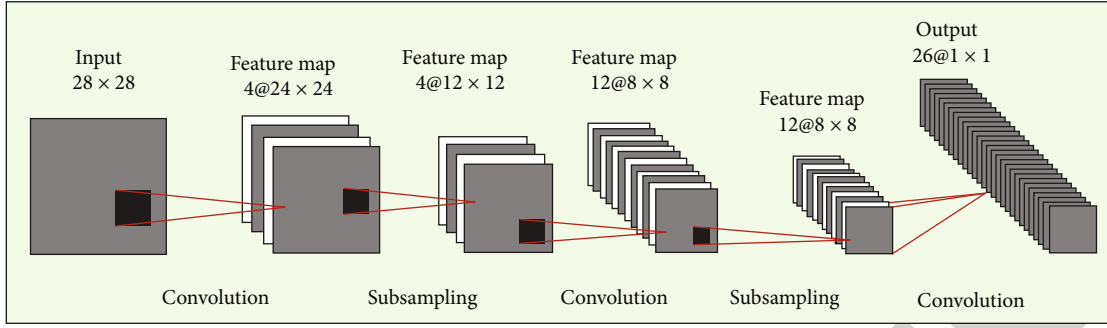


FIGURE 1: convolutional neural network for handwritten character recognition task.

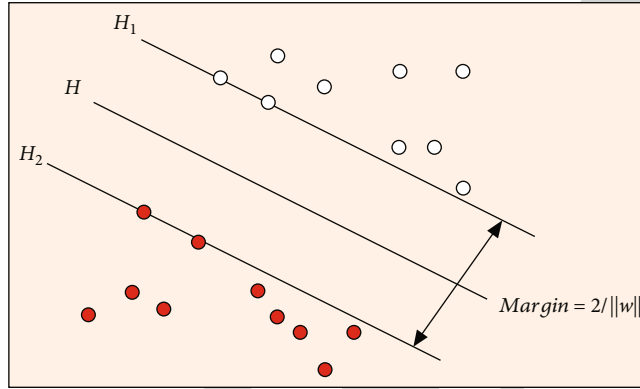


FIGURE 2: Schematic diagram of the optimal classification surface.

classification hyperplane between the two samples. The curves 1H and 2H pass through the sample with the smallest distance to the optimal classification hyperplane in their respective classes. The distance between 1H and 2H is called the classification interval.

The optimal classification hyperplane is the one that not only classifies the samples of both categories correctly, but also maximizes the classification interval. The classification error rate that can be trained by correctly classifying samples of both categories is 0, i.e., the empirical risk is minimized.

Let $(x_i, y_i), i = 1, \dots, n, x \in R^d, y \in \{+1, -1\}$ denote the linearly separable sample set, where +1 and -1 denote the category symbols. The linear discriminant function is denoted by $g(x) = w \cdot x + b$, and the classification surface equation is denoted by $w \cdot x + b = 0$. In order to make all samples in both categories satisfy $|g(x)| = 1$, that is, to make the nearest sample to the classification hyperplane of $|g(x)| = 1$, it is necessary to normalize the discriminant function. At this point, the normalized classification interval is transformed into $2/\|w\|$, and the optimization goal of maximizing the interval is transformed into minimizing $\|w\|$ (or $\|w\|^2$). In order for the classification surface to correctly classify all samples, the classification surface must meet the following conditions:

$$y_i[(w \cdot x) + b] - 1 \geq 0, i = 1, 2, \dots, n. \quad (3)$$

The classification plane that satisfies Equation (3) and minimizes the value of the optimization target $\|w\|^2$ is called

the optimal classification hyperplane. The optimal classification hyperplane is determined based on the sample points on H_1 and H_2 parallel to the optimal classification hyperplane, so these sample points are the support vectors in the support vector machine.

The above optimal classification hyperplane solution problem can be transformed into an easily solvable dual problem by introducing Lagrange optimization methods, i.e., under the constraints:

$$\begin{aligned} \sum_{i=1}^n y_i \alpha_i &= 0 \\ \alpha_i &\geq 0, i = 1, 2, \dots, n \end{aligned} \quad (4)$$

Find the maximum value of the following function for α_i :

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j). \quad (5)$$

Let α^* be the optimal solution, then we have

$$w^* = \sum_{i=1}^n \alpha^* y_i x_i. \quad (6)$$

It can be seen from Equation (6) that there is an equal relationship between the connected weight vector of the optimal classification hyperplane and the weighted sum of

the training sample vectors. The core of the algorithm is to determine the optimal classification hyperplane, that is, to determine the parameters of the classifier function through training samples. To determine the classification hyperplane is essentially to solve a quadratic optimization problem, and the parameters of the classifier are determined by solving the dual problem.

This is a quadratic optimization problem under the inequality constraint, and the optimal solution is unique, and only some values of the optimal solution are non-zero, and the sample points corresponding to these non-zero connection weights are the support vectors. The optimal classification decision function obtained according to the above solution process can be expressed by the following equation:

$$f(x) = \text{sgn} \{ (w^* \cdot x) + b^* \} = \text{sgn} \left\{ \sum_{i=1}^n \alpha_i^* y_i (x_i \cdot x) + b^* \right\}. \quad (7)$$

There are two ways to solve b^* , one is by substituting any support vector into Equation (3); the other is by averaging any pair of support vectors in the two types of samples. Since the corresponding connection weights of the sample points that are not support vectors α_i are zero, the summation process in Equation (7) is actually only for the support vectors, and there is no need to calculate the sample points of non-support vectors.

For the case of linear indistinguishability of training samples, some sample points do not meet the conditions of Equation (3), and then a relaxation variable $\varepsilon_i \geq 0$ can be added to Condition (3) so that Equation (2) becomes

$$y_i [(w \cdot x_i) + b] - 1 + \varepsilon_i \geq 0, \quad i = 1, 2, \dots, n. \quad (8)$$

Obviously the relaxation variable ε_i cannot be arbitrarily large; otherwise, any hyperplane would be eligible. Let

$$F(\varepsilon) = \sum_{i=1}^n \varepsilon_i. \quad (9)$$

When the value of Equation (9) is the smallest, it means that the number of misclassified samples is the smallest and the classification interval is the largest for the linearly divisible case; for the linearly indivisible case, the following constraints need to be introduced:

$$\|w\|^2 \leq c_k. \quad (10)$$

Under the constraints of Conditions (8) and (10), the optimal classification hyperplane for linearly indistinguishable training samples is obtained by finding the minimum value of Equation (9).

The original solved problem can be transformed into a new problem of finding the minimum value of the following function, the constraint of which is Equation (8).

$$\phi(w, \varepsilon) = \frac{1}{2} (w, w) + C \left(\sum_{i=1}^n \varepsilon_i \right), \quad (11)$$

where C is the penalty factor, which is a specified constant. Its purpose is to control the penalty level of misclassified samples so that the training model can balance between the proportion of misclassified samples and the complexity of the algorithm and prevent the complexity of the model from increasing.

Similar to the solution of the optimal classification hyperplane, this optimization problem is transformed into a quadratic extreme value problem to be solved, and the results are close to Equations (4) to (8) obtained in the linearly separable case, but Condition (5) becomes

$$0 \leq \alpha_i \leq C, \quad i = 1, \dots, n. \quad (12)$$

3.2. Convolutional Neural Network-Based Stock Index Prediction Model. Define the input sample length as l , the convolution kernel size as c , and the downsampling reduction as s . The relationship that the above three parameters should satisfy is as follows when the number of convolution layers and the number of downsampling layers is determined to be 2:

$$[(l - c + 1)/s - c + 1]/s = N. \quad (13)$$

When the number of convolution and downsampling layers in the model is 1, the above three parameters should satisfy the following relationship:

$$(l - c + 1)/s = N. \quad (14)$$

Figure 3 shows the range of convolution kernel sizes for different input sample lengths when the network structure contains two layers of convolution.

In the scheme designed in this paper, the output unit of the last layer is actually a probability estimate of the input sample. For the above convolutional neural network, it is not meaningful to focus on the output of the hidden layer, but for other classifiers, the output of the hidden layer can be used as features for prediction or classification. The main connotation of the model in the proposed scheme can be summarized by feature extraction and classification. The mapping and extraction of features at different levels is achieved by convolution and downsampling, and the classification (prediction) function is achieved by a single-layer perceptron. Feature engineering has always been a difficulty in machine learning. The main reason is that different data and different problems have different good features. Good features are often not suitable for signal processing. Even the two tasks are also image recognition tasks, because the data contained are different and the corresponding good features are also different.

The scheme proposed in this paper can extract the features of the samples accurately and effectively. However, classification or prediction by single layer perceptron using the extracted features is prone to overfitting and falling into

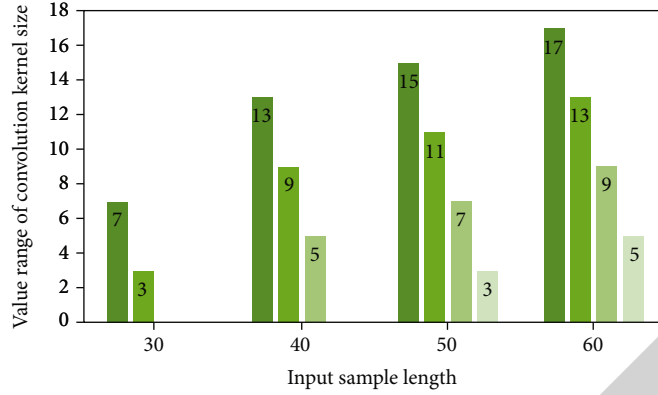


FIGURE 3: Relationship between input sample length and convolution kernel size.

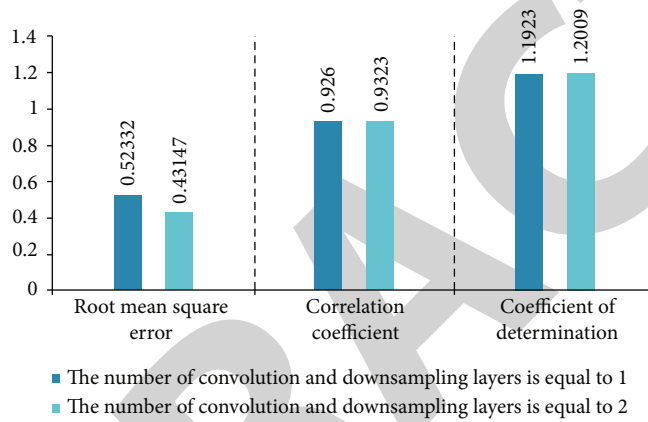


FIGURE 4: The prediction result when the input sample length is 30, and the convolution kernel size is 3.

local optimum. The perceptron is a generalized linear model. The decision boundary of single-layer neural network is linear whether there is activation function or not. Because the single-layer neural network is linear, even the simple nonlinear XOR function cannot be learned correctly. The simplest idea is to manually add some nonlinear features of other dimensions to improve the nonlinear expression ability of the model. This also reflects the importance of feature engineering to the model. The support vector machine shows good performance in classification and prediction, and it avoids the overfitting problem by introducing relaxation variables. The optimization problem of the support vector machine can be treated as a convex quadratic optimization problem, and the optimal solution obtained is the global optimum. Therefore, this paper combines the advantages of both models and proposes a model that combines the advantages of convolutional neural networks and support vector machines (CNN-SVM).

4. Analysis of Simulation Results

In this section, we will study the impact of convolutional neural network related parameters on stock index forecasting results to determine suitable parameters and stock forecasting models. The prediction results use two indicators to judge the quality of the prediction results: the mean square error (MSE) and the correlation coefficient between the pre-

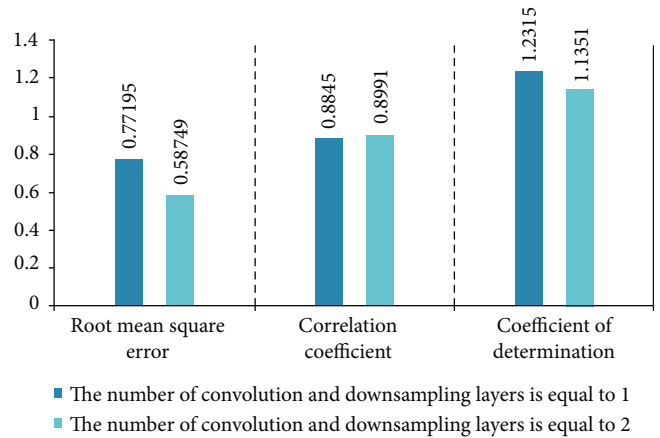


FIGURE 5: The prediction result when the input sample length is 40, and the convolution kernel size is 5.

dicted financial time series and the actual output series according to the proposed scheme. Time series analysis is a basic technology in quantitative investment. Time series refers to the value sequence of a variable measured in time sequence within a certain period of time. For example, if the variable is the stock price, its change with time is a time series. Similarly, if the variable is the return of stocks, its change with time is also a time series. While discussing the

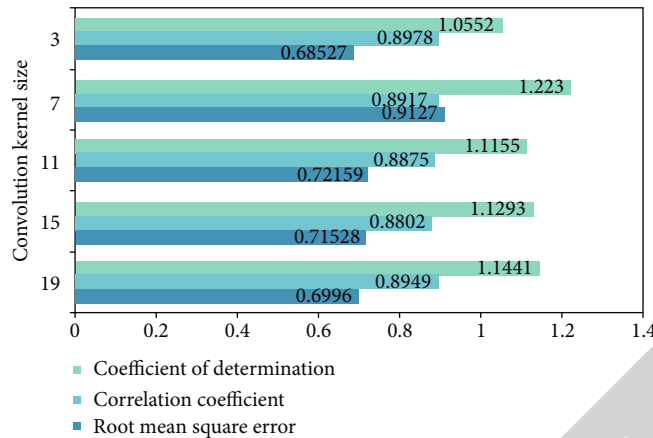


FIGURE 6: Prediction results metrics for different convolution kernel sizes.

impact of specific parameters on the stock index forecast results, the values of other parameters are fixed.

The formula for filtering noise is as follows:

$$m_h(x) = \frac{\sum_{t=1}^T K_h(x - X_t) Y_t}{\sum_{t=1}^T K_h(x - X_t)}, K_h(x) = \frac{1}{h\sqrt{2\pi}} e^{-x^2/2h^2}. \quad (15)$$

The normalization formula is

$$y'_t = \frac{y_t - \min \{Y_t\}}{\max \{Y_t\} - \min \{Y_t\}}. \quad (16)$$

In the simulation experiments in this paper, the cases where the number of convolutional layers and downsampling layers are 1 and 2 are discussed in detail, respectively. According to the length of different input samples, corresponding simulations are also carried out in this paper. For the case where the length of the input samples is divided into 30 and 40, this paper sets 3 and 5 as the corresponding values of the convolution kernel. As can be seen from Figures 4 and 5, it is reasonable to set the number of convolutional layers and downsampling layers in the prediction model proposed in this paper to 2. For stock indices, the features extracted by two convolutional layers are more effective. So all stock prediction models in this paper use two convolutional layers and one downsampling layer.

Considering the limitation of the input sample length on the size of the convolution kernel, 70 is used as the length of the input sample in this experiment, and the two layers of the convolution kernel are 4 and 8, respectively. Next, we ran the prediction comparison experiments for 5 cases with 19, 15, 11, 7, and three kernel sizes and compared the prediction results for the 5 cases. Prediction effect (Figure 6) shows the experimental results of the convolutional neural network stock prediction model in the above five cases.

As mentioned earlier, the stock prediction model designed in this paper uses two convolutional layers and one downsampling layer. On this basis, the number of convolution kernels in the second layer is set to be twice the number of convolution kernels in the first layer. If the num-

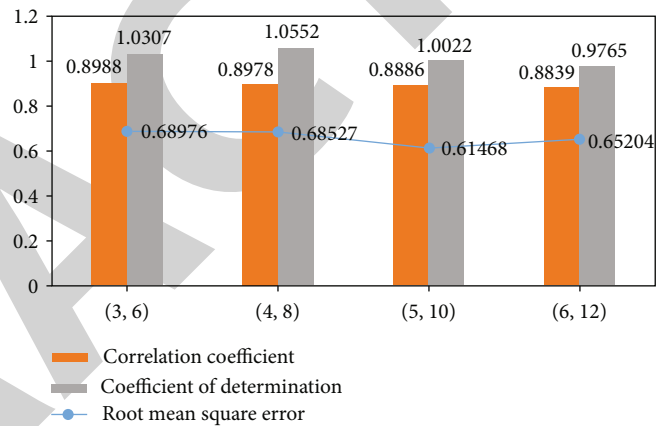


FIGURE 7: Prediction results metrics for different number of convolution kernels.

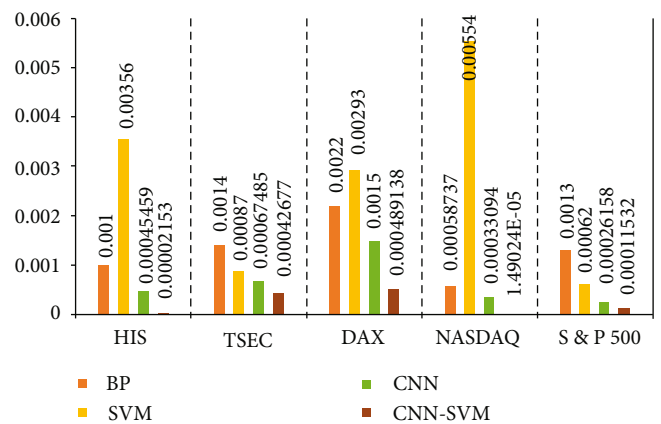


FIGURE 8: Comparison of MSE metrics for different prediction models.

ber of convolution kernels in the first layer is 4 and the number of convolution kernels in the second layer is 8, then it is represented by (4,8). The prediction experiments are carried out in four cases, and the number of convolution kernels is set to (3, 6), (4, 8), (5, 10), and (6, 12). Figure 7 shows the predictions in the four cases comparative results.

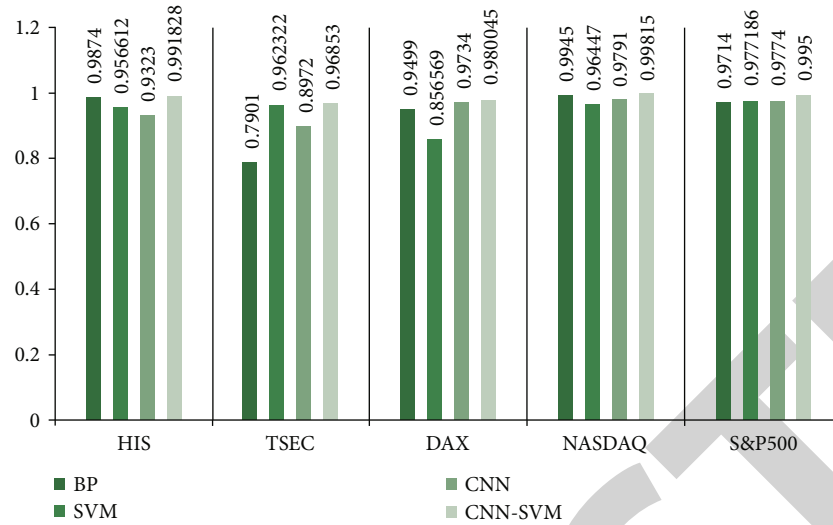


FIGURE 9: Comparison of correlation coefficient indicators of different prediction models.

This paper selects five main indices to train and test the model. Figures 8 and 9 show two metrics of the prediction results of the proposed scheme and the other three comparison schemes on the corresponding datasets. The convolutional neural network stock index forecasting model and the hybrid convolutional neural network-support vector machine forecasting model follow the best trends in the five stock index datasets, and the error sequence for both forecasting models is 0. The four stock index prediction models showed that the best and the worst prediction results of the four stock indexes of HIS, TSEC, DAX, and S&P500 were CNN-SVM, CNN, SVM, and BP network in order. The two forecasting models proposed in this paper are still the best in the forecasting effect of the Nasdaq stock index, and the BP neural network forecasting model is better than the SVM forecasting model. From the correlation coefficient table, the CNN-SVM hybrid prediction model has the best prediction effect on the next trend of all five stock indices, while the CNN prediction model is a hybrid prediction model that predicts the next trend, you can look at the next. According to the above experimental analysis results, the two models proposed in this paper have certain feasibility in stock index forecasting, among which the hybrid forecasting model of convolutional neural network and support vector machine provides the best forecasting effect, only the prediction effect of convolutional neural network is better than that of traditional BP neural network, and the prediction model of support vector machine is supported.

5. Summary and Outlook

The results show that the two prediction models proposed in this paper are still the best prediction models for Nasdaq stock index. BP neural network prediction model is better than support vector machine prediction model. From the correlation coefficient table, it can be seen that CNN SVM hybrid prediction model has the best prediction effect on the next trend of all five stock indexes. CNN prediction model is a hybrid prediction model to predict the next trend.

The innovation of this paper lies in the feasibility of the two models in stock index prediction. The hybrid prediction model of convolutional neural network and support vector machine provides the best prediction effect. Only convolutional neural network has better prediction effect than traditional BP neural network and support vector machine. However, the research has certain limitations. When building a convolutional neural network prediction model suitable for financial time series, the structural parameters of the model are determined through many experiments. There is no theoretical guidance for the selection of parameter values that affect the financial impact behind the prediction. The impact of the results is unknown. Therefore, further validation is needed to predict the value in the next few days to prove the effectiveness 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 author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] W. Dai, "Application of improved convolution neural network in financial forecasting," *Journal of Organizational and End User Computing (JOEUC)*, vol. 34, no. 3, 2022.
- [2] K. Sergey, "Target financial forecasting as an instrument of overcome financial difficulties in present-day global economic environment," *SHS Web of Conferences*, vol. 129, p. 03017, 2021.
- [3] M. Dixon and J. London, "Financial forecasting with α -RNNs: a time series modeling approach," *Frontiers in Applied Mathematics and Statistics*, vol. 6, 2021.

Retraction

Retracted: Prediction Model of Music Popular Trend Based on NNS and DM Technology

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xu, M. Wang, H. Chen, and F. Hu, "Prediction Model of Music Popular Trend Based on NNS and DM Technology," *Journal of Function Spaces*, vol. 2022, Article ID 6104056, 9 pages, 2022.

Research Article

Prediction Model of Music Popular Trend Based on NNS and DM Technology

Yichen Xu ¹, Mingxun Wang ¹, Hao Chen ², and Fan Hu ³

¹College of Air Services and Music, Nanchang Hangkong University, Nanchang, Jiangxi 330063, China

²School of Information Engineering, Nanchang Hangkong University, Nanchang, Jiangxi 330063, China

³College of Science and Technology, Nanchang Hangkong University, Nanchang, Jiangxi 332020, China

Correspondence should be addressed to Yichen Xu; 70640@nchu.edu.cn

Received 15 June 2022; Revised 3 August 2022; Accepted 9 August 2022; Published 28 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Yichen 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.

The popular trend of today's music can be obtained by deep excavation, analysis, and prediction of the audience's preferences. Using huge music library resources and user behavior to form music big data and truly realizing the aggregation of audience preferences determine the popular development trend of music. Therefore, this paper will apply data mining (DM) technology, introduce neural network (NNS) theory, establish a prediction model of music fashion trend, predict and evaluate the music fashion trend according to the selected evaluation index, find the change of music fashion trend in time, and provide decision-making basis for music fashion trend. In this paper, the prediction of music popularity trend based on NNS and DM technology is studied. In the prediction of the number of songs played by 10 artists, the NNS algorithm proposed in this paper reduces the prediction effect from the original 0.074 and 0.045 to 0.044 and 0.032, respectively, and the error rates are reduced by 35.7% and 29.4%, respectively, compared with the learning algorithm and the decision tree algorithm. Among the three methods, the NNS algorithm in this paper has the highest accuracy. Therefore, it can be proved that the model proposed in this paper is more suitable for predicting the trend of music popularity. In the end, it can accurately control the trend of pop music and also realize the aggregation of user preferences to determine the trend of pop music.

1. Introduction

The development of pop music in today's society shows a prosperous situation in which hundreds of flowers contend. Compared with the end of last century, the development of pop music since this century has favored traditional culture, which is one of the development trends of today's pop music. Since the new century, the development of pop music in Hong Kong and Taiwan has become sluggish, and the number of influential singers has decreased. At the same time, the long-term isolation from the mainland has caused pop music in Hong Kong and Taiwan to show a sense of "root seeking" to tap traditional culture. Accurate prediction of music trends cannot only enhance the user experience and increase platform revenue but also improve the popularity of singers. Excavate the singers who will have unlimited prospects in the future, and further enhance the influence and economic benefits of the music platform. Therefore, it

is of great significance for singers, music lovers, and online music platforms to accurately predict the pop trend of music using historical data. Internet music data has the characteristics of diversity and complexity, high dimension, large amount of data, fast change, strong real-time, and obvious time series. The existing music trend models and traditional statistical models are often difficult to effectively analyze the complex relationship between singers, users, songs (downloads, broadcasts, and collections), and other large amounts of data. The effect of deep mining of music data is not ideal.

As one of the products of entertainment consumption, pop music has attracted more and more attention. Whether there are songs that users like in the music platform is one of the necessary factors. With the foundation of users, we can talk about the diversified profit model. From the perspective of users, in addition to personal special preferences, the most popular and popular songs are most likely to be clicked. In the context of today's big data, a large number of pop music

listeners can determine the development trend of pop music to a great extent, which is mainly reflected in the impact of the behaviors of many listeners on major music platforms, such as audition, collection and download, on pop music, which reflects the audience's preferences and preferences for corresponding music [1]. With the further development of major music platforms and music market scale, the gap of music library scale will gradually narrow, and the role of music library scale in user selection will be almost the same [2]. Who can better predict the popular trend of music and show its progressiveness, technicality, and intelligence will win users and market. The popular trend of today's music can be obtained by deeply mining, analyzing, and predicting the preference trend of the audience. Make use of the huge music library resources and user behavior to form music big data. Through accurate big data analysis, we can effectively predict the trend of music, truly realize the aggregation of audience preferences, and determine the pop development trend of music [3].

Therefore, this paper will apply DM technology, introduce NNS theory, establish a prediction model of music fashion trend, predict and evaluate the music fashion trend according to the selected evaluation index, and find the change of music fashion trend in time. DM mainly combines traditional statistical analysis methods with machine learning, genetic algorithm, artificial NNS, visualization technology, and database technology in artificial intelligence [4, 5]. By mining useful information and knowledge, it can predict the future development trend and implement decision-making behavior of enterprises. Therefore, this paper uses NNS to study the prediction of music fashion trend and builds a model [6]. Although the NNS is the most widely used network at present, it still has many problems, the biggest of which is the convergence speed, and its training is difficult to master, especially when the network training reaches a certain level. NNS has been widely used in intelligent detection, nonlinear prediction, pattern recognition, robot control, and other fields [7].

Based on the NNS algorithm, a system for predicting music pop trend is constructed. This chapter designs and implements a music pop trend prediction system based on the NNS algorithm and describes the development and operation environment and structure design of the system in detail [8]. The prediction of pop music trend can be obtained by deep mining and analysis of audience preferences and preferences on the basis of massive user data. According to the actual needs of music pop trend prediction, it is divided into three prediction tasks. The NNS method is used for combined prediction, and the prediction results of single model and combined model are compared [9]. This paper mainly introduces the source and research significance of music pop trend prediction, analyzes the advantages and research status of regression prediction based on NNS and DM technology, explains the main work and contribution of this paper, and combs the organizational structure of the paper.

Research innovation lies in the formation of music big data with huge music library resources and user behavior. By mining the potential rules of user behavior, song information and artist information, select effective information for data cleaning, data integration and data transformation. This paper

studies the prediction of music pop trend based on neural network and data mining technology. When predicting the number of songs played by multiple artists, the neural network algorithm proposed in this paper will reduce the error rate compared with the learning algorithm and the decision tree algorithm.

2. Research Status of Music Pop Trend Prediction

Gan proposed that with the development of science and technology and social productivity, we will predict the popularity and universality of music through information technology. This is a scientific and effective judgment basis for senior musicians who judge pop singers and predict dark horses based on rich personal experience [10]. Zhang proposed that since everyone's favorite music may span several styles, various online radio stations need to collect most people's preferences for different music, and they need to be divided into various types of music according to the style of pop music [11]. KS et al. proposed that Shazam is a music radar software that can identify the music and TV programs playing around you. It realizes song identification by collecting external song fingerprints and comparing them with server-side fingerprints [12]. Scott-Maxwell proposed that in Chinese pop music, Chinese elements tend to increase gradually. For example, traditional opera elements and classical elements appear in pop music, which makes China's pop music get greater development, and the music elements in China's pop music also begin to show a diversified development scene [13]. Liu et al. proposed that the music pop trend prediction competition held by Alibaba cloud is based on the powerful data computing ability of Alibaba cloud platform; through the user's historical behavior data, predict the amount of artists playing in the next stage, and mine the artists who will become the focus of attention, so as to accurately control the pop trend of music in the future [14]. Lim et al., through the analysis of historical user behavior data, predict the playback volume of some artists or songs in the next stage, mine future dark horse singers from massive data, and fundamentally accurately control the development trend of pop music, which is the focus of many music platforms [15]. Canavan and Mccamley proposed that for ordinary users, the prediction of music pop trends can enable users to grasp the popularity of artists on the current music platform, understand the latest and future development momentum, and experience its progressiveness and development [16]. Guerra et al. put forward that the style and form of Chinese pop music are mainly influenced by Europe and America and gradually form a local style on this basis. In recent years, there has been a "Chinese style" in pop music. With the help of this element, musicians with different music styles have jointly expressed a trend closer to Chinese traditional elements, which makes pop music have a unique Chinese style [17]. Jung et al. propose a large number of pop music listeners can determine the development trend of pop music to a large extent, which is mainly reflected in the impact of the behavior of many listeners on the main music platforms, such as audition, collection and download, on pop music, which reflects the audience's preference and preference for corresponding music [18]. Fautley suggests that with the

further development of major music platforms and the scale of music market, the gap between the scale of music libraries will gradually narrow, and the role of the scale of music libraries in user selection will be basically the same [19].

In this paper, the neural network algorithm of data mining technology is used to study it, and the ringing factor of daily listening volume of songs is studied. It realizes song identification by collecting external song fingerprints and comparing them with server fingerprints. Shazam is a music radar software, which can identify the music and TV programs playing around you. Using data mining technology to analyze and study a large number of Internet music, data is one of the most concerned research points of Alibaba, Tencent, NetEase, and other Internet companies. Data mining and data analysis are new research fields of data analysis and processing. On the basis of previous studies, this paper studies the prediction of music pop trend based on neural network and data mining technology. At the same time, the exponential smoothing method and autoregressive moving average model are used to predict the number of song auditions, and the prediction results of different algorithms are compared in order to complete the preskill training of daily listening capacity of songs.

3. DM Process Based on NNS

Before using DM technology to predict the stock price trend, it is necessary to determine the general process of DM. Combined with the characteristics of NNS DM technology, DM based on NNS is determined. NNS not only has very flexible nonlinear modeling ability but also has strong parallel processing, self-learning and adaptive, computational optimization, associative memory, and other abilities [20, 21]. Therefore, NNS has been widely used in many fields, such as intelligent detection, nonlinear prediction, pattern recognition, and robot control. At present, the research of NNS has become more and more mature, and its application is more and more extensive. By understanding the data of users' listening behavior, analyzing, and processing the data and using it, enterprises with Internet music platform can greatly improve their competitiveness in occupying market share [22]. By using DM technology to analyze the user data of music platform, we can get the information about users' listening orientation and hobbies, so it provides a favorable basis for business decision-making. It iterates through multiple cycles of the two processes through DM technology. Each iteration is an epoch. The weight is randomly set before the start. The algorithm passes through the cycle and stops after reaching the conditions.

The NNS algorithm adopts the gradient descent method, and the training is to gradually reach the minimum value of the error along the slope of the error function from a certain starting point. For complex networks, the error function of DM technology is the surface of multidimensional space, and there are many Valley points. The difference method is an approximate numerical solution of differential equations. Specifically, the difference method is to replace the differential with finite difference and the derivative with finite difference quotient, so that the basic equations and boundary conditions (generally differential equations) are approximately expressed by difference equations (algebraic equations), and the problem

of solving differential equations is changed into the problem of solving algebraic equations. In elasticity, the difference method and variation method are used to solve plane problems. This DM technology cycle includes two stages: forward and backward. In the forward phase, neurons are activated in the process of input to output, and the output signal is generated by using the weight and activation function of each neuron until the last layer. In the backward stage, the output signal generated in the forward stage is compared with the actual value, the error between the output signal and the real value is back propagated, and the weight value is corrected to reduce the error. The flow of NNS algorithm is shown in Figure 1.

The NNS algorithm adopts the steepest descent method of data. Theoretically, its training is approaching downward along the slope of the error surface. For a complex network, its error surface is a surface in a high-dimensional space, in which there are many local minima. Once it falls into such local minima, it is difficult to escape with the current algorithm. During the execution of the NNS algorithm, the amplitude of each adjustment of the weights is made by multiplying a term proportional to the network error function or its derivative to the weights by a fixed factor $1r$. In this way, when the error surface is flat, because the deviation value is small, the adjustment range of the weight is relatively small, so that the error function surface needs to be adjusted several times by DM technology. However, in DM, the curvature of the error surface is higher, the partial derivative value is larger, and the adjustment range of the weight value is larger, which may lead to overshoot near the minimum point of the error function, and it is difficult to converge to the minimum point.

The function f expresses the input-output characteristics of neurons. In the model, f is defined as a step function

$$v_i = \begin{cases} 1, & u_i > 0, \\ 0, & u_i \leq 0. \end{cases} \quad (1)$$

If the threshold θ_i is regarded as a special weight, it can be rewritten as

$$v_i = f\left(\sum_{j=0}^n w_{ji}v_j\right). \quad (2)$$

Among them, $w_{0i} = -\theta_i$ and $v_0 = 1$ are continuous functions to express the nonlinear transformation ability of neurons, and s functions are often used:

$$f(u_i) = \frac{1}{1 + e^{-u_i}}. \quad (3)$$

In this context, the purpose of this DM is to establish a NNS model according to the prediction characteristics of music pop trend. On this basis, the effectiveness of the model prediction is proved through empirical research, which shows the advantages of the prediction method based on NNS in the prediction of music pop trend. When the NNS model was published, it did not give a learning algorithm to adjust the connection weight between neurons [23]. However, we can use some common

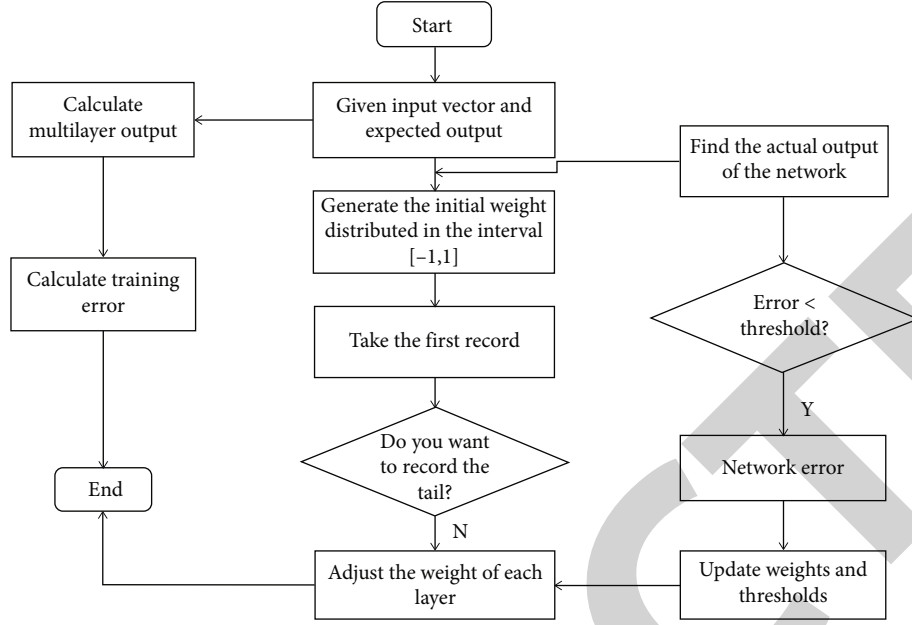


FIGURE 1: Flow chart of NNS algorithm.

algorithms to adjust the neuron connection weight according to the needs to achieve the purpose of learning. NNS is a widely used forward NNS model, which can train a NNS model that meets all training sets as much as possible according to different inputs and outputs. At this time, a large amount of weight modification is required, but in fact, because the derivative value approaches zero, the amount of weight modification is very small, so learning is slow. The same is true for hidden layer neurons. Once learning enters the “false saturation” state, it takes a long time to get rid of this state. The NNS is composed of three or more layers of neurons. The neurons in the same layer are not connected with each other, and only the upper neurons can connect the lower neurons. The NNS model is shown in Figure 2.

The application process of NNS model mainly includes two processes: one is to calculate the output vector according to the forward propagation of the input vector; second, the NNS algorithm reversely updates the weight and width standard between the input layer and the hidden layer and between the hidden layer and the output layer according to the error between the output vector and the expected vector adopts the method of gradient reduction of error function to learn until the sum of error squares between the actual output of the model and the expected output of the model is the smallest [24, 25]. The number of hidden neurons is uncertain, but the determination of its number is crucial. The basic idea of the NNS algorithm is to determine the search direction by gradient method and adjust the weights by using the unit vector in multidimensional space in this direction as the basic modifier, and the concept of basis function is introduced. The definition formula of basis function is as follows:

$$B(k)_w = \sqrt{\sum_{i=1}^{n_2} \sum_{j=1}^{n_2} \left[\frac{\partial E_k}{\partial W_{ij}} \right]^2}, \quad (4)$$

where E_k is the error function of network corresponding mode to (A_k, C_k) and n_2 is the number of neurons in the hidden layer.

There are two commonly used $x_{\text{mid}} = (x_{\text{max}} - x_{\text{min}})/2$ empirical formulas:

$$\begin{aligned} p &= \sqrt{d + m} + \varepsilon, \\ p &= \log_2 d. \end{aligned} \quad (5)$$

During normalization, the data is generally normalized to $[0, 1]$ or $[-1, 1]$, and the normalization formula is

$$\begin{aligned} x_{\text{new}} &= \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}, \\ x_{\text{new}} &= \frac{x - x_{\text{min}}}{2}, \\ x_{\text{new}} &= \frac{x - x_{\text{mid}}}{0.5}. \end{aligned} \quad (6)$$

Normalize the data to between $[0, 1]$ and $[-1, 1]$.

Weight correction from hidden layer to output layer is shown as follows:

$$\Delta w_{ij} = \frac{-\lambda \partial E_k / \partial w_{ij}}{B(k)_w}. \quad (7)$$

Weight correction from input layer to hidden layer is shown as follows:

$$\Delta v_{hi} = \frac{-\lambda \partial E_k / \partial v_{hi}}{\sum_{n=1}^{n_1}}. \quad (8)$$

The approaching progress of the NNS algorithm in the next step and the standard NNS algorithm needs many steps

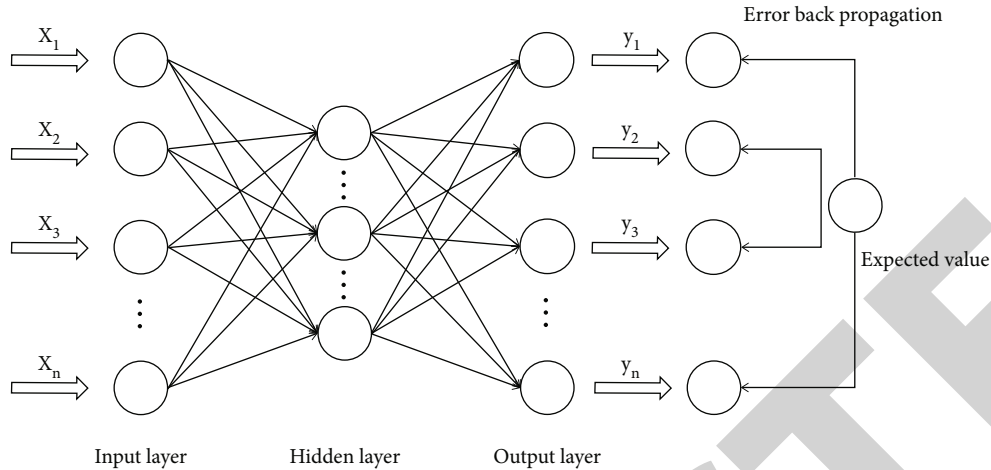


FIGURE 2: NNS model diagram.

TABLE 1: Song record.

Song_id	Artist_id	Public_time	Song_init_plays	Language	Gender
c81f89cf7edd24930641afa2e411b09c	03c6699ea836decbc5c8fc2dbae7bd3b	20210425	0	120	1
cOd7130777c1f1c417e78646946ed909	03c6699ea836decbc5c8fc2dbae7bd3b	20210426	1341	130	1
...

TABLE 2: User operation table.

User_id	Song_id	Gmt_create	Action_type	ds
7063b3dOc075a4d276c5fj6f4327cf4a	effb071415be51f1e845884e67cOf8c	1426405100	2	20210316
Odb66cOdd3993fd3504bb98c3beb15b3	f87ff481d8Sd2fj5335ab602f38a7655	1426416800	2	20210316
...

TABLE 3: Statistics of song click through rate.

Field name	Type	Field length	Remarks
Song_id	Varchar	254	Song ID
Action_type	Varchar	254	User action type
ds	Varchar	254	Date

to achieve. In the attenuation oscillation stage, because it is closer to the minimum error point and the attenuation oscillation decays very fast, the approximation degree achieved by the unit NNS algorithm by attenuating the oscillation every two steps can only be achieved by the standard NNS algorithm after many steps. Under the DM technology, the NNS algorithm is used to analyze the historical data of music and audience, and the more accurate prediction of music pop trend is realized. Based on the existing research, this paper will further explore and innovate the prediction of music pop trend.

4. Research on Prediction of Music Pop Trend

4.1. Data Statistical Processing. This experiment uses the competition data in the Ali music pop trend prediction competition

from April 1, 2021, to September 1, 2021. The data is from the operation record data of real users in the application of Ali music platform. Song_id represents the ID of the song and the unique identifier of the song. Artist_id refers to the ID of the artist and the unique identifier of the artist. Public_time indicates the release time of the song. Song_init_plays indicates the initial number of times the song is played, that is, the initial heat. Language indicates the language type of the song. Gender indicates the artist's gender. This data records the operation records of a large number of users on all songs of 100 artists. The data includes two parts: recording the data of songs and recording the user's operation records of songs. The data of recorded songs are shown in Table 1.

This experiment recorded all songs of 120 artists, with a total of 26,987 records. Each record contains basic information such as song ID, artist ID, song release time, and artist gender. User_id indicates the user id that generates the operation record, which is the unique identifier of the user. Song_id represents the song ID, the unique symbol of the song. Gmt_create indicates the timestamp of the record. Action_type indicates the type of user operation, 1 indicates audition, 2 indicates download, and 3 indicates collection. ds indicates the date when the record was generated. Records of the user's operation of songs are shown in Table 2.

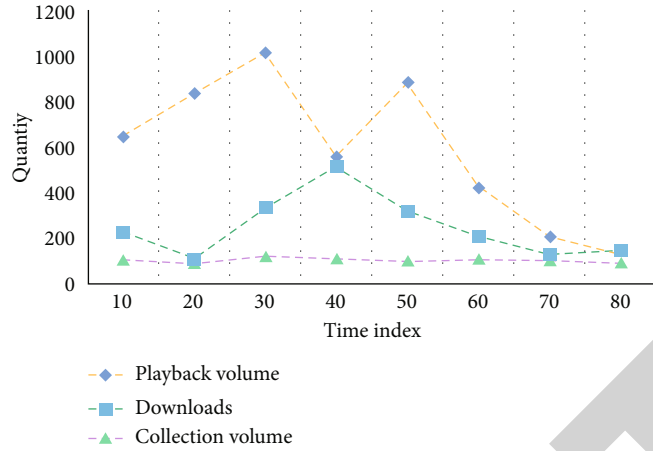


FIGURE 3: Average number of songs played, downloaded, and collected by singers.

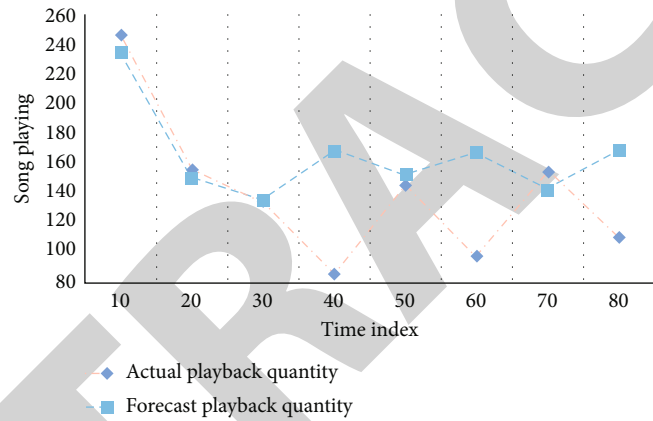


FIGURE 4: Comparison between predicted playback volume and actual playback volume of singer songs.

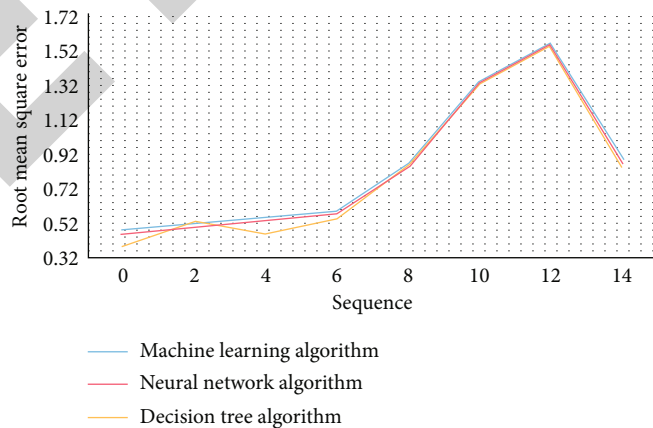


FIGURE 5: Comparison of root mean square error of three algorithms.

Tables 1 and 2 record the operation records of 26,987 songs of 120 artists by all users from April 1, 2021, to September 1, 2021. The data in these two tables cannot be directly used to predict the amount of songs that artists listen to every

day in a certain period of time. It is necessary to first make statistical analysis of the data, then build the corresponding data set, and finally use the prediction-related algorithm to build an algorithm model for prediction.

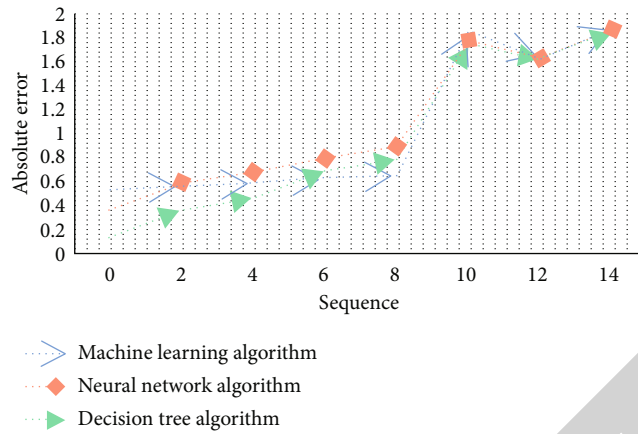


FIGURE 6: Comparison of average absolute error of three algorithms.

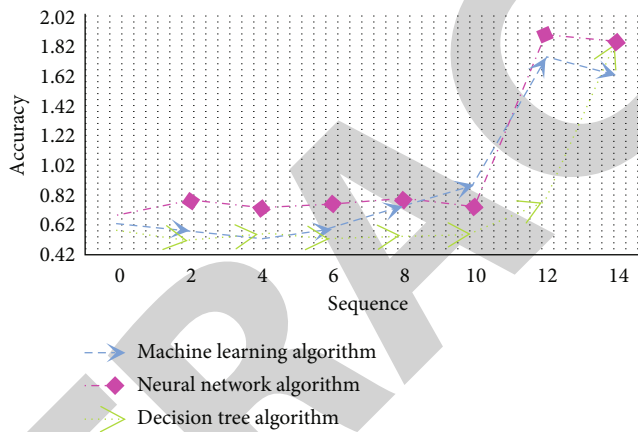


FIGURE 7: Comparison of accuracy of three algorithms.

Count the daily audition, download, and collection of each song. Create a new statistical table in the database. The format is shown in Table 3.

As can be seen from Table 3, use SQL statements to count the data in the above table format and insert the statistical data into the table. For SQL statistical statements, please refer to the following: insert into songsta select count num, DS, song ID, action type from P2 Mars t ianchi user actions group by DS, song ID, and action type. After the above SQL is used to insert statistical data, the table contains 1,663,653 pieces of data.

4.2. Experimental Result. This experiment is aimed at the average number of songs played, downloaded, and collected by singers in the past six months. The test results are shown in Figure 3.

As can be seen from Figure 3, among the songs played, collected, and downloaded by a certain singer, the best one to predict whether a certain singer will be popular in a certain period of time in the future is the playing volume.

According to the song playing data of all singers, the mean value is removed, the variance is normalized and scaled to [-1, 1], and the length of mean filter is set to ave Filter = 4. The

prediction curve shown in Figure 4 can be obtained by using the original playback volume for prediction.

It can be seen from Figure 4 that there is a deviation in the prediction of some singers in the results of the prediction of the original data. The reason may be that the data collected by establishing the prediction model is not comprehensive or accurate, resulting in the deviation of the prediction results. The defect of mathematical model itself. Therefore, the main purpose of this chapter is to predict the daily song audition volume of artists in a certain period of time through the prediction algorithm, so as to find the artists with higher song audition volume according to the pre measurement.

In this experiment, 10 artists are randomly selected in the data set, and the number of artists' songs from July 1, 2021, to September 1, 2021, is predicted by using the machine learning algorithm, decision tree algorithm, and NNS algorithm. The prediction results root mean square error, average absolute error, and accuracy are shown in Figure 5–7.

It can be seen from Figures 5–7 that in the prediction of the number of songs played by 10 artists, the prediction effect of the NNS algorithm proposed in this paper is reduced from 0.074 and 0.045 to 0.044 and 0.032, respectively, and the error rate is reduced by 35.7% and 29.4%, respectively. In addition, among the accuracy of the three methods, the NNS algorithm

in this paper has the highest accuracy. Therefore, it can be proved that the model proposed in this paper is more suitable for predicting the trend of music popularity.

Accurate prediction of music trends can enhance user experience and increase platform revenue. It can also improve the popularity of singers and dig out singers who have unlimited prospects in the future. Therefore, it is of great significance for singers, music lovers, and online music platforms to accurately predict the pop trend of music using historical data.

5. Conclusions

In this paper, the prediction of music popularity trend based on NNS and DM technology is studied. In the prediction of the number of songs played by 10 artists, the NNS algorithm proposed in this paper reduces the prediction effect from the original 0.074 and 0.045 to 0.044 and 0.032, respectively, and the error rates are reduced by 35.7% and 29.4%, respectively, compared with the learning algorithm and the decision tree algorithm. Among the three methods, the NNS algorithm in this paper has the highest accuracy. Who can better predict the popular trend of music and show its progressiveness, technicality, and intelligence will win users and market. The popular trend of today's music can be obtained by deeply mining, analyzing, and predicting the preference trend of the audience. Make use of the huge music library resources and user behavior to form music big data. The potential rules of users' behaviors, song information, and artist information are mined, and effective information is selected for data cleaning, data integration and data conversion. The construction speed of NNS model needs to be further improved. Due to the limited experimental environment, it is impossible to realize parallel computation of NNS algorithm through cloud platform, which leads to the time-consuming training of NNS algorithm, and the whole prediction process is slightly slower. The prediction of pop music trend can be obtained by deep mining and analysis of audience preferences and preferences on the basis of massive user data. However, the research still has some limitations. The study did not discuss and analyze the probability that the prediction deviation and the deviation are different. Including incomplete feature sets, biased training samples, and chaotic data sets. Therefore, further discussion and analysis are needed 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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Nanchang Hangkong University: Research on Artificial Intelligence Composition Based on Deep Neural Network (No. YG2021021).

References

- [1] A. C. North, A. E. Krause, and D. Ritchie, "The relationship between pop music and lyrics: a computerized content analysis of the United Kingdom's weekly top five singles, 1999–2013," *Psychology of Music*, vol. 49, no. 4, pp. 735–758, 2021.
- [2] C. A. Medjesky, "Tween Pop: Children's Music and Public Culture. TylerBickford. Duke UP, 2020. 229 pp. \$25.95 paper," *Journal of Popular Culture*, vol. 54, no. 3, pp. 672–674, 2021.
- [3] A. Vesey, "'A way to sell your records': pop stardom and the politics of drag professionalization onRuPaul's drag race," *Television & New Media*, vol. 18, no. 7, pp. 589–604, 2017.
- [4] J. Kang, J. Kim, M. Yang et al., "Behind the scenes of K-pop fandom: unveiling K-pop fandom collaboration network," *Quality & Quantity*, vol. 56, no. 3, pp. 1481–1502, 2022.
- [5] L. Jansen and M. Westphal, "Rihanna works her multivocal pop persona: a morpho-syntactic and accent analysis of Rihanna's singing style," *English Today*, vol. 33, no. 2, pp. 46–55, 2017.
- [6] A. Goh, M. Thompson, and I. Szeman, "Themed book review: dangerous mediations: pop music in a Philippine prison videoby Áine Mangoang," *Feminist Review*, vol. 127, no. 1, pp. 153–154, 2021.
- [7] E. P. Capistrano, "Understanding Filipino Korean pop music fans," *Asian Journal of Social Science*, vol. 47, no. 1, pp. 59–87, 2019.
- [8] P. A. Sanchez Sanchez, J. C. Zuluaga, D. G. Herazo et al., "Knowledge discovery in musical databases for moods detection," *IEEE Latin America Transactions*, vol. 17, no. 12, pp. 2061–2068, 2019.
- [9] X. Tang, "Fuzzy clustering based self-organizing neural network for real time evaluation of wind music," *Cognitive Systems Research*, vol. 52, pp. 359–364, 2018.
- [10] J. Gan, "Music feature classification based on recurrent neural networks with channel attention mechanism," *Mobile Information Systems*, vol. 2021, Article ID 7629994, 10 pages, 2021.
- [11] W. Zhang, "Research on music genre classification method based on DM technology," *Revista De La Facultad De Ingenieria*, vol. 32, no. 5, pp. 723–732, 2017.
- [12] A. Ks, B. Ena, and A. Ky, "Non-local musical statistics as guides for audio-to-score piano transcription," *Information Sciences*, vol. 566, pp. 262–280, 2021.
- [13] A. Scott-Maxwell, "K-pop flows and Indonesian student pop scenes: situating live Asian pop music in an 'Asian' Australia," *Media International Australia*, vol. 175, no. 1, pp. 20–35, 2020.
- [14] Y. Liu, Q. Tang, X. Zhao et al., "Neural activation of different music styles during emotion-evoking," *Procedia Computer Science*, vol. 49, no. 6, pp. 1546–1560, 2021.
- [15] W. Lim, A. Furnham, and A. McClelland, "Investigating the effects of background noise and music on cognitive test performance in introverts and extraverts: a cross-cultural study," *Psychology of Music*, vol. 50, no. 3, pp. 709–726, 2022.

Retraction

Retracted: Prediction of Concrete Compressive Strength Based on the BP Neural Network Optimized by Random Forest and ISSA

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Chen, D. Zhu, X. Wang, C. Zhou, and X. Chen, "Prediction of Concrete Compressive Strength Based on the BP Neural Network Optimized by Random Forest and ISSA," *Journal of Function Spaces*, vol. 2022, Article ID 8799429, 20 pages, 2022.

Research Article

Prediction of Concrete Compressive Strength Based on the BP Neural Network Optimized by Random Forest and ISSA

Gang Chen,¹ Donglin Zhu ,² Xiao Wang ,³ Changjun Zhou ,⁴ and Xiangyu Chen¹

¹College of Physics and Information Engineering, Fuzhou University, Fuzhou, Fujian 350108, China

²School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China

³Xingzhi College, Zhejiang Normal University, Jinhua, Zhejiang 321000, China

⁴College of Mathematics and Computer Science, Zhejiang Normal University, Jinhua, Zhejiang 321004, China

Correspondence should be addressed to Xiao Wang; tianzhu213@zjnu.cn

Received 16 June 2022; Revised 2 August 2022; Accepted 8 August 2022; Published 27 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Gang 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 modern engineering construction, the compressive strength of concrete determines the safety of engineering structure. BP neural network (BPNN) tends to converge to different local minimum points, and the prediction accuracy is not high in the prediction of the compressive strength of concrete. Therefore, a prediction model based on the BPNN optimized by improved sparrow search algorithm (ISSA) and random forest (RF) is proposed to enhance the generalization ability and prediction accuracy of BPNN for compressive strength of concrete. In terms of algorithm improvement, three improvements are proposed for SSA: Latin hypercube sampling is introduced to initialize the location of sparrows and increase the diversity of sparrows; the somersault foraging strategy is used to enrich the optimal position of producers; and combining with the cyclone foraging mechanism, the position updating process of the scroungers is optimized to obtain a better foraging position. In terms of performance evaluation of the algorithm, the ablation experiment verifies that the three improved strategies have improved effects in SSA, and the performance of ISSA on the CEC2017 benchmark function is better than other peers. In terms of predictive index screening, the important features are selected as the input variables of the model by random forest. The prediction results show that compared with the RF-BPNN model and models optimized by other algorithms, RF-ISSA-BPNN model has the lowest prediction error, and the expected value fits the real value better.

1. Introduction

Machine learning can mine the inherent relationships from a large number of historical data for classification or prediction. However, in addition to deep learning, random forest, support vector machine, and other methods in machine learning, BP neural network is also gradually applied to the prediction of various engineering fields for its great effect. Liang W et al. [1] study and analyze coal ash deformation temperature with linear regression method and FactSage calculation and introduce BP neural network to obtain accurate prediction results. Xu B et al. [2] use static and dynamic methods to simulate in BP neural network, respectively, which ensure the accuracy of road temperature prediction

in different stages. Liu Y et al. [3] prove that BP neural network is practical and feasible in predicting thermal error of five-axis machining center. Dai S et al. [4] propose a prediction model combining multiple regression and BP neural network, which showed good prediction performance in WFFZ height prediction. Similarly, it can also be applied to predict the output pressure of the sensor [5], performance evaluation of manufacturing collaborative logistics [6], depth of concrete carbonation, and amount of steel corrosion [7].

Because BP neural network has the advantages of self-learning, generalization ability, and fault tolerance, it has been widely used by scholars in many fields. However, as the scope of application becomes wider and wider, many

shortcomings are exposed. The traditional BP neural network is a local search optimization method, and the weight of the network is adjusted along the direction of local optimization, which tends to converge to different local minima, leading to the failure of network training. In addition, it is also susceptible to the influence of initial weights and thresholds, and different weights and thresholds result in different training results. Therefore, many scholars use intelligent optimization algorithm to improve its generalization performance and learning ability by optimizing weights and thresholds on the original basis. For example, Dou K et al. [8], Wang H et al. [9], and Supraja P et al. [10] both adopt genetic algorithm (GA) for optimization, while Yuan H et al. [11], Jiang G et al. [12], and Wang W et al. [13] optimize key parameters of BP neural network through particle swarm optimization (PSO) and mind evolutionary algorithm (MEA) with better optimization performance, and this indicates that the performance of the model with BP neural network as the core and intelligent optimization algorithm as the auxiliary is more outstanding than the original BP neural network. In order to seek a breakthrough in the performance of the predictive model, some scholars have done a lot of work on the improvement of the algorithm. For instance, Wu Y et al. [14] adopt the adaptive learning rate to improve the prediction model, and optimize the model with a new improved algorithm integrating GA and SA. Zhang W et al. [15] improve the convergence factor and position update formula in the standard gray wolf algorithm, and used the improved gray wolf algorithm to find the two optimal values in the prediction model, so that the model can meet the requirements of accuracy and real time of short-term traffic flow prediction. Tian H et al. [16] add nonlinear decline factor to the inertia weight of particle swarm optimization, and the IPSO-BPNN model can effectively predict the yield of winter wheat. Wu L et al. [17] introduce the crossover and mutation operation in GA into the improved fruit fly optimization algorithm (FOA) to establish a corresponding GAIFOA-BP model, and the prediction of fatigue life and fatigue consumption by the model can be closer to the actual results.

The above researches only optimize the prediction model by improving the algorithm, but ignore the importance of feature selection that affects the prediction accuracy [18]. In a sense, the contribution of screening good predictive indicators to the improvement of prediction accuracy may be greater than algorithm optimization and combination of models [19]. RF is selected to measure the importance of each characteristic variable [20–22], and a certain threshold is set for screening to find several characteristic variables highly correlated with the dependent variable and eliminate the characteristic variables with low importance, thus reducing the complexity of the prediction model. Similarly, the improvement of the algorithm has also become an important breakthrough to improve the prediction model. Sparrow search algorithm (SSA) has attracted the attention of scholars in recent years, which divides the search population into three roles: producer, scrounger, and scout. The three roles cooperate with each other to find the optimal

value by their position updating mechanism. SSA has more advantages than GWO, PSO, and GSA in terms of search accuracy, convergence speed and stability [23]. However, there are three problems as follows: (1) Randomly generated initial positions may cause the sparrow population to be unevenly distributed throughout the search solution space; (2) in the stage of updating position of producers, the overall trend decreases as the iteration goes on; and (3) in the stage of updating position of scroungers, the value tends to 0 when the population size is relatively large or the sparrow population converges. In terms of algorithm optimization, Latin hypercube sampling firstly is used to replace the formation mode of the original population to enhance the quality of the initial individuals. Secondly, somersault foraging strategy is introduced in the stage of updating the location of producers to enrich the optimal position of producers and expand its search space. Finally, cyclone foraging mechanism is introduced in the stage of updating the location of scroungers to obtain a better foraging position and enhance the escape ability of the local optimal solution. In this paper, the CEC2017 benchmark function is selected for simulation experiments, the feasibility and rationality of the three strategies for improving the algorithm are verified by ablation experiments, and ISSA is compared with 4 classic optimization algorithms and 2 other improved algorithms. The comprehensive ranking of the simulation results shows that the optimization ability of ISSA is superior to other six algorithms. In the prediction of compressive strength of concrete, several features with high importance are screened out as input variables by RF first, and then, the optimal weight and threshold value are found in BPNN by optimization ability of ISSA, and the RF-ISSA-BPNN prediction model is established; finally, three algorithms (PSO, SSA, and chaotic sparrow search algorithm [24](CSSA)) with excellent performance on the CEC2017 benchmark function are selected to establish the corresponding prediction model (RF-BPNN, RF-PSO-BPNN, RF-SSA-BPNN, RF-CSSA-BPNN, and ISSA-BPNN), and they are compared horizontally and vertically. From the predicted data of the final concrete compressive strength, it can be found that the MAE and RMSE values of RF-ISSA-BPNN are smaller than those of the other five models, the predicted data fit the actual data better, and the improvement of SSA and the feature selection based on random forest can improve the prediction model, which can provide a reliable theoretical reference for the safe construction of the project.

In summary, the main contributions of this paper are the following aspects:

- (i) In terms of algorithm optimization, Latin hypercube sampling is used to initialize the population, somersault foraging strategy is proposed to enrich the optimal position of producers, and cyclone foraging mechanism is proposed to obtain a better foraging position
- (ii) Ablation experiments are conducted to verify the feasibility and rationality of the three strategies, and the performance of ISSA algorithm is verified

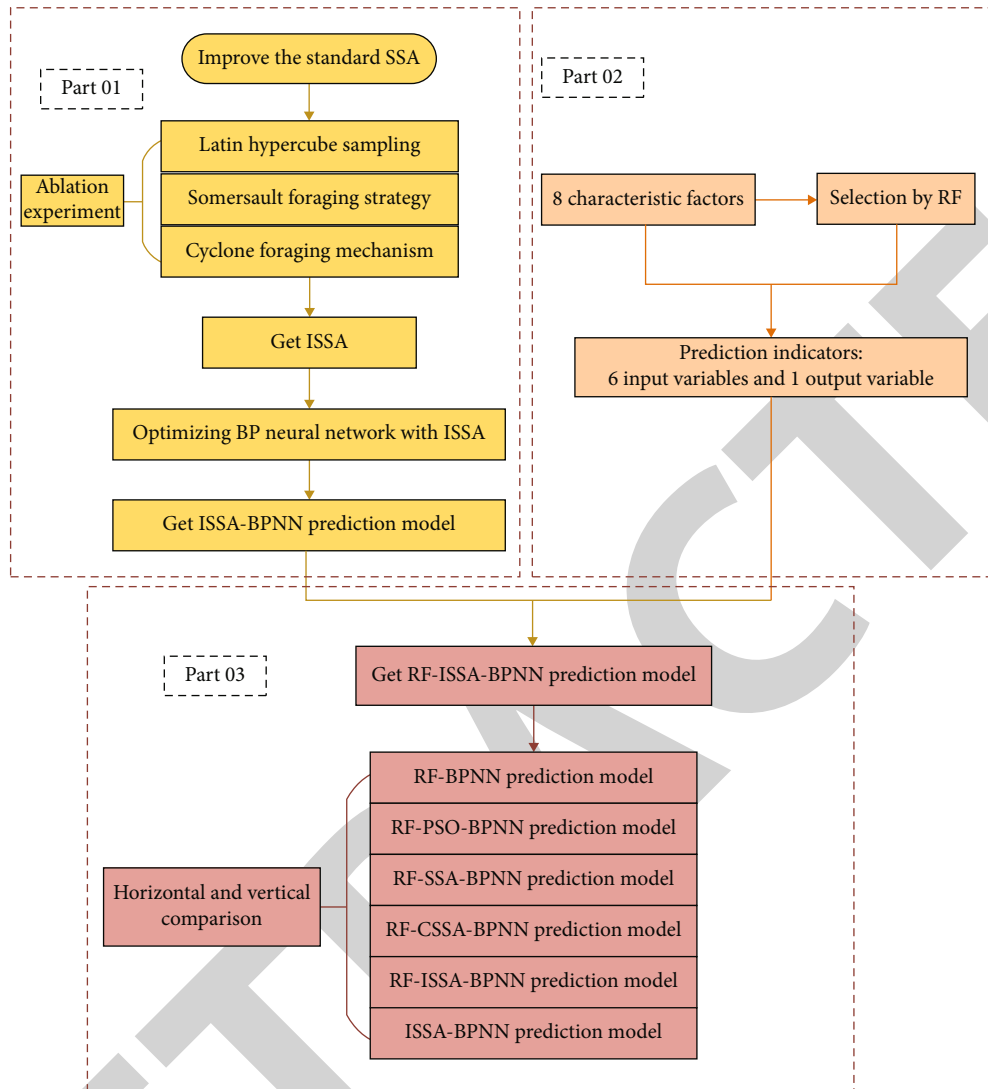


FIGURE 1: Brief flow chart of this research work.

by comparing it with some classical algorithms and two improved sparrow algorithms

- (iii) RF is used to measure the importance of each feature, and 6 features with VIM greater than 0.09 are screened as input variables
- (iv) The RF-ISSA-BPNN prediction model of concrete compressive strength is established and compared with the prediction model without RF and four kinds of prediction models with RF in both horizontal and vertical aspects. In addition, two evaluation indexes (MAE and RMSE) are selected as the measurement standards of the prediction accuracy

The rest of this paper is summarized below. Section 2 is an introduction to the related theoretical background and three improvement strategies of the sparrow search algorithm and gives the ISSA algorithm flow chart. Section 3 is divided into three parts: the ablation experiment, the comparison experiment with other algorithms, and the analysis

of the time complexity of the algorithm. Section 4 mainly includes the introduction of experimental data, the establishment of prediction model, and the comparison of prediction simulation experiment results. Finally, Section 6 is the summary of the whole paper. Figure 1 is a brief flow chart of this research work.

2. Theory

2.1. Related Theoretical Background

2.1.1. BP Neural Network. Figure 2 shows the basic structural framework of BPNN. BPNN is a multilayer feedforward network trained according to error back propagation algorithm [25]. It can learn and realize any complex nonlinear mapping between input and output through training of a large number of data samples [26] and then adjust weights and thresholds continuously through back propagation to minimize the error of output signals. However, it has shortcomings such as sensitive weight setting during fitting. Once

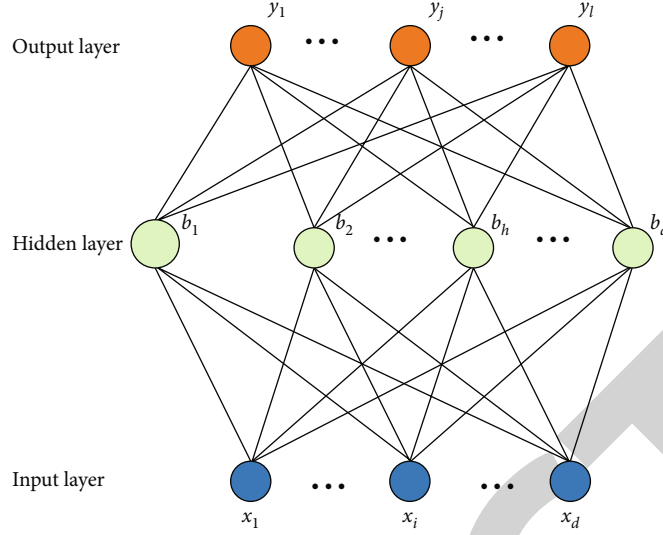


FIGURE 2: Schematic diagram of basic structure of BPNN.

the threshold and weight are set incorrectly, the performance of the model may be greatly reduced.

2.1.2. Random Forest for Feature Selection. In practical applications, there are often dozens of attributes of data set, and even the curse of dimensionality may occur. Therefore, data preprocessing for dimensionality reduction is a critical step in machine learning tasks. Feature selection relies on the feature selection function of the machine learning model itself, and selects the more important features from the input feature variables to achieve dimensionality reduction and simplify the complexity of the model to a certain extent. Generally speaking, there are two purposes for feature selection [27]. The first is to find highly correlated important variables to achieve the purpose of explanation, and the second is to find a small number of feature variables that can make good predictions.

Random forest [28] is not only widely used in prediction, but also in feature selection. It has better robustness and faster learning speed to noise and missing data, and its feature importance can be used as a feature selection tool for high-dimensional data [29]. The measurement indicators of characteristics in random forest mainly include Gini index [30] and out-of-bag data error rate [31, 32]. In this study, RF mainly uses the Gini index to calculate the average impurity and is used as an evaluation index to measure the contribution of each characteristic variable in the compressive strength of concrete. The higher the Gini index, the higher the average impurity, which shows that the importance of characteristic variables is more significant. The Gini index is represented by GI, and the variable importance score is represented by VIM.

2.1.3. Sparrow Search Algorithm. According to the foraging behavior of sparrows, the sparrow population is divided into producers and scroungers. Producers usually store high energy, which determines the direction of the whole population during the foraging process, while the scroungers

update their positions according to the foraging information provided by the producers. As long as sparrows can find a better source of food or have a higher energy reserve, they can become producers, but it is worth noting that the ratio of the producers and the scroungers remains constant in the entire population. The position of the sparrow represents a set of effective solutions in the search space, which is defined as $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,d})$, $i = 1, 2, \dots, n$, where d is the dimension and n is population size. The energy reserve of the sparrow represents fitness value, which is defined as $f(x_i)$. The location of the producers is updated as

$$x_{i,j}^{t+1} = \begin{cases} x_{i,j}^t \cdot \exp\left(\frac{-i}{n \cdot \text{iter}_{\max}}\right), R_2 < ST \\ x_{i,j}^t + Q \cdot L, R_2 \geq ST \end{cases}, \quad (1)$$

where $x_{i,j}^t$ is the position of the i -th sparrow in the j -dimension ($j = 1, 2, \dots, d$) at the t -th iteration. ST ($ST \in [0.5, 1.0]$) is the safety threshold. R_2 ($R_2 \in [0, 1]$) is the alarm value. α ($\alpha \in (0, 1)$) is a random number. iter_{\max} is the maximum number of iterations. L is a d -dimensional row vector with all elements of 1. Q is a random number that obeys normal distribution. When $R_2 < ST$, there is no threat from predators in the environment, and producers can search for food in a wide range. When $R_2 \geq ST$, some sparrows are aware of the presence of predators, and all individuals should move away from their current position to avoid predators.

Scroungers monitor the producers over time, and they often move around producers in the best position and compete with them for resources. The location of the scroungers is updated as

$$x_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{x_{\text{worst}}^t - x_{i,j}^t}{t^2}\right), i > n/2 \\ x_p^{t+1} + |x_{i,j}^t - x_p^{t+1}| \cdot A^+ \cdot L, \text{otherwise} \end{cases}, \quad (2)$$

where x_{worst}^t is the current global worst position. x_p^{t+1} is the best position currently found by the producer. A is a row

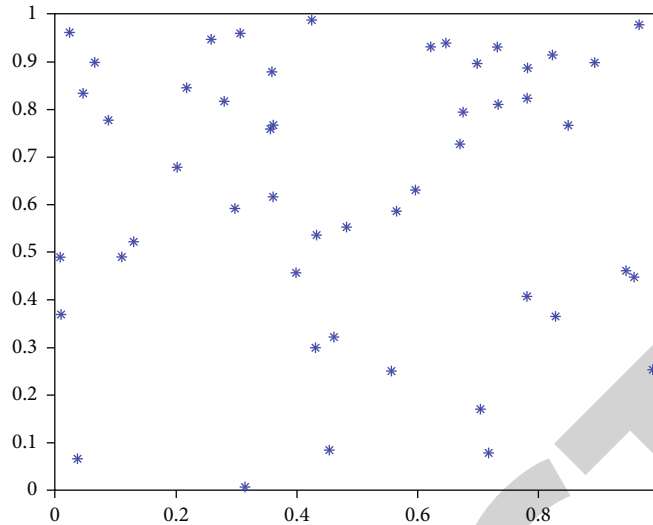


FIGURE 3: Rand function.

vector whose element can only be 1 or -1 , and $A^+ = A^T (AA^T)^{-1}$. When $i \leq n/2$, scroungers search for food around the best location found by the producers. The remaining sparrows are starving and can only fly to other locations to find food.

In SSA, some sparrows will be selected to adopt the reconnaissance and early warning mechanism, so sparrows in different positions of the population will choose different coping methods in the face of incoming danger. When a sparrow is aware of danger, it will actively approach its partners in or around the safety circle to increase its own safety factor. The position of the scouts is updated as in

$$x_{i,j}^{t+1} = \begin{cases} x_{\text{best}}^t + \beta \cdot |x_{i,j}^t - x_{\text{best}}^t| \cdot f_i > f_g \\ x_{i,j}^t + K \cdot \left(\frac{|x_{i,j}^t - x_{\text{worst}}^t|}{(f_i - f_w) + \varepsilon} \right) \cdot f_i = f_g \end{cases}, \quad (3)$$

where x_{best}^t is the current global optimal position. K ($K \in [-1, 1]$) is a random number. ε is the smallest constant to prevent the occurrence of 0 in the denominator. f_w is the current global worst fitness value, f_i is the fitness value of the scouts, and f_g is the current global best fitness value. β represents a step size control parameter that obeys the standard normal distribution. When $f_i > f_g$, the surroundings of the current global optimal position are safe, and the sparrows at the edge of the population realize the appearance of predators and quickly move around x_{best}^t . When $f_i = f_g$, sparrows in the center of the population should change their search strategy in time and seek protection from nearby partners to reduce the risk of being predation.

2.2. Improved Sparrow Search Algorithm

2.2.1. Latin Hypercube Sampling. Population initialization is an indispensable part of swarm intelligence optimization algorithm, and the convergence of swarm intelligence algorithm is easily affected by the distribution of the initial population [33]. The initial population of the traditional

sparrow search algorithm is generated based on random function in the feasible region. It can be found that the randomly generated population is not evenly distributed in Figure 3, which greatly reduces the efficiency of optimization. Latin hypercube sampling (LHS) is adopted to initialize the population to ensure the randomness and uniform distribution of sample points and improve the efficiency of optimization. As shown in Figure 4, random sample points are evenly distributed within the feasible region, which enrich the diversity of the primary population. Among them, population size is 50, dimension is 2, and the interval is $[0, 1]$.

Taking m samples in n -dimensional vector space as an example, the specific steps of LHS [34] are as follows:

- (1) The sample number and dimension of vector space are determined
- (2) Nonoverlapping m equal parts with equal probability is generated in each dimension
- (3) A random number is generated in each cell, so the sampling matrix $[m, n]$ is formed
- (4) A number is randomly selected in each column of the sampling matrix to form a vector

Latin hypercube sampling is used to initialize the population in SSA, so the number of populations is m in the sampling matrix, and the multidimensional decision variables correspond to the n -dimensional vector space. LHS can be used to generate a sampling matrix $[m, n]$, in which each number in each column is generated by different cells and arranged in disorder. Therefore, a population with a wider and more uniform distribution range is formed, and the probability of obtaining a solution with good diversity and convergence is higher.

2.2.2. Somersault Foraging Strategy. In the stage of updating the location of producers, the position of sparrows shows an overall decreasing trend with the progress of iteration when

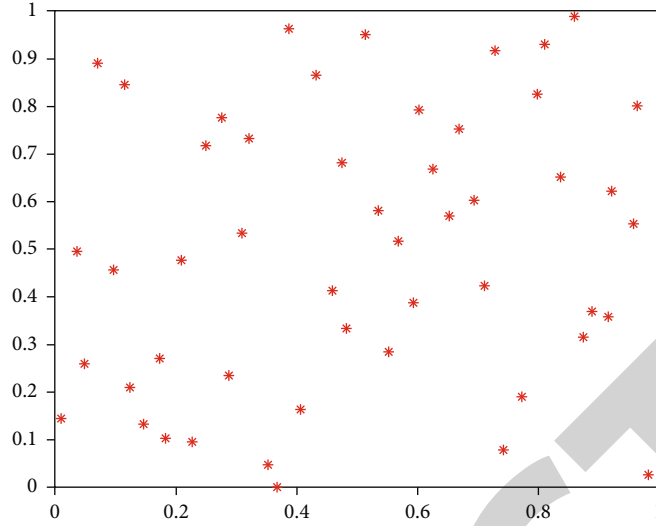
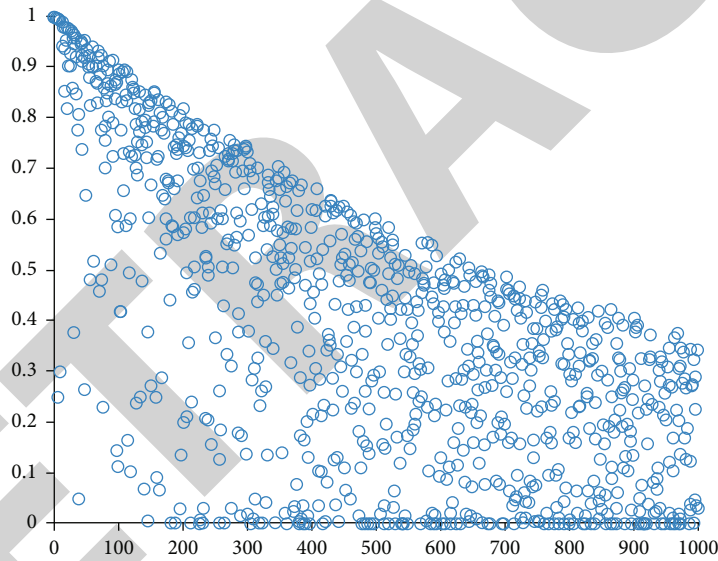


FIGURE 4: LHS.

FIGURE 5: $f(x) = \exp(-x/1000 \times \alpha)$.

$R_2 < ST$. It can be found that the value range of $f(x)$ has changed from $[0,1]$ to $[0,0.4]$ in Figure 5, which means that the diversity of the population may gradually decrease in the later iterations, and the probability of obtaining excellent solutions will also decrease. Therefore, the somersault foraging strategy [35] is introduced to enrich the optimal position of producers, which opened new foraging horizons for the whole population. The somersault foraging strategy is to take the position of the food (the optimal position) as the center point, and the individual always update its position by somersaults around the optimal position, and the expression is

$$x_{i,j}^{t+1} = x_{i,j}^t + S \cdot (r_2 \cdot x_{\text{best}}^t - r_3 \cdot x_{i,j}^t), \quad (4)$$

where x_{best}^t is the optimal position; S is the somersault factor, and the general value is 2; and r_2 and r_3 are two random numbers in $[0,1]$.

It can be seen from Equation (4) that the search area of producers is between the current position and the symmetrical position around the optimal position it currently finds, and it uses the global optimal position as the fulcrum to update its position. But with the iteration of the population, the range of sparrows foraging for somersaults is also shrinking, and all sparrows will gradually approach the optimal position. We use somersault foraging strategy into the stage of updating the location of producers to get an opposite position with the optimal position as the center point, and the optimal position is selected between the current sparrow and the opposite position. If the fitness value at

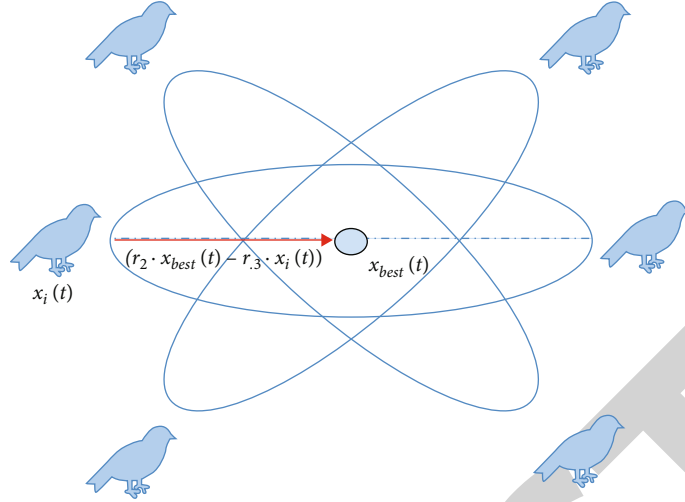


FIGURE 6: Producer's somersault foraging behavior.

the original position is inferior to that at the opposite position during each iteration, the original position is replaced. Otherwise, the original position is retained for the next generation. Different from the reverse learning strategy, the somersault strategy revolves around the optimal solution when updating the position, which makes the algorithm more convergent [36]. Figure 6 is a schematic diagram of the producer somersaulting.

2.2.3. Cyclone Foraging Mechanism. In the stage of updating the location of scrungers, the i -th scrunger is to get rid of the worst position in the current foraging process by virtue of the property of $\exp()$ function, so as to obtain a better foraging position when $i > n/2$. However, its value will gradually tend to 0 when n is relatively large or the sparrow population converges, which means that the scrungers cannot fly to other positions and the population diversity is also lost. Theoretically, all starving sparrows should have the chance to search randomly using food as a reference location. The cyclone foraging mechanism [35] can create such an opportunity for hungry individuals to randomly designate a reference position in the entire search space, which can make the hungry individuals far away from the optimal position to find a new location and improve the global search capability of algorithm. Figure 7 shows the cyclone foraging behavior of sparrows in a two-dimensional space. It can be seen that the sparrows follow the preceding sparrows along the spiral path towards the food. Mathematical expressions are shown in

$$x_{\text{rand}}^d = Lb^d + r \cdot (Ub^d - Lb^d), \quad (5)$$

$$x_i^d(t+1) = x_{\text{rand}}^d + r \cdot (x_{i-1}^d(t) - x_i^d(t)) + \beta \cdot (x_{\text{rand}}^d - x_i^d(t)), \quad (6)$$

$$\beta = 2e^{r_1 \cdot (\text{iter}_{\text{max}} - t + 1 / \text{iter}_{\text{max}})} \cdot \sin(2\pi r_1), \quad (7)$$

where x_{rand}^d is a randomly generated position in the search

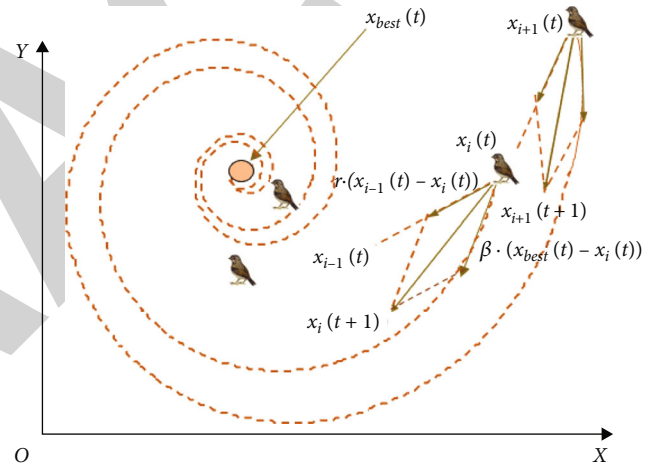


FIGURE 7: Cyclone foraging mechanism of individual sparrows.

space. Ub^d and Lb^d are defined upper and lower limits, respectively. r and r_1 are two random numbers in $[0,1]$. Therefore, the position of scrungers in ISSA is updated as

$$x_{i,j}^{t+1} = \begin{cases} x_{\text{rand}}^d + r \cdot (x_{i-1,j}^t - x_{i,j}^t) + \beta \cdot (x_{\text{rand}}^d - x_{i,j}^t), & i > n/2 \\ x_p^{t+1} + |x_{i,j}^t - x_p^{t+1}| \cdot A^+ \cdot L, & \text{otherwise} \end{cases} \quad (8)$$

2.3. ISSA Algorithm Flow Chart. In summary, the overall flow chart of ISSA implementation is shown in Figure 8.

3. Algorithm Performance Test

3.1. Simulation Environment and Test Function. All simulation tests are performed on a computer with memory: 16 GB DDR4, CPU: AMD Ryzen 5 4600H with Radeon Graphics and operating system: Windows10; the compilation and operation of the program are all carried out in the Matlab2018a environment. Each optimization algorithm is tested to look for the optimal solution on the CEC2017 benchmark function solving problem in this paper, and they are run independently for 30 times to eliminate the

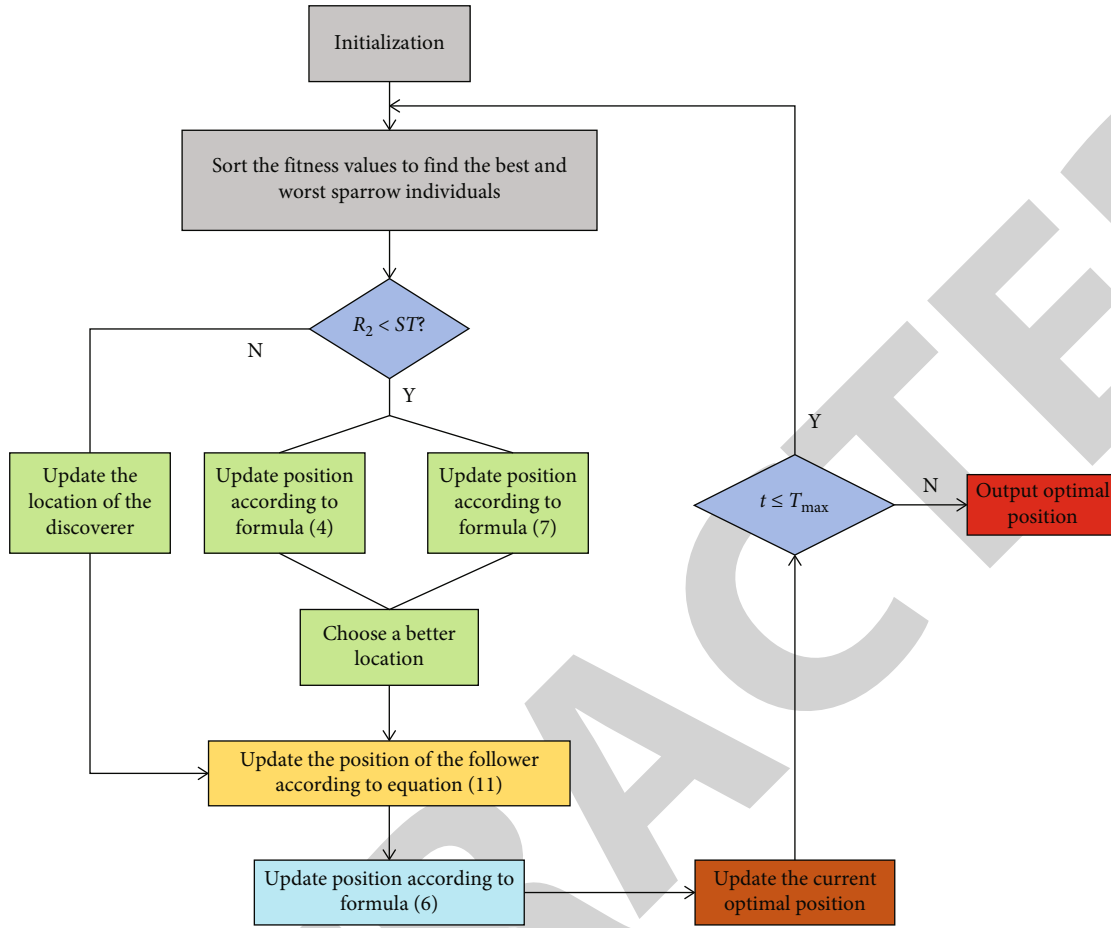


FIGURE 8: ISSA algorithm flow chart.

interference of accidental factors, and the results of mean, standard deviation (Std), and Friedman's ranking test [37] are recorded to obtain the average ranking (Ar) and the final ranking (Fr) on the entire CEC2017 benchmark function. The population size is set as 100, the $iter_{max}$ is set as 200, and the dimension in the benchmark function is set to 30.

3.2. Ablation Experiment. It is worth checking whether the three introduced strategies play a role in the improvement of SSA separately or at the same time, and the ablation experiment [38] is used to investigate the improvement effects of each strategy. SSA1 (SSA with the introduction of LHS), SSA2 (SSA with the introduction of somersault foraging strategy), SSA3 (SSA with the introduction of cyclone foraging mechanism), ISSA (SSA with three strategies introduced at the same time), and SSA are simulated and compared, as shown in Table 1. The algorithm parameters are uniformly set to $PD = 30$, $ST = 0.6$, and $SD = 70$.

From the final overall performance ranking, it can be found that the introduction of each strategy significantly improves the algorithm, and the performance of the SSA with the three strategies added at the same time is better than the SSA with the three strategies introduced separately, which means that LHS, somersault foraging, and cyclone foraging mechanism all play an improved role in SSA.

3.3. Simulation Comparison with Other Algorithms. The CEC2017 test function is often used to test the performance of intelligent optimization algorithms. In order to further evaluate the optimization effect of ISSA, it is compared with SSA, 3 classic swarm intelligence optimization algorithms, and 2 other improved sparrow algorithms; they are as follows: whale optimization algorithm (WOA) [39], artificial bee colony algorithm (ABC) [40], particle swarm optimization algorithm (PSO), chaos sparrow search algorithm (CSSA), and new chaos sparrow search algorithm (NCSSA) [41]. The general parameters are set the same to reflect the objectivity of each algorithm in the process of optimization. Table 2 lists other parameter settings of the seven algorithms.

As can be seen from the final results of Friedman's ranking test in Table 3, the overall ranking of ISSA is the first, followed by CSSA. Therefore, the optimization performance of ISSA on the CEC2017 benchmark function is the best. Compared with the classic algorithms ABC, WAO, and PSO, ISSA is better than WOA on 29 benchmark functions and is better than ABC and PSO on most benchmark functions. However, ABC is better than ISSA in optimization performance only on function F27, and the optimization performance of PSO on functions F4 and F11 is stronger than ISSA. On the whole, it can be seen that optimization

TABLE 1: Ablation experiments.

function	index	SSA	SSA1	SSA2	SSA3	ISSA
F1	Mean	1.3×10^5	1.3×10^5	6.4×10^4	1.4×10^5	9.4×10^4
	Std	5.7×10^4	8.0×10^4	4.3×10^4	7.6×10^4	5.6×10^4
	Rank	4	4	2	1	3
F3	Mean	5.1×10^4	5.1×10^4	4.7×10^4	6.5×10^4	4.8×10^4
	Std	6784.0	1.2×10^4	5714.2	1.5×10^4	9803.3
	Rank	3	3	1	4	2
F4	Mean	511.2	505.8	521.4	501.5	523.5
	Std	16.6	24.7	21.5	27.3	62.3
	Rank	3	2	4	1	5
F5	Mean	676.0	677.5	665.8	683.2	670.8
	Std	45.1	48.8	59.7	31.1	33.9
	Rank	3	4	1	5	2
F6	Mean	626.9	634.7	636.7	629.9	623.5
	Std	7.4	8.8	12.2	9.9	9.7
	Rank	2	4	5	3	1
F7	Mean	1053.8	1017.1	1038.8	1008.2	981.8
	Std	98.7	111.4	86.2	41.1	69.5
	Rank	5	3	4	2	1
F8	Mean	947.3	943.2	942.1	941.3	934.5
	Std	38.8	32.1	31.0	26.9	22.8
	Rank	5	4	3	2	1
F9	Mean	4969.7	4729.5	4301.5	3650.6	3805.2
	Std	940.7	984.3	830.1	1238.4	746.5
	Rank	5	4	3	1	2
F10	Mean	5076.4	4983.1	5465.1	4740.0	4504.6
	Std	683.9	744.3	804.9	661.6	696.8
	Rank	4	3	5	2	1
F11	Mean	1325.6	1314.1	1305.3	1300.9	1302.1
	Std	53.5	70.0	76.0	77.0	63.6
	Rank	5	4	3	1	2
F12	Mean	2.3×10^6	2.0×10^6	1.3×10^6	1.9×10^6	2.4×10^6
	Std	1.4×10^6	9.7×10^5	6.6×10^5	1.0×10^6	1.5×10^6
	Rank	5	4	1	2	3
F13	Mean	3.8×10^4	2.6×10^4	1.6×10^4	4.5×10^4	2.6×10^4
	Std	4.7×10^4	2.4×10^4	2.2×10^4	4.1×10^4	1.3×10^4
	Rank	3	2	1	4	2
F14	Mean	9.5×10^4	9.2×10^4	3.3×10^4	9.3×10^4	3.5×10^4
	Std	7.4×10^4	5.0×10^4	2.1×10^4	7.2×10^4	2.8×10^4
	Rank	5	3	1	4	2
F15	Mean	1.1×10^4	1.7×10^4	1.1×10^4	3.7×10^4	1.3×10^4
	Std	1.0×10^4	1.3×10^4	1.2×10^4	2.4×10^4	1.3×10^4
	Rank	1	3	1	4	2
F16	Mean	2778.1	2728.7	2864.8	2670.1	2610.3
	Std	145.3	259.6	256.1	387.0	293.1
	Rank	4	3	5	2	1
F17	Mean	2300.0	2207.1	2325.4	2245.1	2263.6

TABLE 1: Continued.

function	index	SSA	SSA1	SSA2	SSA3	ISSA
F18	Std	189.4	306.2	290.7	344.7	245.0
	Rank	4	1	5	2	3
	Mean	3.5×10^5	1.0×10^6	9.9×10^5	4.4×10^5	5.0×10^5
	Std	4.7×10^5	1.0×10^6	9.9×10^5	3.2×10^5	8.5×10^5
F19	Rank	1	5	4	2	3
	Mean	2.2×10^4	1.4×10^4	8133.5	1.2×10^4	1.3×10^4
	Std	1.6×10^4	1.3×10^4	5704.1	1.5×10^4	1.2×10^4
F20	Rank	5	4	1	2	3
	Mean	2654.8	2358.4	2769.2	2521.9	2519.2
	Std	160.4	248.2	169.5	128.1	149.4
F21	Rank	4	1	5	3	2
	Mean	2471.8	2460.5	2477.4	2445.1	2440.7
	Std	57.3	41.4	47.0	49.9	30.7
F22	Rank	4	3	5	2	1
	Mean	5352.7	3642.1	2545.9	2938.4	5087.8
	Std	1895.5	1992.0	927.5	1457.3	2009.7
F23	Rank	5	3	2	1	4
	Mean	2817.6	2831.5	2789.5	2858.7	2801.6
	Std	60.5	48.0	35.0	63.8	26.2
F24	Rank	5	3	1	4	2
	Mean	2975.0	2968.0	2991.7	2952.2	2992.5
	Std	63.4	33.1	46.7	37.7	30.3
F25	Rank	3	2	4	1	5
	Mean	2905.6	2894.1	2911.0	2900.2	2897.3
	Std	20.4	13.8	21.5	16.0	14.0
F26	Rank	4	1	5	3	2
	Mean	4675.2	5580.6	5229.6	5384.8	4684.3
	Std	856.7	712.2	971.1	701.3	1332.2
F27	Rank	1	5	3	4	2
	Mean	3235.7	3244.3	3202.5	3237.2	3227.7
	Std	15.1	25.3	7.7	12.5	38.1
F28	Rank	3	5	1	4	2
	Mean	3233.7	3247.9	3246.0	3247.4	3254.9
	Std	22.0	31.0	18.0	20.8	35.6
F29	Rank	1	4	2	3	5
	Mean	3972.0	3981.6	4078.6	4039.0	3912.8
	Std	192.2	191.8	169.7	269.7	193.8
F30	Rank	2	3	5	4	1
	Mean	4.1×10^4	2.6×10^4	3.5×10^4	5.2×10^4	1.9×10^4
	Std	2.5×10^4	2.7×10^4	5.9×10^4	2.6×10^4	1.6×10^4
	Rank	4	2	3	5	1
	Ar	3.552	3.172	2.966	2.621	2.276
	Fr	5	4	3	2	1

performance of ISSA is only slightly worse than the classical algorithm in a few benchmark functions. Compared with SSA, ISSA has better optimization performance on 22 benchmark functions. Compared with the improved algo-

rithms CSSA and NCSSA, ISSA outperforms them on 16 benchmark functions and 26 benchmark functions, respectively, and is slightly worse than NCSSA only on functions F19, F22, and F24. In summary, the global search and local

TABLE 2: Parameter settings.

Algorithm	Parameter settings
ABC	$L = \text{round}(0.6 * \text{dim} * N), a = 1$
WOA	$b = 1, a = 2 - (2t/T_{\max}), P^* = 0.5$
PSO	$c_1 = 2, c_2 = 2, w_{\max} = 0.9, w_{\min} = 0.2$
SSA	PD = 30, ST = 0.6, SD = 70
CSSA	PD = 30, ST = 0.6, SD = 70
NCSSA	PD = 30, ST = 0.6, SD = 70-round((70-50) × (t/200))
ISSA	PD = 30, ST = 0.6, SD = 70

development capabilities of ISSA are better than the other six algorithms, which fully explain the successful introduction of LHS, somersault foraging strategy, and cyclone foraging mechanism into sparrow search algorithm.

3.4. Time Complexity. The time complexity of an algorithm is used to measure the operating efficiency of the algorithm, and it reflects the pros and cons of an algorithm to a large extent. Therefore, ISSA is compared with SSA and other improved algorithms (CSSA and NCSSA) in terms of time complexity as shown in Table 4, where the dimension is denoted by d , and the time to solve the fitness function is denoted by $f(d)$. According to the introduction of the principle of the standard sparrow algorithm in Section 2.1.3, the algorithm is mainly composed of five phases: population initialization, location of producers updated, location of scroungers updated, location of scouts updated, and update of optimal location.

We discuss and analyze the time complexity of five phases of the algorithm, respectively, and it can be obtained through analysis in Table 4 that the time complexity of SSA, CSSA, NCSSA, and ISSA is equal, so the three improved algorithms do not increase the algorithm complexity in exchange for the improvement of performance. Combined with the experimental results on the CEC2017 benchmark function in Section 3.3, it can be seen that ISSA has better optimization performance than CSSA and NCSSA under the same time complexity of algorithm.

4. Prediction of Concrete Compressive Strength Based on RF-ISSA-BPNN Model

4.1. Experimental Data. With the rapid development of cement and concrete production technology, concrete has now become the largest amount of man-made building materials in the world. High-performance concrete (HPC) is a new type of concrete; using modern concrete technology, the use of high-quality raw materials, in addition to cement, aggregates, and water, must be used in a suitable water-cement ratio, mixed with sufficient high-quality mineral admixtures and efficient admixtures. As the academic community generally agreed that strength is the most important indicator of the performance of concrete, so many studies have long been done around how to improve the strength. Concrete as a very common material on modern construction projects, its compressive strength also determines the

quality of construction. Therefore, predicting the compressive strength from the available data is a challenging task. The prediction of compressive strength of concrete is a very complex nonlinear curve, and many factors directly or indirectly affect the compressive strength of concrete. The experimental data for the compressive strength of concrete used in this study consisted of 1,030 sets of data as well as nine properties. Among them, the compressive strength of concrete is affected by age, fine aggregate, cement, superplasticizer, blast furnace slag, water, coarse aggregate, and fly ash. The units of the last 7 factors are kg/m^3 , the unit of age is calculated by days, and the unit of compressive strength of concrete is MPa. The compressive strength of concrete is highly nonlinear with age and ingredients. It can be analyzed from Figures 9 and 10 that the compressive strength of concrete is highly nonlinearly correlated with age and composition. Detailed information about the input properties is mentioned in Table 5.

4.2. Establishment of RF-ISSA-BPNN Model. ISSA shows strong optimization ability in the CEC2017 test function, so ISSA can directly participate in the process of network parameter optimization. The so-called network parameters optimization is to find good weights and thresholds to minimize the global error in network [42], and then, train the optimized model to obtain the final prediction result. The dimensionality is $q \times (l + d + 1) + l$ because the dimension of individual sparrows is decided by the weights and thresholds obtained together.

The specific steps for establishing the RF-ISSA-BPNN model are as follows:

- (1) The sample data are imported, and RF is used to conduct feature selection on the sample data. The number of neurons in each layer, the transfer function, and the number of training times are determined so as to start building the network
- (2) Parameter and sparrow population are initialized, and the sum of the absolute values of the prediction errors obtained by training is set as the objective function, as shown in

$$f(x) = \sum_{i=1}^n |\tilde{x}_i - x_i|. \quad (9)$$

- (3) The optimal solution corresponding to the minimum fitness function value is found by ISSA, the optimal solution obtained by optimization is assigned to weights and thresholds, and then, the training begins. When the training accuracy requirements are met, the final prediction results are output

4.3. Application of RF-ISSA-BPNN Prediction Model

4.3.1. Evaluation of the Importance of Features. In the Python 3.7 operating environment, the importance score V

TABLE 3: Experimental results of 7 algorithms applied to CEC2017 test set.

function	index	ABC	WOA	PSO	SSA	CSSA	NCSSA	ISSA
F1	Mean	4.8×10^9	4.0×10^9	2.4×10^7	1.3×10^5	1.5×10^5	5.9×10^5	9.4×10^4
	Std	8.0×10^8	1.3×10^9	6.2×10^6	5.7×10^4	1.1×10^5	2.6×10^5	5.6×10^4
	Rank	7	6	5	2	3	4	1
F3	Mean	3.0×10^5	2.5×10^5	7.5×10^4	5.1×10^4	4.6×10^4	5.9×10^4	4.8×10^4
	Std	8.8×10^4	8.8×10^4	1.4×10^4	6784.0	7117.0	3913.0	9803.3
	Rank	7	6	5	3	1	4	2
F4	Mean	1900.0	1408.6	485.4	511.2	524.2	528.1	523.5
	Std	259.0	411.5	25.6	16.6	23.3	21.6	62.3
	Rank	7	6	1	2	4	5	3
F5	Mean	785.1	841.0	709.0	676.0	649.9	676.0	670.8
	Std	22.1	54.2	39.6	45.1	43.1	25.3	33.9
	Rank	5	6	4	3	1	3	2
F6	Mean	646.2	674.9	652.0	626.9	624.2	624.2	623.5
	Std	3.4	5.2	5.5	7.4	12.7	4.9	9.7
	Rank	4	6	5	3	2	2	1
F7	Mean	1060.4	1255.8	1098.6	1053.8	945.4	1038.7	981.8
	Std	20.3	117.8	40.2	98.7	90.3	123.7	69.5
	Rank	5	7	6	4	1	3	2
F8	Mean	1084.3	1046.1	966.5	947.3	946.3	944.4	934.5
	Std	17.3	22.5	28.1	38.8	23.7	31.2	22.8
	Rank	7	6	5	4	3	2	1
F9	Mean	1.1×10^4	9696.4	7369.8	4969.7	4221.6	4871.8	3805.2
	Std	1471.7	3438.2	1098.3	940.7	1153.1	789.6	746.5
	Rank	7	6	5	4	2	3	1
F10	Mean	9299.1	7029.2	5620.8	5076.4	4858.0	4767.8	4504.6
	Std	328.5	714.8	437.3	683.9	396.0	432.5	696.8
	Rank	7	6	5	4	3	2	1
F11	Mean	1.5×10^4	7897.1	1271.3	1325.6	1283.9	1324.6	1302.1
	Std	2741.5	3154.0	33.2	53.5	68.0	89.8	63.6
	Rank	7	6	1	5	3	4	2
F12	Mean	8.4×10^8	3.8×10^8	1.1×10^7	2.3×10^6	1.6×10^6	2.5×10^6	2.4×10^6
	Std	2.3×10^8	3.0×10^8	5.7×10^6	1.4×10^6	1.2×10^6	2.6×10^6	1.5×10^6
	Rank	7	6	5	2	1	4	3
F13	Mean	4.5×10^7	1.1×10^7	1.1×10^5	3.8×10^4	4.4×10^4	7.7×10^4	2.6×10^4
	Std	7.9×10^6	9.7×10^6	5.6×10^4	4.7×10^4	6.5×10^4	6.6×10^4	1.3×10^4
	Rank	7	6	5	2	3	4	1
F14	Mean	2.7×10^5	1.9×10^6	1.7×10^5	9.5×10^4	2.0×10^4	8.0×10^4	3.5×10^4
	Std	1.8×10^5	1.0×10^6	9.6×10^4	7.4×10^4	1.8×10^4	7.1×10^4	2.8×10^4
	Rank	6	7	5	4	1	3	2
F15	Mean	3.8×10^6	1.8×10^6	2.1×10^4	1.1×10^4	8246.8	1.6×10^4	1.3×10^4
	Std	9.5×10^5	1.7×10^6	8248.9	1.0×10^4	6448.7	1.4×10^4	1.3×10^4
	Rank	7	6	5	2	1	4	3
F16	Mean	4233.3	4227.7	3348.2	2778.1	2921.8	2741.1	2610.3
	Std	102.8	384.3	435.1	145.3	506.3	213.8	293.1
	Rank	7	6	5	3	4	2	1
F17	Mean	3077.4	3061.4	2321.6	2300.0	2304.7	2383.4	2263.6

TABLE 3: Continued.

function	index	ABC	WOA	PSO	SSA	CSSA	NCSSA	ISSA
F18	Std	179.8	542.8	269.5	189.4	287.1	217.1	245.0
	Rank	7	6	4	2	3	5	1
	Mean	2.2×10^7	1.2×10^7	6.7×10^5	3.5×10^5	6.6×10^5	1.1×10^6	5.0×10^5
F19	Std	9.5×10^6	9.4×10^6	4.2×10^5	4.7×10^5	8.4×10^5	1.6×10^6	8.5×10^5
	Rank	7	6	4	1	3	5	2
	Mean	1.4×10^6	1.7×10^7	4.2×10^4	2.2×10^4	8868.6	1.1×10^4	1.3×10^4
F20	Std	4.8×10^5	1.3×10^7	4.9×10^4	1.6×10^4	1.0×10^4	7202.9	1.2×10^4
	Rank	6	7	5	4	1	2	3
	Mean	3179.1	2790.0	2725.0	2654.8	2515.0	2570.3	2519.2
F21	Std	137.9	174.2	217.2	160.4	203.7	173.0	149.4
	Rank	7	6	5	4	1	3	2
	Mean	2583.3	2637.1	2536.4	2471.8	2466.6	2449.2	2440.7
F22	Std	13.6	83.5	32.5	57.3	56.7	29.0	30.7
	Rank	6	7	5	4	3	2	1
	Mean	1.0×10^4	8467.4	7041.4	5352.7	4247.9	4209.8	5087.8
F23	Std	177.8	734.1	1912.7	1895.5	2115.3	2113.3	2009.7
	Rank	7	6	5	4	2	1	3
	Mean	2977.9	3063.9	3456.9	2817.6	2810.7	2802.0	2801.6
F24	Std	13.0	90.7	221.6	60.5	45.5	51.7	26.2
	Rank	5	6	7	4	3	2	1
	Mean	3130.6	3270.3	3360.5	2975.0	2987.6	2986.9	2992.5
F25	Std	14.5	77.5	59.5	63.4	60.4	41.3	30.3
	Rank	5	6	7	1	3	2	4
	Mean	3342.6	3184.2	2928.3	2905.6	2899.1	2904.8	2897.3
F26	Std	60.6	46.2	18.9	20.4	14.9	17.5	14.0
	Rank	7	6	5	4	2	3	1
	Mean	6507.6	8639.1	6774.7	4675.2	5305.7	5067.9	4684.3
F27	Std	254.1	548.0	2309.6	856.7	1086.6	756.5	1332.2
	Rank	5	7	6	1	4	3	2
	Mean	3200.0	3494.4	3861.9	3235.7	3249.1	3253.5	3227.7
F28	Std	5.4×10^{-5}	123.5	395.0	15.1	17.8	20.1	38.1
	Rank	1	6	7	3	4	5	2
	Mean	3300.0	3748.8	3271.5	3233.7	3261.4	3267.4	3254.9
F29	Std	0.03	256.0	21.4	22.0	50.7	26.6	35.6
	Rank	6	7	5	1	3	4	2
	Mean	5184.5	5206.8	4550.6	3972.0	3983.7	3969.8	3912.8
F30	Std	113.7	640.5	264.7	192.2	175.7	196.1	193.8
	Rank	6	7	5	3	4	2	1
	Mean	1.6×10^6	4.5×10^7	2.1×10^6	4.1×10^4	2.4×10^4	6.0×10^4	1.9×10^4
	Std	8.6×10^5	2.9×10^7	9.5×10^5	2.5×10^4	1.7×10^4	4.5×10^4	1.6×10^4
	Rank	5	7	6	3	2	4	1
	Ar	6.103	6.276	4.931	2.966	2.379	3.172	1.724
	Fr	6	7	5	3	2	4	1

IM of each feature is calculated through RF and sorted. The higher the ranking, the higher the importance of the feature, which also means the more correlated it is with the compressive strength of concrete.

According to the ranking results of V IM in Figure 11, it can be seen that age has the greatest influence on the compressive strength of concrete, while the importance of fly ash is the lowest. This paper uses 0.09 as the threshold to

TABLE 4: Time complexity comparison of algorithm.

Phase	Algorithm			
	SSA	CSSA	NCSSA	ISSA
Population initialization	$O(d + f(d))$	$O(d + f(d))$	$O(d + f(d))$	$O(d + f(d))$
Location of producers updated	$O(d)$	$O(d)$	$O(d)$	$O(d)$
Location of scroungers updated	$O(d)$	$O(d)$	$O(d)$	$O(d)$
Location of scout updated	$O(d)$	$O(d)$	$O(d)$	$O(d)$
Update of optimal location	$O(d)$	$O(d)$	$O(d)$	$O(d)$

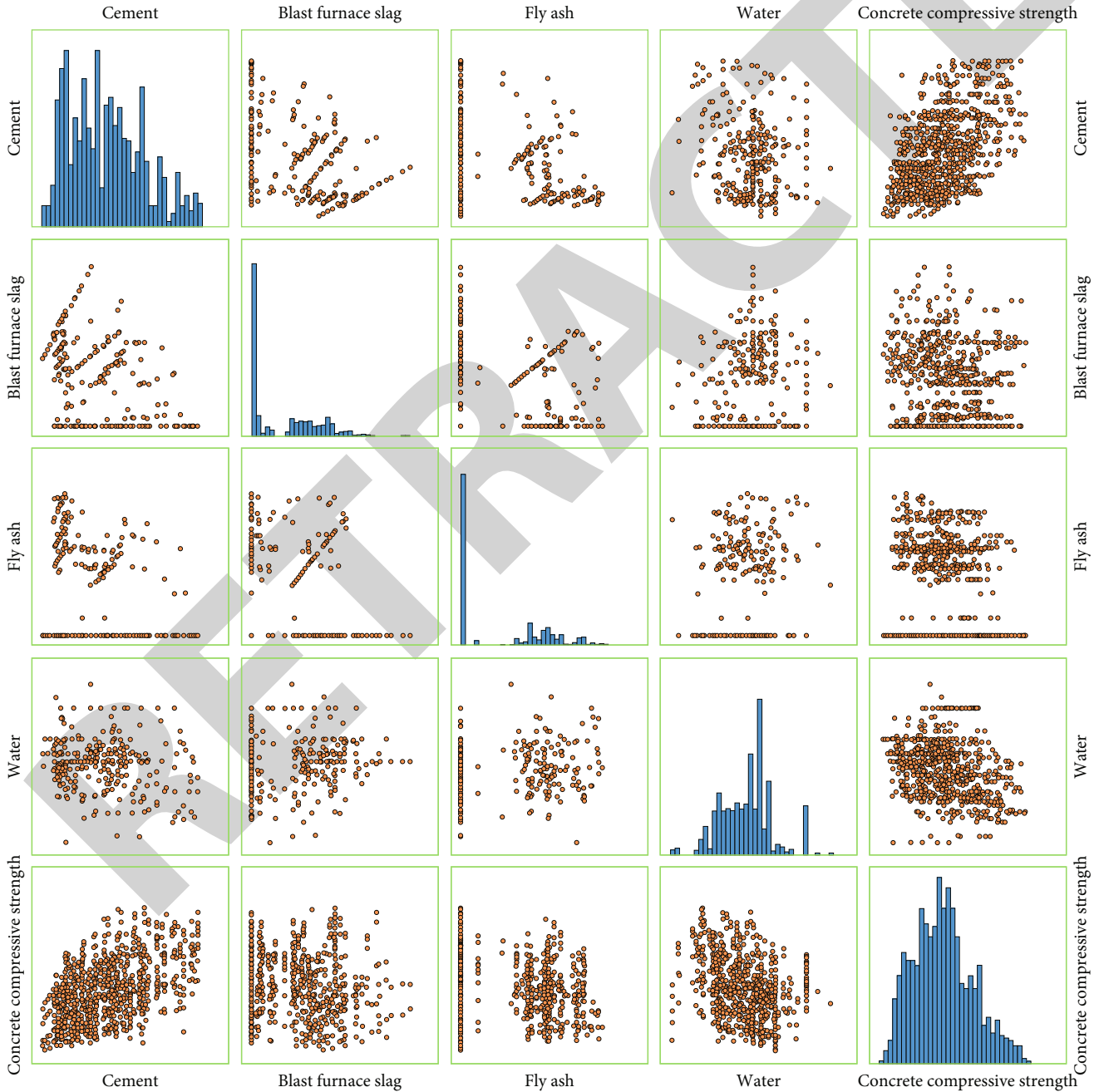


FIGURE 9: Plot of the relationship between output variables and the first 4 input feature variables (cement, blast furnace slag, fly ash, and water).

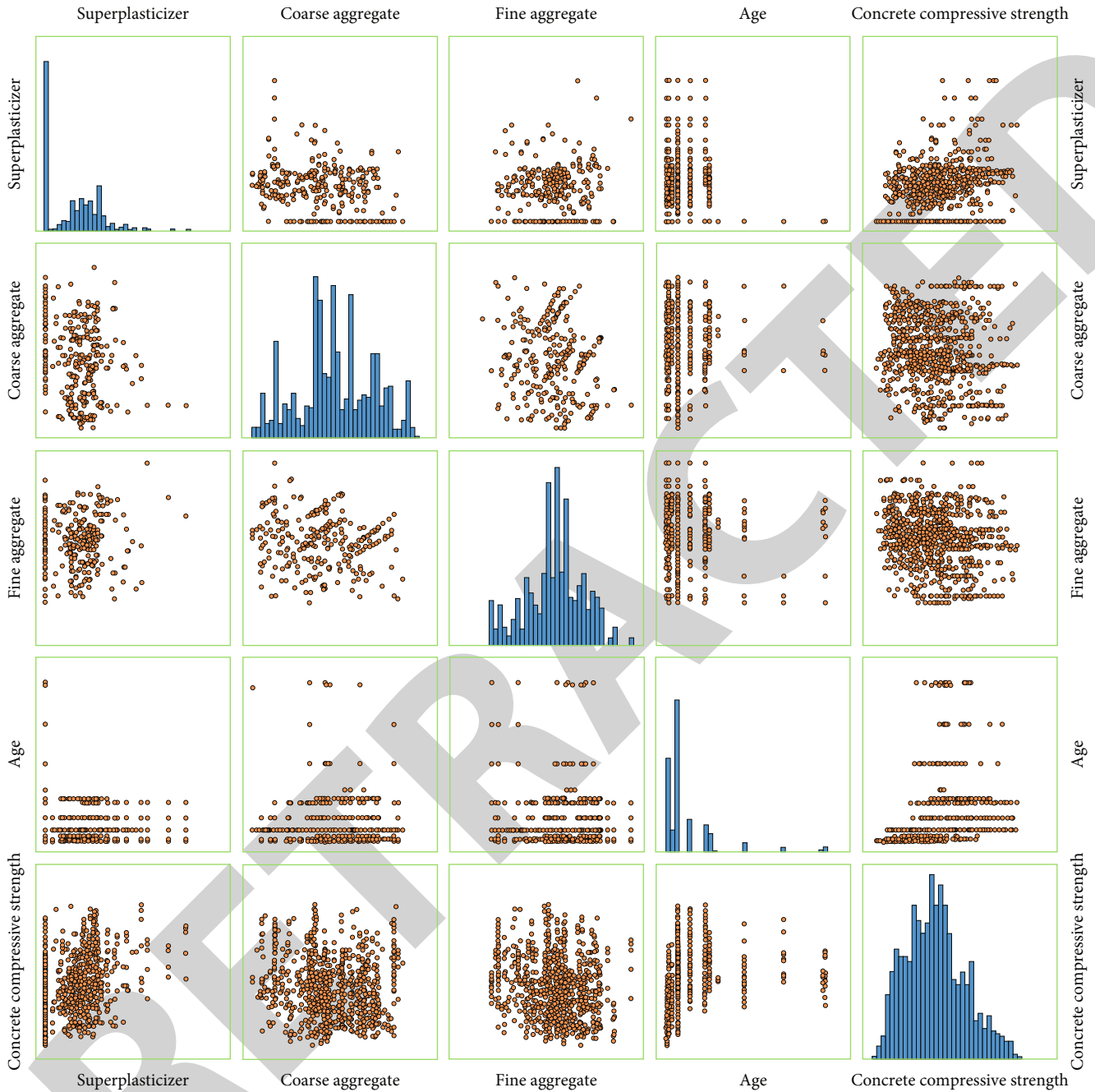


FIGURE 10: Plot of the relationship between output variables and the last 4 input feature variables (superplasticizer, coarse aggregate, fine aggregate, and age).

filter out 6 characteristic factors (age, cement, fine aggregate, coarse aggregate, water, and blast furnace slag) whose V_{IM} is greater than 0.09, as input variables of the prediction model, the two characteristic variables (superplasticizer and fly ash) that have little influence on the compressive strength of concrete are eliminated, so that the model is simplified without losing important features and the calculation efficiency of the model is improved.

4.3.2. *Select an Appropriate Number of Nodes.* The 6 features selected by RF are set as the 6 nodes of the input layer of the BPNN, and the compressive strength of concrete is the only

node of output, but improper setting of node number may lead to poor prediction results in the hidden layer. Therefore, how to select the appropriate number of nodes in the model is a key problem, so as to achieve the best prediction performance. Kolmogorov theorem [43] proposes that the optimal number of points of the hidden layer is generally $2n + 1$, where n is the number of nodes in the input layer. When $n = 6$, the optimal number of points is 13. Figure 12 is the basic structure diagram of BPNN.

4.3.3. *Prediction of Concrete Compressive Strength.* The trained RF-ISSA-BPNN model is used to predict the

TABLE 5: Statistical features of concrete data sets.

Concrete components	Maximum	Minimum	Mean	Std deviation
Cement	540.0	102.0	281.2	104.50
Blast furnace slag	359.4	0.0	73.9	86.27
Fly ash	200.1	0.0	54.2	63.99
Water	247.0	121.8	181.6	21.35
Superplasticizer	32.2	0.0	6.2	5.97
Coarse aggregate	1145.0	801.0	972.9	77.75
Fine aggregate	992.6	594.0	773.6	80.17
Age	365.0	1.0	45.7	63.16
Concrete compressive strength	82.6	2.3	35.8	16.70

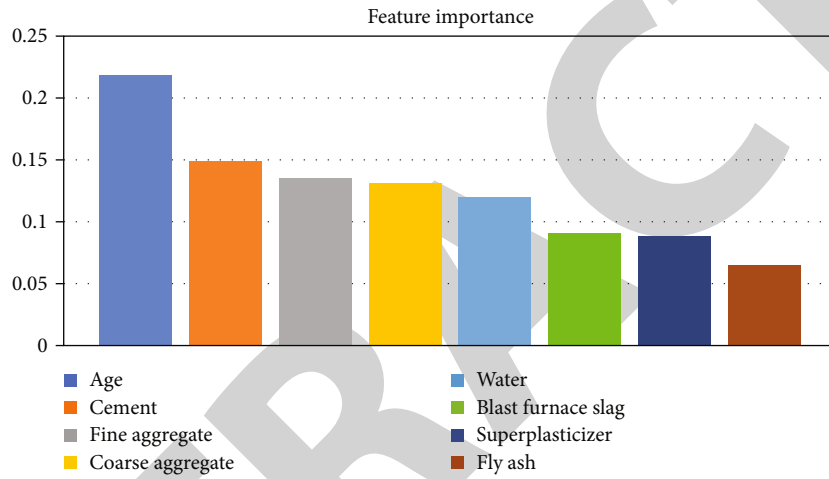


FIGURE 11: Feature importance ranking.

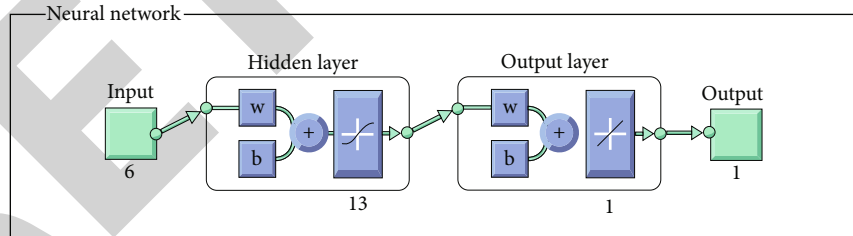


FIGURE 12: The basic structure of BP neural network.

compressive strength of concrete, and the feasibility and superiority of the model are proved. It is longitudinally compared with the prediction model ISSA-BPNN that has not been processed by the RF feature selection process. In addition, the algorithms (PSO, SSA, and CSSA) with excellent performance on the CEC2017 benchmark function are selected to participate in the process of network parameter optimization, and the prediction models of RF-BPNN, RF-PSO-BPNN, RF-SSA-BPNN, and RF-CSSA-BPNN are established for horizontal comparison. In this paper, 1,000 sets of data are used as the training set, and 30 sets of data are used as the test set for simulation and prediction. The

model parameters of BP neural network are set uniformly: The number of training is 100, the accuracy is 0.001, and the learning rate is 0.01. The parameters of the selected optimization algorithm are consistent with those in Table 2, and the simulation experiments are performed in the Matlab2018a compilation environment. The prediction results of concrete compressive strength of each model are shown in Figure 13.

Cross-validation is an important parameter for the evaluation of machine learning algorithms how well prediction capability is with generalization for an independent data set and can be used to assess the model capability to predict

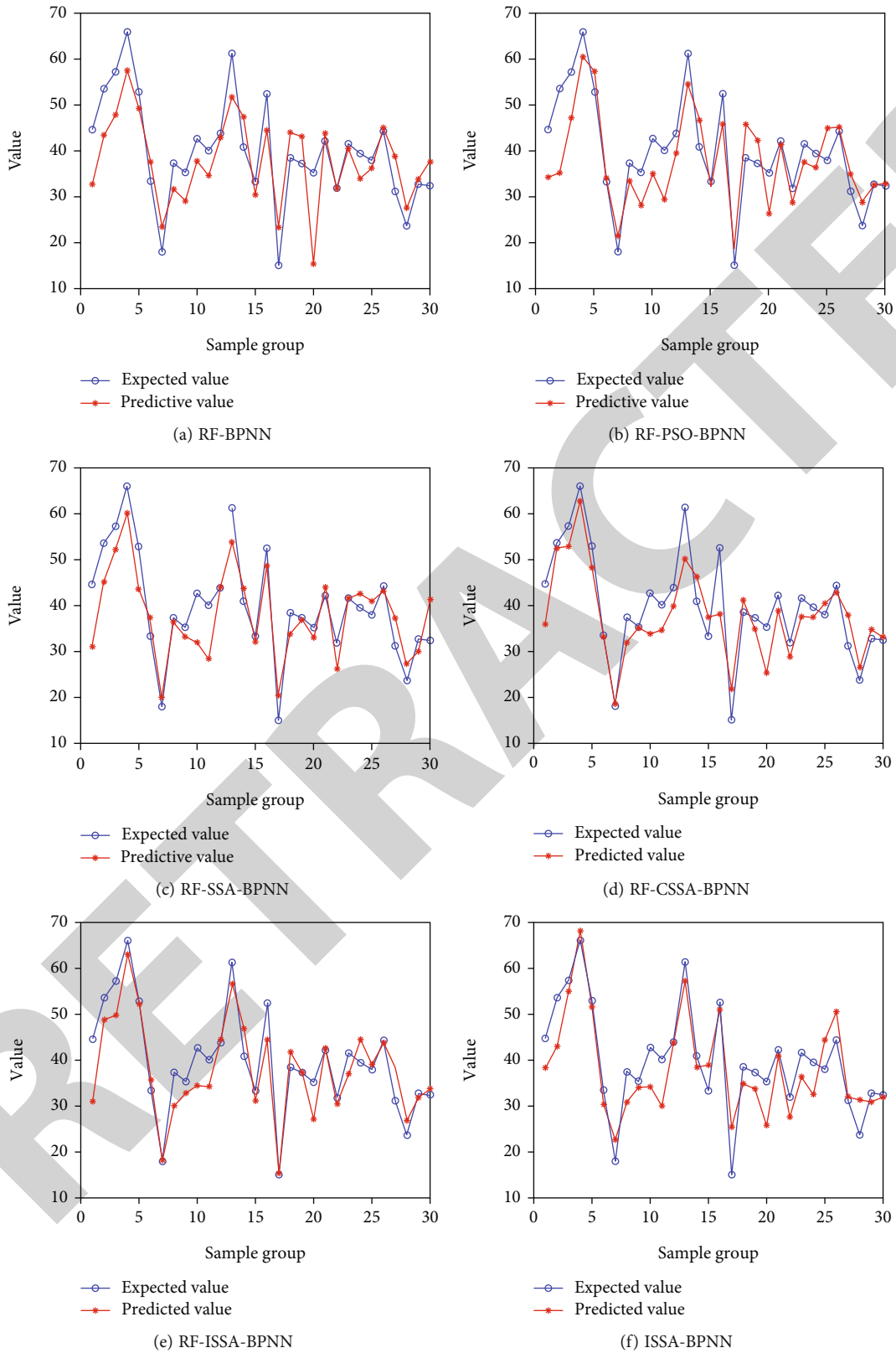


FIGURE 13: The prediction results of the concrete compressive strength of each model.

the new data and to obtain insight about the model capability for prediction of the independent data set [44]. Tenfold cross-validation [45] is a commonly used technique for

cross-validation, and this section discusses the accuracy of the RF-ISSA-BPNN model from the perspective of tenfold cross-validation. The error evaluation indexes can fully

TABLE 6: Predicted results of 6 models.

Index	Model					
	RF-BPNN	RF-PSO-BPNN	RF-SSA-BPNN	RF-CSSA-BPNN	RF-ISSA-BPNN	ISSA-BPNN
MAE	5.6620	4.9291	5.4083	5.1226	4.3798	4.9887
RMSE	7.5048	6.7247	6.8324	6.4402	5.6687	6.5417

reflect the real gap between predicted results, and the actual results, two commonly used error evaluation indexes (MAE and RMSE) are selected as a measure of each model performance, as shown in Equations (10) and (11). The closer the final result of the three is to 0, the better the performance of the model, as shown in Table 6.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N \left| \tilde{y}_i - y_i \right|, \quad (10)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\tilde{y}_i - y_i)^2}, \quad (11)$$

where N is the predicted array length, y_i is the actual concrete compressive strength, and \tilde{y}_i is the predicted concrete compressive strength.

5. Results and Discussion

As can be seen from Figure 13, whether horizontal or vertical comparison, the predicted data of RF-ISSA-BPNN model has the best fitting degree with the actual data, and the prediction performance is better than other models. The validity of the prediction model is verified by the tenfold cross-validation method, as can be seen from the data of the predictive indicators MAE and RMSE in Table 6:

- (1) The RF-ISSA-BPNN model has the lowest values of MAE and RMSE, which means that the error between the actual and predicted values is the smallest, so the RF-ISSA-BPNN model has the highest accuracy in predicting the compressive strength of concrete
- (2) The prediction performance of BPNN model is significantly enhanced after the optimized weights and thresholds, and the prediction results (MAE and RMSE) of RF-PSO-BPNN, RF-SSA-BPNN, RF-CSSA-BPNN, and RF-ISSABPNN models are better than RF-BPNN models
- (3) The significant difference between MAE and RMSE values of RF-ISSA-BPNN and ISSA-BPNN indicates the importance of RF feature selection. MAE is reduced by 0.6089, and RMSE is reduced by 0.873
- (4) The performance of the improved prediction model for SSA has also been further improved. MAE and RMSE of RF-CSSA-BPNN and RF-ISSA-BPNN models are all smaller than those of RF-SSA-BPNN model. It is fully verified that the improvement of the algorithm can also seek a breakthrough in the

performance of the prediction model, which also indirectly confirmed that the three improvement strategies proposed provide key help to the improvement of the original SSA

6. Conclusion

In this study, the standard sparrow search algorithm is improved by integrating LHS, somersault foraging strategy, and cyclone foraging mechanism to solve the problems of premature convergence and insufficient global search ability. In the prediction application of concrete compressive strength, the RF-ISSA-BPNN prediction model is established by finding the optimal threshold and weight by ISSA and using the feature selection function of RF. The main conclusions are as follows:

- (1) The ablation experiment on the CEC2017 benchmark function verifies that the three strategies introduced separately and simultaneously can significantly improve the optimization effect of SSA, which fully demonstrates the effectiveness and feasibility of the three strategies introduced
- (2) Compared with the classic algorithms (PSO, ABC, WOA, and SSA), ISSA has better optimization ability. In the case of the same time complexity of algorithm, the performance of ISSA is more outstanding than other improved algorithms (CSSA and NCSSA).
- (3) The importance of 8 features is sorted by RF, and 6 features with $V_{IM} > 0.09$ are selected as input variables, so that the complexity of the model is reduced. It can be seen that there is a great correlation between age and compressive strength of concrete, which indirectly indicates that the effect of age on the compressive strength of concrete should be emphasized in practical application. The results of the longitudinal comparison between RF-ISSA-BPNN and ISSA-BPNN show that the feature selection of RF can make the model have more efficient running efficiency and higher prediction accuracy
- (4) The RF-ISSA-BPNN prediction model is horizontally compared with other models (RF-BPNN, RF-PSO-BPNN, RF-SSA-BPNN, and RF-CSSA-BPNN), the prediction results show that the MAE and RMSE values of RF-ISSA-BPNN are lower than those of other models, and the predicted data has the highest fitting degree with the actual data. It not only improves the shortcomings that BPNN is low prediction accuracy and easy to converge to different local

minima, but also enhances its accuracy and stability in prediction

Although RF-ISSA-BPNN can greatly reduce the risk of falling into local optimal solutions, the situation of falling into local minima still occurs in the later stage of the search. How to fundamentally eliminate the occurrence of this phenomenon and make the prediction model more characteristic and practical value will be the focus of our future work. In addition, due to the small number of samples, training the neural network model will be affected to some extent. In order to build a more accurate prediction model, a complete and higher quality data is needed. In the next research, the main work is to extend the database and introduce more factors related to the compressive strength of concrete to further improve the generalization ability of the model.

Appendix

Data source: <http://archive.ics.uci.edu/ml/datasets/Concrete+Compressive+Strength>

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 have no conflicts of interest to declare that are relevant to the content of this article.

Acknowledgments

This research was funded by the National Youth Science Foundation of China No. 62002046 and No. 61802040, the project supported by the Zhejiang Provincial Natural Science Foundation of China (No. LQ21F020005), and the Basic Public Welfare Research Program of Zhejiang Province (No. LGG18E050011).

References

- [1] W. Liang, G. Wang, X. Ning et al., "Application of BP neural network to the prediction of coal ash melting characteristic temperature," *Fuel*, vol. 260, p. 116324, 2020.
- [2] B. Xu, H. C. Dan, and L. Li, "Temperature prediction model of asphalt pavement in cold regions based on an improved BP neural network," *Applied Thermal Engineering*, vol. 120, pp. 568–580, 2017.
- [3] Y. Liu, X. Wang, X. Zhu, and Y. Zhai, "Thermal error prediction of motorized spindle for five-axis machining center based on analytical modeling and BP neural network," *Journal of Mechanical Science and Technology*, vol. 35, no. 1, pp. 281–292, 2021.
- [4] S. Dai, B. Han, S. Liu, N. Li, F. Geng, and X. Hou, "Neural network-based prediction methods for height of water-flowing fractured zone caused by underground coal mining," *Arabian Journal of Geosciences*, vol. 13, no. 12, pp. 1–11, 2020.
- [5] Y. Li, J. Li, J. Huang, and H. Zhou, "Fitting analysis and research of measured data of SAW micro-pressure sensor based on BP neural network," *Measurement*, vol. 155, p. 107533, 2020.
- [6] J. Gao, "Performance evaluation of manufacturing collaborative logistics based on BP neural network and rough set," *Neural Computing and Applications*, vol. 33, no. 2, pp. 739–754, 2021.
- [7] Q. Liu, P. Sun, X. Fu et al., "Comparative analysis of BP neural network and RBF neural network in seismic performance evaluation of pier columns," *Mechanical Systems and Signal Processing*, vol. 141, p. 106707, 2020.
- [8] K. Dou and X. Sun, "Long-term weather prediction based on GA-BP neural network," *IOP conference series: earth and environmental science*, vol. 668, no. 1, article 012015, 2021.
- [9] H. Wang, Z. Zhang, and L. Liu, "Prediction and fitting of weld morphology of Al alloy-CFRP welding-rivet hybrid bonding joint based on GA-BP neural network," *Journal of Manufacturing Processes*, vol. 63, pp. 109–120, 2021.
- [10] P. Supraja, V. M. Gayathri, and R. Pitchai, "Optimized neural network for spectrum prediction using genetic algorithm in cognitive radio networks," *Cluster Computing*, vol. 22, no. S1, pp. 157–163, 2019.
- [11] H. Yuan, X. Wang, X. Sun, and Z. Ju, "Compressive sensing-based feature extraction for bearing fault diagnosis using a heuristic neural network," *Measurement Science and Technology*, vol. 28, no. 6, article 065018, 2017.
- [12] G. Jiang, M. Luo, K. Bai, and S. Chen, "A precise positioning method for a puncture robot based on a PSO-optimized BP neural network algorithm," *Applied Sciences*, vol. 7, no. 10, p. 969, 2017.
- [13] W. Wang, R. Tang, C. Li, P. Liu, and L. Luo, "A BP neural network model optimized by mind evolutionary algorithm for predicting the ocean wave heights," *Ocean Engineering*, vol. 162, pp. 98–107, 2018.
- [14] Y. Wu, R. Gao, and J. Yang, "Prediction of coal and gas outburst: a method based on the BP neural network optimized by GASA," *Process Safety and Environmental Protection*, vol. 133, pp. 64–72, 2020.
- [15] Z. Wensheng and J. Z. Haozqi, "BP neural network model for short-time traffic flow forecasting based on transformed grey wolf optimizer algorithm," *Journal of Transportation Systems Engineering and Information Technology*, vol. 20, no. 2, pp. 196–203, 2020.
- [16] H. Tian, P. Wang, K. Tansey, S. Zhang, J. Zhang, and H. Li, "An IPSO-BP neural network for estimating wheat yield using two remotely sensed variables in the Guanzhong Plain, PR China," *Computers and Electronics in Agriculture*, vol. 169, p. 105180, 2020.
- [17] L. Wu, Y. Yang, and M. Maheshwari, "Strain prediction for critical positions of FPSO under different loading of stored oil using GAIFOA-BP neural network," *Marine Structures*, vol. 72, p. 102762, 2020.
- [18] D. Lianjie, C. Degang, W. Ningling, and L. Zhanhui, "Key energy-consumption feature selection of thermal power systems based on robust attribute reduction with rough sets," *Information Sciences*, vol. 532, pp. 61–71, 2020.
- [19] L. Wen and X. Yuan, "Forecasting CO₂ emissions in Chinas commercial department, through BP neural network based on random forest and PSO," *Science of the Total Environment*, vol. 718, p. 137194, 2020.
- [20] J. K. Jaiswal and R. Samikannu, "Application of random forest algorithm on feature subset selection and classification and

Retraction

Retracted: A New Machine Learning Algorithm for Regional Low-Carbon Economic Development Analysis Based on Data Mining

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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, "A New Machine Learning Algorithm for Regional Low-Carbon Economic Development Analysis Based on Data Mining," *Journal of Function Spaces*, vol. 2022, Article ID 5692666, 8 pages, 2022.

Research Article

A New Machine Learning Algorithm for Regional Low-Carbon Economic Development Analysis Based on Data Mining

Xinlei Liu 

School of Economics and Finance of Xi'an Jiaotong University, Xi'an 710061, China

Correspondence should be addressed to Xinlei Liu; beifenglxl@stu.xjtu.edu.cn

Received 28 May 2022; Revised 29 July 2022; Accepted 9 August 2022; Published 25 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Xinlei 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 development of information technology such as the continuous improvement of mobile Internet infrastructure and the performance of computers has made it easy to process and share information. The huge market demand for location-based information services provides huge impetus to the generation and development of mobile terminal positioning technology. Generally speaking, the main causes of climate change can be summarized into two categories: natural climate fluctuations and the impact of human activities which is a major measure taken by China to actively respond to climate change. This is a successful approach to actively explore the rapid development of China's industrialization and urbanization, which not only develops the economy and improves people's livelihood but also responds to climate change and reduces carbon intensity. Firstly, this paper mainly is aimed at the connotation of regional low-carbon economic development mode, studying the basic mode of regional low-carbon economic development, and analyzing the characteristics and applicable conditions of each mode. Secondly, based on the machine learning algorithm of data mining, the main mode selection of regional low-carbon economic development is discussed. Thirdly and finally, when choosing the regional low-carbon economic development mode, comprehensive consideration should be given to the economic development basis, energy structure, resource characteristics, industrial status, development mode, geographic location, and other factors. This paper studies the basic conditions and applicable conditions of regional economic development models. The conclusion shows that from the perspective of regional economic evolution, low-carbon economy can be regarded as the decarbonization process of economic development. It is an economic form combining its own characteristics and an inevitable requirement for the transformation of regional economy from other economic models to low-carbon economic models. And other factors of Selection of regional economic development foundation, energy structure, resource characteristics, industrial status, development mode, geographical location, were also discussed.

1. Introduction

In recent years, with the continuous decline of computer terminal price and the continuous improvement of its performance, smart phones appears and popularizes. The development of information technology, such as the continuous improvement of the Internet and mobile Internet infrastructure, makes it easy to spread and share information. In today's society where data flow and information flow explode, deep data mining has gradually developed into a new trend [1]. Nowadays, mobile terminal positioning technology is widely applied in many fields such as emergency rescue, military positioning, peer network operation, and optimization. The huge market demand based

on location information service makes the emergence and development of mobile terminal positioning technology have great power. The application of data mining technology in daily life can promote the development of social science and technology, improve people's information processing ability, and play an important role in the information age. Under this background, the concept of low-carbon economy and low-carbon development came into being and quickly gained wide attention from governments and academic circles [2]. Different decoupling indicators can be distinguished based on the actual situation, and primary decoupling and secondary decoupling are defined. Primary decoupling refers to the decoupling between natural resource utilization and economic growth. Secondary

decoupling refers to the decoupling of environmental pollution and natural resources from the traditional “high carbon consumption, high pollution and high waste” economy which has caused a serious resource and environmental crisis. Ecological destruction, environmental pollution, and resource depletion are the most widespread impacts of the global ecological crisis, which are directly manifested in the increasingly obvious greenhouse effect and a significant global warming trend. The technique for analyzing and processing large-scale data with the participation of computers and extracting the meaningful laws or patterns implicit in the data is data mining [3].

Human beings are facing increasingly severe resource and environmental situation. In the decades since the mid-20th century, with the outbreak of the new technological revolution, the global social economy has developed at an unprecedented speed. At the same time, the popularization of modern cellular mobile communication network and the establishment of global positioning system have made great progress in mobile terminal positioning technology, and the positioning accuracy has been greatly improved. Using data mining technology can not only diminish the data management costs significantly but also improve the efficiency of communication between them. Although low-carbon economy is put forward in the practice of human response to climate change, its essence is the result of human continuous reflection on the relationship between their own development and resources and environment. It is emphasized that economic development should be coordinated with resources and environment, and ultimately promote the transition from a carbon-based energy economy to a low-carbon and carbon-free energy economy. The key factor of climate change is the increase in atmospheric temperature caused by the emission of carbon dioxide and other greenhouse gases under the influence of human factors, which has been basically reached by the global scientific community. This is an important step taken by China to actively respond to climate change. It will actively explore the successful practices of developing economy, improving people’s livelihood, coping with climate change, and reducing carbon intensity in the rapid development stage of China’s industrialization and urbanization. It is of great significance to accumulate beneficial experience in promoting low-carbon green development in different regions. By calculation, the importance of multiple factors can also be compared [4]. Among the influencing factors of China’s carbon emissions, energy consumption ranks first. Economic growth ranked second and trade ranked third. The predecessor of machine learning method is the theory of statistical learning and optimization, which was born with the advent of computers. So far, a large number of algorithms have been proposed for different disciplines and problems, emphasizes that developed countries should provide adequate and sustained financial, technical, and capacity-building assistance for the implementation of adaptive actions in developing countries.

China’s decoupling of energy and carbon emissions will be in a state of growth linkage for a long time. China has made great efforts, but it still has a long way to go [5]. In China, as the first priority, development must vigorously develop the economy to catch up with developed countries and improve people’s material living conditions. This is obviously based on a large amount of energy consumption [6]. The transfor-

mation of enterprises and the progress of the industry all need the improvement of machine learning algorithms, so as to conveniently process huge information data and realize the development of society [7]. But in fact, data mining is more and more difficult to analyze because of its limited data processing ability, and machine learning algorithm is an interdisciplinary subject, that is, computer simulates human behavior and automatically acquires new skills and knowledge. At present, there are many studies on low-carbon economy and its regional evaluation, so it is necessary to sort them out systematically. Further clarify the connotation, goals, and paths of low-carbon development, and provide clear ideas for future research and practice [8]. Climate warming makes the production and consumption patterns in traditional economics no longer able to meet the requirements of the era of sustainable human development in the future. In the future, the economic development mode must be changed and a new economic model must appear. The low-carbon province and low-carbon city pilot program means that the country should give full play to the enthusiasm, initiative, and creativity of each region to explore a regional low-carbon development model with Chinese characteristics. Therefore, accelerating the pace of transformation and construction of the economic belt or elevating the construction of the economic belt to a national grand strategy will help to effectively connect with the “Belt and Road,” accelerate China’s economic and social development, narrow the gap between the north and the south, and achieve regional coordination in the eastern, central, and western region development [9].

2. Related Work

As the foundation of low-carbon economy, low-carbon industry has not been strictly distinguished by scholars in domestic research. Lin et al. put forward ecological footprint theory, which reveals the regional space of biological productivity needed to produce a certain amount of substances and absorb wastes generated by consuming these substances. It provides a method to judge the ecological carrying capacity of a certain country or region, so as to judge the greenhouse gas emissions generated by human activities, and demonstrate the necessity of developing low-carbon industries with the deterioration of ecological environment [10]. Peng et al. put forward carbon emissions in low-carbon economy will be a new indicator to measure economic development [11]. Zhang et al. put forward reviewing and depicting the scenario of long-term climate stability and decomposing carbon emissions into three factors: carbon dioxide concentration, energy efficiency, and economic activities [12]. Li et al. believe that low-carbon economy is based on technological innovation and institutional innovation. Only by reducing greenhouse gas emissions as much as possible can we achieve coordinated economic, social, and environmental development [13]. Yu et al. established the following: from the perspective of energy, low-carbon economy is regarded as an energy revolution with the goal of slowing down climate change, and its core is the technological innovation and institutional innovation of energy [14]. Fleming et al. established the following: using Kuznets curve to study the trajectory of human development from high-carbon economy to low-carbon economy, came to the conclusion that human

development will pollute the environment at the same time [15]. Kuong et al. put forward the theory of urban mines. This theory believes that human beings can rely on technological innovation and other means to support the use of renewable resources [16]. Kapitonov believes that low-carbon economy should be regarded as the most important economic, social, and environmental revolution of this century. And that its significance will be more far-reaching than the industrial revolution [17]. Yang et al. constructed a method to describe the low-carbon economic development scenario at the urban scale and applied this method to the Shiga region of Japan [18]. Field and Derwent established the humanistic development goal of low-carbon economy and defined it as an economic form in which both carbon productivity and humanistic development have reached a certain level [19]. Wang et al. also emphasized its institutional attribute and believed that low-carbon industry is a new concept driven by institutional innovation and development ideas [20].

3. Basic Model for Regional Development of Low-Carbon Economy

3.1. Emerging Low-Carbon Industrial Cluster Model. This mode is to build new low-carbon industrial clusters, cultivate industrial clusters with low-carbon technology industries as the main body, reduce the production cost of low-carbon industries, and accelerate the knowledge spillover effect and technological innovation among enterprises. The method is simple and clear, and the conclusions obtained can make us intuitively understand the relationship between economic development and resources and environment and have a certain early warning function. Because the independent assumption of naive Bayesian algorithm is not true in most practical situations, its performance is worse than many more advanced machine learning algorithms. The emission of carbon-containing greenhouse gases mainly comes from the burning of fossil fuels, and the increase of carbon-containing greenhouse gases in the atmosphere directly affects the global climate change, especially the global warming, which leads to the global ecological crisis. The annual economic losses caused by global climate anomalies have gradually made people realize the importance of sustainable development. In response to global climate change, nonrenewable energy depletion, and other issues, countries all over the world regard developing new energy and supporting the development of emerging industries such as energy conservation and environmental protection as national development strategies to adjust economic structure and promote economic growth. On the premise of building an innovative country, through introduction, digestion, absorption, and independent innovation, form a group of low-carbon advanced technologies with independent intellectual property rights, and focus on the development, demonstration, and promotion of a group of common and key link technologies that have a major driving effect on the low-carbon economy. Therefore, under normal circumstances, the error of the training sample set with the closest distance to the sample to be classified obtained by using the nearest neighbor method is very large. For example, based on the location information, information such as the distribution of the main residential area of the user, the respective char-

acteristics of the users in different shopping malls, the user's preferred area, and the development of the area can be obtained. Although we are only in the early stages of predicting global warming trends, the ecological responses to near-term climate change are already clearly visible, as shown in Figure 1.

This kind of information has certain reference value for investors, real estate developers, market operators, and the government. At present, we should vigorously promote the regional low-carbon transformation to achieve green and clean development. Only by minimizing the pressure of economic activities on resources and environment can we finally achieve ecological civilization. The merge time interval is 2 s to 10 s for simulation, as shown in Table 1.

Effective simulation of its positioning performance shows the specific simulation results, as shown in Table 2.

When the K value of the positioning method based on K -nearest neighbor method is different, its performance is effectively simulated, which represents the specific simulation results, as shown in Table 3.

Compared with fossil energy, renewable energy has become cost competitive in many countries. Government leadership continues to play a key role in driving the growth of renewable energy, especially wind and solar power. 173 countries around the world have set renewable energy development goals, and 146 countries have issued support policies. Many cities, communities, and enterprises have taken the lead in launching the rapidly expanding "100% renewable energy" initiative, which plays a crucial role in promoting global energy transformation. Since then, more and more scholars began to pay attention to climate change, and its related research and international conferences have become more and more frequent. Global renewable energy structure is shown in Figure 2.

Then, the distance between the dividing lines at the extreme of these two lines can be expressed as follows:

$$d = \left\| \frac{2}{\omega} \right\|. \quad (1)$$

By solving the optimization problem:

$$\begin{aligned} \min_{\omega, b} &= \frac{1}{2} \|\omega\|^2, \\ \text{s.t.} & y_i((\omega \cdot x_i) + b) \geq 1, \quad i = 1, \dots, l. \end{aligned} \quad (2)$$

The dual problem of the optimization problem in the maximum margin method can be found as follows:

$$\begin{aligned} \min_{\alpha} & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j (x_i \cdot x_j) - \sum_{j=1}^l \alpha_j, \\ \text{s.t.} & \sum_{i=1}^l y_i \alpha_i = 0, \\ & \alpha_i \geq 0, \quad i = 1, \dots, l. \end{aligned} \quad (3)$$

The following formula can be used as a degree of misclassification of the training set as a whole:

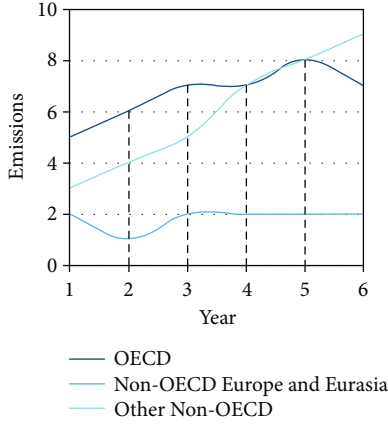


FIGURE 1: 002 emissions around the world.

TABLE 1: Simulation results of data consolidation for locations.

Merge time interval	None	1	2	3	4	5
Positioning accuracy	82.35	82.94	83.6	83.99	84.20	84.26
Positioning time	39.30s	19.65 s	9.80s	6.7 s	4.90s	3.92 s
Merge time interval	5	6	7	8	9	
Positioning accuracy	84.30	84.30	84.15	83.87	83.23	
Positioning time	3.26	2.80	2.45	2.16	2	

TABLE 2: Simulation results of tertiary positioning area size based on K -nearest neighbor method.

Positioning area side length	100 m	200 m	300 m	400 m
Positioning accuracy	80.90	83.35	84.33	84.99
Positioning time	0.90	1.83	3.29	4.91

TABLE 3: Simulation results of tertiary positioning K value based on K -nearest neighbor method.

K value setting	10	20	30	40
Positioning accuracy	83.20	83.88	84.30	84.11
Positioning time	3.25	2.28	3.27	3.30

$$\sum_{i=1}^l \xi_i. \quad (4)$$

The extended maximum interval method to solve the dual problem of the optimization problem is as follows:

$$\text{s.t. } \sum_{i=1}^l y_i \alpha_i = 0. \quad (5)$$

3.2. Traditional Low-Carbon Industry Maintenance Model. Most of the original industries are general traditional industries, such as agriculture, handicrafts, and tourism, and the carbon emission is not very high. The maintenance method is adopted for these industries. Among them, the policy changes

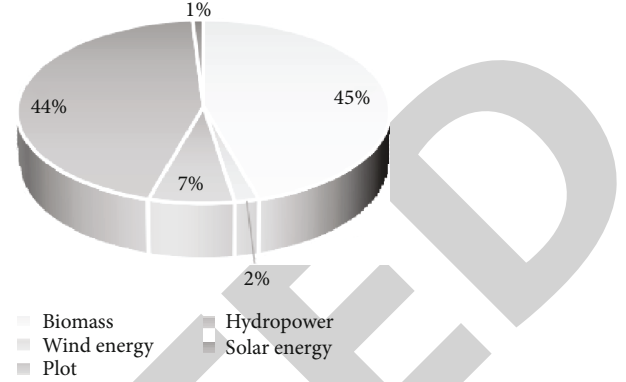


FIGURE 2: Global renewable energy structure.

cannot be quantified, the data in the process of quantifying the indicators of resource-based industries and high-energy consumption infrastructure are not complete, and there is no cointegration relationship in the modeling process of household consumption structure. For some linear nonseparable problems, although linear partition cannot completely separate the two kinds of points in space, it may cause fewer wrong points. This classification problem becomes an approximately linear separable support vector machine. Zhuang, an early scholar engaged in low-carbon economy research in China, believes that low-carbon economy refers to relying on technological innovation and policy measures. Implement the energy revolution and establish an economic development model with less greenhouse gas emissions. In recent years, the total amount and growth rate of energy production and consumption in remote areas are relatively large, as shown in Figure 3.

It is mainly manifested in industry, high-efficiency production, and energy utilization; energy structure and renewable energy production will occupy a relatively high proportion. China has adopted alternative technologies for fossil energy, mainly including alternative technologies for clean energy, renewable energy, and new energy. Improve energy efficiency and reduce CO₂ emissions by reducing energy consumption. Adopt cleaner production and other technologies to improve energy efficiency, especially the clean utilization technology of coal. On the premise of ensuring crop yield, reducing fertilizer consumption plays an important role in reducing CO₂ emissions in the process of fertilizer generation and protecting the environment. Through the adjustment of the internal and external structure of fossil energy, the growth rate of carbon emissions can be effectively slowed down, and China's energy structure will be developed towards low-carbon energy. In other words, western countries have launched a number of mutually supporting algorithm linkage mechanisms and data classification creation mechanisms in the research field of machine learning algorithms, so as to improve the feasibility of data processing task with the help of multiangle model for practical training. With the progress of science and technology, the change of economic structure, and the development of energy consuming industry, the elasticity coefficient of energy consumption will generally show a downward trend, as shown in Figure 4.

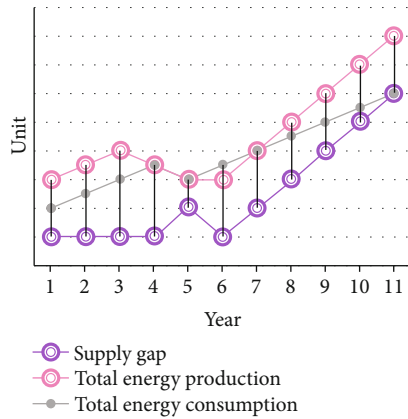


FIGURE 3: Total amount of energy production and consumption.

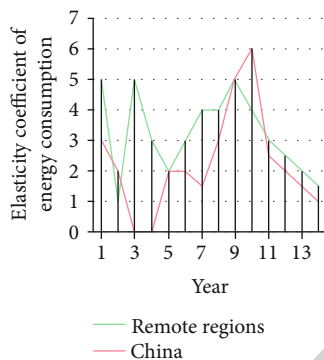


FIGURE 4: Variation trend of elasticity coefficient of energy consumption in remote areas and China.

At this time, it can be known that the partition function with a larger distance between the sample point and the function value will have a far corresponding curve, and it is very likely that it belongs to the relatively positive class of the partition curve with a larger function value. In fact, the nearest neighbor method is a kind of training which is achieved by calculating the distance between two groups of sample point feature vectors in the training set and sample feature vectors to be distinguished. As far as developed countries are concerned, they have achieved a high level of human development and should and must achieve absolute reduction of carbon emissions while maintaining development. Based on the fact of climate change and the requirements of a low-carbon economic model, many scholars have begun to incorporate regional carbon emissions into the indicator system for measuring regional economic development. Its characteristics are that the effect of low-carbon economy can be seen quickly, without too high cost input, and it is relatively easy to implement.

4. Analysis on the Problems Existing in China's Policies to Support the Development of Low-Carbon Industries

4.1. Lack of Systematic Policy System to Support the Development of Low-Carbon Industry. There is lack of coordination between low-carbon policies. The promulgation of rel-

evant policies is due to different industries, different policy subjects, or to some social problems, which leads to the lack of coordination among various policies, and some policies become mere formality and lack of operability. At the national level, China's national will to deal with climate change is constantly increasing, and it has repeatedly issued clear instructions to actively deal with climate change and develop a low-carbon economy. In addition, follow a principle, the principle of minimization. The energy consumption of public buildings is about standard coal, as shown in Figure 5.

In this way, accurate analysis of data mining can be made. The least squares of the outstanding performance of the vector machine of the machine learning algorithm has an irreplaceable role compared with other algorithms. Therefore, the formula is used to calculate the longitude and latitude of the center of the merged area, and then, the data to be positioned for the secondary timing based on the support vector machine is continuously merged. The calculation of sample vector machine and the calculation of decision function vector machine are important links in the later budget. For the design of low-carbon economy comprehensive evaluation index system, based on the existing research, we can further learn from the idea and method of system dynamics. Integrate and optimize subjective and objective information to the greatest extent. Berkeley formulates comprehensive strategies from multiple perspectives. The main strategies adopted in operation and management include smart growth, improving the existing bus frequency and reliability, developing supplementary transportation system, and pricing strategy. The comparative analysis with other provinces and cities is shown in Figure 6. Relationship between building carbon emission and overall carbon emission is shown in Figure 7.

If the area is not suitable for large-scale industry, it can trade emission rights. Get funds and technology from economically developed coastal areas for ecological protection, and develop their own ecological agriculture and tourism, so that both sides can benefit. In recent years, China has implemented a series of policies such as the development of the western region and the transfer of eastern industries to the western region. There is a strong dependence on central investment and infrastructure investment, and investment in resource-based industries and real estate has become the main driving force for economic growth. Investment in high-end machinery manufacturing industry and new technology industry is low, and development is weak. Attribute features are the information extracted from the user's geographic location, the number of microblogs forwarded by the publisher, and other personal information, through the calculation and transformation of this user information. Usually, some quantitative features can be obtained that can reflect the attention relationship between users from a certain aspect. For the low-carbon economic behavior of the same market entity, different policy supports can be found in different policy fields. However, due to the lack of coordination or consistency among different policies, it is difficult to accurately assess the comprehensive value and actual utility of a low-carbon policy during the specific implementation process.

4.2. Inadequate System and Insufficient Policy Support. The development of low-carbon industries is not only conducive to energy conservation and emission reduction, reduction of

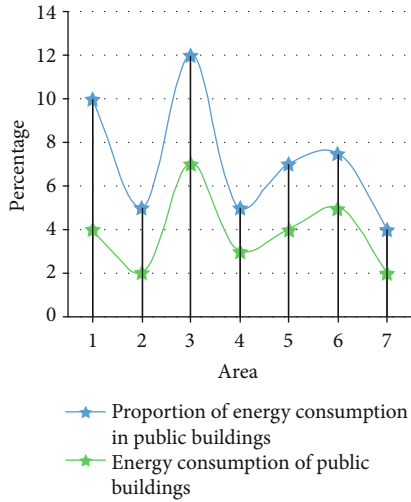


FIGURE 5: Public building energy consumption and its proportion.

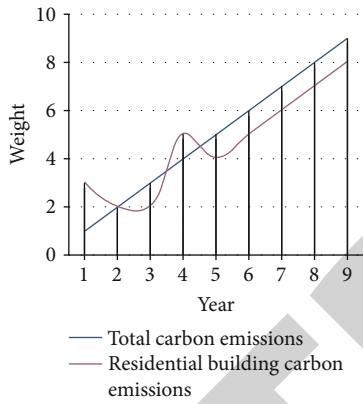


FIGURE 6: Comparison of carbon emissions between residential buildings and residential buildings.

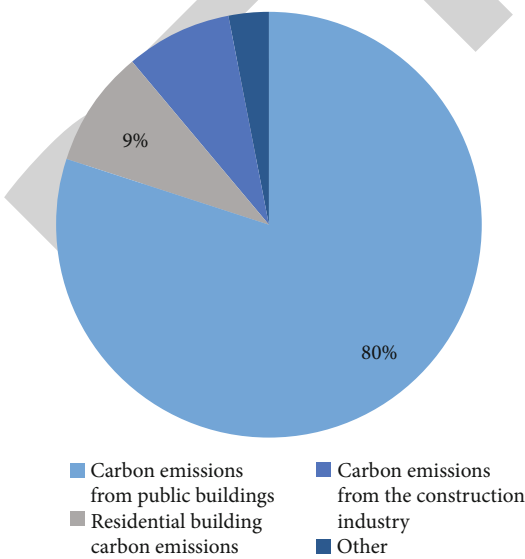


FIGURE 7: Relationship between carbon emission reduction and total carbon emission.

environmental pollution and response to climate change, and the realization of ecological civilization construction, but also conducive to promoting the upgrading and adjustment of China’s energy structure, industrial structure, and market structure, so as to achieve sustainable development. As one of the key areas in the country for energy conservation and emission reduction, remote areas must ensure that the national energy emission reduction targets are completed on time. Now the primary task of the whole region is energy conservation and emission reduction, vigorously eliminating some backward production capacity in energy and other industries and resolutely curbing high energy consumption. Blind development of the industry. Especially in the world’s major developed economies, it has been recognized for a long time that the traditional industrialization path with high energy and high emissions has brought enormous pressure to their own resources and environment, and it is an inevitable trend to change the way of economic development. In this application process, the weights need to be revised. The main method used is to use the error generated by the data calculation to initialize and output the expected value. Due to the increasing interval of merging the positioning data, the positioning accuracy is continuously improved until the highest value appears and gradually decreases. In addition, an appropriate operator should also be selected. Because the genetic BP neural network needs to be improved and optimized, it is required that the selected operator must follow the corresponding proportional algorithm to achieve local optimization. For developed regions, the key is to realize low-carbon economy, optimize industrial structure, and change the mode of economic development. It mainly depends on increasing the consumption of material resources to mainly rely on scientific and technological progress, improving the quality of workers and management innovation. To some extent, the economic structure determines the energy structure and, to a certain extent, the intensity of greenhouse gas emissions. The energy consumption intensity of the secondary industry is far greater than that of the primary and tertiary industries. The adjustment of industrial structure will inevitably promote the transformation of energy structure. Integrating the variable weight with the existing evaluation model can directly reflect the balance between the factors of low-carbon economic development in the evaluated area in the comprehensive evaluation results, which is helpful to improve the evaluation reliability of the existing research. Because climate warming and other changes brought about by it are still happening and unavoidable, we must actively deal with the public health, economy, quality of life, and other related risks caused by environmental change related to climate change. This area is mainly mountainous and sparsely populated, with a large proportion of planting industry and developed tourism. We should choose the maintenance mode of traditional low-carbon industry and increase the development of ecological and tourism industry. To a certain extent, these measures have reversed the situation that the development gap between the East and the West has gradually widened, but the economy in the upper reaches of the economic belt is growing rapidly. At the same time, a large number of infrastructure projects with high energy consumption and material consumption are concentrated. Topic model is an effective method to analyze large text. It analyzes the training

corpus composed of documents with topic annotation. The lack of specific implementation rules or interpretation provisions leads to the lack of clarity and operability of fiscal and taxation and the dilemma that fiscal and taxation policies cannot be effectively implemented.

5. Conclusions

With the outbreak of the new technological revolution, the population has increased dramatically, the urbanization process has been accelerated, and the traditional “high carbon consumption, high pollution and high waste” economy has developed unprecedentedly. It has caused serious ecological and environmental crises such as ecological destruction, environmental pollution, and climate change, especially the increasingly obvious greenhouse effect, and the global warming trend has intensified. At the same time, the scarcity of water resources, the decrease of land area, the depletion of natural and mineral resources, the decrease of biodiversity, and the destruction of ecological environment are all problems to be solved in the context of the continuous increase of world population. Choosing a low-carbon economic development model suitable for China’s national conditions plays a vital role in the formulation of low-carbon economic policies. Especially in the era of big data, with the exponential growth of information, some data mining applications have gradually surfaced, and the problem solving has become more and more complicated and difficult. High accuracy and low computational complexity are obtained based on traditional research methods. As a multifield interdisciplinary subject, machine learning algorithm can effectively improve the positioning speed and accuracy, and the application effect is good. It makes an in-depth analysis of these two types of concepts and points out that the concept of low-carbon economy in a narrow sense emphasizes the characteristics of “phase” and “coordination” between elements of low-carbon economy. From the perspective of regional economic evolution, low-carbon economy can be regarded as a process of decarbonization of economic development. It is an economic form that combines its own characteristics and is necessary for the regional economy to transform from other economic models to low-carbon economic models. There are regional economic development basis, energy structure, resource characteristics, industrial status, development mode, geographical location, and other factors to choose.

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 no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Q. Li, S. Li, S. Zhang, J. Hu, and J. Hu, “A review of text corpus-based tourism big data mining,” *Applied Sciences*, vol. 9, no. 16, p. 3300, 2019.
- [2] C. Louche, T. Busch, P. Crifo, and A. Marcus, “Financial markets and the transition to a low-carbon economy: challenging the dominant logics,” *Organization & Environment*, vol. 32, no. 1, pp. 3–17, 2019.
- [3] I. M. Yusupov, “Scientific and practical experience in studying ecological problems,” *Asian Journal of Multidimensional Research*, vol. 10, no. 5, pp. 563–568, 2021.
- [4] J. He, Z. Li, X. Zhang et al., “Comprehensive report on China’s long-term low-carbon development strategies and pathways,” *Chinese Journal of Population, Resources and Environment*, vol. 18, no. 4, pp. 263–295, 2020.
- [5] N. Watts, M. Amann, N. Arnell et al., “The 2018 report of the Lancet Countdown on health and climate change: shaping the health of nations for centuries to come,” *The Lancet*, vol. 392, no. 10163, pp. 2479–2514, 2018.
- [6] E. Campiglio, Y. Dafermos, P. Monnin, J. Ryan-Collins, G. Schotten, and M. Tanaka, “Climate change challenges for central banks and financial regulators,” *Nature Climate Change*, vol. 8, no. 6, pp. 462–468, 2018.
- [7] I. Lee and Y. J. Shin, “Machine learning for enterprises: applications, algorithm selection, and challenges,” *Business Horizons*, vol. 63, no. 2, pp. 157–170, 2020.
- [8] L. Ionescu, “Transitioning to a low-carbon economy,” *Geopolitics, History, and International Relations*, vol. 13, no. 1, pp. 86–96, 2021.
- [9] J. Ma, Q. Hu, W. Shen, and X. Wei, “Does the low-carbon city pilot policy promote green technology innovation? Based on green patent data of Chinese A-share listed companies,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 7, p. 3695, 2021.
- [10] D. Lin, L. Hanscom, A. Murthy et al., “Ecological footprint accounting for countries: updates and results of the National Footprint Accounts, 2012–2018,” *Resources*, vol. 7, no. 3, p. 58, 2018.
- [11] T. Peng and H. Deng, “Research on the sustainable development process of low-carbon pilot cities: the case study of Guiyang, a low-carbon pilot city in south-west China,” *Environment, Development and Sustainability*, vol. 23, no. 2, pp. 2382–2403, 2021.
- [12] Z. Zhang, H. Xie, J. Zhang, X. Wang, J. Wei, and X. Quan, “Prediction and trend analysis of regional industrial carbon emission in China: a study of Nanjing City,” *International Journal of Environmental Research and Public Health*, vol. 19, no. 12, p. 7165, 2022.
- [13] J. Li, Y. Chen, Z. Li, and Z. Liu, “Quantitative analysis of the impact factors of conventional energy carbon emissions in Kazakhstan based on LMDI decomposition and STIRPAT model,” *Journal of Geographical Sciences*, vol. 28, no. 7, pp. 1001–1019, 2018.
- [14] Y. Yu and Y. Du, “Impact of technological innovation on CO₂ emissions and emissions trend prediction on ‘New Normal’ economy in China,” *Atmospheric Pollution Research*, vol. 10, no. 1, pp. 152–161, 2019.
- [15] M. D. A. Fleming, L. Poruschi, T. Measham, J. Meyers, and M. Moglia, “Economic vulnerability and regional implications of a low carbon emissions future,” *Australian Journal of*

Retraction

Retracted: Designing Landscape of Urban Gardening Based on Optimized Artificial Intelligence Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Liu, "Designing Landscape of Urban Gardening Based on Optimized Artificial Intelligence Model," *Journal of Function Spaces*, vol. 2022, Article ID 7963173, 12 pages, 2022.

Research Article

Designing Landscape of Urban Gardening Based on Optimized Artificial Intelligence Model

Long Liu 

School of Architectural and Artistic Design, Henan Polytechnic University, Jiaozuo, Henan 454000, China

Correspondence should be addressed to Long Liu; ll5618@hpu.edu.cn

Received 26 May 2022; Revised 23 July 2022; Accepted 28 July 2022; Published 24 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Long 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 landscape is driven by innovative design, adhering to the concept of “poetic residence, inheritance, and innovation,” which has long served China’s urban and rural development and the construction of ecological civilization and provided high-quality planning and design services for governments at all levels throughout the country; we construct a particle swarm optimization (PSO) landscape pattern spatial optimization model and solution algorithm to optimize the spatial layout of the landscape for economic development, ecological protection, and integrated scenarios in a city in southwest China. The results show that the PSO-based landscape pattern spatial optimization model and algorithm can use particle position to simulate landscape distribution for spatial pattern optimization. In the development of landscape pattern optimization methods, the landscape pattern optimization model with landscape simulation evolution as the core has shown its advantages. In the target year, the dominant landscape of economic development scenario is urban and orchards, and the landscape pattern shows the distribution characteristics of urban, farmland in the western dam area, and orchards in the eastern mountainous area; the dominant landscape of ecological protection scenario is forest, urban, and rural residential and industrial mining; and the landscape pattern shows the distribution characteristics of urban and rural residential and industrial mining, orchards, farmland in the western dam area, and forest in the eastern mountainous area. The landscape pattern shows that the western dam area is dominated by urban and rural residential and industrial, while the eastern mountainous area is dominated by forests and orchards. The integrated scenario has the highest potential for the future, and its economic, ecological, and comprehensive benefits can be optimized, which is the best spatial layout of the landscape pattern in the study area in the target year.

1. Introduction

In recent years, the urbanization process of rapid development has focused on the safety and smoothness of traffic, with much emphasis on the vehicular experience of the road, and the consideration of the human living environment is also mostly from the perspective of the driver [1]. However, many factors have led to the fact that most of this year’s urban construction has not only failed to achieve the original purpose of smooth traffic flow but also induced negative emotions such as congestion and anxiety among pedestrians due to various traffic congestions and other “urban diseases.” The American scholar Roger Trancik (2008) proposed the concept of “lost space” in “In Search of Lost Space,” arguing that the misuse of automobiles, private interests overriding

public interests in urban renewal, and land use patterns within cities have created a vacuum. The concept of “lost space” is proposed in “In Search of Lost Spaces” by Roger Trancik (2008), who argues that the misuse of cars, private interests overriding public interests in urban renewal, and land use patterns within cities have created a vacuum in cities. “The “human,” “vernacular,” and “imaginative” positive spaces are increasingly neglected and rejected [2]. The traditional “human,” “vernacular,” and “imaginative” positive spaces are increasingly neglected and rejected [3, 4]. Along with the increasing demand and requirements for the quality of human living environment, the public is increasingly looking forward to public spaces that can bring them pleasant experiences. As one of the important public open spaces in the city, street space is an important direct

carrier for people to recognize and experience urban culture and urban imagery and plays an important role in the daily life of residents, far from being only a simple function of transportation [5].

Urban landscaping landscape design is a landscape pattern that can protect and restore biodiversity and achieve effective control and continuous improvement of ecological and environmental problems [6]. Landscape design is the optimization and adjustment of the landscape quantity structure and spatial layout with the help of GIS technology, scenario analysis, spatial optimization models, and methods to form a landscape spatial configuration scheme with the maximum ecological and economic integrated benefits [7]. The suburban area is an area close to the central area of the city, which is closely related to the central area in economic, social, and cultural aspects and has convenient transportation links with it [8]. In recent years, due to unreasonable land use development and urban sprawl, the structure and function of the ecosystem in periurban areas have been seriously damaged, resulting in dramatic changes in regional climate, hydrological processes, biogeochemical cycles, and biodiversity. Pinto and Maheshwari proposed methods for assessing critical river health in river systems around cities [9]; Zhang et al. studied the relationship between landscape structure of woody plant communities and land use intensification, species diversity, and functional diversity in semiarid regions [10]. Forest reduction and pollution by “three wastes” have appeared. Although some places have improved the local ecological environment through environmental remediation, the overall deterioration trend has not been curbed, mainly because there is no scientific planning of land use in ecological construction, leading to blind and inefficient ecological protection, and the contradiction and ecological protection have been intensified. How to build a reasonable landscape safety pattern and balance the contradiction between economic development and environmental protection spatially has become a practical problem that needs to be solved to implement the national ecological civilization construction strategy [11]. Landscape pattern means to build landscape security pattern and realize regional ecological security, and it is an effective way to ease the conflict between ecological protection and economic development, which has important practical significance.

The innovation of the research lies in that on the grid level of landscape type map, a method of spatial optimization of landscape pattern based on PSO is proposed. It is proved that this method can effectively couple the results of landscape quantity optimization of constrained optimization model with the relevant policy and economic and social factors of spatial optimization. The optimization of landscape pattern based on high-resolution grid map is realized theoretically. The results show that the model and algorithm of landscape pattern spatial optimization based on particle swarm optimization can use particle location to simulate landscape distribution for spatial pattern optimization. In the development of landscape pattern optimization methods, the landscape pattern optimization model with landscape simulation evolution as the core shows its advantages. This

paper provides a reference value for the best spatial layout of the landscape pattern in the target year of the study area.

Section 1 describes the research background of urban gardening landscape design and the main structure of this paper; Section 2 introduces the current status of domestic and international research in related fields and summarizes the research significance of this paper; Section 3 introduces the evaluation of landscape suitability based on logistic regression model and proposes the PSO landscape pattern spatial optimization model and algorithm. Section 4 tests and analyzes the scheme proposed in this paper. Section 5 summarizes the research content of this paper and gives an outlook on future research directions.

2. The Related Works

The work of landscaping landscape design generally includes many aspects such as scheduling work, labor and safety management, planning management, financial management, science and technology management, material management, information management, production management, business management, equipment management, infrastructure management, and quality management [12]. Landscape design of greenery and landscaping is still in the development stage and is an important science, and with the development of society, this work is receiving more and more attention. In the development process of greening management information system, 1960s onwards, the United States expounded Taylor, Longwood Chemical, and the University of Tennessee Arboretum began to study the management of plant information under the computer [13]. In the 1990s to the present, with the maturity of computer and information technology, foreign greening applications have become more and more extensive and more refined and in-depth. The National Parks Board of Singapore not only has access to every ornamental plant species, number, and growth status in the country through computer search but also has created a digital information file for all trees in the country. Garden information for each country is an indispensable part of the construction of information technology since the 1960s; the United States led the combination of computer technology and forestry; after nearly half a century of development, the garden information system is not the initial scientific computing tools, but a comprehensive decision-making and information management system; garden research methods and conservation and management techniques have also had a radical change [14].

The methods for optimizing the number of landscape patterns mainly include classical optimization methods such as linear programming, multiobjective programming, and system dynamic models, which are widely used in land use optimization; for example, Kopeva et al. used the land use structure of Montgomery County, Maryland [15]. Hosseinpour et al. developed a multiobjective linear programming model for the Iranian province of Kermanshahan Brimvand watershed land use for optimal allocation [16]. Alkan used SD-MOP model to simulate and optimize the land use structure of western Jilin province in 2020 [17]. With the advancement of computer and GIS technology, a large

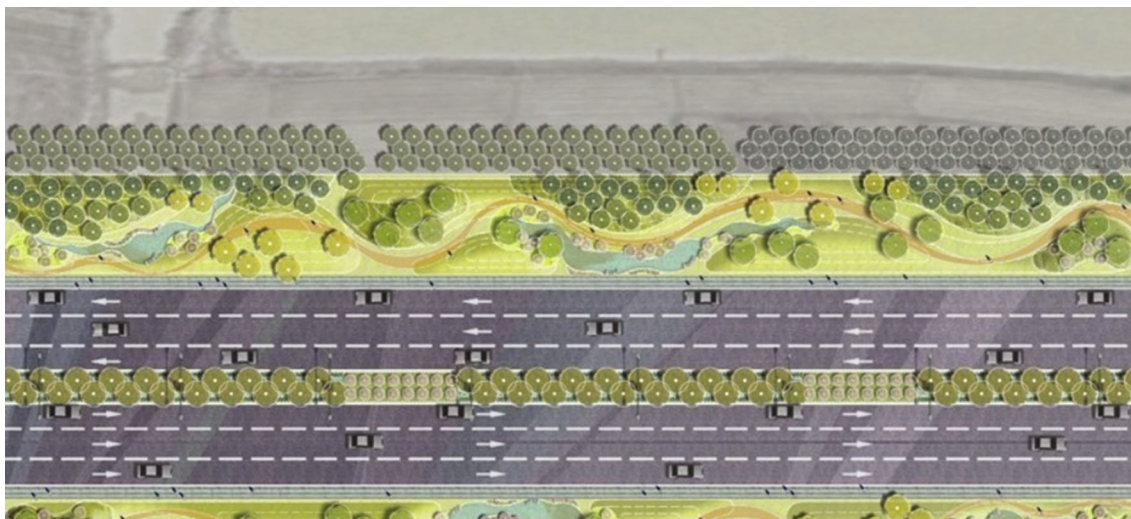


FIGURE 1: Urban green landscape map of A district in a southwestern city.

number of spatial planning decision studies based on meta-cellular automata, scenario analysis method, artificial intelligence optimization method, and integrated optimization method have emerged. For example, Rodrigues et al. used the “ecological location” suitability of metacell to develop probabilistic conversion rules for sustainable land use planning [18]. Padullés et al. applied the Dyna-CLUE model to simulate the spatial distribution of land use for three land demand scenarios in Dianchi watershed [19]. Shan and Sun used a multi-intelligent genetic algorithm to configure the quantitative structure and spatial layout of land use [20]. Alhazaa applied autologistic and CLUE-S integrated models to simulate the land use pattern of Li River watershed in multiple scenarios [21].

The above research has laid the foundation for the development of green landscape design, but most of the modeling methods used in the research are limited to the unilateral optimization of quantitative structure or spatial layout, lacking the organic coupling of the two; some of the modeling methods also lack the integration of policy and economic and social factors and are cumbersome to operate and poorly applied. Therefore, there is an urgent need for a spatial pattern optimization method that can simulate the quantitative and spatial dynamic processes of landscape patterns and reflect the current socioeconomic conditions and development plans of the region with easy application. PSO is an evolutionary algorithm that can parallelize the multidimensional discontinuous decision space, with simple search speed and easy implementation, and has been successfully applied to shopping mall location, soil layout, land use optimization, and other spatial optimization decisions [22], but its application in landscape pattern optimization is still rare. In order to solve this problem and encourage multidisciplinary research on the interaction between urban people and wildlife, Van Dam et al. proposed the extent to which design and planning actions should be consistent. Define the urban ecological content in the context of compact cities [23]. From the perspective of urban design, Abomohra et al. outline the concept of smart city and high-

light the important entry point of smart city landscape design. This paper introduces the strong principles and relationships of smart city landscape design and the role of urban design [24]. Fekete et al. analyzed the sustainable strategy of local bathing beach transformation and landscape protection. The contents of participatory landscape design and water management are analyzed [25].

In view of this, this paper takes a city in southwest China, Area A, as the study area, and firstly conducts landscape suitability evaluation, then applies the constrained optimization method to optimize the quantitative structure of landscape pattern for each scenario, in order to provide a theoretical basis for land use planning, town planning, and ecological construction planning in this area and also provide a methodological reference for other areas to conduct similar studies. It is expected to provide theoretical basis for ecological construction planning in this area and also provide methodological reference for similar studies in other areas.

3. Optimized BP Neural Network Model Based on PSO

3.1. Landscape Suitability Evaluation Based on Logistic Regression Model. Based on the basic principles of landscape suitability evaluation and with reference to relevant researches, this paper selects indicators from socioeconomic factors to build an evaluation index system [23]. Landscape suitability evaluation is the evaluation of the suitability of spatial distribution of a certain type of landscape, which is the basis and foundation of landscape pattern optimization. The study area has both flat dams, mountains, and hills. Referring to similar studies, we choose the elevation, slope, slope direction, and topographic relief to characterize the topographic factors affecting the landscape pattern Figure 1.

Climate is the main factor to determine the distribution of landscape, and the spatial variability of temperature and precipitation in the study area is obvious due to the difference, so the multiyear average rainfall and temperature are

selected to characterize the climate factors affecting the distribution of landscape [24]. Soil type distribution will affect the spatial pattern of landscape to a certain extent and is an important basis for soil type classification, so soil organic matter content is chosen to characterize the soil factors affecting, which has been confirmed in similar studies [25].

High buildings in the city, or excessively high, or close to the visual channel, or poor shape, or color incompatible with the surrounding environment, may cause great landscape impact. Urban roads, especially urban trunk roads, determine not only the pattern of the city but also the main visual space and corridor of the city, which plays an important role in the urban landscape. At the same time, the transformation of old urban areas and the construction of new areas have both the landscape impact caused by the planning and layout and the landscape problems of the building itself. Industrial pollution type construction projects discharge air pollutants and reduce atmospheric visibility. It has a significant impact on the urban landscape and residents' comfort. So the nearest distance to the urban center, the nearest distance to the center of the established town, the nearest distance to the main road, and the nearest distance to the main water area are chosen to characterize the neighborhood factors affecting the distribution of the landscape. In the smaller spatial scale, the influence of socioeconomic factors on the landscape pattern is stronger than that of natural factors; the study area is a suburban area of a large city, influenced by the rapid development of urbanization and industrialization; the population continues to gather in towns and cities; and urban-rural distribution is uneven. Choose population density and per capita gross regional product to characterize the influence of landscape. Therefore, population density and gross regional product per capita are chosen to characterize the socioeconomic factors affecting the distribution of the landscape.

Each type of landscape has two states of "existence" and "nonexistence" in a certain spatial range, which is suitable for analysis by logistic regression model, and the expression of the model is

$$p = \frac{\exp(a + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_\theta X_\theta)}{1 + \exp(a + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_\theta X_\theta)}, \quad (1)$$

where p is the probability value of a certain type of landscape on each raster in the region; the larger the probability value of the raster, the more suitable for the layout of this type of landscape, in other words, the more suitable for the spatial distribution of the landscape, so p also characterizes the suitability of the spatial distribution of the landscape; $p \in [0, 1]$; $X_\kappa (\kappa = 1, 2, \dots, \theta)$ is the influence factor of the spatial distribution of the landscape; a is the constant of the regression equation; and $\beta_\kappa (\kappa = 1, 2, \dots, \theta)$ is the regression coefficient. In this paper, we first extracted the landscape map of farmland and orchard and assigned the value of "1" to the raster of the landscape map and the value of "0" to the raster of the nonexistent landscape map, then extracted the values of each raster of the landscape map and its corresponding raster of the spatial distribution of

13 evaluation indexes. Then, we extracted the raster values of the landscape map and the raster values of the spatial distribution of the 13 evaluation indexes and introduced them into SPSS 19.0 for binary logistic analysis by stepwise regression method and obtained the spatial pattern impact indexes of the landscape and the regression coefficients of each index; finally, we calculated the probability map of the spatial distribution of each landscape according to equation (1) using Python programming, i.e., the landscape suitability evaluation map.

Combining with the landscape pattern situation in the base year (2014), three scenarios of economic development, ecological protection, and integration were designed, and five types of landscape areas, namely, farmland (z_1), orchard (z_2), forest (z_3), urban and rural habitat and industrial and mining (z_4), and water body (z_5), were used as decision variables, as constraints to maximize the economic benefits of landscape utilization and maximize ecological safety and comprehensive benefits, respectively. The optimization model of landscape pattern was established with the objectives of maximizing the economic benefits of landscape utilization, maximizing ecological safety and maximizing comprehensive benefits, respectively, and optimizing the landscape area for different scenarios in the target years (2024 and 2034).

The main objective of the economic development scenario is to produce more goods and provide more services by rational use of limited landscape resources, and its objective function is

$$f(z) = \text{Sup} \left\{ \lim_{K \rightarrow \infty} \sum_{k=1}^K c_k z_k \right\}, \quad (2)$$

where $f(z)$ is the total economic output value of each landscape (million yuan); c_k is the output value coefficient of the k type of landscape ((million yuan) / hm^2); z_k is the area of the k type of landscape; and K is the number of landscape types. The ratio of the total output value of industry and agriculture, forestry, animal husbandry, service, and fishery to the corresponding landscape area in 2014 is used to express the coefficient of landscape output value; then, the objective function of economic development scenario in the study area is

$$f = 5.48z_1 + 6.31z_2 + 2.04z_3 + 539.36z_4 + 3.85z_5. \quad (3)$$

3.1.1. Ecological Conservation Scenarios and Objective Functions. The main objective of the ecological protection scenario is to maximize the regional ecological safety through the rational layout of landscape resources, and its objective function is

$$g(z) = \text{Sup} \left\{ \lim_{K \rightarrow \infty} \sum_{k=1}^K a_k z_k \right\}, \quad (4)$$

where $g(z)$ is the ecological safety index of each landscape and a_k is the ecological safety index per unit area of the k

type of landscape. According to the results of spatial evaluation of ecological safety in the study area, the ecological protection scenario objective function is

$$g = 4.68z_1 + 5.03z_2 + 5.35z_3 + 4.46z_4 + 4.96z_5. \quad (5)$$

3.1.2. Integrating Scenarios and Objective Functions. The main objective of the integrated scenario is to maximize the comprehensive benefits of the regional landscape through the integrated arrangement of various landscape resources, and its objective function is

$$F(z) = \text{Sup} \left\{ \omega_1 \lim_{K \rightarrow \infty} \sum_{k=1}^K v_k z_k + \omega_2 \lim_{K \rightarrow \infty} \sum_{k=1}^K v'_k z_k \right\}, \quad (6)$$

where $F(z)$ is the comprehensive benefits generated by each landscape in the region; ω_1, ω_2 is the objective function weight, respectively; and v_k, v'_k is the standardized k type of landscape production value coefficient and ecological safety index per unit area, respectively.

The total area constraint is

$$\lim_{K \rightarrow \infty} \sum_{k=1}^K z_k = A, \quad (7)$$

where A is the total area of the regional landscape, taken from 55569 hm^2 .

The farmland area constraint is

$$A_L \leq z_1 < A, \quad (8)$$

where A_L is the minimum area of farmland required in the target year. The minimum area of farmland per capita in Area A of a city in southwest China from 1988 to 2014 is 0.0106 hm^2 . Using the population data from 1978 to 2014 in Area A of a city in southwest China, a first-order linear regression model was developed to predict the household population of 680,200 and 742,800 in 2024 and 2034, respectively, and the minimum demand for farmland in 2024 and 2034 was calculated to be 7235 hm^2 and 7901 hm^2 , respectively.

The orchard area is bounded by

$$A_G \leq z_2 < A. \quad (9)$$

The minimum area of orchard is determined according to the principle of guaranteeing the basic demand of fruit for regional residents. The average production of orchard in A district of a city in southwest China from 2000 to 2012 is 12994 kg/hm^2 , and the fruit demand in 2024 and 2034 is 62068250 kg and 67780500 kg , respectively, according to the fruit consumption level of urban and rural residents of 0.25 $\text{kg}/\text{person}/\text{day}$, so the minimum area of orchard in 2024 and 2034 is projected to be 4777 hm^2 and 5216 hm^2 , respectively.

The forest area is bounded by

$$A_S \leq z_3 < A, \quad (10)$$

where A_S is the minimum forest demand area in the target year (hm^2). In order to consolidate the achievement of creating a national ecological zone in Area A of a city in southwest China, the forest planning area in 2024 and 2034 should be no less than the current area 5167 hm^2 .

The urban and rural habitat and industrial and mining area constraints are

$$A_C \leq z_4 < m_1 p_1 + m_2 p_2 + \varphi + \delta. \quad (11)$$

The area of the water body is bounded by

$$A_W \leq z_5 < A, \quad (12)$$

where A_W is the minimum demand area of water bodies in the target year. In order to ensure the regional water security, the planned area of water bodies in 2024 and 2034 should be no less than the current area 1610 hm^2 .

Ecological service value constraints are calculated as

$$\xi_{\min} \leq \lim_{K \rightarrow \infty} \sum_{k=1}^K \varepsilon_k z_k < \xi_{\max}. \quad (13)$$

The minimum and maximum ecological service values in 2024 were estimated to be RMB 669.72 million and RMB 1008.107 million, and in 2034 were RMB 691.80 million and RMB 10086.88 million, respectively, based on the area constraints of each landscape.

Chemical oxygen demand (COD_{cr}) annual load constraint are calculated as

$$\text{COD}_{\min} \leq 10^4 \lim_{K \rightarrow \infty} \sum_{k=1}^K u_k h \text{LC}_k z_k < \text{COD}_{\max}, \quad (14)$$

where COD_{\min} and COD_{\max} are the minimum and maximum annual load of chemical oxygen demand for the target year, respectively (kg); u_k is the runoff coefficient for the k category of landscape; h is the multiyear average rainfall (m); and LC_k is the surface runoff for the k category of landscape COD_{cr} concentration (kg/m^3).

Total nitrogen (TN) annual load constraint is calculated as

$$\text{TN}_{\min} \leq 10^4 \lim_{K \rightarrow \infty} \sum_{k=1}^K u_k h \text{LN}_k z_k < \text{TN}_{\max}. \quad (15)$$

Total phosphorus (TP) annual load constraint is calculated as

$$\text{TP}_{\min} \leq 10^4 \lim_{K \rightarrow \infty} \sum_{k=1}^K u_k h \text{LP}_k z_k < \text{TP}_{\max}, \quad (16)$$

where TP_{\min} and TP_{\max} are the minimum and maximum annual load of total phosphorus in the target year, respectively (kg); LP_k is the TP concentration of surface runoff from the k type of landscape (kg/m³). Based COD_{cr} on the rainfall monitoring data and fertilizer application survey data of major fruits and crops in the study area, the annual load per hectare of farmland, orchard, forest, urban and rural habitat, industrial and mining, and water body landscapes were estimated to be 87.89 kg, 23.12 kg, 16.41 kg, 528.84 kg, and 6.63 kg and TN and TP for 35.08 kg, 22.59 kg, 2.57 kg, 18.70 kg, 2.64 kg, 3.03 kg, 2.66 kg, 0.19 kg, 6.30 kg, and 0.34 kg, respectively. The minimum and maximum annual COD loads in 2024 are 9412702 kg and 17895612 kg and in 2034 are 9481387 kg and 18864542 kg. The minimum and maximum annual TN loads in 2024 are 682316 kg and 3863575 kg and in 2034 are 715596 kg and 3897836 kg. The minimum and maximum annual TP loads in 2024 are 138262 kg and 470088 kg, and in 2034 are 141448 kg and 481631 kg.

Industrial structure constraints are calculated as

$$\zeta_{\min} \leq \frac{a_4 z_4}{a_1 z_1 + a_2 z_2 + a_3 z_3 + a_5 z_5} \leq \zeta_{\max}, \quad (17)$$

where $\zeta_{\min}, \zeta_{\max}$ is the minimum and maximum value of the ratio of the output value of secondary industry to the output value of primary and tertiary industry in the target year. Combining with the macroeconomic situation and the industrial development of a city in southwest A, the growth rates of primary, secondary, and tertiary industries are expected to be 2%, 12%, and 8%, respectively, and then, the values of ζ_{\min} and ζ_{\max} can be projected to be 33.76 and 62.61 in 2024 and 33.76 and 116.98 in 2034, respectively, according to the industrial development statistics of a city in southwest A.

3.2. PSO Landscape Pattern Spatial Optimization Model and Algorithm. The essence of spatial pattern optimization based on landscape type rasterized data is to adjust the image element position around the optimization target, so the key to apply PSO for landscape pattern spatial optimization is to use particle position to simulate the spatial distribution of landscape type raster image elements. The raster map can be regarded as a real number matrix, the raster image elements correspond to the elements in the matrix, and the image element position and attribute value (landscape type code) correspond to the element row number and value, respectively, so the processing of the raster image element position and attribute value is equivalent to the processing of the matrix element row number and value. The matrix is abstracted as a particle; the matrix element value and row number are abstracted as the particle element and position, according to the principle of particle swarm algorithm; no matter how the spatial position of the particle element changes, the particle element itself, that is, the landscape type code, remains unchanged; i.e., the value of several elements composing the matrix is eternally unchanged. The element rank can be optimized by PSO algorithm. When the new matrix corresponds to the landscape spatial pattern,

certain optimization objectives can be made. To achieve the maximum spatial optimization of the landscape pattern under the target. The basic idea of model optimization is shown in Figure 2.

MATLAB is a high-level programming language with matrix as the basic programming unit, which has powerful matrix operations and image processing functions. Therefore, the following key technologies are mainly involved in the design of PSO algorithm for landscape pattern space optimization using MATLAB:

- (1) Let B be a valid element (i.e., the raster value is the element of landscape type code, excluding the value of null -9999 or the element outside the value range of landscape type code) in the matrix $A_{m \times n}$ corresponding to the base year landscape type map, and store the value of vector $(\alpha_1, \alpha_2, \dots, \alpha_N)$, which represents a particle, corresponding to a spatial layout scheme of landscape pattern; $\alpha_j \in B$ is a valid element value, which represents a particle element, and its row number represents the particle element position, then the model optimization process
- (2) Model optimization objectives and constraints. The spatial optimization of landscape pattern is a multiobjective optimization problem; this paper constructs the objective function (fitness function) from two aspects: the maximum suitability of landscape and the maximum spatial compactness, where the suitability of landscape is characterized by the product of various landscape weights and their corresponding suitability, and the spatial compactness is expressed by the landscape patch shape index. The constraints include area, type, and conversion rules. The expressions of model objective function and constraints are

$$\max F(A) = w_1 \sum_{k=1}^K \sum_{l=1}^Q \lambda_k p_{kl} + w_2 \sum_{k=1}^K \sum_{r=1}^R \frac{c_{kr}}{\sqrt{s_{kr}}} \quad (18)$$

$$\text{s.t.} \begin{cases} d_k(A) = D_k^*, \sum_{k=1}^K d_k(A) = \sum_{k=1}^K D_k^* \\ \alpha_j \in [1, 2, \dots, K] \\ 0 \leq T_j(k \rightarrow k') \leq 1 \end{cases} \quad (19)$$

The λ_k weights of farmland, orchard, forest, urban and rural habitat, industrial and mining, and water body in the economic development scenario are 0.1378, 0.2107, 0.0606, 0.5514, and 0.0395, respectively, and the weights in the ecological protection scenario are 0.0706, 0.1366, 0.5071, 0.0350, and 0.2507, respectively, and the weights in the integrated scenario are 0.0930, 0.1613, 0.3582, 0.2072, and 0.1803; p_{kl} is for the suitability value of l image element of k ; c_{kr}, s_{kr} is for the perimeter and area of k landscape patches of r ; w_1, w_2 is for the weights of landscape suitability and

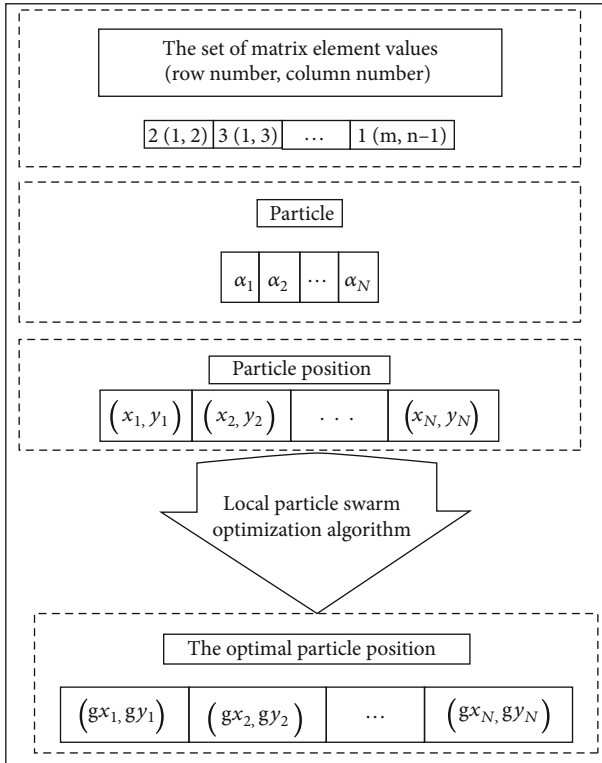


FIGURE 2: Schematic diagram of the PSO landscape pattern spatial optimization model.

spatial compactness in the objective function; k , l , and r for the landscape type, image element, and patch number, respectively; k , Q , and R for the landscape type, a landscape raster, and the number of patches, respectively; $d_k(A)$ is for the landscape pattern; $D_k^* = z_k^*/e^2$ is the number of k -type landscape pixels in the optimization scheme A ; D_k^* is the number of pixels obtained from the optimization of the number of k -type landscape in a certain scenario (z_k^* is the optimal area of k -type landscape; e is the size of pixels); $\alpha_j \in [1, 2, \dots, K] (j = 1, 2, \dots, N)$ is the coding value domain of the corresponding landscape type for pixels; $T_j(k \rightarrow k')$ is the possibility of successful transformation of the corresponding landscape for pixels α_j . Combining with the characteristics of landscape pattern change in the study area, the rules and possibilities of landscape conversion are defined as follows: the interconversion between urban and rural human settlements and industrial and mining and farmland and orchards is completely possible, the interconversion between watershed and farmland is completely possible, the interconversion between forest and farmland and orchards is completely possible, and the interconversion between farmland and orchards is completely possible.

- (3) Initialization of the spatial layout scheme (particles) of the landscape pattern

In order to generate the initial particles with the same number of image elements of each landscape, the spatial allocation of the number of image elements of each land-

scape is carried out based on the landscape type conversion rule and landscape suitability based on the landscape type map of the base period year. The structure of V (see Equation (19)) is an array of cells storing the row and column values of effective elements of each landscape and their corresponding landscape suitability values, $\Delta d_k = |d_k(A) - D_k^*|$ is the absolute difference between the number of base year images of k types of landscapes $d_k(A)$ and the number of optimized images of a scenario, $\text{sort}(\text{Val})$, $\text{sort}'(\text{Val})$ is the ascending and descending order of the variable Val , and $A(x, y) = \alpha$ is the transformation of a landscape image with row number x and y in A into a landscape image with the code α .

4. Analysis of Simulation Results

Combining the regional land use/cover status, the landscape types were classified into five major categories: agricultural land (paddy field, dry land, and watered land), orchard, forest (forested land, shrubland, barren hills, and slopes), urban and rural habitat and industrial and mining (towns, rural settlements, industrial and mining land, special land, and transportation land), and water bodies (rivers, reservoirs, ponds, and ditches), with the study area August 13, 2014, Landsat OLI imagery and ASTERGDEM V2 digital elevation model as the base data, the 2009 1:10,000 land use status map, and the results of the landscape field survey in October–November 2014 as auxiliary data, and the QUEST decision tree classification method was applied to classify the landscape and extract the landscape type map of the areas not involved in landscape pattern optimization.

As shown in Figure 3, the correct prediction rates of the logistic models for each landscape show that the correct prediction rates of the modeled and validated data for farmland are 85.1% and 85.3%, respectively; for orchards, both are 65.6%; for forests, 90.1% and 89.5%, respectively; for urban and rural habitats, 82.7% and 81.7%, respectively; and for water bodies, 97.2% and 97.6%, respectively, indicating that the logistic model has high prediction accuracy for each landscape, and the prediction results have strong credibility.

At the same time, as shown in Figure 4, the ROC values of each logistic model are greater than 0.7. This indicates that the evaluation indicators entering into the regression equation have a better explanation effect on the spatial pattern of each landscape, and these indicators can be used to evaluate the suitability of this type of landscape.

The MATLAB constrained optimization problem solving function $f \min \text{con}()$ was used to solve the landscape pattern quantity optimization model to obtain the optimized area of economic development, ecological protection, integrated scenario farmland, orchard, forest, urban and rural habitat and industrial and mining, and water bodies landscape in 2024 and 2034. The results are shown in Figures 5 and 6.

As can be seen from Figures 5 and 6, compared to the 2014 landscape area, in the economic development scenario, the urban and rural habitat and industrial and mining and agricultural land area increased in the target year, the

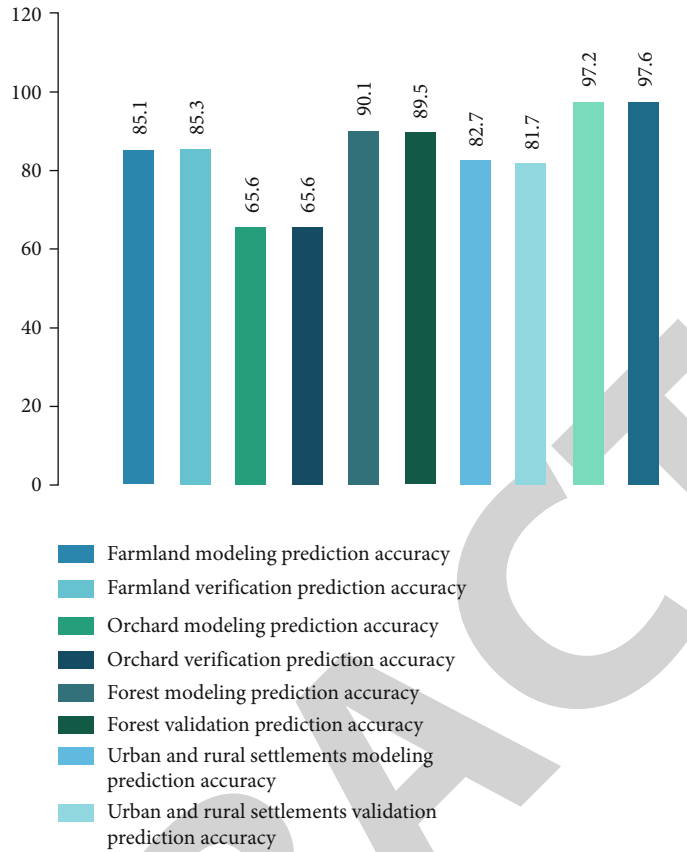


FIGURE 3: Correct prediction rate of the logistic model for each landscape.

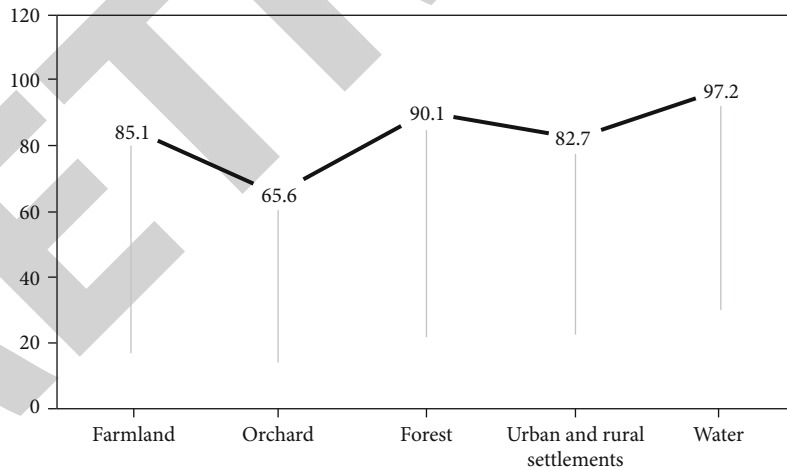


FIGURE 4: ROC values for each logistic model.

orchard area decreased, and the forest and water body area remained unchanged. Urban and rural habitat and industrial and mining increased the most, and orchards decreased the most in both phases of the plan. This is mainly due to the fact that the economic benefits of urban and rural residential and industrial mines are greater than those of orchards, so there is significant increase in urban and rural residential and industrial mines and the significant decrease. The most significant increase is in forests and the most significant

decrease is in orchards in both phases of the plan. This is mainly because the ecological safety level of forests is higher than that of orchards, so the adjustment of large orchards to forests to maximize ecological safety is in line with the actual ecological protection scenario. In the integrated scenario, the area of forests, urban and rural habitats, industrial and mining areas, and agricultural land increased in the target year, while the area of orchards decreased and the area of water bodies remained unchanged. In the two phases of the plan,

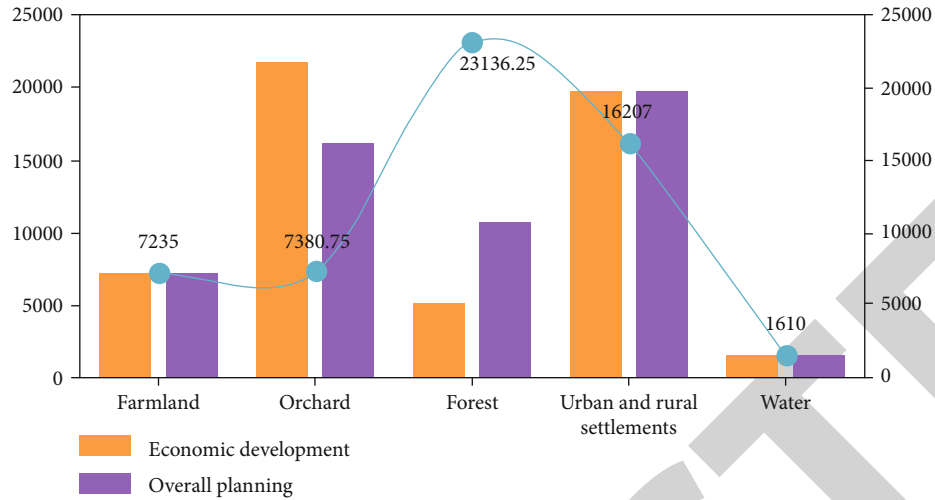


FIGURE 5: Results of landscape area optimization for each scenario in 2024.

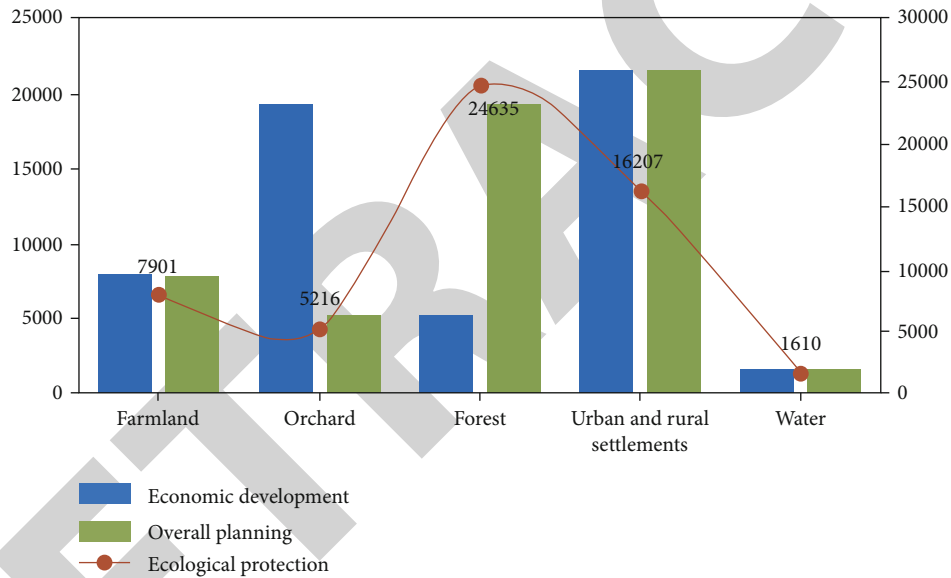


FIGURE 6: Optimization results of landscape area by scenario in 2034.

forests increase the most, urban and rural settlements and industrial and mining areas are relatively second, while orchards decrease the most. This is mainly due to the fact that the ecological safety levels of forests, urban and rural habitats, and industrial and mining economic benefits are greater than orchards, so the adjustment of large orchards to forests and urban and rural habitats in order to increase the comprehensive ecological and economic benefits of the region has resulted in a significant reduction, which is in line with the actual integrated scenario.

The results are shown in Figures 7–9. 2034 economic development scenario has the largest spatial optimization error of 3.47% for forest area and the smallest error of 0.08% for urban and rural habitat and industrial and mining, which indicates that the solution result of PSO landscape pattern spatial optimization model does not fully satisfy

the equation constraint. This is because some particle elements collide in flight; i.e., when a certain type of landscape particle element arrives at an optimized location, another type of landscape particle element arrives at the same optimized location to replace it, resulting in a decrease of the former type of landscape and an increase of the latter type of landscape, resulting in the solution results not fully satisfying the equation constraint. According to the PSO landscape pattern space optimization algorithm, the first particles are initialized and fully satisfy the equation constraint; the reason for breaking the equation constraint is that the corresponding objective function value of the particles breaking the equation constraint is greater than the corresponding objective function value of the first particles; in essence, this makes the landscape pattern space layout further optimized, and a certain error range of the

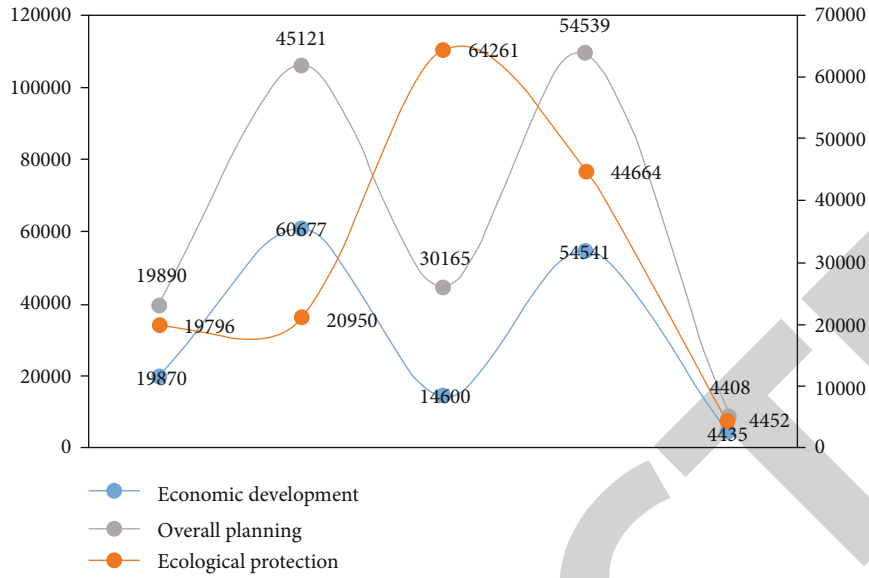


FIGURE 7: Results of the spatial optimization model for PSO landscape pattern in 2024 for each scenario.

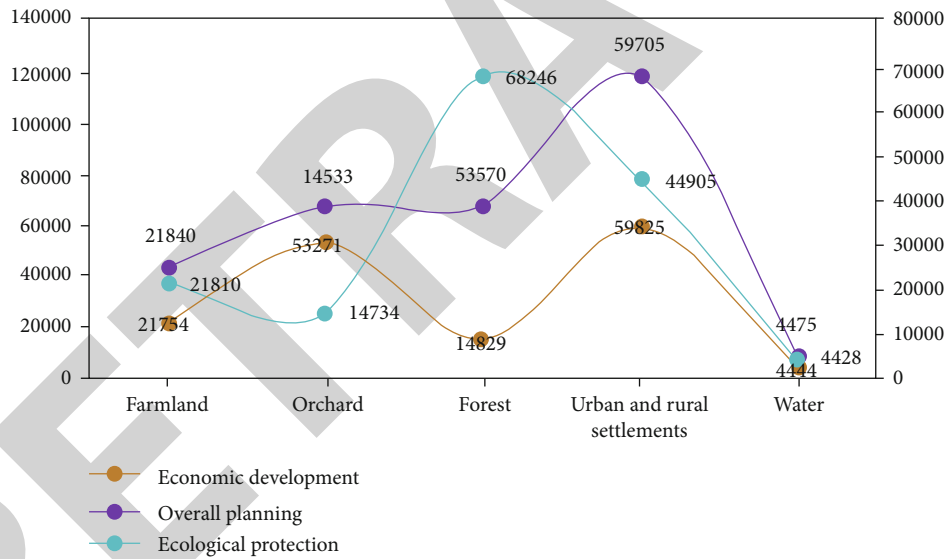


FIGURE 8: Results of the spatial optimization model for PSO landscape pattern in 2034 for each scenario.

equation constraint breakthrough may also be a solution to the number and space optimization objective function coupling problem. Therefore, the model can be applied to optimize the spatial layout of the landscape pattern.

Figure 9 analyzes the landscape pattern distribution characteristic index of PSO landscape pattern spatial optimization results in each scenario from the landscape level, such as overall economic ecology, economic ecology, overall development, protection planning, and other indexes. It provides necessary conditions for further study on the changes of ecological processes and their interactions in the whole region. The index of landscape level is suitable for comparative analysis of landscape pattern and feature changes in

different phases or regions. Since there is only one scene of remote sensing data in this study, this table only reflects the overall landscape pattern features of the study area.

Through the calculation of various landscape indexes in the study area, a comprehensive description of the landscape structure characteristics of the study area is very important for the real-time analysis of the urban landscape pattern structure. The integrated application of the overall optimization method of landscape pattern and the analysis of ecological sensitivity and suitability takes into account the vertical matching and horizontal correlation of landscape units. On this basis, the ecological regulation of the overall spatial pattern provides a more feasible spatial approach for the

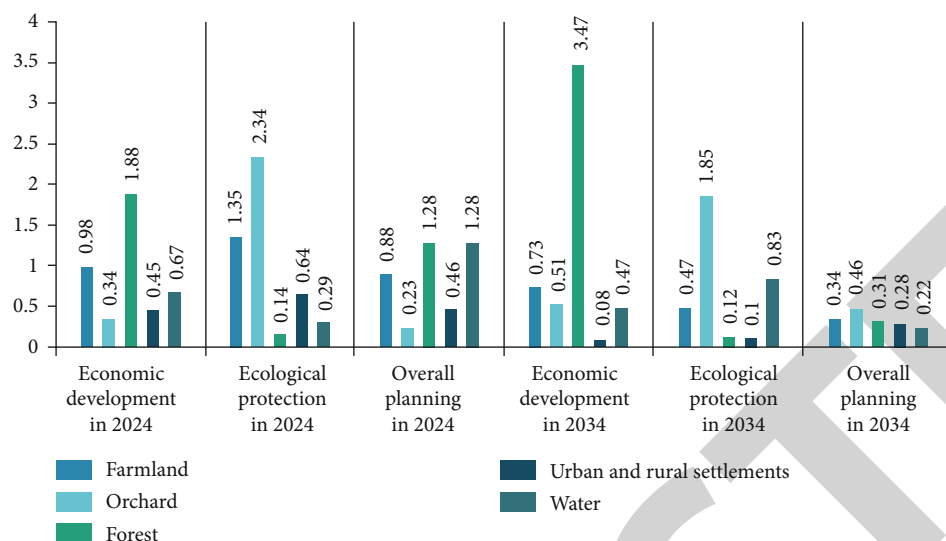


FIGURE 9: Relative errors of spatial optimization results and quantitative optimization results of PSO landscape pattern for 2024 and 2034 for each scenario.

ecological regulation of urban areas and provides support for urban landscape ecological planning.

5. Summary and Outlook

In this study, a model and algorithm of spatial optimization of landscape pattern based on PSO are designed, and the optimization of spatial layout of landscape pattern under three scenarios is realized. The results show that the dominant landscape under the economic development scenario is orchard, and the distribution pattern is farmland in the western dam area and orchard in the eastern mountain area. For the dominant landscape under the ecological protection scenario, the western dam area is composed of urban and rural settlements, industrial and mining areas, orchards, and farmland, and the eastern mountain area is composed of forests. The main landscapes in the integrated scene are forests, urban and rural residential areas, industrial mining areas, orchards, and farmland. The landscape types are forests, industrial mines, and orchards. The distribution pattern is mainly urban and rural settlements, industrial mines, and farmland in the western dam area and forests and orchards in the eastern mountain area. Among them, the economic, ecological, and comprehensive benefits of the comprehensive scenario can be optimized, with the greatest potential in the future. It is the most ideal spatial layout of the landscape pattern in the target year of the study area. In this paper, a method of spatial optimization of landscape pattern based on PSO is proposed on the grid level of landscape type map. It is proved that this method can effectively couple the results of landscape quantity optimization of the constrained optimization model with the relevant policy and economic and social factors of spatial optimization. The optimization of landscape pattern based on high-resolution grid map is realized theoretically, but in application.

However, with the improvement of the grid map resolution or the expansion of the research scope, the amount of

calculation and the running time of the program increase exponentially. This not only puts forward new requirements for computer hardware support but also is the focus of further improvement of this method. Therefore, when making spatial optimization decisions, we should select the appropriate grid map resolution according to the research scope and scale. In addition, considering the coupling between the number of landscape patterns and the objective function of spatial optimization and the efficiency of the algorithm, the particle collision constraint is not added separately, and the relative error of the calculation results is less than 4%, which is a certain limitation for the macroplanning of landscape security pattern at the regional level.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Philosophy and Social Science Planning Project of Henan Province: Study on function evaluation and development countermeasures of urban green Space (No. 2020CYS039); Philosophy and Social Science Planning Project of Henan Province: Environmental Design of Traditional Villages in Northern Henan Province Based on Comfort Degree (No. 2018BYS007); and Humanities and Social Science Project of Education Department of Henan Province of 2019: Study on color comfort of traditional villages in northern Henan province (No. 2019-ZDJH-105).

Retraction

Retracted: Optimization of Ecological Mode of College Students' Innovation and Entrepreneurship Education Based on a Big Data Analysis Model

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Yu, Y. Lu, and H. Wang, "Optimization of Ecological Mode of College Students' Innovation and Entrepreneurship Education Based on a Big Data Analysis Model," *Journal of Function Spaces*, vol. 2022, Article ID 4357354, 8 pages, 2022.

Research Article

Optimization of Ecological Mode of College Students' Innovation and Entrepreneurship Education Based on a Big Data Analysis Model

Miao Yu ¹, Yanju Lu ², and Hongliang Wang ³

¹College of Teacher Education, Baoding University, Baoding, Hebei 071000, China

²College of Tourism, Northwest Normal University, Lanzhou, Gansu 730000, China

³College of Artificial Intelligence, Baoding University, Baoding, Hebei 071000, China

Correspondence should be addressed to Yanju Lu; 2020223022@nwnu.edu.cn

Received 1 June 2022; Revised 25 July 2022; Accepted 28 July 2022; Published 23 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Miao Yu 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.

However, the number of college students enrolling in our country has only begun to increase on a large scale in the 21st century. As educational big data gets more and more attention, it has its own distinct characteristics in the field of education and at the same time consolidates the overall trend of big data development. The countermeasure analysis of innovation and entrepreneurship education under the reform of the education supply side has become a subject to be explored in an all-round way. This paper is mainly aimed at the importance of a big data analysis model to college students, explores the current development status of innovation and entrepreneurship education for college students, finds some problems in the application of big data, and improves the existing problems based on the purpose.

1. Introduction

Big data is often equated with large-scale data information, and this view is very one-sided. Actually, big data is a systematic process integrating information collection, storage, sorting, and analysis, it is a comprehensive processing of data information, and it is an advanced technology with great data value [1]. Behind this series of innovative policies is to serve China's strategy of creating a strong country with double creation and create more jobs. College students, whose goal is to train students to grow into talents, are also promoting the development of innovation and entrepreneurship education in succession, so as to improve the innovation and entrepreneurship ability of graduates. However, it is not easy to bring the great value of educational big data into full play and make use of big data to reform entrepreneurship education. As educational big data gets more and more attention, it has its own distinct characteristics in the field of education and at the same time consolidates the overall trend of big data development. With the further penetration of the concept of "mass entrepreneurship and innovation," the

whole society is affecting college students with the upsurge of entrepreneurship and entrepreneurship and small and microenterprises, and the concept of entrepreneurship and innovation among college students is increasingly strengthened [2]. The adjustment of the macroeconomic structure has put pressure on the reform of education and also pointed out the direction. Innovation and entrepreneurship education is closely related to the reform of the economic system.

College students are the new blood of revitalizing the prosperity of the Chinese nation and the strong main body of promoting the development of innovation and entrepreneurship. The continuous development of Internet innovation and entrepreneurship makes the data and information present typical characteristics of large scale, wide source, and variety [3]. There are great deficiencies in the promotion time of new entrepreneurship education and the students involved. Some college students' innovation and entrepreneurship education has problems such as quick success and lack of experience. In the face of many developed countries, innovation is the driving force of progress [4]. Therefore, to find out the many problems that college

students need to solve in their dual-venture entrepreneurship and to put forward solutions for the “pain points,” only in this way can the college students be encouraged and assisted to start a double business, solve the employment problem, promote innovation, improve the quality of entrepreneurship, and implement the strategy of innovation and entrepreneurship development. Moreover, it can promote the reform of university students’ talent education and its in-depth exploration, so as to meet the objective requirements of modern updating of college students’ educational concepts, modes, and functions in the era of knowledge economy. This will facilitate the development of intellectualized sports industry and promote innovation and entrepreneurship in the future [5].

Starting from the characteristics of college students’ double-venture entrepreneurship, we need to find the exploding point and path to improve the quality of double-venture entrepreneurship to meet the needs of a comprehensive national creation strategy [6]. This is also the manifestation of traditional natural thinking turning to intelligent thinking. It is an opportunity and a challenge for the survival and development of all walks of life. According to the statistics and analysis results of big data [7], it can ensure that scientific decision-making is closer to the actual situation, especially in the field of college students’ double-innovation education. Faced with the competition from some research universities, it is necessary to cultivate a large number of innovative talents, transform innovative achievements, and attract more innovative and entrepreneurial teachers and outstanding enterprises to join the development of college students [8]. Studying the analysis method and application mode of educational big data is conducive to actively responding to the development of China’s big data technology, thus promoting the systematic reform of education [9].

2. Related Work

With regard to the solutions to the predicament of college students’ entrepreneurship and innovation, most scholars put forward their own opinions based on their professional knowledge. Mace opened “New Entrepreneurial Enterprise Management.” Many entrepreneurship scholars define the opening of this course as the beginning and starting point of innovation and entrepreneurship education for Western college students [10]. Feng analyzes the problems existing in the entrepreneurship and innovation policies in Shandong, Shanghai, and other places [11]. Lin established that there is still a lot of room for college students to play their role in such aspects as creating an innovative culture, leading social development, and improving mechanisms. The emphasis on scale and quantity has shifted to the emphasis on structure, efficiency, quality, and innovation, and we should follow the trend and have the courage to reform [12]. Yang proposed to set up entrepreneurship and innovation courses; Baisan Business School has the first undergraduate major named after entrepreneurship education [13]. Wang et al. pointed out the shortage of college students’ entrepreneurship education and also pointed out that the service guarantee system for college students is not perfect

and the social entrepreneurship service mechanism and financing system are not perfect, thus restricting college students’ entrepreneurship [14]. Bai proposed three new measures for the reform of college students’ entrepreneurship education from three perspectives: concept transformation, multiple paths, and system support, positioning of educational goals from improving employment rate to improving employment quality and level, respecting cultural differences in different regions, exploring multiple development paths, and establishing relevant systems to support the development of college students’ entrepreneurship education [15]. Liu et al. did research on the entrepreneurship of female college students, referring to the resistance of opportunity recognition, entrepreneurship financing, the resistance of creating team, the resistance of actually starting a new enterprise, and the resistance of successfully developing the enterprise [16]. Robert Davis founded the student-oriented SIFE, which Smiler explored the US entrepreneurship University paradigm in a highly competitive global entrepreneurship environment with internal and external outcomes, entrepreneurship, contact mechanisms, and support systems [17]. Wang analyzed the impact of the reform on the supply side of education on the private college students. Combining with the characteristics of the private college students themselves, she tried to find out the focus and development direction of the reform on the supply side of college students [18]. Zhang believes that the ability and consciousness of individual dual-venture entrepreneurship affect dual-venture entrepreneurship; in addition, we need to optimize the innovation and entrepreneurship ecosystem through improving policies [19]. Most scholars explore the angle of innovation and entrepreneurship education, but few mention issues such as the quality of college students’ dual-venture entrepreneurship and the success rate of college students’ dual-venture entrepreneurship, which leaves a large space for this study.

3. Significance of Applying Big Data in Double-Creative Education for College Students

3.1. Value of Application. Its educational model has also begun to change subtly, and innovation and entrepreneurship education has been transformed from system to ecological construction. For example, teachers need to fill in the students’ social practice, physical examination, and some data related to students’ moral sentiment for the evaluation of students’ comprehensive quality [20]. Through the analysis and application of educational big data, it is possible to grasp the detailed problems in the students’ learning process and realize the effective reform of the contemporary educational model and educational system. Students have not yet formed a complete and unified understanding of the developmental purpose of this education, as shown in Table 1.

In the teaching of its courses, there are mainly the present situations of dull content and methods, imperfect knowledge structure, weak linkage of professional learning, and low teaching level of teachers in this respect, as shown in Table 2.

TABLE 1: Overall understanding of college students' innovation and entrepreneurship education.

Survey content	Student number	Questionnaire	Proportion (%)
Depends on your career choice	380	790	49.37
All students	130	790	16.46
Know a lot	90	790	11.39
General emphasis	600	790	66.67
Respond to the government's call	280	790	35.44

TABLE 2: Teaching of innovation and entrepreneurship education for college students.

Survey content	Student number	Questionnaire	Proportion (%)
Electives	470	790	59.49
Novel frontier	50	790	6.33
Disconnected from each other	695	790	87.97
Entrepreneurship training skills	240	790	30.38
Important role	140	790	17.72

TABLE 3: Activities attended by college students.

Survey content	Student number	Questionnaire	Proportion (%)
Provided	275	790	34.81
Knowledge lecture	325	790	41.14
Skills competition	125	790	15.82
Corporate internship	80	790	10.13
Project planning training	130	790	16.46

Innovation and entrepreneurship education for college students has the status quo of unbalanced development of activities and unclear cognition, as shown in Table 3.

Accelerate the construction of maker space, and form an innovative service carrier integrating innovation and entrepreneurship, online and offline integration, incubation, and investment. The carriers of innovation and entrepreneurship mainly include the following: (1) the carriers of public service platforms, various professional technical service platforms, and scientific research institutions that gather to provide services for scientific and technological innovation. (2) Entrepreneurial carrier (incubator) mainly refers to the carrier that gathers scientific and technological start-ups and promotes the transformation and incubation of achievements. (3) Innovation and entrepreneurship are entrepreneurial activities based on innovation, which is different from either pure innovation or pure entrepreneurship. (4) Therefore, in the concept of innovation and entrepreneurship, innovation is the foundation and premise of entrepreneurship, and entrepreneurship is the embodiment and extension of innovation. In a narrow sense, it is based on the reorganization, reengineering, and integration matching of related theory, elements, technology, and capabilities. The satisfaction survey of university innovation and entrepreneurship education teachers in the questionnaire survey is shown in Figure 1.

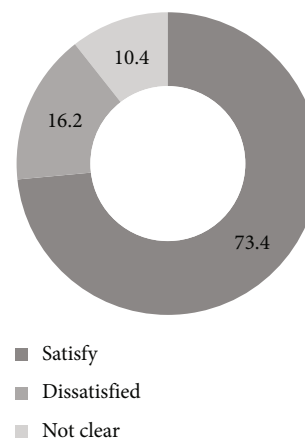


FIGURE 1: The satisfaction of teachers in innovation and entrepreneurship education.

In order to realize a kind of practical teaching that promotes self-physical and mental development, sound personality cultivation, and intelligent space improvement in the new production relationship, college students can choose to take a break from school to start a business and count their entrepreneurial achievements into their students' grades, making the schooling system for entrepreneurship

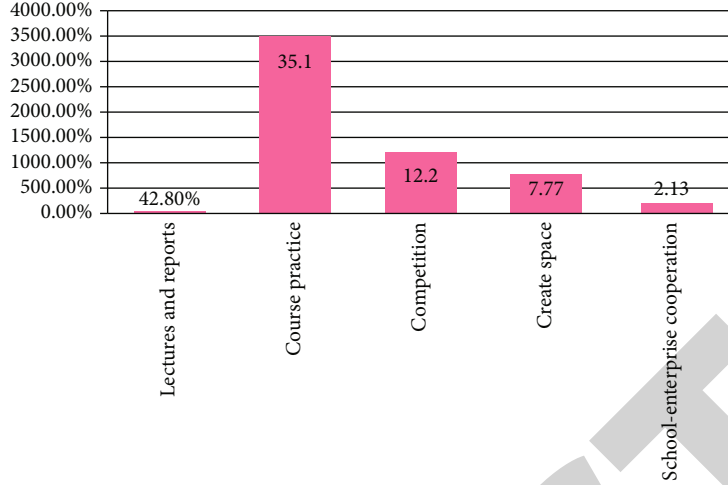


FIGURE 2: Social practice of innovation and entrepreneurship education for college students.

more flexible. From the data survey, the way students participate in social practice is mainly used in lectures, presentations, and courses. There is still a lot of room for progress towards enterprises and related competitions, as shown in Figure 2.

For example, big data needs to collect data through “crawlers,” process data through big data algorithms, and flexibly change strategies according to the results of data processing. Through practical projects or products, it provides a good platform for entrepreneurship education, so that students and entrepreneurs can give full play to their respective advantages on this platform, thus improving their innovation ability. In the modeling and analysis of educational data, the current focus is on the establishment and qualitative analysis of theoretical models, and there is a lack of specific analysis and model construction for the educational teaching process and teaching and learning behavior. The core of the iterative upgrading of modern education is the surge in demand for data analysts. Scientific strategic planning is needed, so that teachers can have a better understanding of students. The current teaching methods are developing in this direction, and information sharing between different schools is strengthened. It is guaranteed that each point in the common space has an equal probability of being included in the sample, which is

$$\begin{aligned} x_i &= \min_{v \in D} \{ \text{dist}(p_i, v) \}, \\ y_i &= \min_{v \in D, v \neq q_i} \{ \text{dist}(q_i, v) \}. \end{aligned} \quad (1)$$

Calculate statistics:

$$H = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i + \sum_{i=1}^n y_i}. \quad (2)$$

The contour factor of object o is defined as

$$s(o) = \frac{b(o) - a(o)}{\max \{a(o), b(o)\}}. \quad (3)$$

For each sample i , calculate the classes it should belong to:

$$c^{(1)} = \arg \min_j \|x_i - \mu_j\|. \quad (4)$$

For each cluster j , the centroid of the class is recalculated:

$$u_j = \frac{\sum_{i=1}^n \text{sign}(c^{(1)} = j) x_i}{\sum_{i=1}^n \text{sign}(c^{(1)} = j)}. \quad (5)$$

$$J(c, \mu) = \sum_{i=1}^n \|x_i - \mu_j\|.$$

Prediction strength is defined as

$$ps(k) = \min_{1 \leq j \leq k} \frac{1}{n_{kj}(n_{kj}) - 1} \sum_{i \neq t \in A_{te}} (D[C(X_{tr}, k), X_{te}]_{it} = 1). \quad (6)$$

That is, the research on establishing mathematical model and mining analysis based on real data is rare. Targeted teaching and counseling will separate students from traditional learning concepts and attitudes. Deepening students' understanding and memory of the content of “double creation education” is conducive to strengthening students' innovative thinking and entrepreneurial ability and promoting students' formal development in entrepreneurial activities.

3.2. Application Function. Entrepreneurship and entrepreneurship education grasp the development opportunities in the era of big data and give full play to its technical advantages, which can realize the mining of big data technology, systematic analysis, and visualization of various teaching

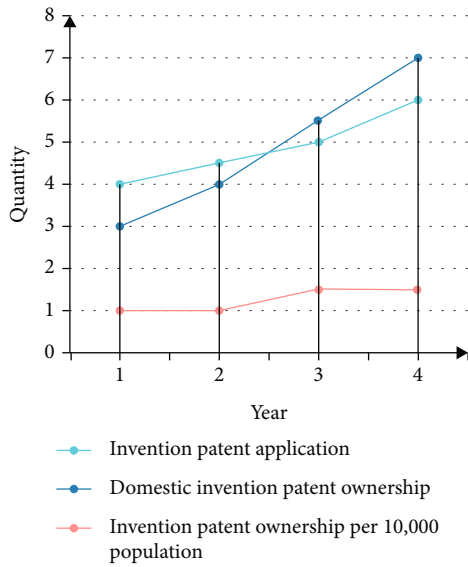


FIGURE 3: Data map of the Chinese patent.

data and provide accurate educational resources for the cultivation of students' practical skills. Of course, this has always been a prominent problem in China's entire education. Educational big data analysis involves multiple interdisciplinary fields, including educational science based on educational theory and learning theory and computer science based on big data and artificial intelligence technology. Based on the analysis of large data, it provides personalized service for the target and makes learning plans according to the specific learning situation and the preferences of the object, so as to stimulate the learning interest of the object. According to the statistics table of comprehensive operation of crowdsourcing space published by the Torch Center of the Double Foundation Department, it is mentioned that "college students' entrepreneurship refers to the entrepreneurship team or enterprise founded by college students". "Among them, college students refer to those who have not studied or graduated for more than two years in institutions of higher learning." Scientific and technological entrepreneurship is creative, a breakthrough in existing technology, and the application of new technology. Intellectual achievements are the most scientific, complete, and core manifestation of scientific and technological innovation, as shown in Figure 3.

From a broad perspective, it is considered that entrepreneurship education is to set the basic qualities and multidimensional skills of entrepreneurship as the integrated goal of entrepreneurship talent cultivation and to teach them in practice, so that their career development and career activities can be effectively transformed into creative social labor. At the same time, what responsibilities students, families, government, and society take in the whole process of education, what obligations they practice, and the external driving force of entrepreneurship education is clear distribution of rights, responsibilities, and benefits. Who is leading college students to set up education? However, the lack of deep utilization of big data by most college students directly restricts

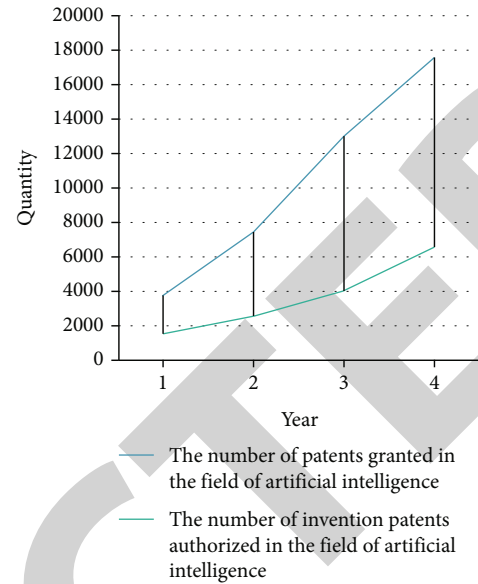


FIGURE 4: Patent-related data in the field of artificial intelligence in China.

the educational effect. The existing innovation and entrepreneurship education system cannot fully realize intelligent and automatic information collection and processing. Universities and scientific research institutes have made great contributions to basic algorithms and basic hardware, while enterprises have more patent achievements in vertical applications and have obvious advantages, as shown in Figure 4.

College student makers are also gradually paying attention to makerspaces to find a starting point for technological entrepreneurship and where to go after six months of graduation, as shown in Figure 5.

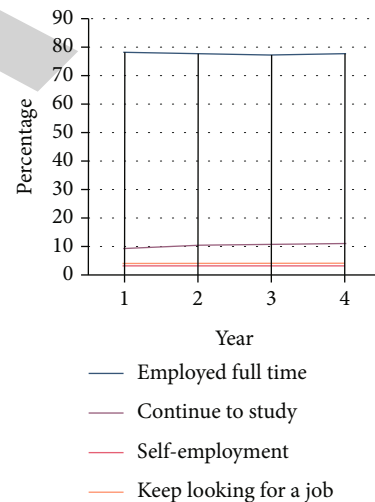


FIGURE 5: Graduates' whereabouts six months after graduation.

the educational effect. The existing innovation and entrepreneurship education system cannot fully realize intelligent and automatic information collection and processing. Universities and scientific research institutes have made great contributions to basic algorithms and basic hardware, while enterprises have more patent achievements in vertical applications and have obvious advantages, as shown in Figure 4.

College student makers are also gradually paying attention to makerspaces to find a starting point for technological entrepreneurship and where to go after six months of graduation, as shown in Figure 5.

Starting the transformation to an innovative country, the internal double-venture entrepreneurship, the educational mechanism, and other innovative entrepreneurship continue to accelerate the reform. Using the statistical functions of large data, we can evaluate the effectiveness of double-

innovation education, help to set up a double-innovation education environment that suits the characteristics of students efficiently, and ensure the effectiveness and pertinence of double-innovation education according to the training objectives of talents.

4. Practical Strategy of Introducing Big Data Technology into College Students' Double-Innovation Education

4.1. *Improve the Dual-Innovation Education Application System.* To deepen the reform of education evaluation in the new era, we should effectively combine education big data with education evaluation methods and make innovative applications. Starting from the concept of measurement and evaluation and the innovative application of statistical analysis methods, a unified evaluation standard and comparable evaluation system are created. And scientifically use equivalent technology to solve the comparability problem of education big data. Specifically, we can develop evaluation tools, build a large education database, pay attention to potential variables and covariates, and innovate and apply equivalent technology. We will develop a dynamic discipline development scale, train and train professionals, and build a database sharing platform.

Under the concept of data entrepreneurship and innovation, the education application system for college students' innovation and entrepreneurship is the education data collection layer and can also be used to analyze previous technical materials. In addition, like the diversity and hierarchy of ecology, college students in different countries and regions should develop their own models of diversity education and training, none of which is suitable for all regions. The willingness of college students to start a scientific and technological business is increasing, as shown in Figure 6.

Based on homework evaluation and examination data, judge students' academic situation and predict students' academic achievements. In the actual situation of college students' pioneering work. It involves not only the application of advanced science and technology in the technical field but also our practical exploration of new fields and innovative involvement and development in all walks of life. Innovation and entrepreneurship education, as a social practical activity of linking education with production, fully endows the theory of people's all-round development with reality and concreteness. It not only breaks through the professional teaching barrier through the interdisciplinary integration, but also enriches everyone's theoretical knowledge, thus promoting the comprehensive development of human intelligence. Among those who choose to start their own businesses half a year after graduation from undergraduates and graduates of higher vocational colleges, more than half of them have given up entrepreneurship for various reasons, and the survival rate of entrepreneurship is less than half, as shown in Figure 7.

The combination of big data and innovative entrepreneurship can be multifaceted. For example, in the start-up and operation of enterprises, big data can accurately target

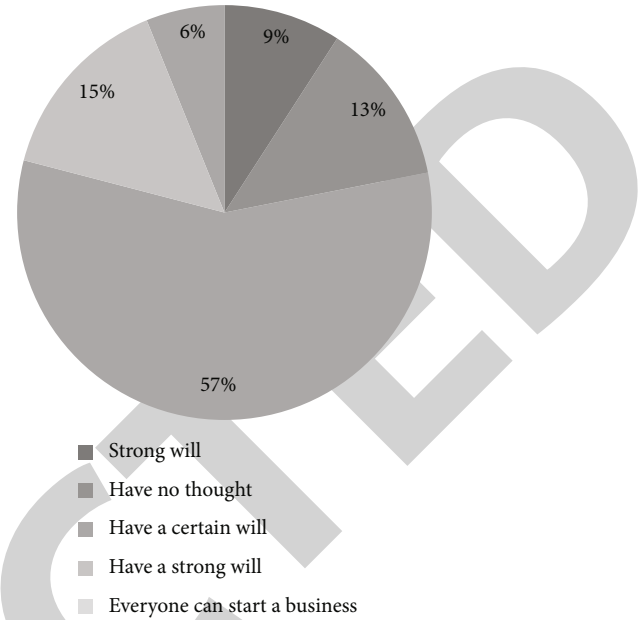


FIGURE 6: Business intention.

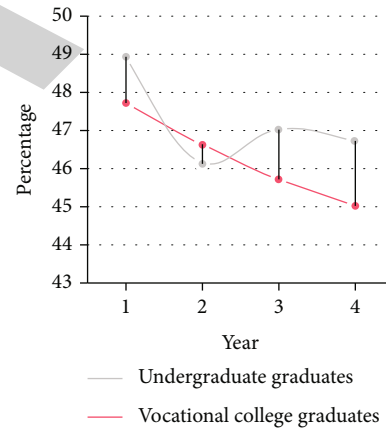


FIGURE 7: The proportion of self-employed entrepreneurs who insist on starting their own business half a year after graduation.

customers with user portraits. Students need to design advertisements with large data user portraits to link customer needs with enterprise products. With the cooperation of college students, teachers and students, cooperative enterprises, and college students' administrators, resource integration, information sharing, technology push, and project promotion are realized, which provides technical support for project management and innovation of dual-innovation education mode.

4.2. *Build a Dual-Creation Practice Platform.* First, build an online network platform led by big data. Use big data to collect entrepreneurship and innovation information, simulate the actual needs of students and users, and strengthen the user's operability and experience. At the same time, with the advantages of low cost and diversified sharing, it provides more convenience for students. Just like an ecosystem,

there are no completely independent individuals, and there are always more or less connections between different individuals, that is, mutual symbiosis, which cannot be analyzed solely for a single factor of education. It has opened up a new research perspective of educational big data analysis: taking information-based classroom teaching as a specific research field of educational big data mining and analysis, building a systematic modeling and analysis framework, and forming big data analysis models and methods. Then, students log in to the teaching platform, receive personalized preview resources pushed by teachers, prepare for class, and complete relevant assessments before class. Because the education they receive is more specialized, they specialize in a certain field of subject knowledge, and they have a higher level of education, and the technical content of college students is relatively high when they start a business. Its essence is to inherently require the establishment of people's subject status and equal development of their own abilities, so that they can not only freely choose the direction, road, and way of self-development. Subjective initiative is the unique ability and function of people to reflect and transform the world. Correct consciousness can guide people to take correct actions and promote the development of things. At the same time, it can transcend the real life itself by virtue of its own subjective initiative. In addition, big data also has the function of public opinion analysis. In innovation and entrepreneurship practice, students can assume that they are the public relations personnel of a large enterprise. Then, they can design a public opinion analysis system that combines corporate brands with big data. Through online and offline interaction, build a three-dimensional platform for double-innovation practice, which is more conducive to promoting the development of students' double-innovation ability in practice, so as to achieve the ideal effect of double-innovation education.

5. Conclusions

This paper studies from the perspective of big data, from the aspects of system construction, technology application, concept change, etc., to expand the path to improve the quality of dual-innovation education and promote the closer relationship between big data and dual-innovation education. In the changing environment of economy, politics, and culture, the traditional university education and operation mode are also changing and innovating with the times. In the new era, China vigorously advocates innovation and uses the development strategy of innovation and entrepreneurship to guide all kinds of social subjects to give full play to their subjective initiative, stimulate the vitality of social innovation, gather "innovation," and deal with the dual-innovation revolution. The application of big data in innovation and entrepreneurship education can timely discover students' personality differences. Innovation and entrepreneurship education requires teachers to keep pace with the development of the times, stand at the forefront of the development of the times, let students learn the latest innovation and entrepreneurship knowledge and tools, and let students create a new future.

However, the current teaching content giving too much emphasis on theoretical knowledge, operability, and practicality is not strong. Although many colleges and universities have set up this course, most of them only study theoretical knowledge and lack a platform for practical operation, and students' entrepreneurial activities only stay at the level of paper work. Therefore, further research is needed.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Research Fund for Transformation and Development of Baoding University in 2021 project: Research on ideological and political work and campus culture construction of Baoding universities (No. 2021J01).

References

- [1] M. Guerrero and D. Urbano, "Effectiveness of technology transfer policies and legislation in fostering entrepreneurial innovations across continents: an overview," *The Journal of Technology Transfer*, vol. 44, no. 5, pp. 1347–1366, 2019.
- [2] Y. Wenhua, "College students employment and entrepreneurship guidance system based on big data," *Microcomputer Application*, vol. 37, no. 9, p. 4, 2021.
- [3] D. Hahn, T. Minola, G. Bosio, and L. Cassia, "The impact of entrepreneurship education on university students' entrepreneurial skills: a family embeddedness perspective," *Small Business Economics*, vol. 55, no. 1, pp. 257–282, 2020.
- [4] J. Huan, "The organic integration of ideological and political education and innovation and entrepreneurship education for college students in the era of digital economy—comment on "empowering the digital economy: the enlightenment of big data innovation and entrepreneurship"," *Chinese Science and Technology Papers*, vol. 16, no. 2, p. 1, 2021.
- [5] J. E. Souto and A. Rodriguez-Lopez, "Entrepreneurial learning in an experiential and competences training context: A business plan in Bachelor thesis," *The International Journal of Management Education*, vol. 19, no. 3, p. 100513, 2021.
- [6] J. Ping, "Research on employment and entrepreneurship guidance strategies for college students based on big data," *Journal of Jiamusi Vocational College*, vol. 36, no. 9, p. 3, 2020.
- [7] L. Xiaohong, "University libraries' competitive intelligence service for college students' innovation and entrepreneurship—based on the background of big data," *Journal of Guiyang University: Social Sciences Edition*, vol. 15, no. 5, p. 4, 2020.
- [8] M. Israr and M. Saleem, "Entrepreneurial intentions among university students in Italy," *Journal of Global Entrepreneurship Research*, vol. 8, no. 1, pp. 1–14, 2018.

Retraction

Retracted: The Effect of Temperature on Intestinal Flora Imbalance Based on Time Series Prediction Model

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Zhu, Y. Xue, and S. Yang, "The Effect of Temperature on Intestinal Flora Imbalance Based on Time Series Prediction Model," *Journal of Function Spaces*, vol. 2022, Article ID 6122406, 11 pages, 2022.

Research Article

The Effect of Temperature on Intestinal Flora Imbalance Based on Time Series Prediction Model

Jing Zhu ^{1,2}, Yafei Xue ^{1,3} and Shuyuan Yang ^{1,2}

¹School of Computer and Information, Hohai University, Jiangsu, Nanjing 211100, China

²College of Computer and Information Engineering, Xinjiang Agricultural University, Xinjiang Wulumuqi 830052, China

³Department of Information Science and Technology, Nanjing Normal University Zhongbei College, Jiangsu, Nanjing 210046, China

Correspondence should be addressed to Jing Zhu; 150207040009@hhu.edu.cn

Received 9 June 2022; Revised 30 July 2022; Accepted 1 August 2022; Published 22 August 2022

Academic Editor: Miaocho Chen

Copyright © 2022 Jing Zhu 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 series of effects caused by temperature change are the biggest problems faced by biological systems. These irregular environmental characteristics have brought new challenges to people and even animal groups. The immune function of human intestinal flora has a great protective effect on other organs and the body environment. It can help the human body carry out intestinal digestion, food absorption, nutritional metabolism, and so on. Based on the above situation, this paper uses the time series prediction model to study the factors affecting the imbalance of intestinal flora in the process of temperature change. Firstly, biological experiments are carried out with animals to simulate the human environment. Based on the sequence information of historical temperature change parameters, a temperature prediction device based on time series model is proposed. The effects of air factors and carbon dioxide content on the prediction results are evaluated by statistical analysis. Secondly, in order to ensure the accuracy of the experimental data, the neural network algorithm is used to optimize the model, and the white blood cell count is used to analyze the influence of temperature change on the intestinal flora structure of the two organisms. Finally, the experimental results are applied to the human environment to analyze the research results. The results showed that with the irregular change of temperature, the number of intestinal flora and internal colony structure also changed. The richness index in the normal temperature environment is relatively large, which can effectively explain the high richness of intestinal microflora in the experimental population. Further analyze the subjects and distinguish the animals according to sex. The number of Bacteroides in the intestinal flora of male animals was higher than that of female animals, but this phenomenon disappeared immediately after physical ligation. In addition, the intestinal flora abundance of female animals is higher than that of male animals, and the metabolic level is faster.

1. Introduction

The internal environment of the human body is composed of various flora. Intestinal flora can ensure the digestion and absorption of food and the promotion of metabolism in daily life [1]. With the irregular change of external temperature, climate poses a great threat to the survival and development of organisms. How to explore the influence of temperature on the number, structure, and state of flora is a common research topic for scholars all over the world [2]. Therefore, based on the above situation, the time series prediction model has become one of the means to solve the current problems. It can calculate and process massive

data, obtain effective information, and accurately predict the development trend of the research object [3]. At present, the generated data exist in the form of a certain sequence, and the conventional sequence reflects the prediction function in the statistical model. According to the development process of time series, this quantitative estimation can solve the problem of information diversity [4]. However, the data processed by the time series model is relatively large, and the internal structure of the system is relatively complex; so, the prediction function needs to be continuously improved [5].

The time series prediction model has two forms: single structure and combined structure. According to the attribute division, it can be determined as combined and noncombined

state [6]. Each prediction model has its own obvious problems, which need to be further optimized and improved in combination with various big data technologies. The first thing to be solved is the accuracy of the model processing data. Whether the prediction results are accurate or not can directly affect the effectiveness of the experimental results [7]. We can select data and restructure the model structure for different experiments. Secondly, as for the noncombinatorial prediction model, the improvement of fuzzy time series algorithm mainly tends to the establishment of fuzzy matrix [8]. Finally, whether the prediction model is reasonable and effective and whether it can be directly applied to solve problems. We found that the multiangle parameter representation of the data can judge the prediction level and then complete the detection of the prediction model [9]. According to the above analysis of the time series prediction model, we also need to combine various biological experimental data in the study of intestinal flora imbalance. The temperature environment simulation scene is used to control the external humidity, internal humidity, temperature supercooling, temperature overheating, and other conditions during the experiment [10]. Liu et al. explore the changes of food intake and physical fitness data of the subjects in the animal experimental group with each degree of temperature rise. Environmental elements such as oxygen content and carbon dioxide content in the environment are added to the prediction model to study their impact on the temperature prediction curve [11]. Sheng et al. use the time series prediction model to simulate the experimental environment, create irregular changes in temperature under various conditions, and analyze the internal correlation of the experimental data. Thus, the experimental environment is changed to ensure the accuracy and effectiveness of the research results [12]. Huaixiang and Kai using the spotted lizard and mouse as the two-way comparative data of animal experiments and the effects of temperature changes on the number, structure, and state of intestinal flora were explored, respectively [13]. Finally, combine the experimental results with the characteristics of human intestinal flora, further sort out and analyze the research results, and judge the specific factors affecting the imbalance of intestinal flora.

2. Application Status of Time Series Prediction Model in Various Countries

The field of time series research is proposed by British scientists, which is transformed from the regression variable model [14]. Subsequently, Swedish scholars made targeted improvements to the automatic regression mobile operation to form a time series average model in a flat mode, which can basically realize the decomposition and classification of data calculation [15]. With the increasing number of research papers on this discussion, the time series model also became a hot research content in various countries at that time [16]. Among them, the average moving time series that can process dynamic data has become a representative in the prediction model because of its wide application [17]. Zeng et al. proposed a modular temperature prediction method based on the autoregressive moving average model. The ARMA method requires historical and current temperature

data of IGBT module, obtains temperature time series through uniform sampling, and constructs the autoregressive (AR) model [18]. With the advent of the era of big data, the proportion of data in the experiment is increasing. The complexity of the time series structure in the process of processing dynamic disordered data makes the experimental results change more. However, the traditional time series model has computational defects due to the increase of data [19]. Therefore, we optimize the structure of the model, and the combined prediction model gradually replaces the original sequence. The combined forecasting model is to integrate multiple calculation methods into the same structure. Analyze and predict the sample set by using the function vertical and horizontal expansion [20]. The improved calculation results of this combination mode have a significant improvement in performance. That is to discretize the data calculation process of the experimental object, obtain the best solution of the calculation steps, and finally form a complete combination.

The time series prediction technology is applied to the observation of historical objects, and the historical development law is captured and inferred according to the time change. The prediction model has no requirements for the background environment of information; so, it can be applied to any object analysis process. The applicability of the time series prediction model is also widely used in river trend, transportation, population migration, urban power consumption, consumption level, etc. [21]. In the economic field, researchers use time series models to judge the trend of capital curve and amount index and use the characteristics of these data to determine the next investment direction of enterprises [22]. According to the different problems to be solved, the long-term and short-term time series prediction models are established. For the same stage, different objects, and environments, the scope of long-term prediction is different from that of short-term prediction. Taking the economic field as an example, the long-term forecast is only for changes with a cycle of more than two years. The short-term prediction model can be applied to power consumption [23]. In order to predict the residential power consumption of the city at the same time, the length of this time period is usually between a few minutes and a few hours. Some scholars have different requirements for prediction application. They divided the model into point structure and interval structure and used probability calculation to assist the prediction model [24]. Point structure is to combine the behavior track after machine training with the dynamic change of time, analyze the change of the sample object at a certain time in the future, and carry out certain control and guidance according to the analysis results. However, the point structure will be unstable in practical application, and it is easy to be disturbed by many factors, which makes the experimental results difficult to convince the public [25]. Then, in the application of interval structure, the calculation range is extended, and the activity interval of the sample object at a certain time is obtained, so as to calculate the reliability of the data. Based on the above research status, this paper uses the time series prediction model to predict the development trend of temperature environment.

Then, the imbalance of intestinal flora under different temperature environment was studied.

3. Influencing Factors and Correlation Analysis of Intestinal Flora Imbalance

3.1. Construction of Experimental Temperature of Intestinal Flora Based on Time Series Prediction Model. The human environment contains many kinds of cells and flora. Flora participates in the normal activities of organs and human body and plays an important role in texture health. Intestinal flora is one of the main factors affecting human dietary behavior. With people's daily intake of different nutrients and foods, intestinal flora will also be affected accordingly and then begin to adapt to the internal environment. Over time, the internal structure and species of intestinal flora have changed greatly. Many young people's daily eating habits are often too cold and overheated. Doctors believe that the alternation of cold and hot food is easy to harm gastrointestinal health. This habit of eating too cold and overheating will always lead to the imbalance of human spleen and stomach. From the perspective of temperature, the root cause is the imbalance of intestinal flora caused by the rapid change of the intestinal environment with temperature. In view of the above situation, we use the time series prediction model to simulate an environment suitable for biological feeding and study the regulation of internal temperature.

The integrated biological food room is a complex simulation scenario, and the temperature of the whole room can be regulated and monitored. In the process of biological experiment, both supercooling and overheating of the environment will cause different changes. The traditional model uses the combination of statistics and algorithm to predict the temperature, but without the help of natural wind, the experimental temperature change will produce an irregular state. Because it is based on the feedback principle, for the process with large delay, such as temperature control, the stability time is too long. On the other hand, many modern control algorithms need to build models in advance. This is costly for high-dimensional and multivariable industrial processes. Considering all kinds of interference, it is difficult to obtain ideal results. In addition, some scholars use humidity series for temperature analysis, and the calculation result curve is static, which cannot effectively calculate the dynamic data. This vector regression prediction method can only judge the humidity value of the next stage and estimate the experimental temperature according to the humidity. Although this method shows the linear transformation of the data, the accuracy of the experimental results is not properly controlled, and the accuracy is relatively poor. In this paper, the time series model is used to dynamically predict the time change of the experimental environment. The optimization of the neural network algorithm is added to the condition of multi module, and the influence of various characteristics in the experimental environment on the results is studied. The data series of experimental temperature are affected by space, geographical environment, relative humidity, oxygen content, carbon dioxide

content, and so on. The process of building a prediction model for temperature data at the same time is as follows:

As can be seen from Figure 1, first, process the temperature data sequence, use the neural network model and data processing technology, and judge the reliability index of the model. Finally, the experimental results are evaluated according to the single variable of the missing function. In the data acquisition and processing, the multisource remote monitoring system is used to collect the values of humidity, oxygen content, and carbon dioxide content in the experimental environment, preprocess the data, and remove the wrong data. The experiment also adds the cyclic unit network algorithm, that is, the improved neural network algorithm. Compared with the data disappearance problem of traditional algorithms, it increases the efficiency of the training model and reduces the decline of parameter accuracy and too complex problems caused by too much data. The internal structure and expansion structure of the cyclic element algorithm are as follows:

As can be seen from Figure 2, the state of each midway transformation of transmission data will be recorded. The judgment conditions can be updated according to the input results of the current node, and the calculated data can be reset by using the cyclic control structure. Finally, it can reduce the data participation and the complexity of the prediction model. The specific application formula of cyclic unit network is as follows:

$$r(t) = \text{sigmoid}[U_r h(t-1) + W_r x(t)], \quad (1)$$

$$z(t) = \text{sigmoid}[U_z h(t-1) + W_z x(t)], \quad (2)$$

$$\tilde{h}(t) = \text{ReLU}\{U[r(t) \odot h(t-1) + Wx(t)]\}, \quad (3)$$

$$h(t) = [1 - z(t)h(t-1) + z(t)\tilde{h}(t)]. \quad (4)$$

In the formula, U_r and W_r , respectively, represent the hidden state of each control condition and the weight value of input data. Hide and reset the changes at each time in the future. In the data processing stage, association rules are combined to calculate and extract the correlation of each feature in the sample data. Variance or least square fitting is used in regression, which has the disadvantage of unstable scoring of correlation characteristics. The average accuracy is reduced by disrupting the order of features to see the impact on the accuracy of the model. Stability selection is a new method based on the combination of quadratic sampling and selection algorithm. The principle is to run the feature selection algorithm on different data subsets, repeat constantly, and finally summarize the feature selection results. The effect is generally good. The number of feature points contained in the set and association rules jointly construct the spatial region. The proportion of valid data is calculated as follows:

$$\text{support}(X \implies Y) = \frac{|T; X \cup Y \subseteq T, T \in D|}{|D|}. \quad (5)$$

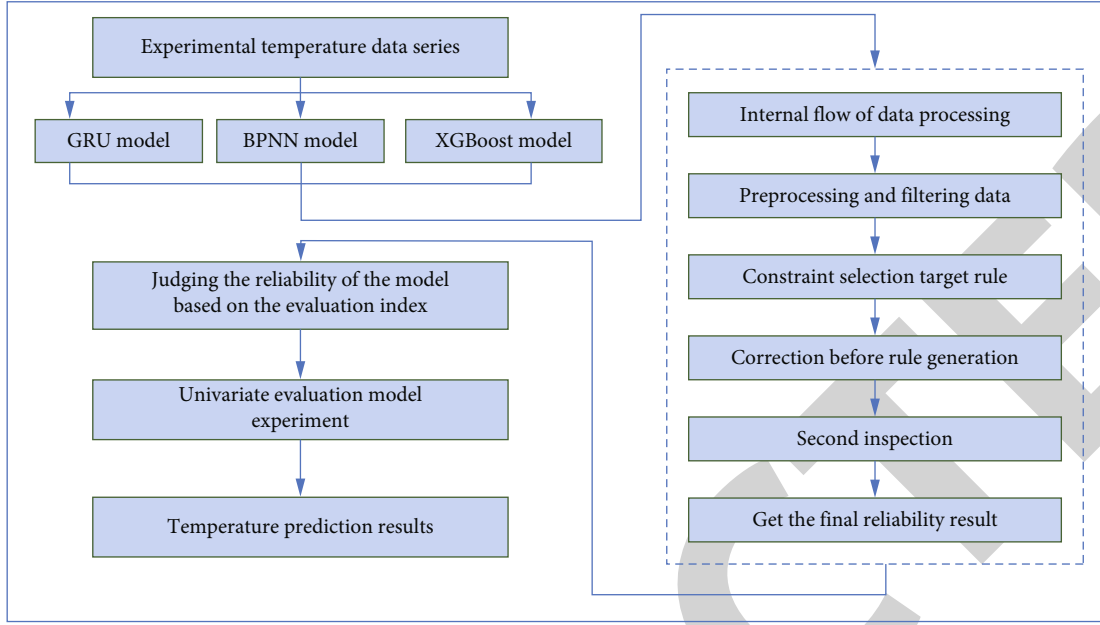


FIGURE 1: Flow chart of the prediction model.

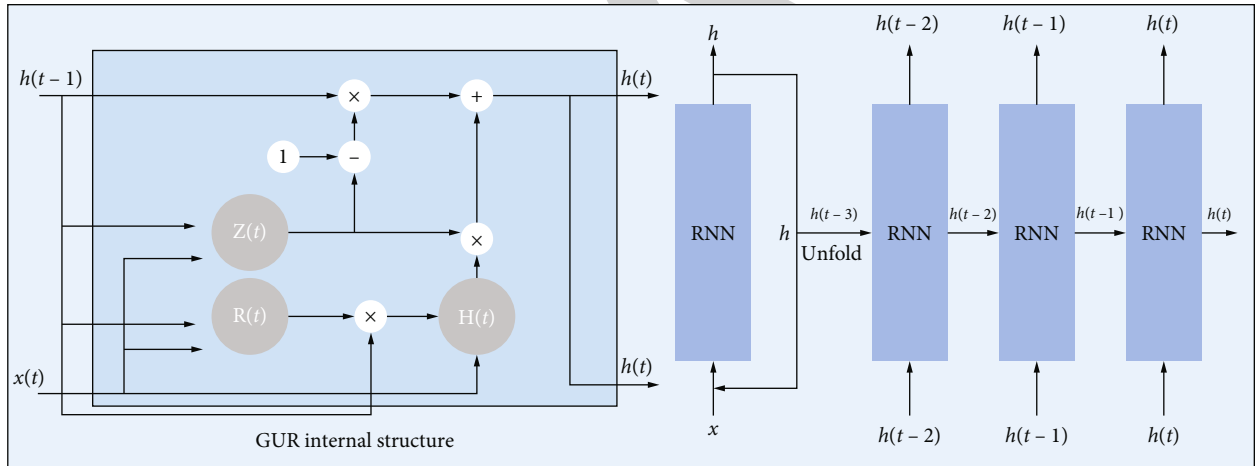


FIGURE 2: Internal structure and expanded structure diagram of the cyclic element algorithm.

In the formula, $|D|$ represents the length of the sample set, which can express the correlation strength of association rules. The credibility of association rules in the object set is expressed as

$$\text{confident}(X \implies Y) = \frac{|T; X \cup Y \subseteq T, T \in D|}{|T; X \subseteq T, T \in D|}. \quad (6)$$

When a key sample is specified, the function of data processing is to find the calculation result with the highest reliability. Therefore, we can also divide the experimental data into two parts for interpretation. The first part is to find the minimum confidence position of all data in the set and mark this part as the preferred area. The second part is to use the rules generated by the preferred region to find all

nonempty sets for each sample subset. If the following formula is satisfied, the calculation result is the best credible:

$$\frac{\sup \text{port}(M)}{\sup \text{port}(m) \geq \min \text{conf}}. \quad (7)$$

The generated rule tree can also be represented by the decision tree algorithm. The decision tree algorithm with gradient rise is used to optimize the performance of the system and further improve the effectiveness of calculation speed. This improved algorithm can realize the processing of parallel computing in the structure of the time series prediction model and establish relevant branch and leaf nodes for exclusion calculation in the part with scarce data.

The sample object function is expressed as

$$Obj(\Psi) = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^g \Omega(f_k), \quad (8)$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda w^2 \quad (9)$$

In the formula, \hat{y}_i represents the predicted value of data, y_i represents the actual value, and l represents the loss change in the calculation process. The child nodes in each tree structure can be expanded and calculated by mathematical formula:

$$L^{(t)} = \sum_{i=1}^n [l(y_i, \hat{y}_i) + f_t(x_i)] + \Omega(f_t) \quad (10)$$

Make a second-order expansion of the above formula:

$$L^{(t)} \cong \sum_{i=1}^n \left[l(y_i, \hat{y}_i^{(t1)}) + g_i f_t(x_i) + \frac{1}{2} g_i^2 f_t^2(x_i) \right] + \Omega(f_t), \quad (11)$$

which is further simplified as

$$g_i = \partial_y^{(t-1)} l(y_i, \hat{y}_i^{(t1)}) \quad (12)$$

Finally, we predict the actual results based on the time series prediction theory. Assuming that there are relevant rules for several influencing factors, the node function responsible for the input data is as follows:

$$O_1^i = \mu_{A_i}(x_1), \quad i = 1, 2. \quad (13)$$

Multiply the predicted results:

$$W_i = \mu_{A_i}(x_1) \mu_{B_i}(x_2), \quad i = 1, 2. \quad (14)$$

Finally, all rules are output uniformly to simulate the prediction data:

$$y = \frac{\sum_i W_i' f_i}{\sum_i w_i}, \quad i = 1, 2. \quad (15)$$

We compare the experimental data from the processing speed of the traditional prediction method, adding cyclic neural network and association rule decision tree algorithm, as shown in Figure 3.

It can be seen from Figure 3 that the calculation speed of the traditional method is significantly reduced in the face of complex space-time conditions. Finally, the prediction model using cyclic network combined with the decision tree algorithm can simplify the computational complexity and improve the computational speed in the overall data processing. Next, according to the above calculation process, we can get a combined time series prediction model based on neural network

algorithm and related rules. The internal structure and basic structure of the model are shown in Figure 4.

It can be seen from Figure 4 that after the rules are obtained by the combined prediction model, the spatiotemporal correlation is first established to arrange the relationship and influence of the areas near the sample. Then, the number of units is obtained in the input layer and the competition calculation layer, the association rule algorithm is used to normalize, and finally, the experimental prediction data is obtained. In the environmental construction, cooling and warming instruments are also prepared. Before the experiment, the number and distribution position of animal intestinal flora are recorded in detail. The experimental time is three days, and the total test perimeter is 72 hours. The time consumed by replacing the power supply is not included in the cycle. Finally, more than 17000 effective data were collected as experimental samples. It includes indoor temperature, oxygen, relative humidity, carbon dioxide, and number of intestinal flora in animal samples. Upload the data to the prediction model to further explore the state of animal intestinal flora within three days.

3.2. Analysis of the Relationship between Temperature and Animal Intestinal Flora Based on Thermometer Imaging and Sample Control.

The climate disorder is largely caused by human beings. The energy consumption and pollution emissions in people's daily life have an impact on the law of temperature change. The concentration of carbon dioxide in the air is increasing, and the greenhouse effect and sudden cooling affect all fields. The sudden rise or fall of temperature does great and irreversible harm to human beings and even biological groups. Many wildlife researchers have found that with the change of temperature, the warm season and frequent precipitation season have migrated, and the growth law of many animals has been in an uncoordinated state. At the same time, temperature also affects the reproduction and alternation of flora in human body and animals. Among them, intestinal flora can protect human gastrointestinal absorption, promote digestion, and improve people's immunity. We monitored and studied the number, structure, and state of intestinal flora under extreme temperature changes. At first, the culture method was used to set the control conditions for the number of microorganisms, so as to obtain the numerical changes. However, most microorganisms cannot be effectively analyzed, resulting in the final results are not convincing. Subsequently, it was found that using the time series prediction model to build the experimental environment, using temperature imaging to judge the changes of animal body temperature, and analyzing the structure of intestinal flora according to flux sequencing technology had a significant effect. Therefore, we only need to design different animal control groups to expand the credibility of the research results.

In this experiment, 30 female experimental rats were selected, with an average weight of about 200 grams and an error of no more than 5 grams. The equipment selects infrared temperature imaging instrument and uploads it to the computer for image comparison. The temperature changes of stomach and intestine of experimental rats were

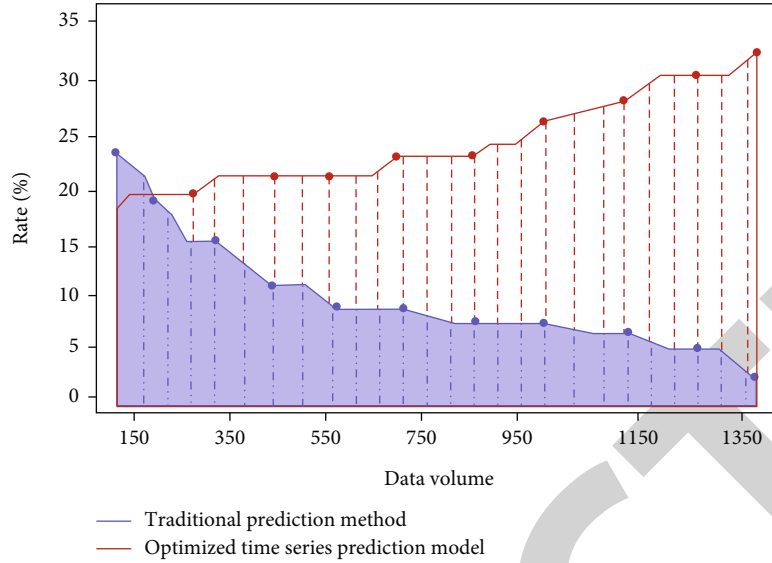


FIGURE 3: Comparison of processing speed between traditional prediction method and optimized time series prediction model.

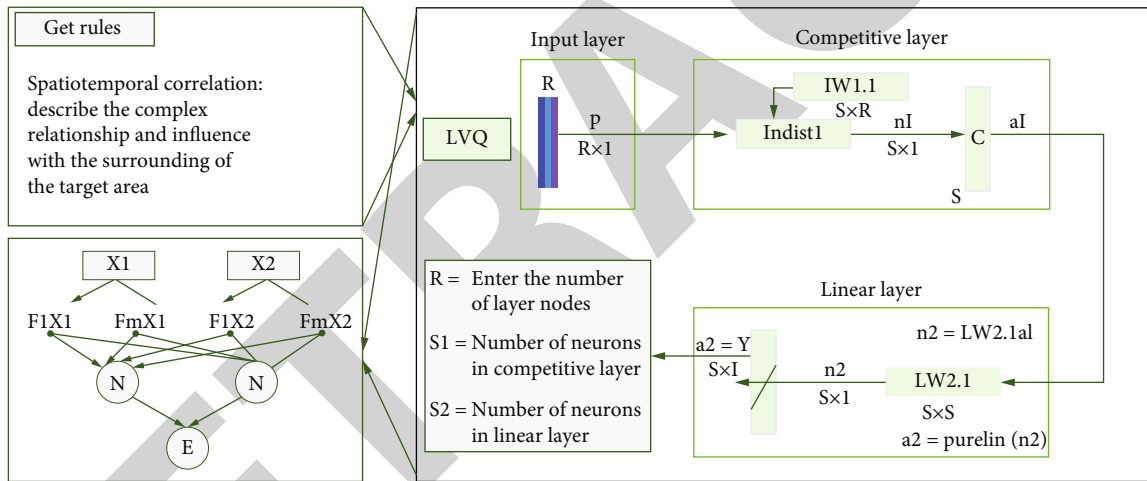


FIGURE 4: Internal structure and basic structure of the model.

recorded every hour according to the experimental cycle. The rats were divided into warm water group and cold water group. The experimental group was perfused with warm water every three hours, and the control group was perfused with cold water. Then, the change of the average body weight of the experimental rats was recorded in a cycle of three days, as shown in Figure 5.

As shown in Figure 5, the body weight of the experimental rats in the warm water experimental group did not change significantly after one cycle, while the average body weight of the experimental rats in the cold water control group decreased significantly, and the body fat rate increased slowly. From the observation of behavioral activity area, it is known that the rats in the cold water group have a small amount of activity and often gather for heating. Next, the spotted lizard samples were numbered and recorded in the same environment. Measure the length, weight, body fat rate, and other data of each object. First, maintain constant

temperature for two days and nights and then conduct temperature control experiment. The experimental group was in a constant temperature state and maintained the normal activities of the spotted lizard under the maintenance conditions. The warm group was heated with a heating lamp and set as the control group. During this period, the temperature was monitored for 24 hours with one week as the experimental cycle. The excreta of the lizard was collected and analyzed daily. The relevant results were obtained according to the temperature of the excretion place of the spotted lizard and the changes of intestinal flora and immune ability collected from the excreta. After flux sequencing of all samples, 546178 pieces of data were obtained, and the sequence information with a similarity of more than 97% was classified and processed at the same time. Count the types of intestinal microflora microorganisms in the experimental samples and annotate all types into the ranking to obtain the following statistical chart.

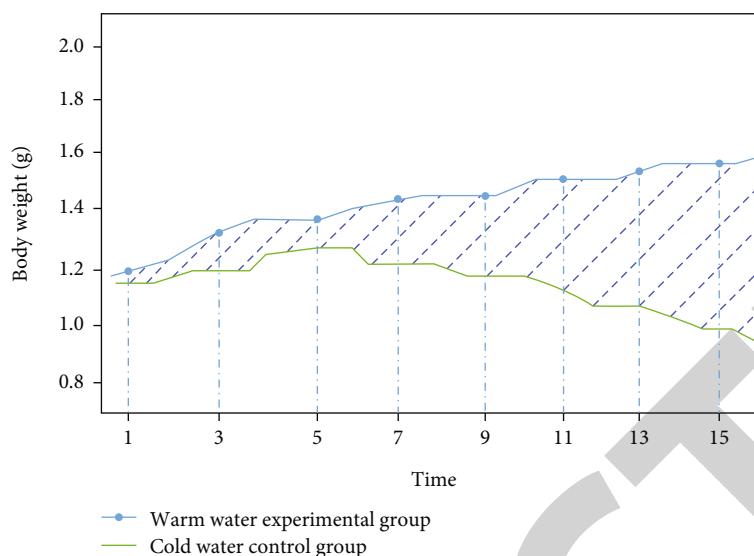


FIGURE 5: Comparison diagram of average weight change of experimental rats.

It can be seen from Figure 6 that in the six boundary method of biological classification, we only use the first five, the horizontal represents the number of each spotted lizard, and the vertical represents the biological classification attribute of the flora. The experimental subjects of spotted lizard showed impatience after the continuous increase of temperature, and the number of activities changed from initial increase to decrease. In the subsequent capture of experimental subjects, it was found that the spotted lizard in the warming control group had weak resistance and low spirit. There is also a surplus of food put in a fixed amount every day. The experimental results further showed that temperature had an effect on the number, structure, and distribution area of intestinal flora.

4. Analysis of Research Results on the Construction of Temperature Environment and Its Impact on Intestinal Flora Imbalance

The construction of the experimental environment adopts the plane roof design, the floor adopts the nonventilated full seam type, and the ventilation port is the mechanical air supply equipment. The wireless monitoring system is used as a data acquisition tool. Record the quantity, number, body shape, weight, and other data of each experimental object in detail. There are several abnormal information in the time series prediction model. It may be missing during the collection process. This includes data loss caused by sudden change of power supply, equipment instability, and other factors. The anomaly detection problem of time series is usually expressed as outliers relative to some standard signals or common signals. Although there are many exception types, we only focus on the most important types from the business perspective, such as unexpected peak, decline, trend change, and grade conversion. Generally, the anomaly detection algorithm should mark each time point as abnormal/non abnormal. Or predict the signal of a certain point and mea-

sure whether the difference between the real value of this point and the predicted value is large enough, so as to treat it as an anomaly. Using the latter method, we can get a visual confidence interval, which is helpful to understand why exceptions occur and verify. We know that this information is invalid; so, we use mathematical interpolation to predict and fill the abnormal results. According to the characteristics of temperature time series, the effectiveness of time series is detected by the cyclic unit neural network algorithm. Firstly, the mathematical statistics method is used to detect the running stability of the sequence, and the neural network model is added to predict the temperature change law to select the final result. We compare the average error index of the actual observation value and the observation value before and after the optimization of the time series prediction model, as shown in Figure 7.

It can be seen from Figure 7 that with the increase of prediction time, the optimized time series model can basically match the actual observation values with less error. Therefore, in the experimental temperature prediction, the time series prediction model optimized by the neural network algorithm can improve the effectiveness of the calculation results. The data collected from the experimental rats and lizards were stored in their respective databases. The sequencing screening method is used to check one by one from top to bottom. The experimental rats are selected as an example, and the infrared imaging of animal body temperature is dynamically represented by multimedia animation software, as shown in Figure 8.

It can be seen from Figure 8 that compared with the experiment on the first day, the body surface temperature of the experimental mouse control group showed obvious blue at the end of monitoring. In the normal temperature group, the temperature imaging is normal, indicating that there is no error in the experimental process, which is consistent with the actual expected results. Next, the number of intestinal flora of the sample object is detected. The flora abundance index and average size index were used to judge

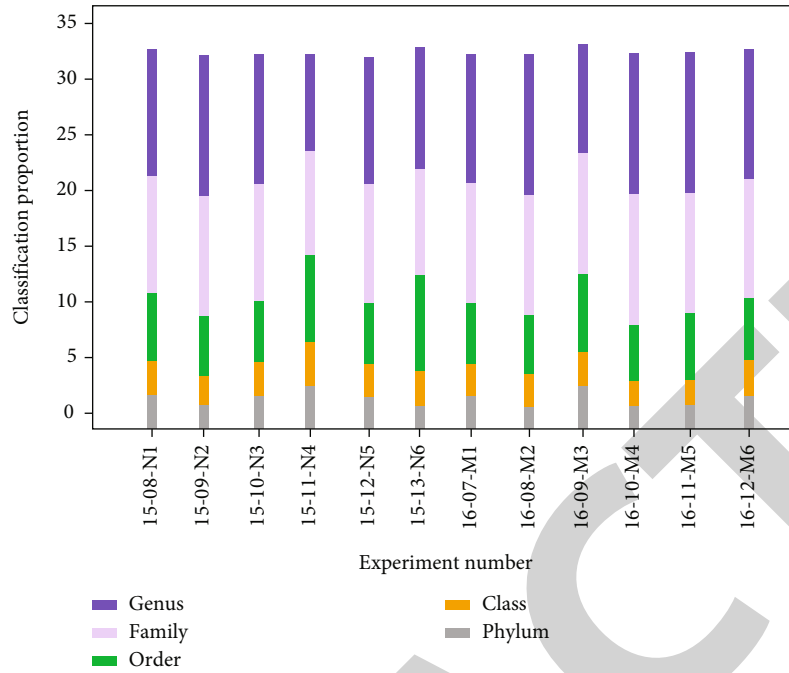


FIGURE 6: Annotation to statistics in sorting.

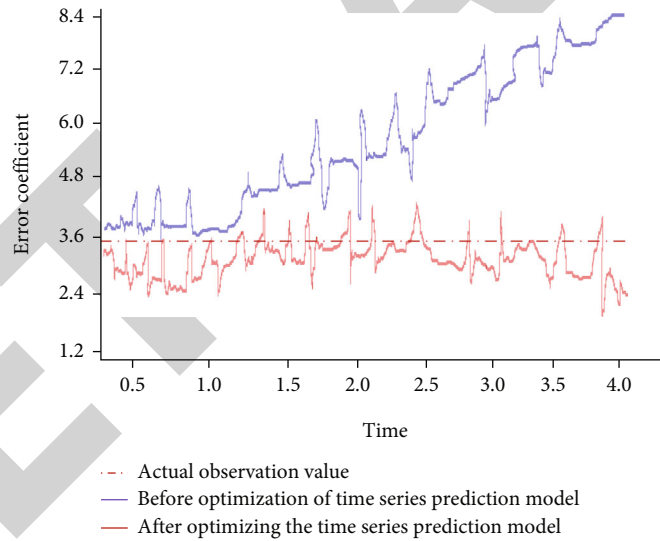


FIGURE 7: Variation of average error between actual observation value and time series prediction model before and after optimization.

whether there was imbalance. The excretion test samples of two animals under normal temperature can be classified into 32 to 35 families and genera, and the low-temperature control group and high-temperature control group can only be allocated to 30 species. According to the statistics of microbial diversity, under normal conditions, the mouse flora is mainly composed of 29.6% *Corynebacterium*, 46.4% *Bacteroides*, and 12.3% *dauricum*. After cooling with cold water in the experimental cycle, the number of *Corynebacterium*, *Bacteroides*, and *dauricum* in mice decreased to 7.7%, 13.6%, and 80%, respectively. Under normal conditions, the flora of lizard is mainly composed of 46.9% *Bacter-*

oides, 20% *Proteus*, and other flora. After warming, the content of *Bacteroides* decreased to 43%, and *Proteus* increased to 25%. The intestinal tract of mice inhabits a complex and dynamic microbial community, in which bacterial groups dominate. The intestinal normal indigenous flora of adult mice is very similar to that of mammals at the high classification level. The intestinal microorganisms of mice are mainly facultative anaerobes, including *Streptococcus*, *Lactobacillus*, and *Escherichia coli*. The microbial flora of mice is composed of a mixture of microorganisms from different partitions of the gastrointestinal tract. There are four dominant microbial system groups in feces, two of

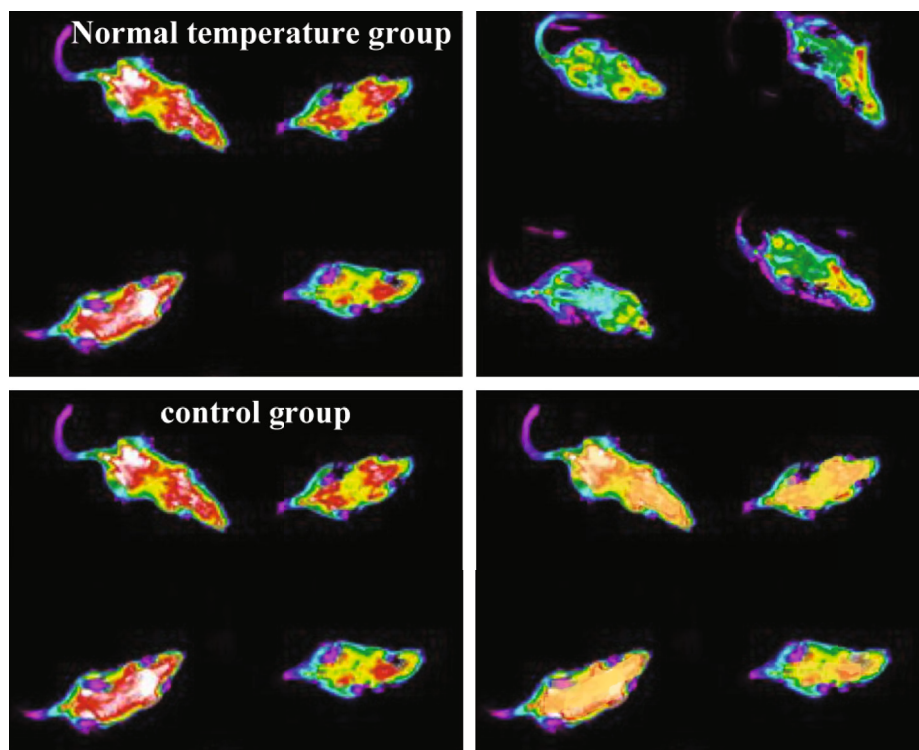


FIGURE 8: Infrared imaging of body temperature of experimental animals.

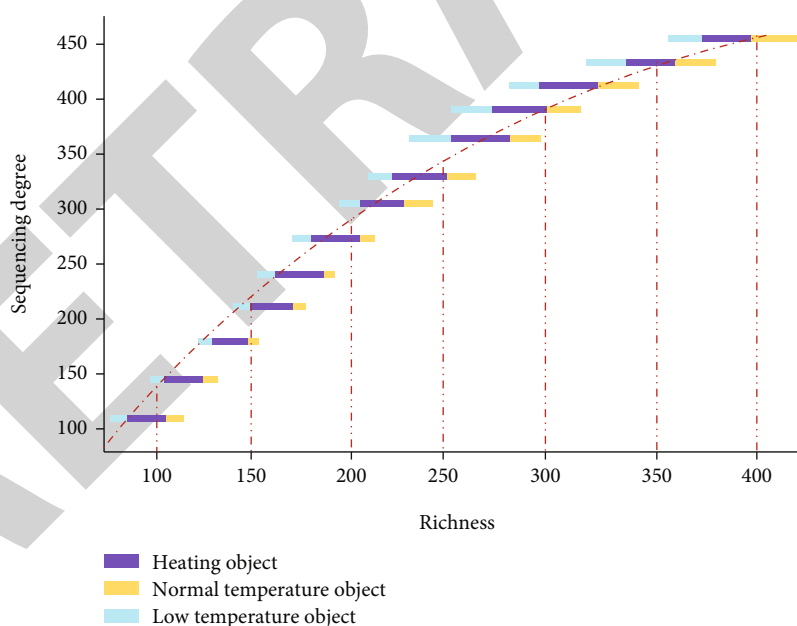


FIGURE 9: Richness changes of diversity sequences of experimental animals.

which are related to intestinal flora. Fecal sample analysis can obtain the diversity of most microorganisms in the cecum. The composition of microbial flora in fecal samples and cecal contents of mice is similar in nature but different in quantity. Cecal contents can reflect cecal microorganisms of chicken, while fecal samples cannot. The calculation results show that the number and structure of animal intestinal flora are relatively stable at room temperature, in which

Bacteroides and actinomycetes account for a high proportion. It shows that the bacteria detected above are the core flora in the intestine. With the change of temperature into supercooled and overheated state, the intestinal microbial content of the experimental subjects changed significantly. Among them, the soft membrane flora of mice and the deformable flora of spotted lizard all increased too much. It can be directly determined that the intestinal flora of

animals is in a state of imbalance. In order to judge whether the sequence richness change of the diversity contained in the tested object can meet the research needs, we use sparse curve for detection. The optimized data sequences of mice and lizards are randomly selected. The abscissa represents the degree of sequencing, and the ordinate represents the richness, as shown in Figure 9.

As can be seen from Figure 9, the objects in the heating experiment are marked with red, the normal temperature group is marked with yellow, and the low temperature group is marked with blue. With the increase of the sequencing degree of sample data, the similarity of each group of data is very high, showing a gentle development trend, indicating that the range and quantity of experimental objects meet the detection standards. Among them, the richness index in the normal temperature environment is relatively large, which can effectively explain that the richness of microbial content of intestinal flora in the experimental population is high, which meets the conditions of routine experimental research. The experimental subjects were further analyzed, and the animals were distinguished according to gender. The number of Bacteroidetes in the intestinal flora of male animals was higher than that of female animals, but this phenomenon disappeared directly after physical ligation. In addition, the abundance of intestinal flora of female animals is higher than that of male animals, and the metabolic level is also faster.

5. Conclusion

The human intestine contains abundant colony groups, which participate in normal physiological activities and play an important role in diet, exercise, and metabolism. Intestinal flora can promote the human body's absorption of food nutrition, help the metabolic cycle, improve immune function, and improve health indicators. Because people's eating habits are gradually irregular and often eat in the state of temperature imbalance, they often have symptoms of gastrointestinal discomfort. In this paper, the time series prediction model is used to study the effect of temperature on the imbalance of intestinal flora. Firstly, neural network algorithm and association rule algorithm are used to optimize the traditional time series prediction model and change the internal complex structure of the model, so as to simplify the prediction process and improve the operation speed. Secondly, the temperature infrared imaging instrument was used to monitor the body surface temperature of experimental rats and lizards, and the experimental group and control group were set up in the simulated experimental environment. Explore the changes of animal behavior and the number of intestinal flora in excreta under normal temperature, low temperature, and high temperature. Flux sequencing was used to detect the richness index of intestinal flora and judge the structural components of intestinal flora at various temperatures. Finally, the specific effects of temperature changes on the diversity of animal intestinal flora were studied under the condition of microbial richness. Under the condition of high temperature, the behavior of *Corynebacterium* increased first and then decreased. The results showed

that the appetite of animals decreased under the condition of high temperature. It shows that the change of temperature has an obvious effect on the number and internal structure of intestinal flora. The human environment is similar to the environment of the experimental object, which can directly affect the experimental results in the study of human intestinal flora, and has a certain credibility. However, the research still has some limitations. Due to the lack of simulation verification, this paper only discusses the simulation results of mouse experiments. The human environment is similar to the environment of the experimental object, which lacks certain scientificity. Therefore, further research and analysis are needed 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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] X. Donghui, "Rul time series prediction model based on kecanarx," *Battery*, vol. 51, no. 6, pp. 582–586, 2021.
- [2] Z. Xiaofang, X. Wang, H. Wang, K. Lijing, G. Gao, and Z. Liying, "Human intestinal flora metabolizes amino acids to produce short chain fatty acids and its effect on cell permeability," *Chinese Journal of food*, vol. 21, no. 7, pp. 60–67, 2021.
- [3] Z. Cheng, N. Tingli, X. Wang et al., "Changes of intragastric temperature gradient and its effect on intestinal flora in experimental rats," *Chinese Journal of traditional Chinese medicine*, vol. 34, no. 7, pp. 3251–3253, 2019.
- [4] X. Yang, Y. Ding, Y. Qing, and K. Qiaoqin, "Infrared time series prediction based on weighted combination model," *Computer engineering and design*, vol. 42, no. 11, pp. 3076–3083, 2021.
- [5] S. Ping, W. Jian, Y. Huajian, L. Zenghui, and C. Shijiang, "Effect of sudden temperature change on intestinal flora of Japanese medical leech," *Journal of Microbiology, China*, vol. 752, 2021.
- [6] X. Wang, L. Xin, and Y. Xiao, "Adaptive frequency domain model for multivariate time series prediction," *Computer science*, vol. 48, no. S2, pp. 204–210, 2021.
- [7] Q. Yu and T. Xian, "Nonlinear time series prediction algorithm based on combined model," *Computer Engineering and Science*, vol. 43, no. 10, pp. 1817–1825, 2021.
- [8] Q. Guo, L. Dong, Z. Lei, and W. Chuyuan, "Time series prediction model based on time point process," *Computer engineering and science*, vol. 43, no. 7, pp. 1299–1307, 2021.
- [9] W. Xiaofeng, L. Xiaoyan, and J. Yanan, "Integrated combined forecasting model based on time series model selection," *Statistics and decision making*, vol. 37, no. 9, pp. 5–8, 2021.
- [10] F. Wang, W. Xiawen, J. Qian et al., "Effect of temperature on intestinal flora structure of *Procambarus clarkii*," *Jiangsu Agricultural Journal*, vol. 38, no. 1, pp. 157–164, 2022.

Retraction

Retracted: Application of Optimized Support Vector Machine Model in Tax Forecasting System

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xin, "Application of Optimized Support Vector Machine Model in Tax Forecasting System," *Journal of Function Spaces*, vol. 2022, Article ID 6212579, 10 pages, 2022.

Research Article

Application of Optimized Support Vector Machine Model in Tax Forecasting System

Yu Xin 

Emerging Economic Formats Research Institute, Shandong Management University, Shandong, Jinan 250357, China

Correspondence should be addressed to Yu Xin; 14438119970177@sdmu.edu.cn

Received 6 June 2022; Revised 27 July 2022; Accepted 5 August 2022; Published 22 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Yu Xin. 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.

Tax forecast has an important impact on financial budget and tax plan. The amount of tax data is greatly increased, the difficulty of tax forecast is improved, and the accuracy is always difficult to keep up with the development demand. Analyze the application of optimized support vector machine model in tax prediction system. Based on the simple analysis of the research situation of tax prediction and the research status of data mining algorithm in tax data classification and prediction, this paper constructs a tax prediction model. Aiming at the problem of too many influencing factors of tax prediction, this paper puts forward the use of principal component analysis to extract the main factors, reduce the dimension of tax data, and reduce the difficulty of analysis. Support vector machine is used to realize tax prediction, aiming at the problem of parameter optimization proposed to optimize the parameters. Finally, the prediction accuracy is evaluated by comparing the error between tax prediction value and real value. The results show that the algorithm used in this paper can optimize the parameters of support vector machine. The tax prediction results show that the predicted value is similar to the real value curve. After grid search optimization, the introduction of principal component analysis reduces the redundancy and improves the prediction accuracy.

1. Introduction

Tax forecast analysis is also called “tax trend analysis.” It refers to an analysis that predicts the trend of future tax revenue and provides decision-making services for leaders according to the mastered historical laws of taxation, using the continuously reflected relevant materials and data, combined with the current economic tax sources and investigation and research materials. This method is applicable to tax planning analysis, economic tax source investigation analysis, and tax accounting analysis. It is an indispensable analytical method for formulating and inspecting tax policies and measures. Especially before the changes of national tax policies, such as the new or suspended collection of a certain type of tax, as well as the provisions on tax increase, tax reduction or exemption, and other major tax measures, the impact of tax revenue should be predicted and analyzed. In the information age, tax analysis and prediction has become an essential content of tax management [1]. Tax data is a kind of time series data, which is not only affected by the economic environment, but also affected by cultural back-

ground, political background, and other factors. These data are nonlinear and unstable [2]. Tax forecasting can timely discover the tax change trend and collection; has an important impact on strengthening the organization’s income, resource allocation, and management decision-making; and can help relevant personnel formulate tax policies. Therefore, the accuracy of tax forecasting is very important [3]. At present, there are many researches on tax prediction, from linear analysis and regression analysis to various data mining algorithms, but the accuracy is not high [4].

Based on this background, this paper studies the application of optimized support vector machine in tax forecasting system. In Section 1, we briefly analyze the current tax forecasting research and briefly introduce the chapter arrangement of this research; in Section 2, we mainly introduce the research of support vector machine and prediction algorithm and summarize the shortcomings of the current research. In Section 3, we construct the tax prediction. Aiming at the problem of many tax influencing factors, the principal component analysis method is introduced to analyze the correlation of the main tax factors and extract the main

indicators. In Section 4, we simulate and analyze the support vector machine tax prediction model constructed in this paper. After preprocessing, the tax data is brought into the tax prediction model and compares and analyzes the tax data before and after parameter optimization, and the accuracy of tax prediction is significantly improved after the parameter optimization of support vector machine. In the last section, we summarize the whole paper.

The innovation of this paper is that the support vector machine method is used in tax prediction, and the principal component analysis method is introduced, which reduces the influencing factors of tax and improves the prediction accuracy. Aiming at the problems that are used to optimize its parameters, the regularization parameters and radial basis kernel function parameters are applied to tax prediction to improve the accuracy of data analysis.

2. State of the Art

The importance of tax forecasting has been recognized. In terms of relevant research, it has also developed from early qualitative research to current quantitative analysis. Tax forecasting is more scientific and normative [5]. Aprilia and Agustiani adopt K-means clustering analysis algorithm in account data classification, which improves the effect of clustering analysis [6]. In the analysis of tax data, Sun et al. proposed a tax dynamic clustering algorithm, which automatically determines the number of clusters and modifies the cluster center according to the frequency of internal attribute values. TDCA algorithm is a tax dynamic clustering based on weak convergence sequence coefficient judgment function and tax domain knowledge, and the clustering effect is good [7]. In the research of tax risk assessment, Didimo et al. proposed the Maldivo method. Through the information diffusion strategy, it is possible to expand the collection of risky taxpayers and display the output to analysts through the network visualization system [8]. In the analysis of financial data, Li constructed financial related indicators and financial report recognition model. The accuracy, precision, recall, and F value show that the performance of the improved BP neural network has been improved [9]. Mwanza and Phiri used intelligent mining algorithm in the research of tax data to realize the outlier algorithm of fraud detection, continuous monitoring based on distance and outlier query based on distance, which improved the accuracy of abnormal data analysis [10]. Miller used KH coder's data mining technology to model individual tax behavior and used unsupervised machine learning text mining and modeling technology to help conduct large-scale analysis of tax behavior methods and find the problems of avoidance and tax evasion in time [11]. Battiston et al. used machine learning algorithm for tax analysis in their research and proposed a loss function to analyze the difference between ideal tax and display [12]. In his research, Hao et al. proposed an algorithm for financial risk prevention and carried out special data preprocessing on convolutional neural network. Combined with the requirements of digital inclusive financial risk method, he constructed a digital inclusive financial risk prevention model

[13–15], so as to timely find financial abnormalities and carry out risk early warning.

To sum up, it can be seen that there are many researches on tax prediction at home and abroad, but different algorithms have their own limitations. Linear regression analysis can only analyze the linear relationship, ignoring the nonlinear relationship of tax influencing factors. The structure of neural network algorithm itself is complex and requires high sample size, so the accuracy of tax prediction is not very high. Support vector machine has its own optimization in the application of prediction, can get the global optimal solution, the data can be analyzed in high-dimensional space, the complexity is not high, and has certain advantages in tax prediction. However, this algorithm needs to determine the regularization parameters and radial basis function parameters, and other methods need to be adopted for parameter optimization in application.

3. Methodology

3.1. Design of Tax Forecasting System. Tax prediction uses various impact indicators of tax revenue to analyze and introduces prediction theory, data mining algorithm, and model to predict. Affected by the economic environment, social environment, and science and technology, the tax changes greatly, and the formed data has a large amount of redundant data, which has great correlation. The traditional economic prosperity index cannot fully reflect the real situation of economic development. In response to these problems, tax big data can give full play to the advantages of complete and dynamic tax coverage, find more accurate and sensitive synchronization indicators and leading indicators in the big data set (for example, take the value-added tax as one of the synchronization indicators and the value-added tax on imported goods as one of the leading indicators), and further adopt the machine learning method to compile the composite index. The economic prosperity index based on tax big data can meet the technical requirements of the above two key steps, and it can more accurately reflect the real situation of economic development to a certain extent. Compared with the traditional economic prosperity index, the economic prosperity index based on tax big data shows its progress in two aspects. First, at the level of computing methods, it has become a trend to apply machine learning methods to study economic problems, such as ridge regression and lasso in machine learning.

Considering the multidimensional and nonlinear characteristics of tax data, it is difficult to find the change characteristics and trends of tax data by a single-factor analysis. Therefore, it is necessary to find the main influencing factors from a large number of data.

In this paper, the main factors affecting the prediction, eliminate the redundancy between various factors, and then establish a model to predict tax. This method makes use of the advantages of nonlinear function prediction and the ability of principal component extraction. The prediction accuracy can be improved. The prediction model is shown in Figure 1.

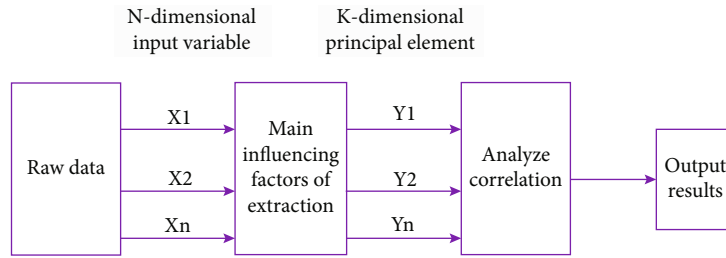


FIGURE 1: Design of tax forecasting model.

Reduce the scalar index, and retain the amount of information of the index. This algorithm rotates the n -dimensional spatial coordinates and processes the results without changing the sample data. The principal components obtained are uncorrelated [16]. The introduction of this algorithm can reduce the dimension, eliminate redundant information, and combine it to simplify the architecture and improve the prediction performance. Assuming that the p -dimensional random vector is $X = (X_1, X_2, \dots, X_p)^T$; these samples form a sample matrix, and the samples are transformed. The formula is

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, \quad (1)$$

where i and j are positive integers. After processing, standardized samples are obtained. Calculate the coefficient matrix for standardized samples, and the formula is

$$R = (r_{ij})_{p \times p} = \frac{Z^T Z}{n-1}, \quad (2)$$

where i and p are positive integers, and the value range is $(1, p)$. Calculate the characteristic equation $|R - \lambda I_p| = 0$ of the sample correlation matrix to obtain the characteristic root and determine the value of m according to the following formula:

$$\frac{\lim_{m \rightarrow \infty} \sum_{j=1}^m \lambda_j}{\lim_{p \rightarrow \infty} \sum_{j=1}^p \lambda_j} \geq 0.95. \quad (3)$$

After calculation, the information utilization rate reaches more than 95%. For each λ_j , j value, calculate the equation $Rb = \lambda_j b$ to obtain the eigenvector. The variables become principal component factors after orthogonalization. The principal component factors obtained from the comprehensive evaluation are weighted to obtain the final evaluation result value.

3.2. Optimized Support Vector Machine Model. Among the common machine learning algorithms. This algorithm is based on statistics and can be used for classification, as well as regression analysis. The core lies in the minimization of structural risk, which needs to adjust the confidence range

and empirical value and improve the generalization ability through the proportion between the two [17–19].

When using support vector machine for prediction analysis, it is assumed that there is a training sample set, expressed by $\{x_i, y_i\}$, and the mapping condition from input to output is calculated to meet $f(x) = y$. The regression function is established by using support vector machine algorithm, and the formula is

$$y = f(x) = w \times x + b. \quad (4)$$

The nonlinear mapping is expressed as

$$y = f(x) = w^T \times \phi(x_i) + b. \quad (5)$$

The objective function minimization is used to determine the regression function. In the application of support vector machine, in order to avoid the problem of space mapping with different dimensions, it is necessary to introduce kernel function and turn it into a linear analysis problem. Kernel function is to calculate the form of inner product when the data is indeed mapped. For linearly nonseparable data, the data can be mapped to another space to make the data perfect or almost perfect linearly separable. The same is true when we need to map data to a high-dimensional space. Instead of providing an accurate mapping to SVM, we provide the dot product of pairing all points in the mapping space. In the specific application, it is assumed that the data in the low-dimensional space is not linearly analyzable, and the kernel function is used to map these data to Gao Wei to realize the linear analysis of sample data [20]. For support vector machine algorithm, the change of kernel function parameters directly affects the performance, and the influence degree of different kernel functions is also different. We need to choose a reasonable kernel function. At present, there are many kinds of kernel functions. Generally, functions that meet Mercer rules belong to kernel functions. Assuming that a subset of the input space is represented by X and the feature space is represented by H , there is a mapping $\varphi(x)$ from space to feature space, so that all x belong to the input space. All functions that can satisfy $K(x, z) = \varphi(x)\varphi(z)$ are kernel functions. Polynomial kernel functions are

$$k(x_i, x_j) = (x_i x_j + 1)^d. \quad (6)$$

The sigmoid kernel function is

$$k(x_i, x_j) = \tanh [b(x_i x_j) + c]. \quad (7)$$

RBF kernel function is expressed as

$$K(x, y) = \exp \left(-\frac{\|x - y\|^2}{\delta^2} \right). \quad (8)$$

In the support vector machine algorithm, RBF kernel function is selected. This kernel function has the characteristics of wide convergence domain and can handle a variety of samples, which is widely used [21–23]. RBF kernel function only involves one parameter, so its performance is more stable and easy to apply. In the analysis of this paper, assuming $g = 1/\delta^2$, finding the best parameter is to find the best g . The change of G value directly determines the mapping function, which affects the complexity of spatial distribution.

After RBF kernel function calculation, the matrix can be transformed into

$$\begin{bmatrix} A, E \\ E^T, 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix}, \quad (9)$$

where $A = k + V$ and e represent a matrix with all elements of 1. SVR regression analysis model can be obtained, and the formula is

$$f(x) = \lim_{N \rightarrow \infty} \sum_{i=1}^N a_i k(x, x_i) + b. \quad (10)$$

It can be seen that after processing, there are only two parameters to be optimized: the regularization. The function of the regularization parameter C is to adjust the range of information and reduce the risk. By adjusting the ratio of the two, the generalization ability of the learning machine can be adjusted. If the regularization parameter C is too large, it can only reduce the empirical risk. If it is too small, it will increase the empirical risk. δ mainly controls the width of Gaussian function and data range. If δ value is too large, the risk will increase. If δ value is too small, the structural risk will increase. In the optimization analysis of these two parameters, the following algorithm is adopted.

This algorithm is used for real number solution. It has strong universality, simple principle, few parameters, and strong collaborative [24]. This algorithm is based on the foraging of birds. Each individual is regarded as a particle. It is assumed that the position of the particle is represented by (C, g) , the speed of the particle is represented by v_i , and the best position of the particle is represented by p_i . The particle adjusts its position through its current position, neighbor position, and empirical position and realizes the continuous updating of the position through the equation:

$$v_{ij}^{t+1} = wv_{ij}^t + c_1 r_1 (p_{ij}^t - x_{ij}^t) + c_2 r_2 (p_{ij}^t - x_{ij}^t), \quad (11)$$

where w represents the inertia factor, which is a positive number, c represents the acceleration constant, and r is the random transformation number, ranging from 0 to 1.

When optimizing parameters by least square method, the optimization performance is closely related to regularization parameters and radial basis function kernel function [25–27]. When using particle swarm optimization algorithm, it can adjust its position and rely on its own experience. Therefore, the direction of iteration is stronger, and the global search ability is improved. Support vector machine algorithm is used to regression analyze the training set, find the best parameters, and then bring the training set into the best parameter model [28–30]. In the optimization using the least square method, the fitness parameter is the content that needs to be determined. Considering that each particle reflects a set of parameters, in the particle swarm optimization design, the fitness function is used to change the fitness. The fitness function selects the mean square error function, and the formula is expressed as

$$\text{MSE} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_i)^2, \quad (12)$$

where n represents the training sample set. Through calculation, it can be obtained that the fitness is a positive integer. The smaller this value is, the higher the fitness is. When the fitness can meet the accuracy requirements, the optimal solution is the optimal parameter.

Although the particle swarm optimization algorithm has obvious advantages, it also has some problems. The search accuracy is insufficient to ensure the global optimal solution. It has strong dependence on specific empirical parameters and lacks independence. Therefore, in the research, the genetic algorithm is used to further optimize the parameters, and the regularization parameters and radial basis kernel function optimized by the least square method are genetically coded. The range of regularization parameters is 0–1. Using genetic algorithm to realize parameter optimization, this problem can be simply understood as chromosome, binary coding the parameters, randomly generate SVR parameter values, select the parameters with high fitness according to the survival of the fittest principle, and get better parameters through cross coding. With the further evolution of genetic algorithm, the parameters with the highest fitness can be obtained. The application of the value range of regularization parameters and kernel function parameters is determined first, and then they are binary coded. Then, the mean square error function is used as the fitness function, and the chromosome is decoded to generate the random initial population. Judge whether the population optimization performance meets the optimization requirements. If so, output the optimal parameters and establish the support vector machine model. If not, continue the optimization algorithm and update the population until the requirements are met.

Genetic algorithm has a large search range; it can compare multiple individuals in parallel. It has simple operation and strong expansibility. However, this algorithm is difficult

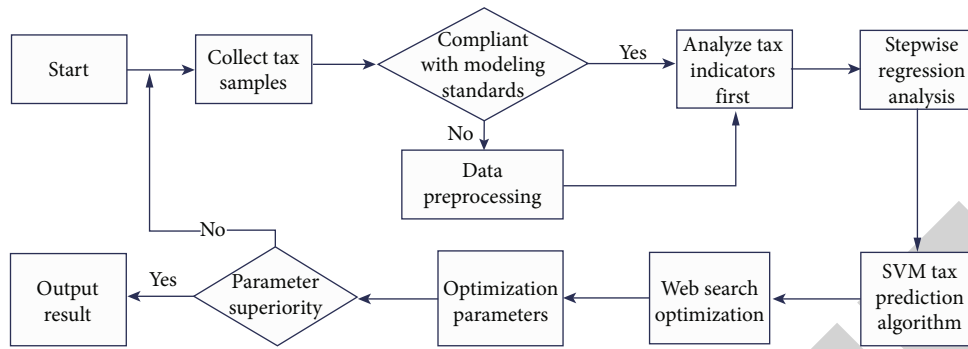


FIGURE 2: Network search optimization prediction process.

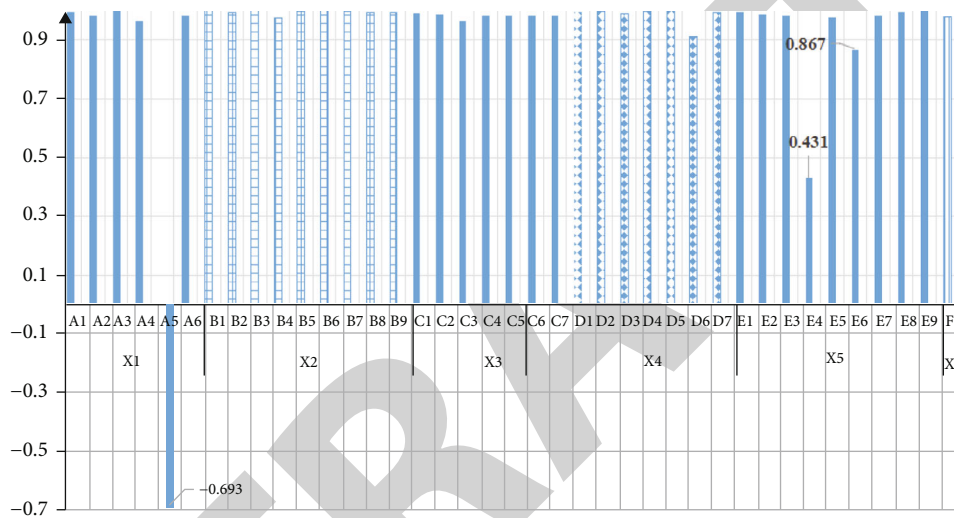


FIGURE 3: Correlation coefficient analysis between tax and secondary indicators.

to implement in programming, and it needs to continue to calculate after obtaining the optimal solution. The selection of parameters will affect the crossover rate and mutation rate. Moreover, the genetic algorithm cannot effectively feedback the network information, and the search speed is slow. Therefore, the method based on grid search is further used for parameter optimization. Grid search algorithm can search all possible values in the range. Compared with particle swarm optimization algorithm, it is more comprehensive and needs less parameters to be optimized. In the application, the network parameters are generated according to the parameter range and step size of the kernel function and the parameter direction, so as to ensure that all parameter sequences can be analyzed and shrink. Determine the training samples and test samples, test them at the same time, output the errors corresponding to all parameters, and judge whether the error results meet the required accuracy. Reset the parameters and step size, carry out the second search process, and continue this step to find the parameters with the highest accuracy. The specific process is shown in Figure 2. The tax index sequence is represented by X, and the C range and search step are set. This algorithm can avoid blindness. After obtaining the optimal structure, the tax forecast can be determined.

4. Result Analysis and Discussion

4.1. Data Source and Preprocessing. In the tax forecast analysis, the influencing factors need to be determined first. At present, in the domestic statistical yearbook, taxes are divided into six categories. Therefore, these six categories of taxes are taken as the primary research indicators in the research, namely value-added tax (x1), business tax (x2), consumption tax (x3), personal income tax (x4), enterprise income tax (x5), and tariff (x6). The vertical coordinate in Figure 3 is the correlation coefficient of tax. The relevant indicators contained in these taxes are used as secondary indicators, and the secondary indicators are coded. The value-added tax (x1) contains 6 secondary indicators, with the GDP code of A1, the total import value code of A2, the industrial added value code of A3, the retail sales code of A4, the proportion code of industrial added value of A5, and the added value code of A6. The units of these secondary indicators are 100 million yuan. There are 9 secondary indicators of business tax. The code of highway freight volume is B1, the code of highway passenger volume is B2, the code of added value of construction industry is B3, and so on. The code of business income of catering owners above the limit is B9. There are 7 secondary indicators of

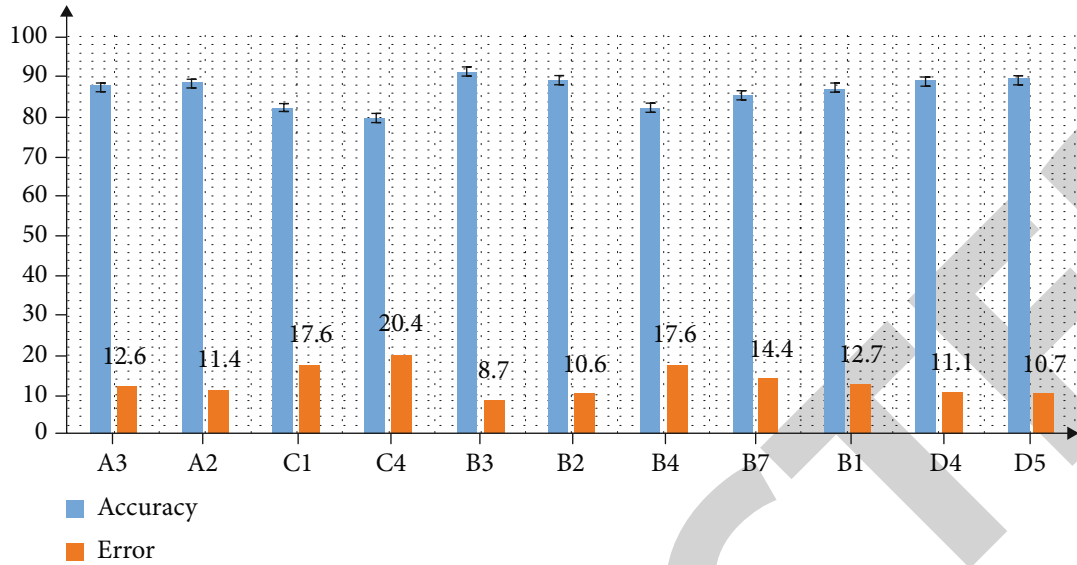


FIGURE 4: Accuracy and error analysis.

consumption tax, represented by *C*, 7 secondary indicators of personal income tax, represented by *D*, 9 secondary indicators of enterprise income tax, represented by *E*, and the secondary indicator of tariff is the total import and export volume, represented by *F1*.

There are often many indicators and insufficient historical data in economic forecasts. Using neural network to predict such a problem is a typical small sample prediction of large-scale neural network. This situation weakens the generalization ability of neural network. In this case, scholars at home and abroad usually divide these indicators into several subsystems according to their relationship to alleviate the problem of excessive network scale. However, the division of subsystems is very complex, which relies too much on the analysis of the operation mechanism of economic system, and cannot fundamentally solve the problem of multi-index and small sample complex system prediction. In fact, these indicators are often relevant. Therefore, it is necessary to reduce the number of indicators (dimensionality reduction) on the premise of minimizing information loss. Principal component analysis, as a dimension reduction processing technique, starts from reducing the number of input nodes of neural network. It can fundamentally reduce the scale of neural network and improve the generalization ability of neural network in multi index small sample problems.

Considering that there are too many influencing factors of tax prediction, in the correlation analysis, set the limit of correlation coefficient as 0.9 and eliminate the indicators with small correlation. Taking the tax revenue data of China Statistical Yearbook as the data sample, the correlation analysis is carried out by using statistical software. Among the six primary indicators, the correlation with tax is 0.992, 0.991, 0.984, 0.985, 0.994, 0.997, and 0.982, respectively. From the correlation analysis results, it can be seen that the primary indicators have a great impact on tax, and the correlation coefficients exceed 0.95. Continue to analyze

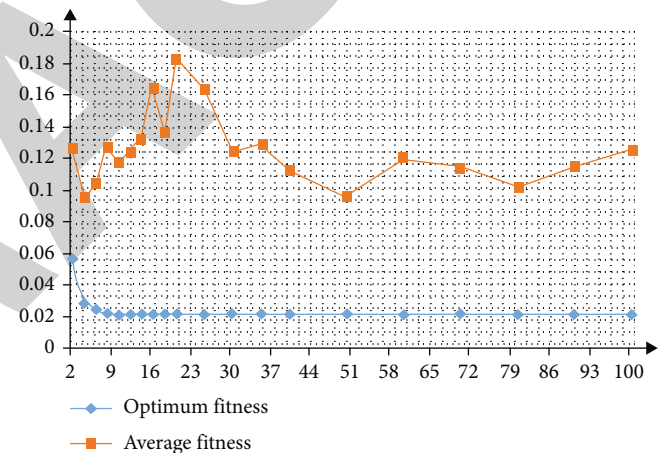


FIGURE 5: Particle swarm optimization simulation analysis.

the correlation between primary indicators and secondary indicators, as shown in Figure 3. The correlation coefficient between VAT and its secondary indicators is quite different. Except for A5 indicator, the correlation coefficient of other indicators is more than 0.98. The correlation coefficient between business tax and its secondary indicators exceeds 0.99, and all indicators will be analyzed in the next step. The correlation coefficients of consumption tax, individual income tax and their secondary indicators are relatively high, all above 0.96. There is a certain difference in the correlation coefficient between enterprise income tax and its secondary indicators. The correlation coefficient of E4 indicator is only 0.431, excluding this indicator. The correlation coefficient between tariff and secondary indicators is 0.976. It can be seen from the data in the figure that most indicators are highly correlated with their respective tariffs, but there are also negative correlation indicators. Eliminate the negative correlation indicators and the secondary indicators with correlation coefficient less than 0.9.

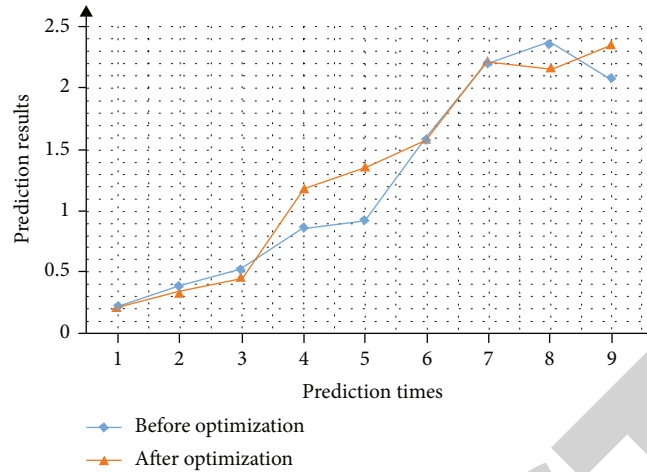


FIGURE 6: Prediction results of particle swarm optimization algorithm.

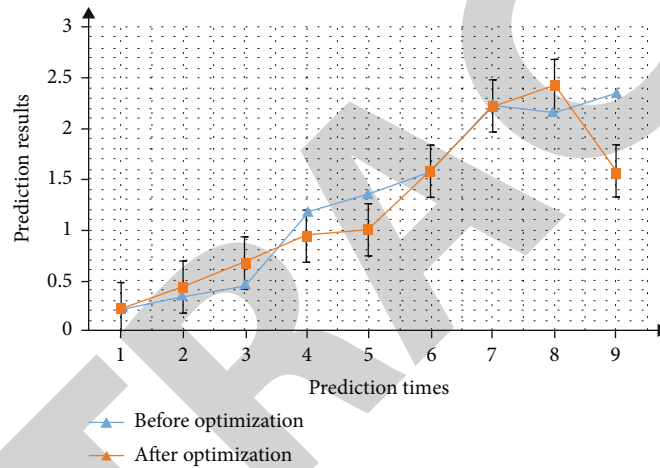


FIGURE 7: Prediction results of genetic algorithm.

Through correlation analysis, excluding the indicators with small correlation coefficient, there are still many remaining tax indicators, so stepwise regression analysis is also needed. Set a dependent variable, analyze the effect of other independent variables on it, and then sort and select the factors with great influence. The original data is collected and processed for simulation analysis. The accuracy and error results are shown in Figure 4. The prediction accuracy has been improved after stepwise regression analysis.

The data dimensions of different indicators are different, so the tax cannot be predicted directly. Therefore, the data need to be standardized. The formula is

$$\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad (13)$$

where x_i represents the data of the indicator column, x_{\max} represents the maximum value of the data of the indicator column, and x_{\min} represents the minimum value. After this formula, the range of data is controlled within 0~1.

When the parameters are optimized by the least square method of particle swarm optimization, the range of param-

eters C and δ^2 is 0~100, with 100 iterations and 20 population evolution. Through experimental simulation, the optimal value of C is 1.032, and the optimal value of δ^2 is 0.01. In order to evaluate the effectiveness of tax forecast, the average absolute percentage error and mean square percentage error are used for analysis. The formulas are as follows:

$$\begin{aligned} \text{MAPE} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \bar{y}_i}{y_i} \right|, \\ \text{MSPE} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sqrt{\sum_{i=1}^n \left[\frac{(y_i - \bar{y}_i)}{y_i} \right]^2}, \end{aligned} \quad (14)$$

where MAPE represents the average absolute percentage error and MSPE represents the mean square percentage error.

4.2. Simulation Analysis of Support Vector Machine Optimization. In the simulation analysis, the regularization parameter range is set to 0.1~100, kernel parameter range is set to 0~10, the cross validation is 5 times, the maximum population iteration is 100 times, and the population size is 20. And the results are shown in Figure 5. And the optimal

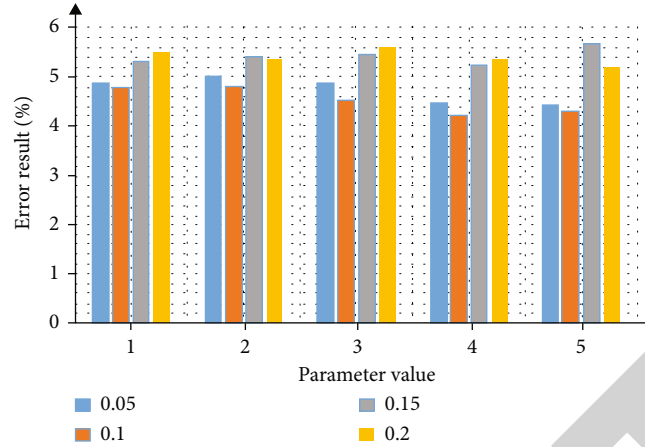


FIGURE 8: Error results under different parameters.

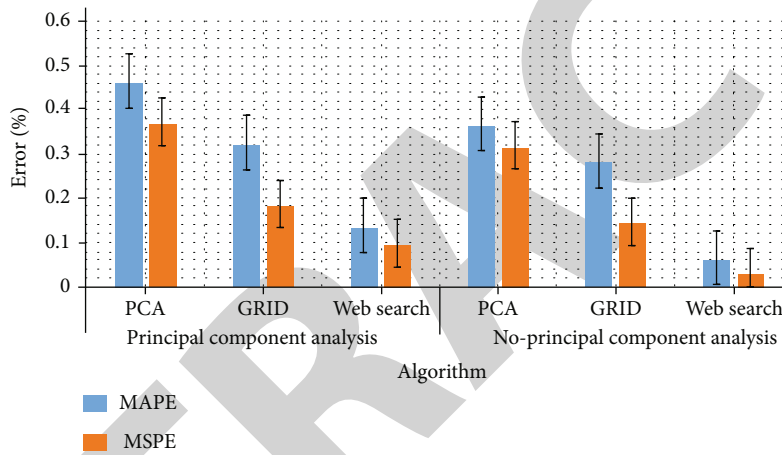


FIGURE 9: Analysis of application results of principal component analysis.

fitness basically remains unchanged after 20 iterations, indicating that the optimal value of regularization parameter is 1.184 and the optimal parameter of radial basis kernel function is 0.1.

The statistical results are shown in Figure 6. It can be seen that the predicted value is highly consistent with the actual value, indicating a high prediction effect.

Then genetic algorithm is used to optimize. The range of regularization parameters is 0~100, the range of radial basis kernel function parameters remains unchanged, the interactive verification is 5 times, and the maximum iteration is 299 times. Similarly, MATLAB software is used for simulation analysis. It is found that after 4 times, the average fitness function remains unchanged, and the best fitness function also tends to be stable. The best regularization parameter value is 4.457, and the best radial basis kernel function parameter is 0.0175. The obtained parameters are applied to the support vector machine prediction model, and the prediction results are shown in Figure 7.

In grid search optimization, the number of samples will affect the simulation results. Combined with the actual situation, the parameter defaults to $1/n$. Select the values around this value for simulation analysis. The test results are shown

in Figure 8. As can be seen from the figure, when the parameter value is 0.1, the accuracy is relatively high.

Under the optimal parameter model, the application effect of principal component analysis is compared and analyzed. The results are shown in Figure 9. It can be seen from the data in the figure that after parameter optimization, the model algorithm error decreases, and the model prediction results of the three methods are improved. After the prediction results of the principal component analysis model, the average absolute percentage error also decreased significantly, indicating that the tax prediction of the data after the redundancy reduces the data dimension and improves the prediction accuracy.

5. Conclusion

Tax revenue is closely related to people's life. Accurate prediction of tax revenue is of great significance to regulate local policies. Considering the tax changes and influencing factors, it is difficult to accurately judge it by using linear analysis alone. Based on this, this paper constructs a support vector machine model in tax prediction and analysis. This machine learning algorithm has great advantages in dealing

with nonlinear problems. Considering that the regularization function and radial basis kernel function of support vector machine have a great impact on the tax prediction results, other algorithms are used to optimize these two parameters, and the application effect of principal component analysis method and the accuracy of tax prediction after parameter optimization are analyzed. The experimental results show that the principal component analysis method can reduce the redundancy and reduce the data dimension. The model optimized by the algorithm is applied to tax prediction. The predicted value has a high degree of fit with the actual value, the prediction effect is good, and the accuracy is significantly improved.

However, the independent variable and dependent variable in the actual tax problem are linear. Taxation is affected by many aspects, and its system is complex and uncertain. The relationship between independent variables and dependent variables is mostly nonlinear, and the prediction accuracy still has room for optimization. The following contents need to be further modified in future research. Only one kernel function is selected in this paper. There are many kinds of kernel functions. Without analyzing other types of kernel functions, parameter settings will also directly affect the analysis results. When searching and optimizing parameters in the grid, you can also improve the parameter search process, take tax samples as the search scope, and further narrow the parameter search scope.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Shandong Province Key R&D Program (Soft Science Project): Research on the vertical price stabilization and income increase mechanism of the whole industrial chain of geographical indication agricultural products based on rural revitalization (No. 2021RKY06110).

References

- [1] L. M. Mauler, "The effect of analysts' disaggregated forecasts on investors and managers: evidence using pre-tax forecasts," *Accounting Review*, vol. 94, no. 3, pp. 279–302, 2019.
- [2] A. Habib, M. M. Hasan, and A. Al-Hadi, "Financial statement comparability and corporate cash holdings," *Journal of Contemporary Accounting & Economics*, vol. 13, no. 3, pp. 304–321, 2017.
- [3] O. H. Kwang-Wuk and K. I. Eun-Sun, "Effect of tax-related information on pre-tax income forecast and value relevance," *Journal of Asian Finance Economics and Business*, vol. 7, no. 1, pp. 81–90, 2020.
- [4] J. L. Ye, "The effects of analysts' tax expense forecast accuracy on corporate tax avoidance: an international analysis," *Journal of Contemporary Accounting and Economics*, vol. 17, no. 2, article 100243, 2021.
- [5] A. H. Miller, L. Alkindi, and A. Alblooshi, "Using a database approach, with big data and unsupervised machine learning to model tax behavior in the expatriate community," *Solid State Technology*, vol. 63, no. 2, pp. 379–389, 2020.
- [6] H. D. Aprilia and D. Agustiani, "Application of data mining using the K-means algorithm in rural and urban land and building tax (PBB-P2) receivables data in Bantul Regency," *Journal of Physics: Conference Series*, vol. 1823, no. 1, pp. 12–19, 2021.
- [7] F. Sun, Z. Wang, and Z. Li, "A tax dynamic clustering method based on weak convergence sequence coefficient," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 5, pp. 216–223, 2017.
- [8] W. Didimo, L. Grilli, G. Liotta, L. Menconi, F. Montecchiani, and D. Pagliuca, "Combining network visualization and data mining for tax risk assessment," *IEEE Access*, vol. 8, pp. 16073–16086, 2020.
- [9] S. L. Li, "Data mining of corporate financial fraud based on neural network model," *Computer Optics*, vol. 44, no. 4, pp. 665–670, 2020.
- [10] M. Mwanza and J. Phiri, *Fraud Detection on Big Tax Data Using Business Intelligence, Data Mining Tool: a Case of Zambia Revenue Authority*, [Ph.D. thesis], University of Zambia, Zambia, 2016.
- [11] A. H. Miller, "Data modeling and visualization of tax strategies employed by overseas American individuals and firms," *Individuals and Firms*, vol. 29, pp. 5–28, 2019.
- [12] P. Battiston, S. Gamba, and A. Santoro, "Optimizing tax administration policies with machine learning," *Working Papers*, vol. 7, pp. 436–447, 2020.
- [13] Y. Hao, "Digital inclusive finance risk prevention based on machine learning and neural network algorithms," *Journal of Intelligent and Fuzzy Systems*, vol. 2, pp. 1–11, 2021.
- [14] Y. He, K. F. Tsang, Y. C. Kong, and Y. T. Chow, "Indication of electromagnetic field exposure via RBF-SVM using time-series features of zebrafish locomotion," *Sensors*, vol. 20, no. 17, p. 4818, 2020.
- [15] I. Unceta, J. Nin, and O. Pujol, "Risk mitigation in algorithmic accountability: the role of machine learning copies," *PLoS One*, vol. 15, no. 11, pp. 241–286, 2020.
- [16] P. F. Dai, X. Xiong, and W. X. Zhou, "A global economic policy uncertainty index from principal component analysis," *Finance Research Letters*, vol. 40, no. 4, article 101686, 2020.
- [17] Y. Zhu, L. Zhou, C. Xie, G. J. Wang, and T. V. Nguyen, "Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach," *International Journal of Production Economics*, vol. 211, no. 5, pp. 22–33, 2019.
- [18] P. Li, Y. Peng, P. Jiang, and Q. Dong, "A support vector machine based semiparametric mixture cure model," *Computational Statistics*, vol. 35, no. 3, pp. 931–945, 2020.
- [19] D. Zhao, J. Ding, and S. Chai, "Systemic financial risk prediction using least squares support vector machines," *Modern Physics Letters B*, vol. 32, no. 17, article 1850183, 2018.
- [20] J. Zhu, P. Sun, Y. Gao, and P. Zheng, "Clock differences prediction algorithm based on EMD-SVM," *Chinese Journal of Electronics*, vol. 27, no. 1, pp. 128–132, 2018.

Retraction

Retracted: The Implementation of Unified Application Data Collection and Analysis in Tobacco Enterprise Platform Design

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Jin and W. Lou, "The Implementation of Unified Application Data Collection and Analysis in Tobacco Enterprise Platform Design," *Journal of Function Spaces*, vol. 2022, Article ID 3933465, 9 pages, 2022.

Research Article

The Implementation of Unified Application Data Collection and Analysis in Tobacco Enterprise Platform Design

Yong Jin  and Weidong Lou 

Department of Information, China Tobacco Zhejiang Industrial Co., Ltd., Hangzhou, Zhejiang 310007, China

Correspondence should be addressed to Yong Jin; jinyong@zjtobacco.com

Received 1 June 2022; Accepted 29 July 2022; Published 18 August 2022

Academic Editor: Miaocho Chen

Copyright © 2022 Yong Jin and Weidong Lou. 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 deepening and development of tobacco manufacturing informatization construction, tobacco manufacturing enterprises have the characteristics of strong production capacity, high degree of production automation, and advanced enterprise-level management information system. In order to cooperate with the market and brand strategy of the tobacco industry, the State Tobacco Monopoly Administration has carried out serious strategic thinking on informatization and customized the overall plan and strategic goals. In recent years, China's informatization construction has made great progress, and both industries and enterprises have felt various benefits from the application of information technology. In the deployment and development of mobile services, it is often local and municipal mobile companies that, in order to meet the different needs of the market, immediately carry out corresponding services, resulting in one service for each municipal subsidiary in the same province or one platform for a class of services. Situation is good for quickly meeting user needs at that time. The exchange and communication of enterprise personnel information are becoming more and more frequent, and the collection of enterprise information through mobile has become a new way of enterprise information collection. The work orders, process parameters, documents, notifications, and instructions generated by the planning layer are sent to the production control layer to respond, guide, and trigger the production site events. The ultimate goal of mobile data collection is to gather the collected data into the enterprise production data center, which plays an important role in mobile devices, mobile performance, mobile planning, mobile balancing, mobile prediction, and so on. Based on the unified mobile application data collection method, this paper analyzed and analyzed the design of application architecture in the tobacco enterprise platform and further described the business functions, security performance, and system characteristics of the platform.

1. Introduction

Tobacco enterprises, as an early enterprise in informatization construction, have carried out beneficial informatization exploration in various fields involved in business for many years, which has played a good role in supporting and promoting the development of business and made remarkable achievements. With the continuous deepening and development of information construction of tobacco manufacturing, tobacco manufacturing enterprises show the characteristics of strong production capacity, high degree of production automation, and advanced enterprise-level management information system. To promote the modernization of the tobacco industry with information technology

is to use the most advanced information technology to fully penetrate the production, circulation, and management process of the tobacco industry. Change the traditional production and operation mode, accelerate the adjustment of the tobacco industry and product structure, and improve the management system of the tobacco industry. Meet the needs of the market and consumers, improve the ability to adapt to the market, and enhance the overall competitive strength. The development of industry informatization is closely related to the joint reorganization of enterprises, the integrated group management of assets, and the synergy of industrial and commercial information, but it is also dynamic. It can be said that the success of its construction is an important symbol of the strength of the industry.

Therefore, it is of great practical significance to study and explore the implementation strategies of industry informatization construction. The reason is that the wide range of tobacco R&D business leads to the problems of large volume, multifarious types, and scattered storage of tobacco R&D data. It is difficult for industry experts who are proficient in tobacco R&D business to effectively mine the value by using the abovementioned complex data [1]. In order to cooperate with the market and brand strategy of the tobacco industry, the State Tobacco Monopoly Bureau has made a serious strategic thinking on informatization and customized the overall plan and strategic objectives. In order to build a resource-saving and environment-friendly society, earnestly do a good job of saving energy and reducing consumption in tobacco production, and strive to achieve economical, clean, and safe development, it is necessary to build a mobile data collection system for tobacco companies. In recent years, China's informatization construction has made great progress, and both industries and enterprises have felt various benefits from the application of information technology. The WAP business is also growing at a rapid pace [2], in the context of the deepening of China's reform and opening up and the domestic and international environment of economic globalization. Like all developing enterprises, the problem faced by tobacco enterprise informatization is a problem in development and improvement, and all solutions point to the data center system based on data warehouse.

According to the characteristics of business, data, and users of the tobacco R&D system, a layered technical architecture is adopted to build a user-oriented tobacco R&D system big data cloud platform, encapsulate data collection and storage, and standardize data management. The most direct function of mobile data collection is to provide complete mobile information to relevant mobile management personnel, assisting mobile management personnel to quickly query related data and perform corresponding management operations. In the deployment and development of mobile services, it is often the case that local and municipal mobile companies, according to the different needs of the market, immediately launch corresponding services in order to meet the needs. It is beneficial to quickly meet user needs at that time [3]. In recent years, with the market competition and gradual opening of the domestic telecommunication industry, various telecommunication services have made great progress. The establishment of the data center brings great convenience to the use of standard information and data of the same caliber for all functional departments and related personnel of the enterprise and creates conditions for the effective use of information and data mining [4]. At the same time, the tobacco purchasing data collection system is the basis of MES implementation. It is mainly responsible for the collection, analysis, processing and continuous, automatic and complete transmission of data, equipment status, personnel, events, time, and other information in the production control layer to provide basic information support for scheduling and commanding production. Monthly report of consumption, tobacco leaf sales, tow purchase and consumption, cigarette paper purchase and consump-

tion, production process industry process business data, and cigarette sales link business data. Reducing the technical requirements of large data analysis for users is an important exploration and attempt for tobacco developers to make use of data-driven tobacco R&D business.

In recent years, the tobacco industry has developed unprecedentedly, and it has become the most advanced tobacco production and processing base in Asia with large production scale. Second, mobile operators have been vigorously promoting the development of various value-added services. Different from many traditional services, the development and promotion of new services are often a process of exploration. The exchange and communication of enterprise information are becoming more and more frequent. Collecting enterprise information through mobile mode has become a new way of enterprise information collection. For tobacco enterprises, data center system can realize economic operation analysis, marketing data analysis of provincial companies, marketing data analysis of branch companies, monopoly data analysis, financial data analysis, human resources analysis, etc. [5]. The work orders, process parameters, documents, notices, and instructions generated by the planning layer are sent to the production control layer to respond, guide, and trigger the production site events. Simplify the data analysis process. Let the "non-data expert users" of the tobacco research and development system focus more on the tobacco research and development business itself and avoid users facing complex big data analysis technologies and complex tobacco research and development raw data. Analyze the data structure and design the function of the data acquisition subsystem in the unified business platform, which is responsible for the aggregation of data resources for the scattered data. However, informatization is a process of "informatization." According to the model of informatization construction, tobacco companies have gone through the process of building a large number of process-oriented software for control and have begun to enter the stage of data value-oriented integration and analysis [6]. The ultimate goal of mobile data collection is to gather the collected data into the enterprise production data center, which plays an important role in the work of mobile devices, mobile performance, mobile planning, mobile balance, and mobile forecasting [7].

2. Related Work

Since the implementation of the national tobacco monopoly system, China's tobacco industry has made great achievements in "meeting the consumer demand, improving the quality of tobacco products, increasing the national financial accumulation and supporting the development of national enterprises." Bauer established the idea that consumers' rationality is relatively low and plays an important role. He believes that any action of consumers will produce results that he cannot be completely sure of [8]. Ross put forward arbitrage pricing theory. The western financial management theory is coming to maturity. Financial management has developed into a management activity that integrates financial forecasting, financial decision-making, financial planning,

financial control, and financial analysis, with financing management, investment management, working capital management, and profit distribution management as its main contents and occupies a core position in enterprise management [9]. Miller and Mordelli Yanni put forward the famous principle that the capital structure of a company has nothing to do with dividend policy in an efficient securities market [10]. Copeland Kou proposed that the classification of consumer goods into convenience goods, shopping goods, and specialized goods is based on the analysis of consumer behavior in three aspects [11]. Barnes has established various information technologies, mainly based on network technology, provide broader and more advanced technical means and methods for financial management of enterprises, and enterprises can keep the technological lead based on flexible technology [12]. Dr. E.F. Cold put forward the data model and theory of relational database. After long-term business development and demonstration, relational database has been widely used. Since then, the storage and utilization of data have entered a new era [13]. Dean published "Capital Budget," which changed the focus of financial management from the emphasis on external financing to the rational allocation of funds within the company, which made a qualitative leap in the company's financial management [14]. Mead published the first book devoted to the financial management of corporate fundraising [15]. Fama and Miller published the book "Financial Management," which is a collection of Western financial management theories, marking the maturity of Western financial management theory [16]. Mr. Chares designed what was then called an integrated data store, the design laid the foundation for a meshed, hierarchical data model, seen as the early days of relational databases in database history [17]. Since the implementation of the national tobacco monopoly system, China's tobacco industry has made great achievements in "meeting the consumer demand of residents, improving the quality of tobacco products, increasing the national financial accumulation, and supporting the development of national enterprises."

3. The Way of Mobile Information Collection

3.1. SMS. In recent years, short message service has penetrated into the daily life of the public in a variety of ways. With this model, it is easy to store the original detailed data from each data source. And flexible storage of historical data is with enough flexibility classes to meet all requirements. Due to many objective factors such as inaccurate counting of electrical system, shift, clearing of power-off data, completion of work order, trial production, etc., the data collected continuously has a great error [18]. Therefore, the platform generally uses a hierarchical architecture, providing simple data analysis applications and interface services for top-level "non-data expert users." It encapsulates the underlying data aggregation, data storage, subject data services, and other functional modules, as well as the algorithm library and analysis services. The whole enterprise may already have some mobile systems and centralized field control systems, and there will be many scattered mobile meters [19] in the above range. Although data services will occupy an increas-

ing proportion in telecommunication services, voice services are still the most important and basic services for mobile communication systems. The information office automation system weaves a set of efficient and smooth information interconnection system within enterprises and institutions, which greatly promotes the development of productivity of enterprises and institutions. The resulting changes in business process, business processing rules, and business information flow are the three major factors of data acquisition system reconfiguration. It is also a problem that must be considered and solved to build a new data acquisition system [20]. Database integrity constraint is the guiding principle to ensure the data quality in the preprocessing domain of data warehouse. Data warehouse technology is a combination of customer service and business processing information. The unified framework is a data warehouse that provides a scientific basis for enterprise managers to formulate strategies; develop market analysis market; benefit evaluation, public relations, and personal image design; and make decisions by using analysis and mining technology. By collecting data samples from different systems within the enterprise and external data sources, the consistency of data is maintained, and it is easy to be accessed by users. At the same time, these data are organized to facilitate analysis. The specific data quality dimensions include accuracy, uniqueness, compatibility, validity, integrity, and completeness. The relationship between them is as shown in Figure 1.

The service layer includes four modules: automatic data aggregation, data storage management, thematic data service, and data analysis service, which can automatically collect data related to tobacco R&D and realize the storage management of complicated data. Add a new interface, and introduce the construction concept in the system into the middle layer, so that the process can be realized without ignoring the interface. The specific content is shown in Figure 2.

Engineering inspection is not only to observe the change of the structure but also to obtain quantitative data reflecting the structural performance. Only when reliable data are obtained can correct conclusions be made on the structural performance and the purpose of inspection be achieved. Only reliable data can be used as the basis for judging structural performance or establishing calculation theory. If there are no measuring instruments installed on mobile devices or nodes included in the mobile accounting system, the related data will not be available in the data collection system. As the most basic operation unit in the whole mobile communication industry, the local and municipal mobile operators adopt different equipment and business platforms from different manufacturers in the long-term business process. Training material enterprise training information view is as shown in Tables 1, 2, and 3.

This is reflected in the different customized interface for user-oriented business in each municipal mobile. The different devices lead to the different storage modes and formats of user data and the development of value-added business provided to operators. In the information age, in the face of fierce competition in the market, enterprises must grasp the pulse of the market, timely, fast, accurate, efficient, and

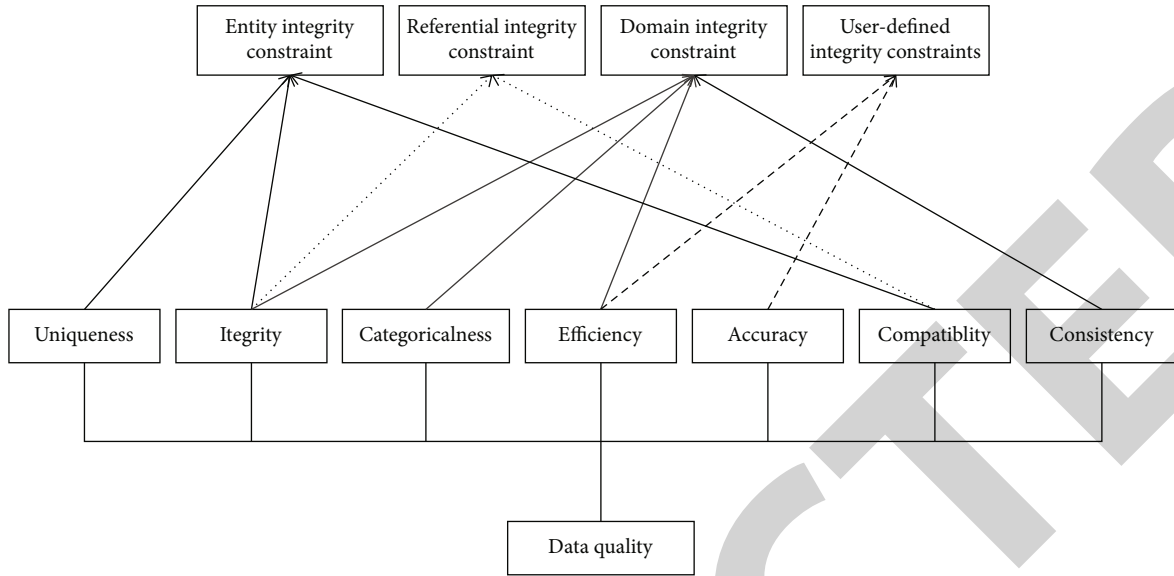


FIGURE 1: Data quality dimension of data warehouse preprocessing domain.

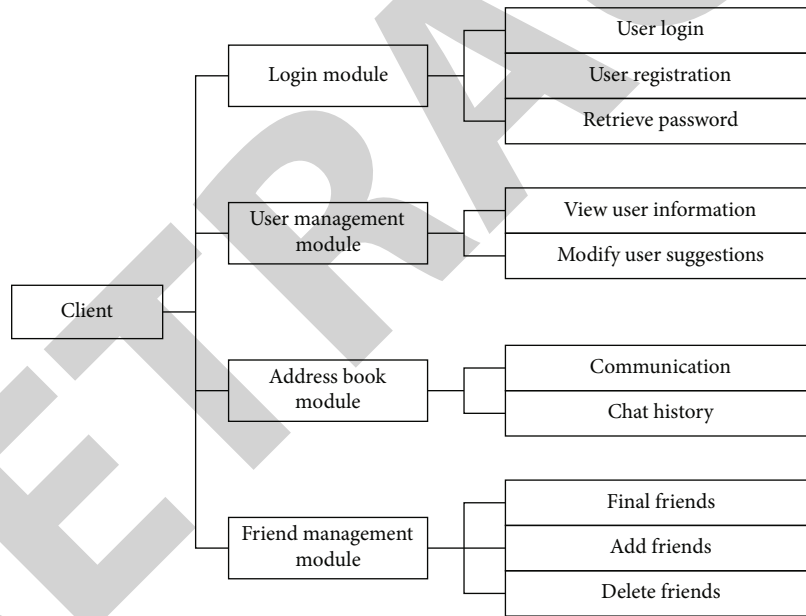


FIGURE 2: Client connection structure diagram.

low-cost in order to stand out. Obtaining effective real-time information from the complicated mass information is also the main problem that puzzles all departments of the enterprise. However, the disadvantage is that users have to edit short messages according to the format, which is restrictive to some extent. Seasonal variation index calculation formula:

$$S_k = \frac{\bar{x}_k}{\bar{x}}, k = 1, 2, \dots, m. \quad (1)$$

Prediction model:

$$\hat{Y}_{T+1} = T_{T+1}S_{T+1}, \quad (2)$$

$$T_t = b_0 + b_1t. \quad (3)$$

General exponential seasonal factor adjustment model:

$$T_T = Ae^{r(t=1)}, \quad (4)$$

$$\hat{Y}_{t+1} = T_{T+1}S_{T+1}. \quad (5)$$

And its parameter test and prediction formula:

$$\varphi_p(B)\Phi_p(B^2)(1-B)^d Y_t = \theta_q(B)\Phi_Q(B^t)^{e_t}, \quad (6)$$

TABLE 1: Training information view use case specification table.

Use case name:	Training material enterprise training information view
Brief overview:	Training supervisors/corporate leaders and corporate employees Training supervisors, corporate leaders, and corporate employees can view corporate training information and training materials. Information exists in the form of news, while materials exist in the form of documents or videos.
Preconditions:	The user has logged in to the management system and logged in to the interface.
Basic event slip:	After the training supervisor, enterprise leaders and enterprise employees enter the training management interface, and they can browse the training information and materials.
Exception event flow:	The system pops up an error message, and the system returns to the previous operation interface and gives a prompt message
Postconditions:	The system gives a prompt interface for successful operation and returns the operation record
Remark:	None

TABLE 2: Specifications of use cases for enterprise training information release.

Brief overview:	Training supervisors and business leaders need to organize and summarize the content of the original training and then release the training materials needed by the enterprise, as well as training notices and other information
Preconditions:	The user has logged in to the management system and logged in to the interface
Basic event flow:	After the training supervisor and enterprise leaders enter the system, they can release training information or upload training materials through the information release interface.
Exception event flow:	The system pops up an error message, the system returns to the previous operation interface and gives a prompt message
Postconditions:	The system gives a prompt interface for successful operation and returns the operation record
Remark:	None

TABLE 3: Rearranged material use case specification table.

Use case name:	Rearrange material
User role:	Training supervisor, business leader
Brief summary:	The uploading of training information and training materials cannot be uploaded by anyone. Hey, after the materials are uploaded, they need to be approved by the leader before they can be uploaded. If the materials that do not meet the requirements or WeChat will be returned for revision, the materials need to be uploaded here. Refresh and upload or publish
Preconditions:	The user has logged in to the management system and logged in to the interface
Basic event flow:	For the documents or information that do not meet the requirements, the training supervisor will reorganize the materials after revising them according to the revision opinions and upload them for review.
Exception event flow:	The system pops up an error message, the system returns to the previous operation interface and gives a prompt message
Postconditions:	The system gives a prompt interface for successful operation and returns the operation record
Remark:	None

$$T_T = a \times \text{RMSE} + \text{MAPE} \times D_1 + c \times \text{MaxAPE} + \text{MapAE}, \quad (7)$$

$$\hat{Y}_{T+1} = T_{T+1} S_{T+1}. \quad (8)$$

3.2. WAP Method. In our concept, corporate website is a window to display corporate products and a means to use the Internet as a mass media to publicize companies. This allows business personnel to use some front-end display tools to freely combine data items and customize reports and charts according to their own needs. Enterprises and institutions urgently need an office automation system that can be used at anytime, anywhere, and at any time, so that company managers and business personnel can associate

with the internal systems of enterprises and institutions as they want, no matter where they are. The production activities of tobacco factories are always affected by shift hours, shifts, production days, and work orders. The division of these time periods and the events that occur at alternate time periods are the key to statistical data. The data is sorted when stored, the location of the file is retrieved when new data arrives, the data of the original file is copied and integrated with the new data, the filename of the new file is the same, and the timestamp is updated to the current time. Install the data collection computer, configure the corresponding communication board card in the computer according to the communication mode supported by the meters, and collect the data of the meters into the computer.

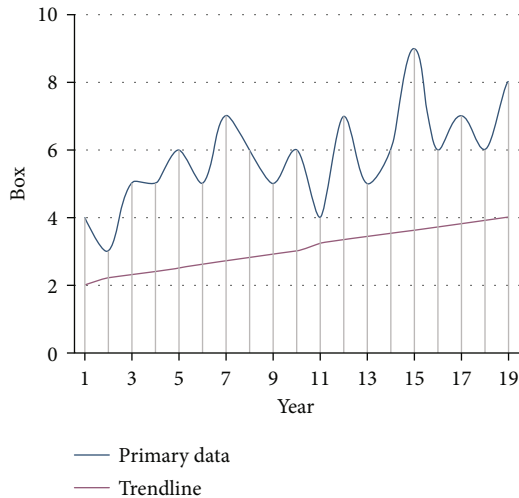


FIGURE 3: Sales trend chart.

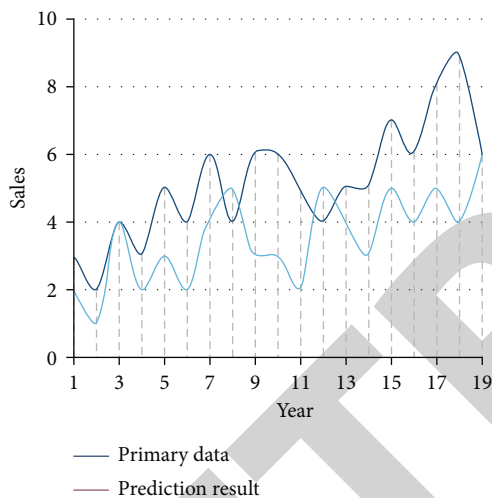


FIGURE 4: Monthly sales forecast.

It is clear that the overall sales volume increases over time. As shown in Figure 3, the monthly sales forecast is shown in Figure 4.

At present, the mobile communication market is developing rapidly, and the second generation digital communication system has basically replaced the first generation analog communication system. And it is stepping into the third generation broadband digital communication system, the number of mobile users is also growing rapidly, and the competition in the mobile communication market is becoming increasingly fierce. By doing so, it can be confirmed that the business requirements are correctly understood and in the next stage. Data warehouse designers get reliable, business-driven data structure, which greatly reduces the cost of maintaining the logical and physical structure of data warehouse in the short, medium, and long term. Through the fast wireless data network and terminal, tobacco company staff can not only extend the office automation system to mobile phone terminals but also manage orders in real time. If it is processed in the program, its run-

ning speed cannot meet the needs of users. Therefore, encapsulating statistical operations in stored procedures can greatly improve the efficiency of system operation. This design can effectively reduce users' requirements for big data analysis technology, simplify the operation of big data analysis process, and enable "non-data expert users" to use the service conveniently. After patrolling and observing on-site equipment and instruments, operators manually record relevant mobile data on site and then manually enter the mobile data acquisition system. WAP enterprise website can be like other applications, without operator restrictions, as long as the mobile phone with WAP function. You can log on the WAP website of the enterprise to view the information at any time, so the WAP website of the mobile phone is an important step in the realization of the wireless information construction of the enterprise.

4. Platform Application and Effect

4.1. Instance Overview. The application process and results of the large data cloud platform of tobacco R&D system in Yunnan for users are verified by an example. All data collection points will produce millions or even tens of millions of record values in a day, making it difficult to efficiently process ordinary databases of records of this magnitude. Real-time databases use a variety of optimization and compression storage algorithms. However, there is bound to be the problem of relying on virtual operator, which will restrict the business content and service quality, making it inconvenient to carry out its own characteristic business, and all information must be transferred through service providers, so the security of enterprise information cannot be guaranteed. Actual and predicted sales results for a certain two years are as shown in Figure 5.

Forecast the results in the next two years, as shown in Figure 6.

It is a process of identifying and selecting the system to be developed and determining the system development time according to the industry's best business practices based on the enterprise's strategic objectives, organizational structure, business processes, and information technology. Through the existing network environment of tobacco companies and the existing network environment and security measures of telecommunication operators, the service guarantee of telecommunication-level security can be fully guaranteed. To strengthen the management's ability to coordinate and schedule production, tobacco companies are required to process and analyze the collected data in the data collection system, so as to provide direct data support for management decisions. Through the application system, you can view all parts of the mobile data at any time. At the same time, information such as equipment running status, equipment failures, mobile flow information, mobile consumption distribution and statistics, pipe network transmission, and mobile utilization comparison can be viewed through a pre-determined data statistical analysis system. When the location is updated and exchange data, after receiving the location update request, check the user's roaming authority, update its own location information about the user, notify

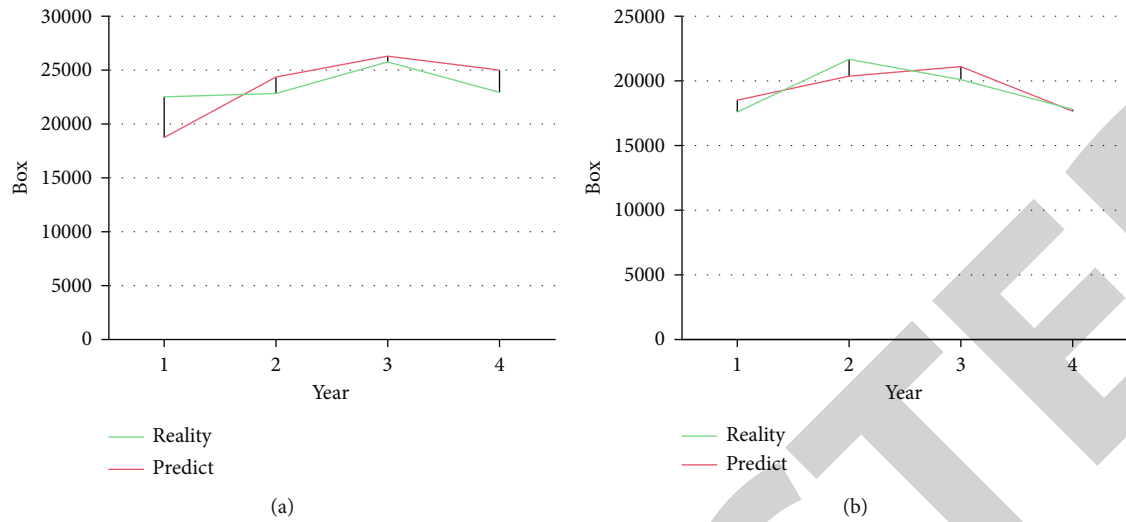


FIGURE 5: Actual and predicted results.

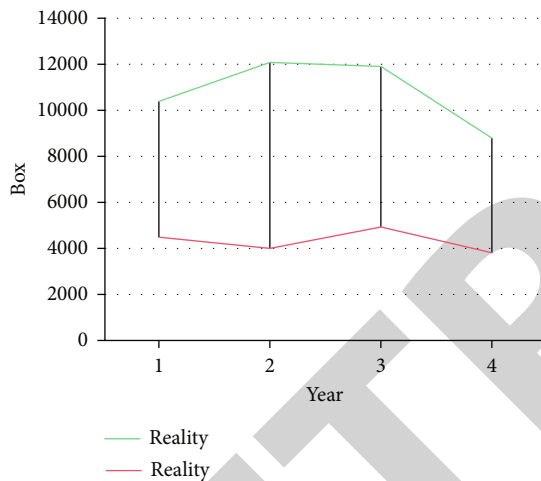


FIGURE 6: Expected results in the next two years.

the original to delete the user's data, and then transfer the user's data to the new one. The layers are independent of each other, interact with each other using standard interface, and have the feature of weak coupling, which makes the system have good scalability and stability. IVR voice platform can collect information manually and display information and voice to mobile phone users. Business architecture is the current or future period of time for an enterprise. The overall way of business management is reflected, and the impact of business architecture on information systems is very important. Information systems and related information processes serve business architecture requirements. Understanding the actual situation of tobacco quality monitoring, real-time analysis of the use of laboratory equipment, use saturation, etc. to ensure the stability of tobacco product quality and style is a typical large data analysis application scenario.

4.2. Instance Implementation. The results of physical and chemical experiments on tobacco and its raw materials, fla-

vors and spices, materials, etc. were obtained through the experimental analysis process. View the wear and tear data of moving in the production, conversion, transfer, and use link of each part according to custom conditions. These two processes have been monotonous in the past, but as the demand for personalized services becomes stronger and stronger, the number of functions required for the system to achieve is increasing and more complex, and a relatively independent and flexible business system and accounting system need to be built. Database servers are the preferred target for attackers and must be secured against internal, external, network, and application-level attacks. The integration architecture is focused on the collaborative integration of data, application, process, and interface of each application system in the enterprise, ensuring a complete technical carrier of integrated management and a perfect integration technology method. Because the information system currently used in tobacco industry mostly uses wired network communication and field operation mode, once people leave the office, they cannot get or submit the corresponding real-time information, causing delays in information and decision-making. Second, the real-time requirement of users, data incremental information is transmitted to the database in seconds. The objective factors of multiequipment and multi-information collection points in tobacco factories determine that there must be a large number of data records in the database. The system should ensure that a large amount of data is counted, and data analysis reports are generated within the valid time required by users. By using directory service, all information and resources on the network appear as an orderly and unified directory structure in front of network administrators, users, and applications. Databases usually contain the most sensitive and confidential data, such as personal details, customer details, orders, or credit card details of the human resources department. However, in practical application, the query access flow on the business side is very large, a single thread will make the system inefficient, and even the query will lose data. Therefore, how to perform multi-threaded parallel

queries between the server and the client to process multiple requests has become the focus of research. In addition, you can further check the saturation and contribution of laboratory equipment. At the same time, the stability of physical index and chemical index of tobacco products can be obtained by analyzing each test index.

5. Conclusions

With the rapid development of data technology, data analysis and application are increasingly important in the daily operations of enterprises. A resource management statistical analysis system was established, including mobile comprehensive energy and water consumption statistical analysis module, mobile consumption proportion map module, and production and nonproduction hydropower map module. The implementation of the data acquisition system makes the information of the planning layer and the production automatic control layer of the tobacco production enterprise realize two-way circulation, and the status of the production activities can be displayed intuitively. It effectively promotes the management and scheduling of production by enterprises and provides a foundation for the realization of agile manufacturing in modern tobacco enterprises. The basics of the overall design planning and principle introduction gradually start from the dimensions of technology use introduction, network, hardware, software architecture, design and implementation methods, and working principles. How to sort out the massive data generated in the R&D business and revitalize the data assets is a topic currently faced by the industry. Combining the advantages of centralized and noncentralized service discovery mechanisms provides effective service discovery mechanisms to ensure business continuity when individual peripherals join or exit in groups. The two are different in essence. Under the centralized mode, the unit combination and generation curve are determined based on the security constraints, which is a transaction mode closely related to the power grid operation and unified optimization of various transactions. The actual user queries entered by the business tier are broken down into subqueries for each data source, and the results of each subquery are returned to the business tier application in a unified view. Thus, data collection and unified view are achieved, which facilitates the development of upper telecommunication business. With the data warehouse-based data center system of tobacco enterprises as the research object, the research on system background, system design, and implementation process has great theoretical research value and important application value.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Z. Haitao and X. Xiang, "Big data analysis platform for tobacco industry research and development applications," *Computer Applications and Software*, vol. 38, no. 6, p. 6, 2021.
- [2] Y. Lin, X. Wang, and Y. Xu, "Semi-supervised human resource scheduling based on deep presentation in the cloud," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, pp. 1–9, 2020.
- [3] S. Yingqiang, "A brief analysis of the practical application of Jingzhou tobacco's visualization platform. Mobile," *Information*, vol. no. 11, p. 4, 2020.
- [4] A. González, J. Olazagoitia, and J. Vinolas, "A low-cost data acquisition system for automobile dynamics applications," *Sensors*, vol. 18, no. 2, p. 366, 2018.
- [5] K. Benke and G. Benke, "Artificial intelligence and big data in public health," *International journal of environmental research and public health*, vol. 15, no. 12, p. 2796, 2018.
- [6] Z. Jun, C. Yan, and Y. Yang, "Design of ship mobile data collection and analysis system in the internet environment," *Ship Science and Technology*, vol. 42, no. 8, pp. 158–160, 2020.
- [7] J. Huang, Z. Duan, J. Kwok et al., "Vaping versus JUULing: how the extraordinary growth and marketing of JUUL transformed the US retail e-cigarette market," *Tobacco Control*, vol. 28, no. 2, pp. 146–151, 2019.
- [8] S. Jian, L. Xiao, and L. Shuo, "Design and application of energy monitoring system for tobacco industry," *Computer Knowledge and Technology: Academic Edition*, vol. 18, no. 9, p. 4, 2022.
- [9] K. B. Moon, J. H. Jeon, and H. Choi, "Construction of SARS-CoV-2 virus-like particles in plant," *Scientific reports*, vol. 12, no. 1, pp. 1–7, 2022.
- [10] T. Cheng, G. Zhao, M. Xian, and C. Xie, "Improved cis-Abienol production through increasing precursor supply in *Escherichia coli*," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020.
- [11] I. Balke and A. Zeltins, "Recent advances in the use of plant virus-like particles as vaccines," *Viruses*, vol. 13, no. 3, p. 270, 2020.
- [12] P. Srinivasan and C. D. Smolke, "Engineering a microbial biosynthesis platform for de novo production of tropane alkaloids," *Nature communications*, vol. 10, no. 1, pp. 1–15, 2019.
- [13] J. Jin, Y. Xu, P. Lu et al., "Degradome, small RNAs and transcriptome sequencing of a high-nicotine cultivated tobacco uncovers miRNA's function in nicotine biosynthesis," *Scientific reports*, vol. 10, no. 1, pp. 1–11, 2020.
- [14] X. G. Xu, Y. Guo, Y. Fu, Z. W. Wang, and X. D. Zhao, "Non-linear discrete system model of tobacco supply chain information," *Nonlinear Engineering*, vol. 11, no. 1, pp. 316–323, 2022.
- [15] H. J. Jeong, C. W. Kang, and B. H. Kim, "Development of a Quality Management System based on a Platform for Customized Application of Small & Medium Manufacturing Enterprises," *Journal of the Korean Society for Precision Engineering*, vol. 35, no. 10, pp. 973–985, 2018.
- [16] J. Ye, Y. Ding, X. Qi, J. Xu, X. Yang, and Z. Zhang, "Geographic and position-based variations in phyllospheric bacterial communities present on flue-cured tobacco," *Applied Microbiology and Biotechnology*, vol. 105, no. 24, pp. 9297–9308, 2021.
- [17] M. Kermani, D. L. Carnì, and S. Rotondo, "A nearly zero-energy microgrid testbed laboratory: Centralized control strategy based on scada system," *Energies*, vol. 13, no. 8, p. 2106, 2020.

Retraction

Retracted: Mathematical Model Construction of Teaching Evaluation in Colleges and Universities Based on Convolutional Neural Network under the Background of Big Data

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Tan, "Mathematical Model Construction of Teaching Evaluation in Colleges and Universities Based on Convolutional Neural Network under the Background of Big Data," *Journal of Function Spaces*, vol. 2022, Article ID 7064287, 8 pages, 2022.

Research Article

Mathematical Model Construction of Teaching Evaluation in Colleges and Universities Based on Convolutional Neural Network under the Background of Big Data

Changming Tan 

Department of Management, Anshan Normal College, Anshan, Liaoning 114005, China

Correspondence should be addressed to Changming Tan; changmingtan@mail.asnc.edu.cn

Received 27 May 2022; Revised 25 July 2022; Accepted 28 July 2022; Published 17 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Changming Tan. 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 teacher evaluation, the main purposes of university evaluation are divided into the final purpose of ensuring teaching quality and the formative purpose of promoting professional learning. In order to effectively help teachers teach students according to their aptitude and realize personalized learning. This paper mainly analyzes and studies the mathematical model of college teaching evaluation based on convolutional neural network under the background of big data. The research analyzes the teaching evaluation under the background of big data and evaluates the teaching of big data participants. The teaching evaluation data can promote the development of education and teaching. Data acquisition processes and functions in the context of big data. Ensure the comprehensiveness and accuracy of data collection. At the same time, the function of teaching evaluation results is analyzed. The evaluation and formulation of mathematical model program in this study make full use of education big data information to provide scientific support for education decision-making.

1. Introduction

Nowadays, education has entered the stage of industrialization, and big data has become the main way, spreading in many fields such as educational technology, distance education, and teaching evaluation, and has become an important research hotspot. From the evolution of textbooks, it can be seen that new media textbooks are not simple digitization of paper textbooks, but a redesign of teaching materials. In the commercial field, help e-commerce to analyze customer characteristics and realize personalized recommendation. Higher education is the foundation of talent training in China, and undergraduate education is the main part of higher education. As the main position of talent training, it undertakes the main task of talent training [1]. The evaluation work must implement the national education policy and make use of books and periodicals to educate students in political, ideological, moral, cultural, and scientific knowledge, etc. With the help of certain means and tools, such as smart wearable devices, the characteristics of students can be collected without the knowledge of the students, which can

effectively help teachers teach students in accordance with their aptitude and achieve personalized learning [2]. However, the e-learning system developed to achieve this goal of online education has lost many functions of education itself, such as the function of personalized education and the function of tracking learners' emotional feedback. The continuous improvement of the degree of informatization has also gradually transformed the learning method from the traditional learning method to the information-based and digital learning method. Mathematical models are designed to provide targeted teaching plans based on students' individual characteristics and to provide students with precise and personalized educational services, thereby promoting learning and improving performance. Evaluation tools play an important role in how to effectively implement learning process evaluation. The development of information technology, especially the rise of smart education, big data full sample collection, and high-efficiency data processing capabilities are to realize the collection, storage, and analysis of learning process resources. Then, through the personalized evaluation and recommendation of the learning process,

it provides technical support to promote the improvement of students' learning process evaluation. Therefore, building a personalized and accurate evaluation model based on big data in the smart education environment is of great significance to promote the all-round development of students and stimulate the vitality of classroom life.

The mathematical model in the era of big data relies on the support of information technology and big data analysis technology and depicts the personalized characteristics of students by analyzing and mining students' basic information data and learning behavior data. With the goal of promoting learning and improving performance, a targeted mathematical model scheme is given according to the individual characteristics of students. With the further deepening of the construction and application of smart campuses, changes in the teaching environment, course content and teaching mode, the application of various learning spaces, big data platforms, mobile learning, blended learning, and other learning modes are carried out [3]. It can be seen that the school teaching evaluation is an important educational institution of the school. It is not only the function of borrowing and repaying but also serving the teaching and scientific research of teachers and students. It is also possible to set up an early warning line for learning conditions, conduct specific observations on different students, and conduct timely counseling work, which can reduce the appearance of problem students to a certain extent and promote the healthy development of students. However, the training of convolution neural network models requires large-scale data and strong computing power. Because the CPU lacks parallel computing power, it is inefficient to train and run convolution neural networks. Therefore, the GPU with strong parallel computing power is the most popular general processor for training and running convolution neural networks [4]. Composite textbooks add information technology to the original paper textbooks, so that paper textbooks and teaching resources are connected, so that paper and digital learning resources are all presented in the hands of students. In summary, the mathematical model in the era of big data is of great significance for promoting teaching accuracy, learning individualization, environmental intelligence, and management intelligence.

To implement process management and process evaluation based on big data, the collection of process data is the first step of evaluation. Make its due contribution to improving the quality of the whole nation and the healthy growth of the next generation. Teaching evaluation should not only publicize and guide students to read all kinds of books and periodicals but also provide all kinds of reference materials for teachers' teaching. At present, with the development of teacher evaluation, its main purpose is divided into the summative purpose of ensuring teaching quality and the formative purpose of promoting professional learning [5]. For the judgment of learners' learning experience and data collection, most online education platforms get feedback from users through questionnaires. But the results are often incomplete answers or even misleading information. Open-source hardware refers to the computer and electronic hardware designed in the same way as open-source software. It includes a large number of electronic devices, such as vari-

ous development boards and components with different uses [6]. Providing support for education and teaching decision-making through data-driven is the main feature of mathematical models in the era of big data. The summative purpose is to provide a certain basis for teachers' rewards and punishments through teacher evaluation [7]. The formative purpose is to evaluate the improvement and progress of teachers' abilities in various dimensions through teachers [8]. Based on this, developing university learning evaluation based on educational big data is an important part of solving teaching problems in Chinese universities, and how to build a university evaluation system based on educational big data has also become one of the important contents of the development of university informatization.

2. Related Work

Intelligent education has become a new hot topic in the research of educational informatization. According to their own educational informatization development, many schools gradually turn from theory to practice and implement it into curriculum teaching by using the education cloud platform supported by big data in the context of smart education. Further research has been carried out on the personalized learning of junior middle school students. Carry out personalized teaching research and Taiwan application research through accurate evaluation and personalized learning application research under intelligent education. In order to help students improve their academic performance and core literacy, traditional learning evaluation focuses on the external value of knowledge. The evaluation is mostly based on a certain period of time during which students participate in learning, and the evaluation dimension is limited to the mastery of pure objective knowledge. Science curriculum is a new curriculum with great reform and innovation, which needs to start from improving the scientific literacy of every student. Establish an evaluation system that is compatible with science curriculum with diversified evaluation, comprehensive evaluation content, and diverse evaluation methods. With the dramatic increase of students' process data, learning evaluation has turned to process-based learning evaluation based on educational big data, which runs through the whole process of students' learning, instead of taking a quantitative evaluation of students as the evaluation result. Zhibin designed big data platform, based on the theory of emotional computing, which can identify six basic emotions: happiness, fear, sadness, surprise, anger, and disgust [9]. Ning and Lin put forward a neural network tree facial expression automatic classification method with good recognition effect [10]. Oliveira et al. explored the construction of a blended teaching quality evaluation system, combined with the blended teaching process, to evaluate students' learning activities and learning results and to stimulate students' learning initiative [11]. Youyin et al. have established the learning platform that constructs a learning evaluation index system, which reflects students' learning situation and evaluates students' learning effect from four dimensions: learning goal, interaction, emotional attitude, knowledge, and skills [12]. Chen and Nielsen put

forward that a mathematical model based on big data is constructed from three dimensions: the establishment of teaching objectives, the design of teaching process framework, and teaching evaluation and prediction [13]. Gao et al. have discussed the learning evaluation of universities under the flipped classroom mode under the background of big data and constructed the index system of student evaluation from four dimensions: before, during, after, and learning results, which mainly involves four aspects: learning resources, learning interaction, knowledge feedback, and learning achievement [14]. Yinzhen et al. put forward a mathematical model supported by information technology, which includes four links: accurate target determination, development of materials and teaching process, counting and drawing performance, and data decision [15]. Scholars such as Yangxin et al. established the following: on the basis of facial expressions, other physiological signals such as heart rate, blood pressure changes, and skin electrical conduction were combined. In this way, the emotional changes of learners are obtained, and the full use of learners' emotional feedback in learning is discussed to improve the learning experience [16]. Scholars such as Houhe et al. believed that the FILT-WAM learning framework was proposed by combining facial expression recognition with speech expression. In the learning process, hardware devices such as cameras and microphones are used to obtain learners' facial expressions and voice information [17]. Weibo et al. built a data-driven learning support design model that includes two subjects, four dimensions, and four links. Therefore, it is necessary to systematically analyze the learning evaluation in colleges and universities, propose a learning evaluation index system based on educational big data, and reconstruct the educational evaluation system [18].

3. Teaching Evaluation in the Context of Big Data

3.1. Participants in Big Data. The application of big data in education calls for theoretical innovation while promoting profound changes in education. However, the traditional thinking of the dichotomy of subject and object holds a critical attitude towards the educational application of technology. In this regard, on the basis of clarifying the presence of teaching evaluation in the context of big data, this paper justifies the rationality of the evaluation supported by educational big data from two aspects of method and value and points out the direction of theoretical development. First, it expounds the evaluation methods and principles based on educational big data from three dimensions: the integration of evaluation into the process of teaching activities, the data integration teaching evaluation system, and the intelligent evaluation of human-computer collaborative achievements; secondly, around the value cognition of education big data evaluation, it describes that evaluation promotes the evolution of teaching ecology and cultivates subject consciousness. This promotes the continuous generation of learning value and demonstrates the significance of intelligent assessment in promoting human development in the context of education big data. In practice, the teaching big data in col-

leges and universities is directly carried out in various education and teaching links and activities, including the teaching process and community activities, and the data covers a wide range. Moreover, the period of data collection is also very short, and the amount of data is huge. Theoretically, the evaluation activities have two purposes. The first is to verify the teaching effectiveness and provide a basis for the appointment and promotion of school proofreaders, that is, rewarding and punishing teacher evaluation [19]. Therefore, in the autism teaching scene, teaching evaluation is an important research data carrier of information technology. Convolutional neural network is the mainstream algorithm of teaching evaluation, and it is used in the scheme of identifying students' behaviors and actions in related work [20]. Although the log is the simplest form, it is the most basic form of user behavior and the richest form of data. For example, if a user clicks on some information, it will be stored in the click log. In recent years, teaching evaluation methods based on convolutional neural networks have achieved outstanding results in teaching evaluation-related competitions. And it is also widely praised in the industry, so it is based on the convolutional neural network to realize the teaching resource integration scheme. And on this basis, in order to improve the teaching effect, promote teaching and learning as the goal, and provide targeted teaching guidance for students. Therefore, students' learning should return to the real situation and solve specific practical problems; and the learning process is not only the acquisition, processing, and storage of learning content. College students can learn about the latest and higher-level technologies in today's society through teaching evaluation, and they can connect with the times and understand the technological development of the times. Under the influence of the long-term "separation of teaching resources," the difficulty of obtaining such teaching resources by free means can be imagined. The management mechanism of different colleges and universities also makes the integration of teaching resources difficult. The purpose of teacher evaluation should be to give feedback on teachers' work, to achieve the purpose of improving and perfecting their own abilities, so as to improve teachers' abilities to promote the achievement of future goals. The neural network simulates the neurocognitive mechanism, and by learning the subfeatures of the object, it can still recognize the object even when the object is displaced and deformed. These are the elements that can really help the teaching evaluation work to exert greater value, can promote the development of college teaching management, and can also help students learn management more smoothly. The calculation process is as follows:

$$\begin{aligned}
 W_i &= \frac{w'_i}{\sum_{i=1}^n w'_i}, \\
 W_i &= \sqrt[n]{\prod_{i=1}^n a_{ij}}, \\
 \lambda \max &= \frac{1}{n} \sum_{i=1}^n \frac{(AW)_i}{W_i}.
 \end{aligned} \tag{1}$$

TABLE 1: Second-level indicator judgement matrix under learning attitude after aggregation.

Learning attitude	Necessary activity initiative	Self-directed learning initiative	Explore initiative
Necessary activity initiative	1	0.7030	0.5375
Self-directed learning initiative	1.4225	1	1.7000
Explore initiative	1.8601	0.5883	1

TABLE 2: Content after aggregation master the judgment matrix of next level indicators.

Content control	Knowledge learning situation	Deep processing feedback
Knowledge learning situation	1.0001	1.0623
Deep processing feedback	0.9413	1.0000

TABLE 3: Judgment matrix of secondary indicators under professional competence after aggregation.

Professional ability	Problem solved	Innovative practice	Innovative practice
Problem solved	1	0.4793	0.3114
Innovative practice	2.0855	1	0.5564
Collaborative interaction	0.2640	0.3160	0.1223

3.2. *The Influence of Teaching Evaluation Data on the Development of Education and Teaching.* In the traditional teaching mode, the teacher is the only subject of evaluation and acts as the standard setter and judge in the evaluation. The new curriculum requires diversified curriculum evaluation, which will be a new challenge for teachers who have worked for many years, but it does not mean giving up teacher evaluation. On the contrary, the new curriculum raises the evaluation standard of teachers to a new standard. In ordinary classroom teaching, we can set questions with different difficulties according to the level of students, so that every student can get the opportunity to perform.

In the context of big data, many teaching evaluation data are very diversified and multilayered, which means that the data that can be collected in the context of big data is relatively more comprehensive and accurate. It has helped teachers and students in different degrees. By consistency detection, the consistency of the judgement matrix can be accepted, and the relative index weights of the secondary indicators can be trusted, as shown in Tables 1–3.

For example, the method based on logic, but in reality, it is considered that the possibility of using this method to realize artificial intelligence is almost zero, because so far, no good algorithm has been found to learn and calculate logical symbols. Different from other types of neural networks, convolutional neural networks greatly reduce the overall complexity of the network by sharing weights and connecting locally and avoid the complex preprocessing of input data required by the traditional network. Only when learners actively participate can they get more positive emotional behaviors, thus promoting learning efficiency. Therefore, online education is also increasingly focused on the learning process. However, some of the troubles that humans can solve with instinct are difficult for computers, and this has become a problem that artificial intelligence needs to solve,

including image recognition, natural language processing, and speech recognition. Because the teaching scheme has certain after-effects, it takes a certain period of time for a teaching scheme to have an impact on the teaching effect, and it is not immediate. The output of the convolutional layer is as follows:

$$\text{Feature}_j^i = \phi \left(\sum_{i \in M_i} \omega_{ij}^l \otimes_{x_i}^{l-1} + b_j^l \right). \quad (2)$$

Sigmoid expression:

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}. \quad (3)$$

Expression for tanh function:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (4)$$

Guidelines are defined as follows:

$$\begin{aligned} \text{BIC} &= -2 \ln g(\hat{\theta}_k | y) + k \log n_s, \\ C^p &= \frac{\text{RSS}_k}{\hat{\sigma}^2(n2k)} \end{aligned} \quad (5)$$

Therefore, learning includes network nodes and connections. Nodes are external entities used to form cognitive networks and are all kinds of available information sources. The purpose of students passing through this position is very strong. One is to go to teaching evaluation, and the other is to go in and out of the school gate, so this position cannot well intercept a group of people. However, the classification

of teaching resources in some colleges and universities is unscientific, and related types of teaching resources websites generally only aim at teaching courseware and teaching materials, which cannot provide students and teachers with the resources they need to learn. Although the concept of artificial neural network was put forward by scientists when the computer was first born, it was considered unrealistic due to the limitation of computer hardware at that time. The loss curve and training accuracy curve base on convolutional neural network are shown in Figure 1.

At present, all colleges and universities have built campus networks and accumulated a large amount of online behavior data. Making full and effective use of these data can provide certain decision-making for the education and teaching of colleges and universities. In order to achieve this goal, some decision-making information related to education and teaching can be obtained by analyzing the characteristics of data mining and big data technology, so as to provide powerful decision support for school leaders and teachers in education and teaching management. With the rapid development of science and technology, the deep integration of big data and education will become an inevitable trend. Educational big data will promote the reform and development of traditional education. Big data technology can help people analyze education and teaching from different angles and levels and explore new laws of education and teaching. In the future, big data will set off a new educational revolution, such as innovating the ways and methods of students' learning, teachers' teaching, and educational policy-making.

4. Data Acquisition Process and Function in the Context of Big Data

4.1. *Data Collection Should Pay Attention to Comprehensiveness and Accuracy.* Only when the data of teaching evaluation are valid and accurate can teaching evaluation play its due role and effect. Therefore, it is necessary to change the result of linear calculation into the result of nonlinear function that fits the probability theory. In the classroom scene, teachers and students have one-on-one classroom teaching during the course, and cameras are used to record the whole classroom teaching process. After class, teachers can evaluate students' learning status in the course and adjust the teaching activities after class by watching the recorded videos. In the traditional classroom, teachers can communicate emotionally with students in many ways, such as one-on-one classroom questioning, interactive discussion, and eye contact between teachers and students. Teachers obtain students' emotional information in all directions, and emotional feedback is real-time and diversified. Especially in the face of some comprehensive problems that need to be solved manually, in addition to spending a lot of energy, it may take researchers decades to complete. Specifically, combined with the characteristics of mathematical models in the era of big data, indicators that reflect students' characteristics and behavioral performance are extracted from the data. Observe the correlation coefficient of the trained model and the relationship between the true and predicted values, as shown in Figure 2.

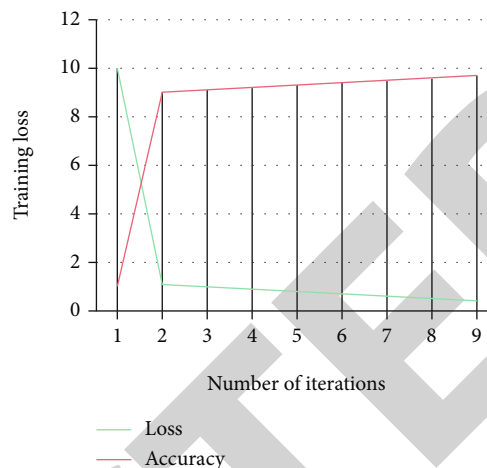


FIGURE 1: Loss curve and accuracy curve of convolution neural network.

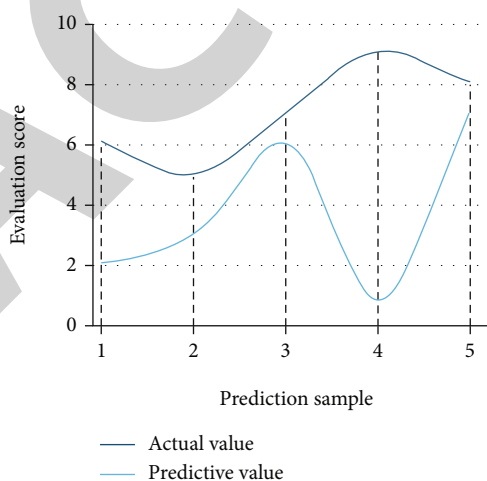


FIGURE 2: Comparison of teacher ability prediction results.

Establish a model to describe the relationship between these indicators and teaching effect, and then complete the evaluation of mathematical model scheme. In the information age, learning evaluation takes promoting students' development as the core concept of evaluation and takes students' learning process and learning results as the evaluation object. Taking the investigation of students' ability as the core content of evaluation, it emphasizes the diversified evaluation subject of the combination of teacher evaluation, peer evaluation, and self-evaluation. And show them the latest scientific and technological knowledge achievements of this era, expand college students' knowledge, and improve college students' professional quality and professional skills. These universities and related majors are increasingly unable to meet the development requirements of information technology. Static data and dynamic data are complementary to each other. Either data alone is not comprehensive enough. Only the combination of these two data can be considered accurate and comprehensive. It is a method to make decisions for complex systems that are difficult to quantitatively

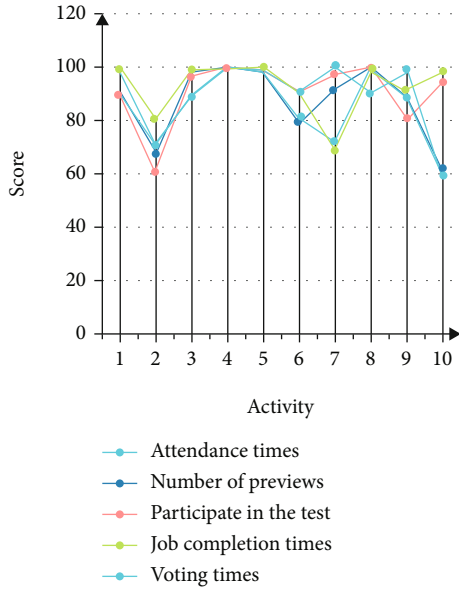


FIGURE 3: Visualization results of students’ learning attitudes and necessary learning activities initiative.

analyze. Originally, the analytic hierarchy process (AHP) is to decompose a problem into various components according to the goal it wants to achieve. The method of controlling the particle size of loop unrolling can completely expand the inner circulation without expanding the outer circulation, because the deeper the loop, the more times it will be executed when the system is running. Expanding the inner loop can significantly improve the running speed of the system. For example, if the examination result of a certain course is not satisfactory, it may be a teaching problem, and these problems will be reflected in the data.

4.2. *The Role of the Results of Teaching Evaluation.* Because teaching evaluation can timely understand and adjust the learning situation of students and teachers’ teaching situation, the importance of teaching evaluation is unquestionable, and the results of teaching evaluation can help teachers to adjust the teaching content. The difference between the convolutional neural network accelerators implemented by the two is that the slice granularity of loop unrolling is different, resulting in different hardware consumption and final performance of the convolutional neural network accelerator. This kind of space isolation can easily lead to the lack of students’ emotions. An open English learning study shows that in the process of big data, learners’ self-control is easily shaken, their adaptability to new learning methods is relatively weak, and their emotions fluctuate greatly. According to the existing activity data of students, the scores of students’ various learning activities are displayed, and the final scores of students can be obtained by weighting the scores of various activities. Some students’ activities are shown in Figures 3–6.

The main problem that deep learning is eager to solve is to organically combine the extracted simple features through automated methods to make them more complex features

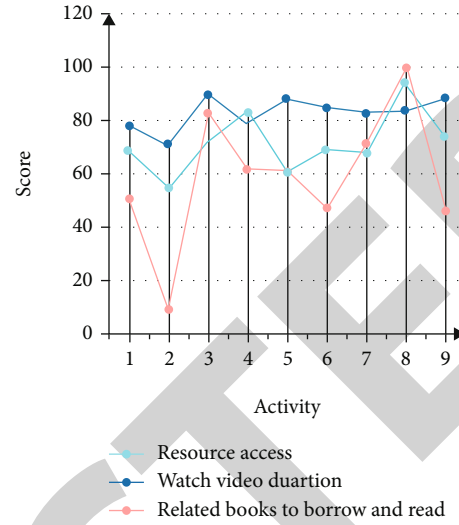


FIGURE 4: Visualization results of students’ learning attitude and self-directed learning activity initiative.

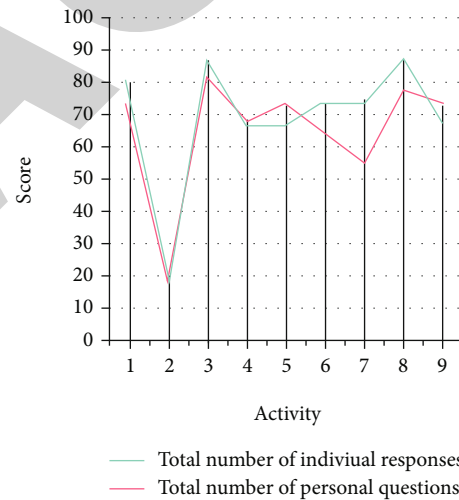


FIGURE 5: Visualization results of students’ learning attitudes and necessary learning activities initiative.

and use these combined features to solve problems. In addition, the more variables the model contains, the higher its complexity will be. If irrelevant variables are introduced into the model, it will unnecessarily increase the complexity of the model and is not conducive to a good interpretation of the model. The “process” of process evaluation is oriented relative to the “result.” Process evaluation is not an evaluation that only pays attention to the process but not the result, let alone simply observing the performance of students. The self-study classroom is for students to borrow and read. Students can read, extract, and mark without taking out the teaching evaluation, which provides convenience for students to read. At the same time, the teaching management department can use the results of teaching evaluation to solve hidden problems in teaching activities. At the same time, the teaching evaluation can also find out the situation

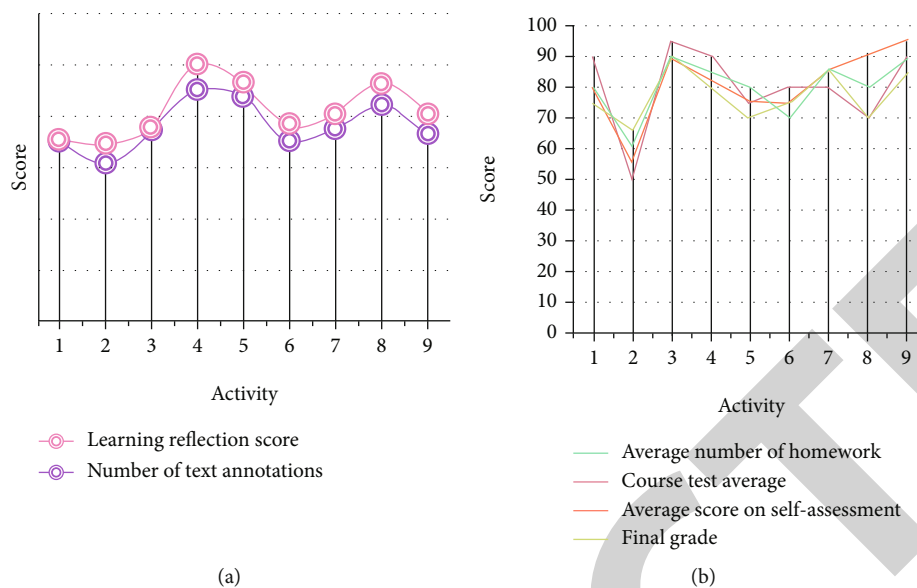


FIGURE 6: Visualization results of student content mastery.

in the evaluation indicators and can also understand and investigate this and put forward some suggestions for improvement. Promoting teaching is mainly in the transformation of hardware and investment in order to build a platform suitable for the development of students' basic skills of English majors, so as to meet the training of students' basic skills of English majors. In terms of the construction concept of language laboratory, the college has changed from the traditional construction of training places to the construction of modern independent learning center integrating teaching, self-study, and examination. Create an atmosphere in the environment and promote the construction of campus culture.

5. Conclusions

In the context of big data, Chinese colleges and universities can use information technology and big data ideas to carry out teaching evaluation and comprehensively use a variety of data mining and analysis techniques to mine and analyze teaching evaluation data and provide timely feedback. For example, through a large amount of data, the overall learning status of students can be monitored more comprehensively, to fully grasp the students' learning situation and to achieve a balanced and personalized teaching. The current online education system does not have the function of learners' emotion recognition, resulting in the phenomenon of "emotional lack" in online classrooms, which causes many learners to have no interest in learning and lack of concentration. When the learner uses the recognition program to carry out learning, the learning task can be completed in a shorter time, and the learning efficiency is higher. Mathematical models for building mathematical models in the era of big data are designed. Research the evaluation and formulation of mathematical model programs, make full use of educational big data information, and provide scientific sup-

port for educational decision-making. At the same time, it also helps teachers improve teaching and enhance teaching quality, so that students can better understand their learning level, and promote the reform of teaching evaluation in colleges and universities.

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 no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] E. Cunningham, "Artificial intelligence-based decision-making algorithms, sustainable organizational performance, and automated production systems in big data-driven smart urban economy," *Journal of Self-Governance and Management Economics*, vol. 9, no. 1, pp. 13–41, 2021.
- [2] T. Zhou, R. Law, and P. C. Lee, "“What motivates me?” Motivation to conduct research of academics in teaching-oriented universities in China," *Journal of Hospitality, Leisure, Sport & Tourism Education*, vol. 31, article 100392, 2022.
- [3] C. Wang, B. a. Jiang, and S. Zhong, "Research on the application of big data in the teaching quality evaluation system in colleges and universities," in *2021 13th International Conference on Education Technology and Computers*, New York, 2021.
- [4] Z. Xin, "Construction of college teaching quality evaluation system under the background of big data," *New Curriculum Research*, vol. 6, 2021.

Retraction

Retracted: The Evaluation Model of College English Diagnostic Exercises Based on Machine Learning

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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 H. Wu, "The Evaluation Model of College English Diagnostic Exercises Based on Machine Learning," *Journal of Function Spaces*, vol. 2022, Article ID 9185827, 12 pages, 2022.

Research Article

The Evaluation Model of College English Diagnostic Exercises Based on Machine Learning

Qi Wang ¹ and Hao Wu ²

¹Department of Public, West Anhui Health Vocational College, Lu'an, Anhui 237005, China

²Electromechanical Department, Anhui Chuzhou Technician College, Chuzhou, Anhui 239000, China

Correspondence should be addressed to Qi Wang; wangqi9828@ustc.edu

Received 2 June 2022; Revised 29 July 2022; Accepted 1 August 2022; Published 17 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Qi Wang and Hao Wu. 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 learning is an important way for college students to learn English independently. The evaluation information provided by the previous online teaching platform is more summative evaluation, which cannot make students have a more intuitive and comprehensive understanding of their English learning status and lack of personalized guiding suggestions. Therefore, this paper combines data mining technology with machine learning to build an English diagnostic exercise model that can analyze students' learning status, the correlation between knowledge points and question types, and predict English achievement, so as to provide students with more comprehensive analysis data information. The experimental results show that the evaluation model of college English diagnostic practice based on machine learning has the classification results of learning state with finer granularity, effectively analyzes the association rules of knowledge points and question types, and has high prediction performance. It can help students fully understand their English learning status, provide students with personalized analysis data and effective guiding suggestions, and enhance students' English application ability, improving CET-4.

1. Introduction

Since the reform and opening up, China has made quite gratifying achievements in English teaching. However, there are still some problems and deficiencies in our teaching mode and teaching methods. China does not pay enough attention to English teaching, especially the cultivation of English application ability of higher vocational students. There are phenomena such as “unable to understand and speak clearly,” and the students' English level is very low. The application ability of English in daily communication is seriously insufficient, and the basic function of English as a tool of interpersonal communication cannot be fully played. At present, the English teaching of higher vocational students in China focuses on exam-oriented education. Students spend a lot of time in memorizing words and reciting grammar but have little training in English listening and speaking. For a long time, it has developed into “deaf English” and “dumb English.” English proficiency and application ability is one of the necessary basic conditions for

many enterprises to recruit fresh college students. CET-4 and CET-6 are also one of the important examinations that college students need to pass before graduation. The results of CET-4 and CET-6 can largely reflect the English foundation and application ability of college students, as well as the effect and level of college English teaching [1]. College teaching and middle school teaching have different emphases. English teaching and examination pay more attention to the comprehensive application ability after mastering basic knowledge and skills, which is like a mountain in front of the door for students with poor English foundation [2]. English self-study and practice in spare time is particularly important for college students. The rise of online education and learning has opened up a new way of self-study for college students [3]. Compared with traditional college English teaching, online teaching and learning methods have many choices and wide resources. Students can arrange learning time and content according to their actual situation to meet students' personalized English practice requirements [4]. In addition, online teaching teachers can realize one-to-one or

one to a few students' teaching and have relatively sufficient time to solve and correct the problems and deviations existing in students' English learning [5]. E-learning has relatively high requirements for learners. Students need to clarify their learning objectives, fully understand their current learning status, and clearly plan their learning plans and contents [6]. Although many e-learning platforms can provide students with English performance tests, the whole e-learning framework process is different from the traditional teaching mode in essence. In most cases, only the status of obtaining scores, question type answers and the analysis of question gains and losses [7, 8]. Based on this, students can only understand their own gains and losses in a certain type of English questions but cannot master their own deficiencies in knowledge points and English skills and lack of targeted and guiding data information. The application of big data technology and machine learning technology can carry out data mining for the massive information data of online English learning system, enhance the analysis of English learners' personalized information and question type relevance information, provide English diagnostic evaluation and guiding suggestions, reduce detours in students' learning process, reduce learning effect deviation, and improve learning efficiency and effect [9].

The innovative contribution of this paper is to establish a diagnostic evaluation model of English in online English learning system by combining data mining, association rules, and random forest model. The model includes student learning state evaluation module, English information correlation analysis module, and English performance prediction module. The corresponding performance experiments, teaching application comparative experiments, and experimental results are analyzed. The evaluation model of college English diagnosis practice based on machine learning has a finer granularity of learning state classification results, effectively analyzes the association rules of knowledge points and problem types, and has high prediction performance. It can help students fully understand their English learning situation, provide students with personalized analysis data and effective guidance and suggestions, improve students' English application ability, and improve college English level.

2. Development and Research Status of Diagnostic Exercise Evaluation Model

The educational process is a circular process. Both educators and learners need to obtain evaluation information in the teaching process as the evaluation standard for the current teaching and learning effect [10]. Educational evaluation is phased and summative. Suggestions that can help students find and solve problems in the process of students' learning play the role of phased evaluation, and the evaluation of students' final achievements and achievements is summative evaluation [11]. There are differences between the two kinds of evaluation for students' learning guidance. The combination of the two can provide students with systematic evaluation information and help students comprehensively understand their own learning situation. Diagnostic evalu-

ation is to provide students with learning status analysis, learning effect feedback, problem analysis, supervision, and early warning of students' learning status and assist students in formulating effective learning plans [12, 13].

The development of network education has changed people's learning methods and ideas. Many learners who are not satisfied with the current situation continue to improve their abilities and expand their knowledge base through network teaching, which also makes the network education platform pay more and more attention to the data information services provided for learners. Diagnostic practice system has become the focus and hotspot of the research of network autonomous learning platform [14]. Some scholars proposed an adaptive learning system based on diagnostic evaluation, hoping that the system can evaluate the learning ability level and learning status of different learners [15]. Based on this, some scholars have carried out system optimization and established a learning system that can recommend exercises according to learners' current learning state. Learners can more truly show the learning effect through exercises with appropriate difficulty [16, 17]. Some companies have introduced the concept of diagnostic evaluation into teaching to form a new teaching model. Through the application experiment of teaching model, it is found that students are easier to show the advantages and disadvantages of their learning and mastery in this model. Educators provide personalized guidance according to students' specific conditions and strengthen students' weak links [18]. Other scholars have studied the computer adaptive test system according to the English learning situation of college students to provide diagnostic evaluation for college students [19]. With the diversified development of college students' needs for English evaluation, some scholars have studied the English learning system facing mobile devices to break the restrictions of learning conditions and realize personalized learning [20]. Other scholars have proposed that English learning is a gradual process, and leapfrog learning has poor effect on most students. Therefore, they classify English reading according to different difficulty levels and provide reading content with corresponding difficulty according to the specific situation of students [21]. At present, the research on English learning system with diagnostic evaluation function is still in the development stage, and the realization of many functions is not ideal. Personalized data analysis and targeted test paper generation module need to be further improved.

3. Construction of College English Diagnostic Practice Evaluation Model Based on Machine Learning

3.1. Overall Framework of Diagnostic Exercise Model. In the current college English evaluation system, there is often a phenomenon that students are only evaluated by their examination results. The evaluation of English teachers' teaching mainly focuses on the completion of the syllabus and students' examination results. This is not conducive to the improvement of students' comprehensive ability and

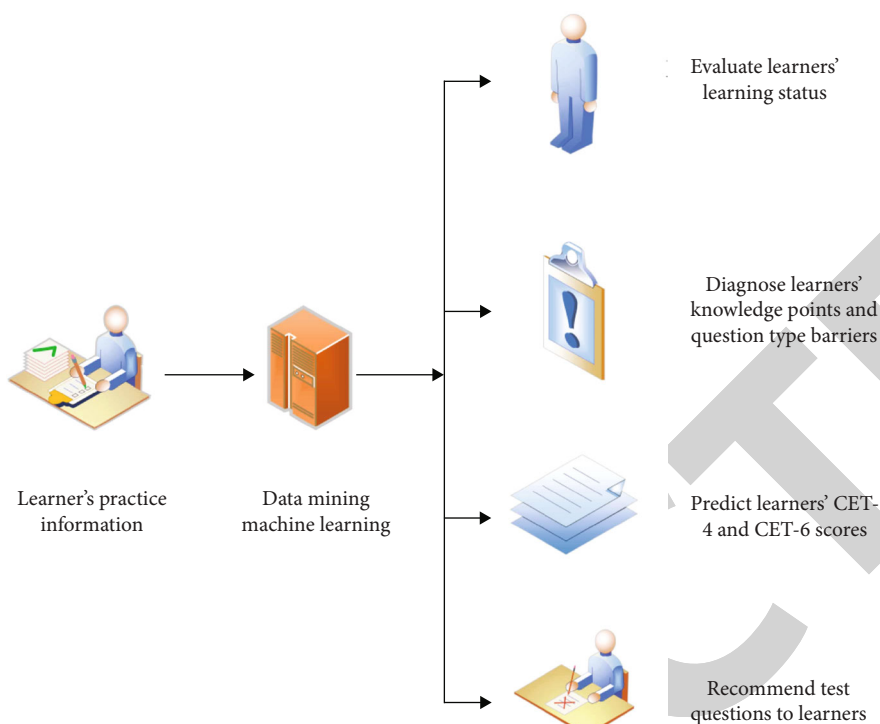


FIGURE 1: Block diagram of college English diagnostic practice evaluation model based on machine learning.

teachers' innovative teaching methods. In view of this phenomenon, we should build a diversified diagnostic system of college English, which is of great significance for reforming this unscientific and unfair teaching method. CET-4 and CET-6 are examinations to evaluate college students' English ability from many aspects, paying attention to students' basic English and comprehensive application ability [22]. However, different students have individual differences in learning ability, understanding ability, reading ability, and application ability; that is, each college student has different problems in English learning status, mastery level of knowledge points, exercise practice status, and so on. Therefore, the evaluation model of college English diagnostic practice should provide targeted, effective, and diversified English practice modes and guidance according to the actual situation of college students. However, the early diagnostic evaluation system still maintains the traditional teaching evaluation mode and ignores the actual state of college students, students can obtain less learning evaluation information and single guidance suggestions, and the direction of learning improvement is not clear [23]. Therefore, this paper introduces the evaluation factors of students' learning status into the model to help students understand their actual situation on the basis of evaluating the current learning ability of college students. The model also analyzes the correlation between English knowledge points and question types to predict students' scores in college English tests. According to students' practice data and feedback data information, focus on adjusting students' English learning arrangements, and pay attention to students' personalized factors. In addition, predict and compare the students' performance,

dynamically show the changes of students' state, analyze the negative factors, and achieve the purpose of supervision and early warning. Figure 1 shows the block diagram of college English diagnostic practice evaluation model based on machine learning.

As can be seen from the above figure, the workflow of the model mainly includes four stages. The first stage is to extract, process, analyze, and calculate the data related to college students' English tests and exercises. The second stage takes students as the core of data information analysis, makes a relatively fair and overall learning state judgment from the test score and students' learning stability, and sends early warning information to students with poor stability and slow overall learning process. In the third stage, analyze the knowledge points and question types of students' possible problems through data correlation and relevant algorithms, effectively find the shortcomings of students, and practice and improve their corresponding abilities. The fourth stage is to predict the score of CET-4 and CET-6 according to the existing student test and practice data. Based on this, students can comprehensively and intuitively understand their current mastery of CET-4 and CET-6.

3.2. Student Learning State Evaluation Module. Generally, the S-P table analysis method will be selected to evaluate students' learning status. The two-dimensional indicators for analyzing students' learning status are the scores of English knowledge points and question types and the classification status evaluation of students' attention coefficient. Let the probability that the n student gets a score be m and

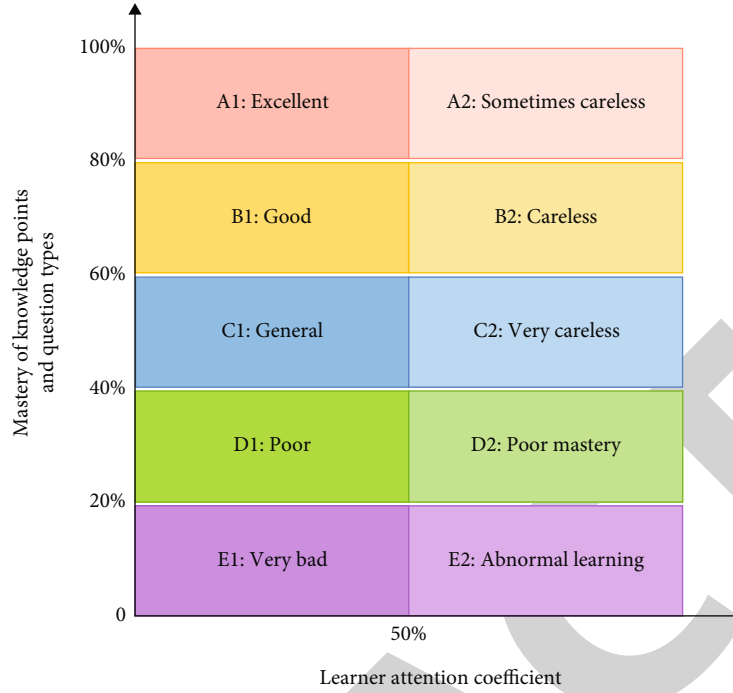


FIGURE 2: Results of students' English learning status score differentiation.

write y_{nm} , and the calculation of attention coefficient is as shown in

$$CS_n = 1 - \frac{\sum_{m=1}^a (y_{nm})(y_m) - (y_n)(\eta)}{\sum_{m=1}^a y_m - (y_n)(\eta)}, \quad (1)$$

where the total score obtained by students is y_n , the number of correct answers to knowledge points is y_m , the average value of correct answers is η , and the attention coefficient is CS_n .

The mastery status of knowledge points and question types is calculated according to

$$S_{\text{master}} = \frac{N_k}{N_t}, \quad (2)$$

where students' mastery status is recorded as S_{master} , N_k represents the number of knowledge point question types whose score rate is higher than the average, and N_t represents all knowledge point question types.

There are three evaluation results of S-P table analysis method, but such evaluation granularity is relatively large. In practical application, the classification of students' status is a little rough and inaccurate. Therefore, this paper refines the classification of students' learning state on its basis, as shown in Figure 2. The scoring results of refined learning state are divided into five types, each of which is divided into two cases. The first category of the level is when the students' learning state and attention coefficient remain stable at the current level. The second category is when the students' learning state is unstable, but the attention coefficient

remains at the current level. In this way, students' learning state can be divided into ten levels to increase the accuracy and personalized effect of evaluation.

The model stores the information of students' knowledge points and question type scores as matrix data and processes the data into a data form that can be used by the evaluation algorithm. Set the threshold value as the average score rate of knowledge points and question types. The data not less than this value is 1; otherwise, it is 0. Judge according to formula (3),

$$\text{Line}(x_{nm}) = \begin{cases} 0, & x < \frac{\sum_{n=1}^a x_{nm}}{a}, \\ 1, & x \geq \frac{\sum_{n=1}^a x_{nm}}{a}. \end{cases} \quad (3)$$

3.3. Knowledge Point Type Relevance Analysis Module. CET-4 and CET-6 test the comprehensive application of students' English knowledge and ability. There is correlation between English words, sentences, grammar, and other knowledge points, as well as between the test questions, which cannot be analyzed according to the independent situation. Therefore, this paper integrates data mining technology and association analysis technology into the analysis module for association analysis. Firstly, the massive original data is processed into effective analysis data; that is, the data due to invalid operation, lack of information, and abnormal errors are excluded, and the repetitive English practice results are screened out. The difficulty of English exercises in the model is different. The average score rate of a certain question type is taken as the judgment standard; that is, the

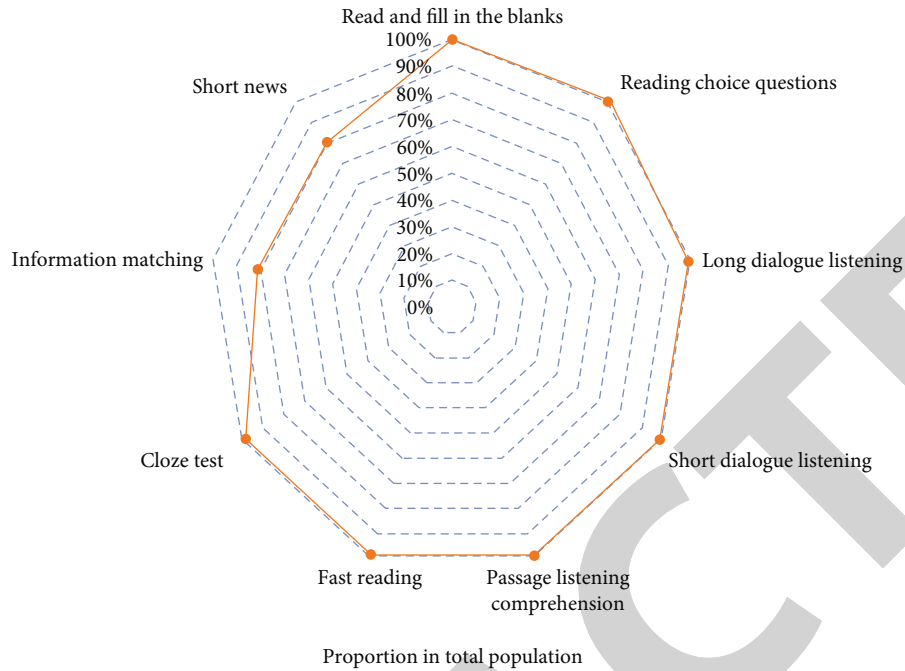


FIGURE 3: Comparison of the number of learners of each English test type before the update of English test type.

students' scores are divided into two types according to the standard, as shown in

$$r_{\text{correct}(nm)} = \begin{cases} F, r_{\text{correct}(nm)} < \overline{r_{\text{correct}(nm)}}, \\ T, r_{\text{correct}(nm)} \geq \overline{r_{\text{correct}(nm)}}, \end{cases} \quad (4)$$

where the probability of students' question type score is recorded as $r_{\text{correct}(nm)}$ and its average value is expressed as $\overline{r_{\text{correct}(nm)}}$.

The form and content of CET-4 and CET-6 in colleges and universities are not invariable. It is adjusted according to the actual needs and the overall quality of modern college students. Figure 3 shows the comparison of the number of learners of each English test type before the update of the English test type. The number in the figure shows that before the adjustment of the test type, some students were relatively weak in information matching and short news practice. The real English questions before adjustment are still of great help to students' practice. Therefore, it is necessary to deal with the English question type data in layers and analyze the data relevance, respectively.

There are two rule evaluation indicators for data association analysis. One is support, as described in

$$S(X \longrightarrow Y) = \frac{P(X \cup Y)}{P(I)}. \quad (5)$$

Another confidence is shown in

$$C(X \longrightarrow Y) = P(Y|X) = \frac{P(X \cup Y)}{P(X)}. \quad (6)$$

3.4. Students' English Score Prediction Module of CET-4 and CET-6. Hidden data may lead to the deviation of mining results, so data cleaning is particularly important. After the completion of data cleaning, a series of processes such as data integration, transformation, and specification are carried out at the same time. This process is data preprocessing. Statistically speaking, the hidden data may produce biased estimates, so that the sample data cannot well represent the population, and most of the data in the implementation contain missing values, so how to deal with missing values is very important. The processing includes two steps, namely, the identification of hidden data and the processing of hidden values. Commonly used methods include deletion method, replacement method, and interpolation method. After the preprocessing of the original data, the module needs to further mine the hidden features between the data, that is, the score of students' English questions and their gender. Students of different genders have different thinking performance, which has a great impact on their grades. The scoring rate is calculated as shown in

$$r_{\text{score}} = \frac{N_r}{N_t}. \quad (7)$$

The number of correct questions answered by students is N_r .

The prediction model of this paper selects the random forest model and combines it with multiple linear regression model to optimize the performance. Decision tree is the basic unit of random forest model. Its construction process includes decision tree generation and pruning. The information gain rate of decision tree is the factor

to be considered when selecting attribute division, and its formula is shown in

$$r_{\text{Gain}}(A) = \frac{\text{Gain}(A)}{S(A)}, \quad (8)$$

where the information gain is represented by $\text{Gain}(A)$ and the split information value is represented by $S(A)$, as shown in

$$\text{Gain}(A) = S_{\text{Info}}(D) - S_{\text{Info}(A)}(D), \quad (9)$$

$$S_A(D) = - \sum_{m=1}^U \frac{|D_m|}{|D|} \times \log_2 \left(\frac{|D_m|}{|D|} \right). \quad (10)$$

Random forest model adds the algorithm of random selection of attributes to the training of unit decision tree and completes the construction of unit decision tree with bagging integration. It shows the advantages of relatively easy implementation conditions and strong performance in practical application. The best classification index of random forest is selected according to the principle of Gini coefficient. The calculation formula is shown in

$$\text{Gini}(t_n) = 1 - \sum_{m=1}^2 [P(Y_m|t_n)]^2, \quad (11)$$

where the probability of dividing node t_n is expressed as $P(Y_m|t_n)$.

Bagging algorithm is derived from autonomous sampling method, which is used to enhance the generalization ability of unit decision tree and improve the prediction accuracy of random forest model. The random forest model completes the prediction and estimation through out of bag data, and its accuracy prediction is shown in

$$Q_n(x, y) = \frac{\sum_{n=1}^p I(h_n(x) = y, (x, y) \in O_n(x))}{\sum_{n=1}^p I(h_n(x), (x, y) \in O_n(x))}, \quad (12)$$

where the denominator represents the number of unselected samples and the numerator represents the number of correctly divided unselected samples.

The prediction results are shown in

$$Y_f = \arg \max_y \sum_{p=1}^{\text{ntree}} \delta(h(x, \theta_p) = y), \quad (13)$$

where ntree is the number of decision trees and $\delta(\bullet)$ is the indicative function.

The model generalization error formula and interval function are shown in

$$\text{PE}^* = P_{x,y}(\text{mg}(x, y) < 0), \quad (14)$$

$$\text{mg}(x, y) = \text{av}_p \delta(h_p(x) = y) - \max_{m \neq y} \text{av}_p \delta(h_p(x) = m). \quad (15)$$

In the formula, the average value of the correctly predicted sample is recorded as $\text{av}_p \delta(h_p(x) = y)$, the maximum value is recorded as $\max_{m \neq y} \text{av}_p \delta(h_p(x) = m)$, and the generalization error is represented as PE^* .

The expression of multiple linear regression equation is shown in

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_p x_p + v. \quad (16)$$

In the formula, the target variable is y , the variable is x_1, x_2, \dots, x_p , the parameter to be solved is $\alpha_m, m = 0, 1, 2, \dots, p$, and the random error is v . The parameters of the demand solution are estimated through the observation matrix vectors of the target variable and the variable, i.e., y_n and $x_{1n}, x_{2n}, \dots, x_{pn}$. If the estimated parameter value is recorded as $\hat{\alpha}_m, m = 0, 1, 2, \dots, p$, the multivariate linear sample regression equation is shown in

$$\hat{y}_n = \hat{\alpha}_0 + \hat{\alpha}_1 x_{1n} + \hat{\alpha}_2 x_{2n} + \dots + \hat{\alpha}_p x_{pn}, \quad (17)$$

where \hat{y}_n is the estimated value.

The difference equation between the estimated value and the actual value of the target variable is shown in

$$e_n = y_n - \hat{y}_n = y_n - (\hat{\alpha}_0 + \hat{\alpha}_1 x_{1n} + \hat{\alpha}_2 x_{2n} + \dots + \hat{\alpha}_p x_{pn}). \quad (18)$$

RMSE and RMSLE are the verification indicators of the prediction module, and their calculation formulas are shown in

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^a (x_{\text{obs},n} - x_{\text{mdl},n})^2}{i}}, \quad (19)$$

$$\text{RMSLE} = \sqrt{\frac{1}{i} \sum_{n=1}^a (\log(x_{\text{obs},n} + 1) - \log(x_{\text{mdl},n} + 1))^2}. \quad (20)$$

In the formula, the predicted value represents $x_{\text{obs},n}$ and the actual value represents $x_{\text{mdl},n}$.

4. Experimental Results of College English Diagnostic Practice Evaluation Model Based on Machine Learning

There are some differences between machine learning and deep learning. First, the application scenarios are different. The application of machine learning in fingerprint recognition, feature object detection, and other fields has basically met the requirements of commercialization. Deep learning is mainly used in word recognition, face technology, semantic analysis, intelligent monitoring, and other

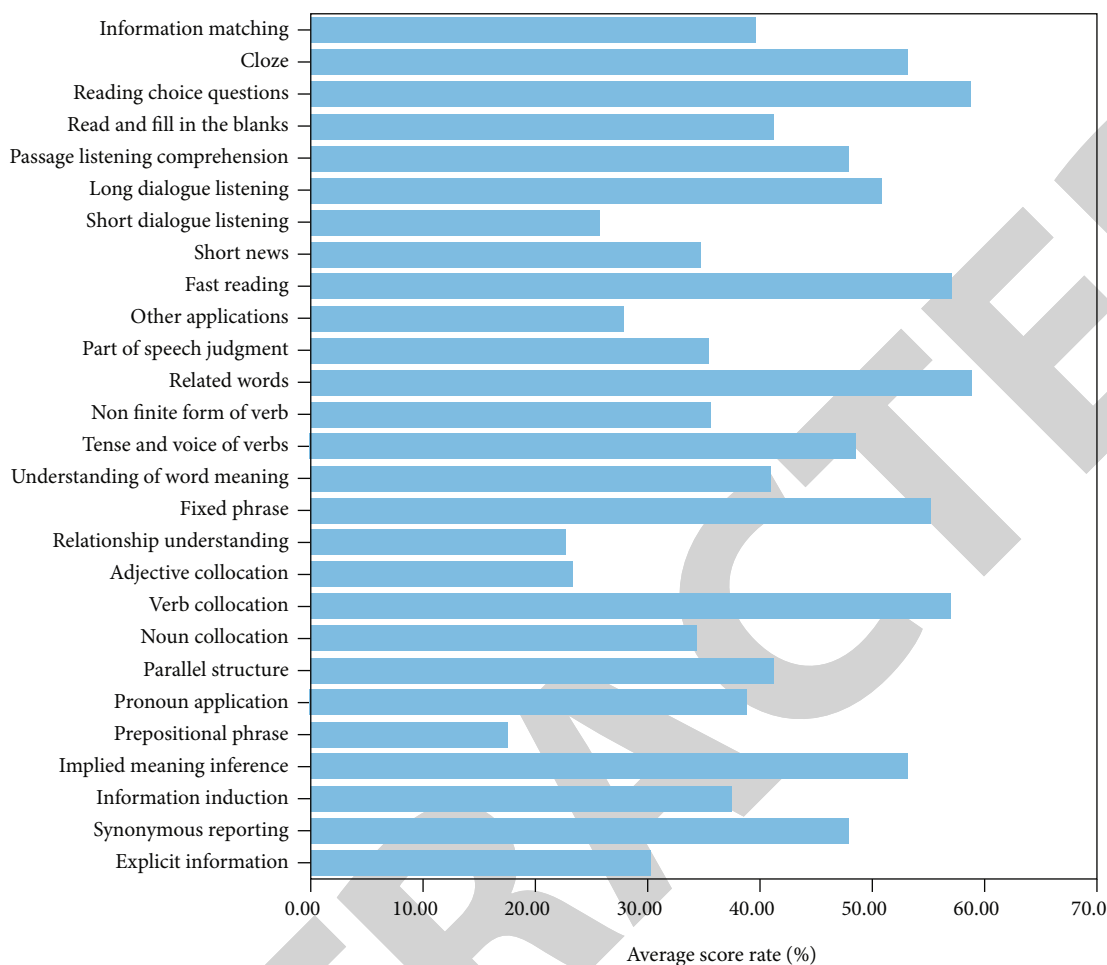


FIGURE 4: The average score of all students participating in the experiment before the experiment.

fields. At present, it is also rapidly deployed in intelligent hardware, education, medical, and other industries. Secondly, the required amount of data machine learning can adapt to a variety of data, especially in scenarios with a small amount of data. If the amount of data increases rapidly, the effect of deep learning will be more prominent, because deep learning algorithms require a lot of data to be perfectly understood. The execution time is different. The execution time refers to the amount of time required by the training algorithm. Generally speaking, deep learning algorithm requires a lot of time to train. This is because the algorithm contains many parameters, so it takes longer time to train them than usual. Relatively speaking, machine learning algorithms take less time to execute.

The application of the evaluation model of college English diagnostic practice based on machine learning was aimed at evaluating students' current English ability, providing targeted and personalized suggestions for students, and improving their English test scores. Therefore, this paper selects two classes of unified major in a university for a comparative experiment of English performance. Class one is an experimental class; that is, the model of this paper is introduced into English teaching. The second class is the control

class, whose English scores before the experiment are similar to those of the experimental class, and both classes are CET-4-nonassessed classes. Before the comparative experiment, the performance of the three modules of the college English diagnostic practice evaluation model based on machine learning is tested to ensure the stability and reliability of the comparative experiment. Take the exercises or homework that students usually do as the training data and the examination organized by the teacher as the verification data, and the final exam we take is called the test data. These three parts of module data are all existing data, in which the training data is only used for training. The goal of students (learning algorithms) is to continuously improve performance on this part of data. The teacher will use verification data to monitor students' learning and then adjust students' learning methods. When the teacher thinks that the student can no longer improve, he uses the test data to simulate the unknown data to conduct the final performance test on the student.

As shown in Figure 4, it is the average result of the score rate of English knowledge points and question types of all students participating in the experiment before the experiment. The data results show that the distribution range of the score rate of students participating in the experiment is

	Fast reading	Short news	Short dialogue	Long dialogue	Passage listening	Banked cloze	Cloze	Information matching	Reading choice
No.1	F	F	F	F	F	F	F	F	F
No.2	F	F	F	T	T	T	F	T	F
No.3	T	T	F	F	F	F	F	F	T
No.4	F	T	F	F	T	F	T	T	F
No.5	F	F	T	T	F	F	T	F	T
No.6	F	F	F	F	F	T	F	T	F
No.7	F	F	F	F	F	F	F	F	F
No.8	T	F	T	F	F	F	T	T	F
No.9	F	T	T	T	T	T	F	T	F
No.10	F	F	T	F	T	T	F	T	F

FIGURE 5: The result of binary Boolean of the score rate of ten students' CET-4 questions.

relatively concentrated, mainly between 30% and 59%. If students are classified according to the traditional S-P table analysis method, the results obtained are too concentrated, which is not conducive to the further analysis of students' learning state. The refined classification table can more refine the classification of students and better show the differences of student groups.

Ten students are randomly selected from the experimental class, and the binary Boolean result of the score rate of level 4 question type is shown in Figure 5. The results showed that before the comparative experiment, most of the ten students did not reach the average value, and even some students scored lower than the average value.

Further analyze the relevance of the judgment data in Figure 5, and analyze the relevance of the question types with high and low scoring rate under the same setting conditions. The analysis results are shown in Figure 6. Two layers are included in T and F analyses, and the minimum reliability is 90%, excluding the data with low probability.

Select the one with the highest confidence and support of association rules in the four analysis layers of T and F, as shown in Figure 7. The support degree of association rules reflects the probability that the front and back association items, i.e., item a and item b exist at the same time, and the confidence degree indicates the possibility that item a also exists within the existence time of item b. It can be seen from this that in T analysis, when the three question types of information matching, rapid reading and reading comprehension all reach the average standard at the same time; the probability is 41.89%. When the first two question types are not less than the average value, the probability of reading comprehension score rate not less than the average value is

96.52%. By further analyzing the relationship between reading comprehension and information matching, the probability that the scores of the two types are not less than the average at the same time is 46.85%. When the score of information matching question type is greater than or equal to the average, the possibility of reading comprehension also reaching this condition is 96.04%. In the F analysis, the first level analyzes the correlation among short passage, long dialogue, and short dialogue listening questions. The probability that the scores of the three questions are lower than the average score standard is 45.97%. When the scores of the first two types of questions are lower than the average standard at the same time, the probability of short dialogue listening comprehension is also lower than the average standard, up to 99.57%. In the second layer, the probability that the scores of reading selection, word filling, and information matching cannot meet the average standard at the same time is 32.89%. When the scores of the first two types do not meet the average standard, the probability that the scores of information matching cannot meet the condition is 95.81%. To sum up, among the CET-4 and CET-6 question types, the question types belonging to the same category have high internal relevance. When students have low accuracy of some question types in the same category, the probability of low accuracy of other question types in this category is very high. In addition, the relevance of the same type of questions is high, which reflects that no matter how the CET-4 and CET-6 question types are updated, students can be familiar with the changes of new questions and enhance their English application ability through the practice of changing the former question types.

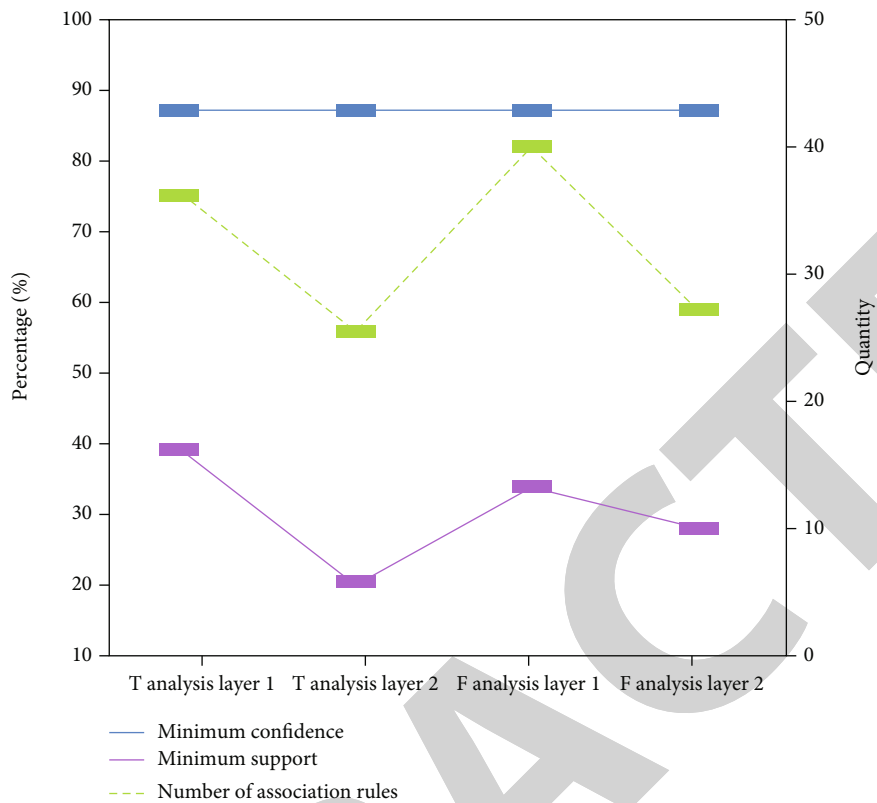


FIGURE 6: Model association rule analysis output.

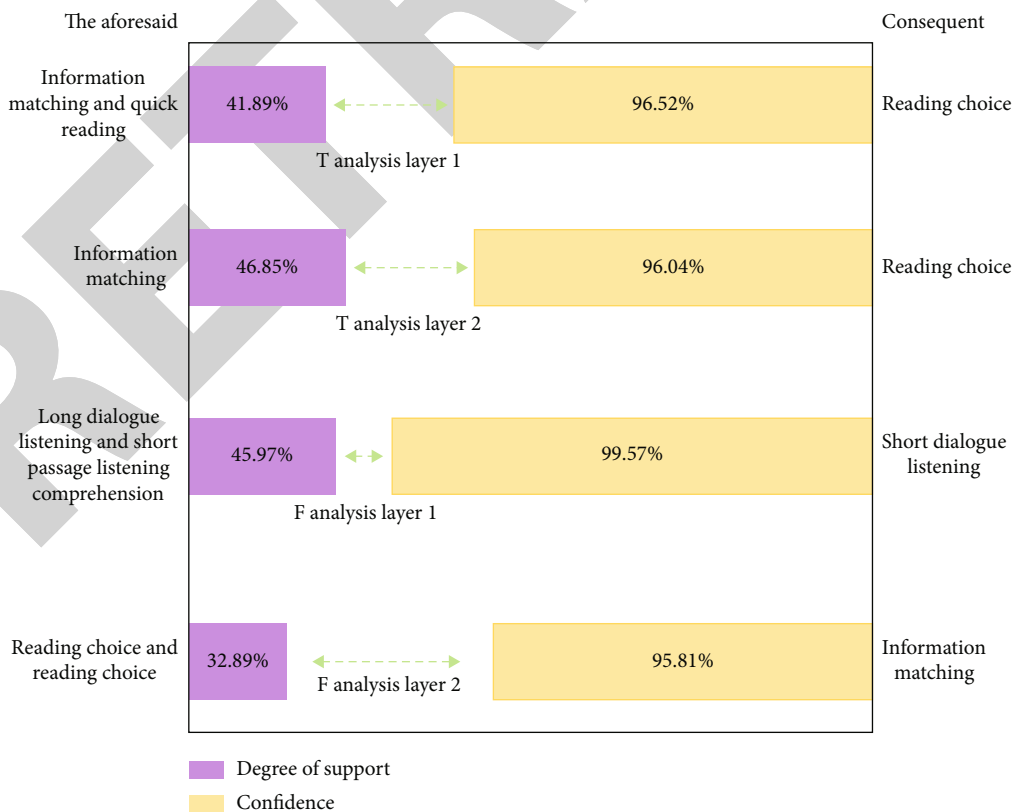


FIGURE 7: The association rule results of question types with the highest confidence and support among all analysis levels.

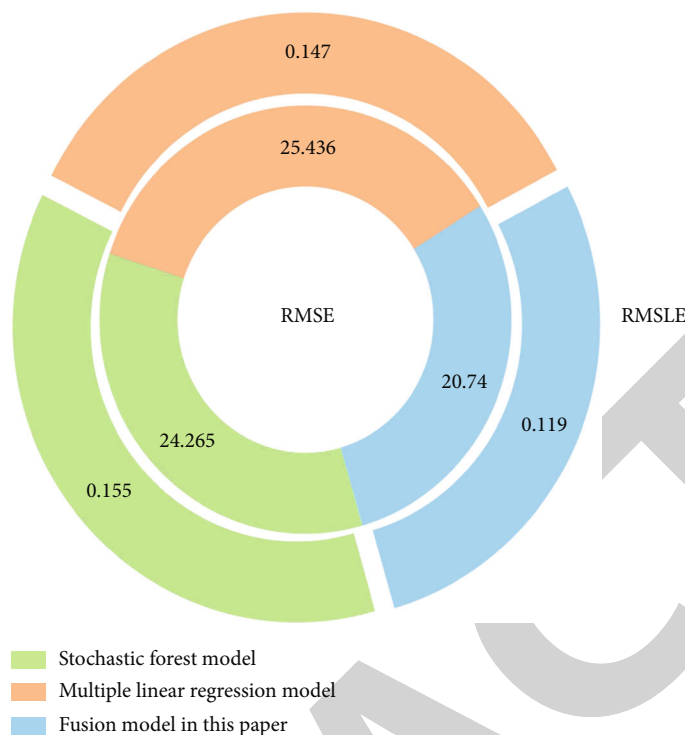


FIGURE 8: Comparison results of RMSE and RMSLE evaluation indexes of the three models.

Figure 8 shows the comparison results of RMSE and RMSLE evaluation indicators of the three models. The comparison results show that the prediction module is only established based on random forest model or multiple linear regression model, and there is little difference between the evaluation index results of the two. After the fusion of the two models, the evaluation index values of the model change obviously, which shows that the model in this paper has a great improvement in the two indexes and enhances the prediction accuracy of the model.

According to the above experimental results, the evaluation model of college English diagnostic practice based on machine learning can divide students' current English learning state in a fine granularity, better realize the correlation analysis between knowledge points and question types, and provide more accurate English performance prediction results for the comparative experiment. In the comparative experiment, the experimental class and the control class will be tested on CET-4 before the experiment, that is, before school and after the experiment, that is, at the end of the semester. In addition to the zero CET-4 test in the experimental class, the college English diagnostic practice evaluation model based on machine learning also predicts the students' English test scores before the second test. The results are shown in Figure 9. The results in the figure show that before the experiment, the English scores of the two classes are basically the same, and most of them are between 280 and 390, which shows that the application ability of CET-4 of the students in the two classes is weak, and most of them still fail to meet the passing standard. After the experiment, the results of the students in the experimental class have greatly improved the results of the students in

the control class, the number of students with low scores has been significantly reduced, and the number of students with up to standard scores is also higher than that in the control class. In addition, there is a close relationship between the actual score of CET-4 in the experimental class and the predicted score of college English diagnostic practice evaluation model based on machine learning, and there is a relatively large error between the actual score and the predicted score of only a few students. This shows that the model can effectively help students analyze the current English learning state, provide effective suggestions and help for students according to the analysis data, enhance students' weak English application ability, systematically practice English question types, and quickly improve students' learning efficiency and adaptability to different question types. The results of the prediction model fit the actual results, more intuitively reflect the students' English learning progress and overall level, promote the students to check and make up the deficiencies, strengthen the students' self-awareness, and achieve the purpose of improving their English performance.

Figure 10 shows the evaluation results of the experimental class students on the evaluation model of college English diagnostic exercises based on machine learning. The evaluation content includes the students' satisfaction with model diagnostics, learning state analysis, knowledge point diagnosis, and question type diagnosis, with a full score of five. The evaluation results show that the average performance scores of the model are more than 3.5 points, indicating that most students are satisfied with the application performance of the model, which indirectly reflects that the diagnostic effect of the model can meet the needs of most students and

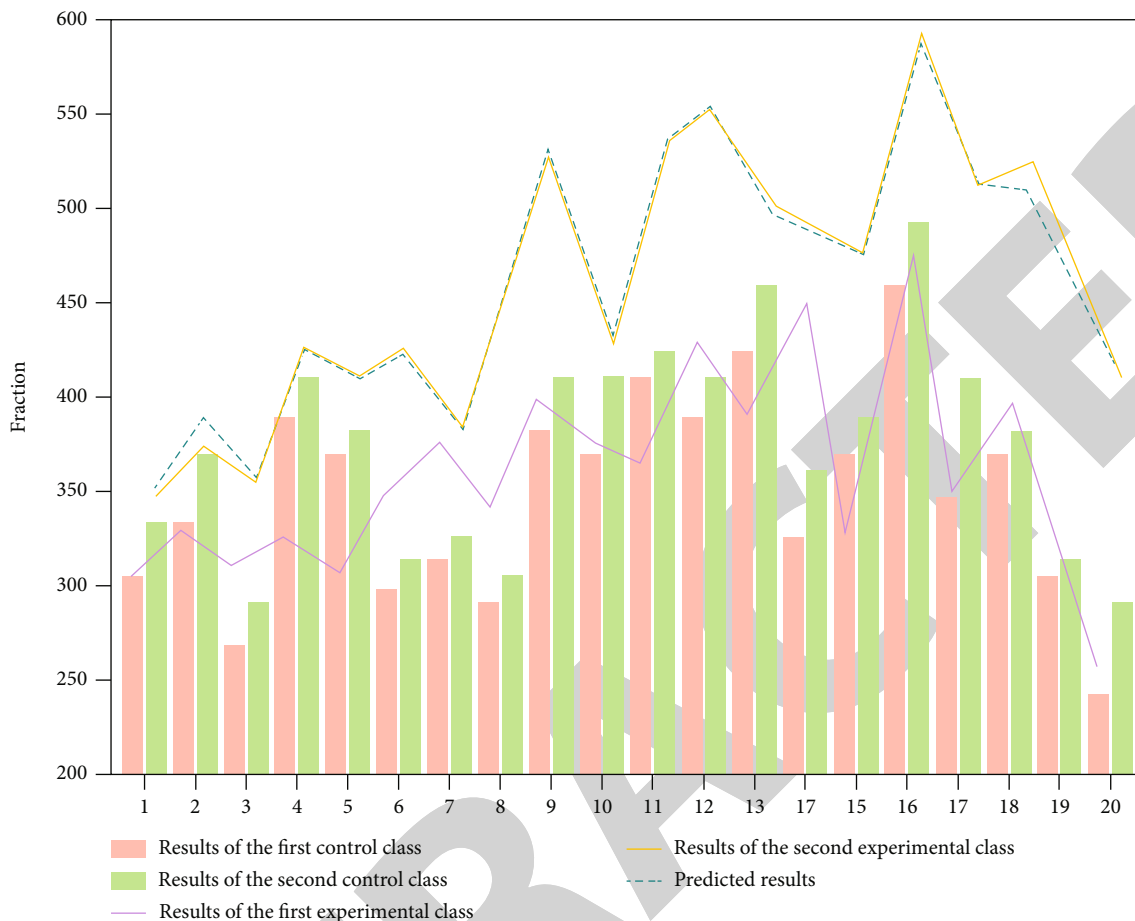


FIGURE 9: The application of test scores before and after the experiment in the experimental class and the control class and the student model in the experimental class predict the results of English scores.

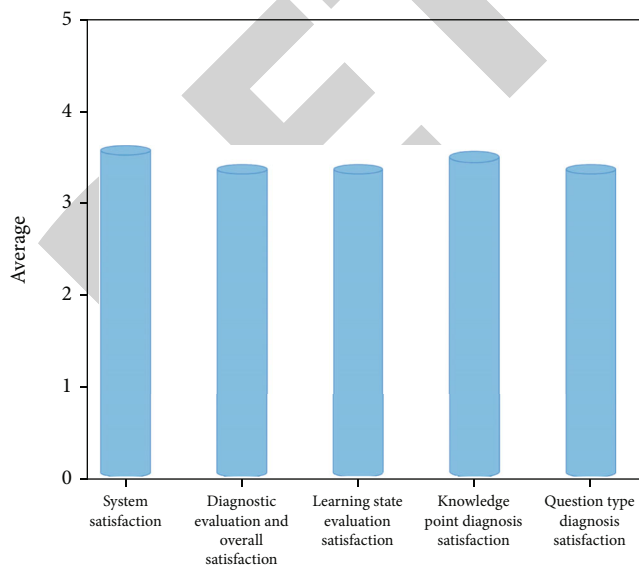


FIGURE 10: The evaluation results of students in the experimental class on the evaluation model of college English diagnostic exercises based on machine learning.

achieve the purpose of the experiment in practical application.

5. Conclusion

This paper combines data mining, association rules, and random forest model to build an English diagnostic exercise model to analyze students' English learning status, knowledge mastery, problem type correlation, and English performance prediction. The experimental results show that the evaluation model of college English diagnosis practice based on machine learning can refine the classification of students' learning status and show the differences of different students' English learning status. The evaluation index of machine learning based college English diagnostic practice evaluation model performs well and has high accuracy. The results of teaching comparison experiments show that the model can help students understand their English learning as a whole, better plan their learning plans, and improve their learning efficiency. Correlation analysis found that the problem types and knowledge points belonging to the same category have strong correlation, and the gain and loss points have a greater impact. The research provides students with intuitive learning state feedback information through English score prediction close to the actual value and realizes

Retraction

Retracted: Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Shen, "Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model," *Journal of Function Spaces*, vol. 2022, Article ID 1078057, 9 pages, 2022.

Research Article

Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model

Shuli Shen 

School of Public Administration, Zhejiang Gongshang University, Hangzhou 310018, China

Correspondence should be addressed to Shuli Shen; shenshuli@mail.zjgsu.edu.cn

Received 7 June 2022; Revised 30 July 2022; Accepted 2 August 2022; Published 17 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Shuli Shen. 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.

China's real estate market is developing rapidly, but the house price is abnormal. The nonlinear relationship between housing characteristics and real estate value is difficult to calculate, resulting in the difficulty of house price prediction. Based on this, the relationship model between characteristic variables and house prices is constructed by using the machine learning method. At the same time, genetic algorithm is used to screen the specific values. The experimental results show that the optimized model converges in 56 iterations; in the application test, the research model found that in 90% of the test samples, the error between the predicted value and the actual measured value shall not exceed 10%. Experiments show that the genetic algorithm is effective in optimizing the BP house price valuation model and improves the calculation efficiency and valuation accuracy of the valuation model.

1. Introduction

Real estate economic activities consist of real estate production, real estate circulation, and real estate consumption. Among them, real estate production is the direct production process of real estate, and real estate circulation is the reproduction and realization process of real estate. Real estate consumption is the realization of the purpose of real estate production and the continuation of some direct production processes in the consumption link. The circulation of real estate production, circulation, and consumption needs corresponding economic operation mechanism and economic system conditions. The real estate economy can be investigated from three aspects: microeconomics, meso-economy, and macro-economy. All economic behaviors taking real estate enterprises as economic units and the operation of real estate development and management projects belong to the micro-level. The real estate sector economy and the real estate regional economy basically belong to the meso-economy. As a kind of industrial economy, the real estate sector economy has many characteristics and special laws beyond its commonness. Real estate regional economy mainly refers to the real estate economy of each province,

city, and district, which belongs to a department of regional industrial structure. Considering the whole national economy as an object belongs to the macro level. The real estate economy is not only the object of national macro-control but also an important part of the macro-economy. Real estate not only includes the material form of buildings and land but also has legal and economic characteristics due to the real rights generated by entities and economic activities generated by transactions. Real estate value is generally the sum of building value and land value, but due to its legal and economic characteristics, region, population, consumption psychology, policy, and even buyer gender have also become important factors affecting the change of real estate value [1]. At the same time, real estate transactions often drive the development of other investment and trading markets, increasing the risk of market professional portfolio, especially when purchasing investment real estate and using extensive leverage. Therefore, maintaining the healthy and stable market is an important measure for China's economy and people's livelihood [2]. The value of real estate is ultimately reflected in the price paid by the real estate demander to the supplier. Therefore, building a valuation model based on predicting the transaction price of real estate is a common evaluation method in the trading market.

Traditional asset appraisal has objective disadvantages in terms of data volume and appraisal funds. Asset appraisers have limited ways to obtain information and data and can only consult and learn through field investigations or cases filed by their own companies. Moreover, the traditional asset evaluation methods inevitably have a lot of subjectivity, which has a certain impact on the objectivity and fairness of the evaluation results. How to eliminate the subjectivity in the evaluation process to the greatest extent and establish an objective evaluation system to the greatest extent is also a major requirement of the economic and social development in the new era. The research method of this paper is to use a specific algorithm to confirm the predicted value. Finally, the algorithm value after the selection is obtained, design the fitness function to specify the iteration scale of the algorithm, and finally classify and input various indicators affecting house prices. Through the above optimization, the real estate valuation model is constructed.

The innovation of the research is that it not only considers the use of genetic algorithm to improve the fitness function design of valuation model and the parameters of neural network threshold but also inputs the characteristic variables affecting real estate value. Before the experiment, the research classifies various characteristic factors into location indicators and building indicators and quantifies and normalizes these variables, so as to avoid the negative impact on the model caused by the non-uniform unit and large difference of numerical magnitude between variables.

The second part briefly describes the theoretical achievements of domestic and foreign researchers on real estate market valuation and investment in recent years, as well as the application experience of genetic algorithm and neural network in different industries. In the third part, based on the characteristic variables of real estate transactions, a neural network model for valuation is proposed, and the optimization process of genetic algorithm for neural network is studied. The last part is the training and testing experiment of the model. The effectiveness of the experimental results is analyzed by using algorithm tools. In terms of measurement accuracy of experimental results, different models are judged differently, and different results are obtained.

2. Related Works

In recent years, although the real estate market is in the development trend of continuous expansion, the real estate price has been in an unstable state due to the influence of real estate speculation, investment, and other factors. Therefore, more and more people begin to pay attention to the real estate market, and professional scholars have put forward different views on the prediction of real estate and house prices, and studied buyers' expectations and investment potential for sustainable residential real estate by capturing four variables: economy, society, environment, and system. The results show that countries adhere to the sustainable development goal (SDG) as the economic strategy. Due to the changing market environment, this paper makes a highly sensitive detection of the powerful environment [3]. Dobrowskien et al. constructed an evaluation model of real estate

sustainability index (resi) at the technical level based on the multi-criteria decision-making method. Experiments show that the model can promote the consideration of sustainable investment in new technologies [4]. Hodoshima et al. analyzed the real estate investment trust market in Tokyo based on four performance indicators: internal risk aversion rate (Irta), sharp ratio, Sotino et al. found that Irta is more sensitive to potential risks and more relevant to risk averse investors [5]. It is found that the total income rate of return initially fluctuated around 5%, but it showed a downward (upward) trend. The results show that the annualized actual total return after deducting costs ranges from about 2.3% of residential real estate to 4.5% of agricultural real estate [6]. D'acci quantifies regional characteristics based on market comparison method. The relationship between factors in different positions was studied. After calculation, the results show that there is a great relationship between the increase of land price and distance [7]. Ulbl et al. constructed the batch valuation model of apartment market based on the generalized additive model method. Experiments show that the model can accurately analyze the impact of apartment market heterogeneity on apartment transaction price [8]. Wang et al. introduced the application trend and classification of batch evaluation according to the group evaluation models and methods from 2000 to 2018 and emphasized the 3I trend. Many experiments in the past have shown that different spatial expressions are very different based on different Internet environments [9]. He et al. designed and described a method of generating finite difference (rbf-fd) based on multi quadratic radial basis function of local RBF format. The results show that rbf-fd method is effective and stable in terms of pre-real estate valuation accuracy [10]. Sabina et al. applied fuzzy logic theory to estimate the land value of agricultural land market. The results show that the real estate is in the process of initial development. The research method of this paper has certain utility. This increases the opacity of the market [11].

Chen et al. prove the accurate prediction effect and stable operation performance of the model [12]. Wang et al. measured the relationship between pupil response and objective pain based on the artificial neural network machine learning method optimized by genetic algorithm. The results show that pupil response and machine learning algorithm are a promising objective pain level evaluation method, which can improve the patient's experience of measuring pain in long-range medical care [13]. Xie et al. believe that genetic algorithm does not guarantee that you can obtain the optimal solution of the problem, but the biggest advantage of using genetic algorithm is that you do not have to understand and worry about how to "find" the optimal solution. As long as you simply "deny" some individuals who perform poorly. This is the essence of genetic algorithm [14]. Jiang et al. taking the gearbox of a tractor as the engineering background, the BP neural network based on genetic algorithm is used to diagnose the gearbox fault. Statistics show that about 60% of gearbox faults are caused by gearbox faults, so only gear faults are studied here. Here, several characteristic quantities in the frequency domain are selected. Gear faults in the frequency domain are more obvious in the edge band

at the meshing frequency [15]. Combined with the analysis of the corresponding training results, using genetic algorithm to optimize the connection weight of BP neural network is more effective than the traditional algorithm, but it also needs to try the network structure. In other words, it is to select the appropriate number of hidden layer neurons [16]. Moreover, using genetic algorithm to optimize the weight and structure of neural network at the same time is more intelligent. It is possible to find the appropriate initial weight and excellent network structure. However, if the data is cumbersome, the search speed will slow down [17].

To sum up, relevant scholars have a variety of research methods on real estate valuation and housing market investment. Based on the different development of real estate market in different regions, the valuation methods adopted also change accordingly. In some foreign regions, the real estate transaction information is not transparent on the Internet or the network transaction is not developed. Therefore, the methods of some foreign scholars are still in the market comparison method of collecting offline transaction information, while China's transaction market has tended to be offline and online parallel. Therefore, the evaluation method also needs to be improved. The model of genetic algorithm optimized neural network has been shown in medicine, geological survey, engineering, and so on. In view of this, this paper studies the nonlinear machine learning method based on BP neural network optimized by genetic algorithm and discusses the impact of location variables and construction variables on house prices, hoping to provide a stable valuation idea for China's real estate trading market.

3. BP Neural Network House Price Valuation Model Based on Genetic Algorithm

3.1. Establishment of BP Neural Network Algorithm Model. There are three traditional real estate valuation methods: Among them, market comparison method is the most widely used because of the frequent transactions. The marketing mode of the real estate industry has also changed from offline dominance to online and offline combination. The traditional market comparison method also has too large deviation in valuation results due to personnel interference and scattered transaction information. Therefore, genetic algorithm (GA) is introduced to optimize the back propagation (BP) neural network threshold. The nonlinear machine learning method of BP neural network is used to replace the traditional calculation to establish the general model of real estate value evaluation [18]. The basic topology of BP neural network is shown in the figure below.

As shown in Figure 1, the weight optimization among the three depends on machine learning adjustment and is not affected by personal experience [19]. At the same time, the nonlinear mapping functions so BP neural network is used to construct house price evaluation model. Firstly, describe the forward propagation model in the network structure. Set the hidden layer nodes as n , and the calculation formula of hidden layer output is shown in formula (1).

$$z_b = f_1 \left(\sum_{a=0}^m v_{ba} \cdot x_a \right). \quad (1)$$

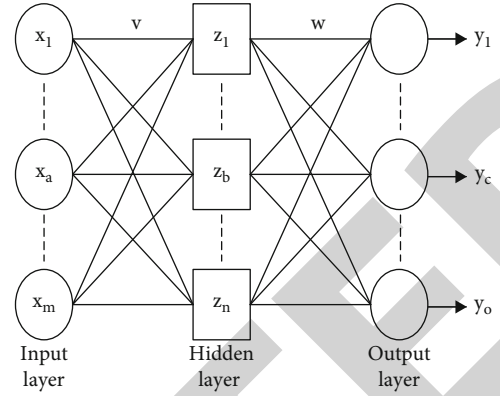


FIGURE 1: Structure and parameter of BP neural network.

In formula (1), x_a represents the a input layer parameter, z_b is its corresponding hidden layer parameter, f_1 is the hidden layer transfer function, indicating the transfer after mapping the input layer variables of weighted x , and v_{ab} indicates the weight between input layer x_a and hidden layer z_b . Similarly, the output formula of the output layer node is shown in equation (2).

$$y_c = f_2 \left(\sum_{b=0}^n w_{cb} \cdot z_b \right). \quad (2)$$

In equation (2), y_c is the calculated output layer signal, $c = 1, 2, 3 \dots o$ and f_2 are the output layer transfer function, and w_{cb} is the weight between the hidden layer z_b which is defined by mean square error, which is defined as the following equation (3).

$$E_g = \frac{1}{2} \sum_{c=1}^o (t_c^g - y_c^g)^2. \quad (3)$$

In equation (3), E_g indicates the error between the trained output of the g learning sample and the expected output t . According to the error function, the global error formula is equation (4).

$$E = \sum_{g=1}^G E_g = \frac{1}{2} \sum_{g=1}^G \sum_{c=1}^o (t_c^g - y_c^g)^2. \quad (4)$$

In equation (4), E represents the global error of all G learning samples, and the weight is adjusted according to the obtained error. The model of weight adjustment of hidden layer is described in equation (5) below.

$$\Delta v = \sum_{g=1}^G \sum_{c=1}^o \eta (t_c^g - y_c^g) f_2'(S_c) w_{bc} f_1'(S_b) x_a. \quad (5)$$

In equation (5), Δv represents the adjusted value of the hidden layer weight, η represents the learning efficiency constant, f_1', f_2' represents the initial transfer, S_c represents the net input of the c node of the output layer, and its calculation

formula is $S_c = \sum_{b=0}^n w_{cb} \cdot z_b$; S_b represents the net input of the b , and its calculation formula is $S_b = \sum_{a=0}^m v_{ba} \cdot x_a$. Finally, the neuron weight adjustment formula of the output layer is shown in equation (6).

$$\Delta w = \sum_{g=1}^G \sum_{c=1}^o \eta (t_c^g - y_c^g) f_2'(S_c) z_b. \quad (6)$$

3.2. Construction of GA-BP Real Estate Valuation Model. This model's advantage is the treatment of the nonlinear relationship between influencing factors and results, but its disadvantage is that it is easy to fall into local minimization and the calculation convergence speed is slow. The former problem comes from the randomized initial weight of machine learning, and the second problem is the sawtooth phenomenon common to gradient descent method [20]. The genetic algorithm and genetic algorithm are introduced to optimize the structure of the two BP networks, as shown in Figure 2.

As shown in Figure 2, encode all individuals of the population in the initial solution space, set the definition domain of variable x as $[j, k]$, and the accuracy requirement is 10^{-5} , then the definition domain $[j, k]$ can be divided into $(k - j) \times 10^5$ equal length intervals, all intervals are represented by binary strings, and the length of the string is set as l , then each point in the definition domain can be represented by $(k_{l-1}k_{l-2} \cdots k_0)$. After the conversion from the above real variable space to binary bit string is completed, convert binary string into decimal system, and its formula is shown in formula (7).

$$\begin{cases} x' = \sum_{i=0}^{l-1} k_i \times 2^i \\ x = j + x' \frac{(k-j)}{(2^l - 1)} \end{cases} \quad (7)$$

The first line of equation (7) is the decimal formula for binary string conversion, and the second line is the calculation formula for the value of the variable. Based on binary coding and supplemented by gray coding, the mathematical expression of its transformation is shown in equation (8).

$$j_i \begin{cases} k_i, & i = 1 \\ k_{i-1} \oplus k_i, & i > 1 \end{cases} \quad (8)$$

In equation (8), j_i represents gray string corresponding to binary string k_i , and \oplus symbol represents modulo two addition operation. Because the individual's adaptability to the environment determines whether the individual can be retained in the new group, the fitness is the basis for forming optimization.

$$Fun = \sum_{i=1}^N a_i \cdot E_g + a \cdot E. \quad (9)$$

In equation (9), Fun is the fitness function constructed by the weight coefficient method, a is the proportion of all sample errors in the fitness function, and a_i is the proportion of each

sample in the fitness function. The third step of genetic algorithm optimizing neural network depends on the operation of selection, and the individuals with the best performance form a new population. The selection operation is repeated until the size of the population reaches the specified size. The mathematical model is shown in equation (10).

$$p = \frac{f_i}{\sum_{i=1}^N f_i}. \quad (10)$$

The function of crossover behavior is to maintain the nature of excellent individuals and obtain new excellent individuals through hybridization. The offspring production of single point crossover can be expressed as formula (11).

$$\begin{cases} x' = [x_1, x_2, \dots, x_k, y_{k+1}, y_{k+2}, \dots, y_n] \\ y' = [y_1, y_2, \dots, y_k, x_{k+1}, x_{k+2}, \dots, x_n] \end{cases} \quad (11)$$

In equation (11), the gray strings of the two parents are $x = [x_1, x_2, \dots, x_n]$ and $y = [y_1, y_2, \dots, y_n]$, respectively, the intersection point of the two is k , and there is only one chromosome breakpoint for single point crossing. When the string length of the parent is n , there are $n - 1$ kinds of crossing results. In the two-point crossover, if the parent string length is n , there are $((n - 1) \times (n - 2)) / 2$ kinds of crossover results, and the mathematical description of the two-point crossover is equation (12).

$$\begin{cases} x' = [x_1, x_2, \dots, x_i, y_{j+1}, y_{j+2}, \dots, y_j, x_{j+1}, x_{j+2}, \dots, x_n] \\ y' = [y_1, y_2, \dots, y_i, x_{j+1}, x_{j+2}, \dots, x_j, y_{j+1}, y_{j+2}, \dots, y_n] \end{cases} \quad (12)$$

In equation (12), the gray string of the parent is $x = [x_1, x_2, \dots, x_n]$ and $y = [y_1, y_2, \dots, y_n]$, and the two intersections are set to i, j and $1 < i < j$ at the same time. The introduction of cross behavior can not only retain the advantages of developing individual groups but also reduce the number of iterations of the model and accelerate the convergence. Finally, the mutation behavior is introduced to expand the diversity of the population. Its specific operation is to randomly select individuals, give the mutation probability, and change their chromosome coding. At the same time, the structure of BP neural network is optimized based on empirical formula (13).

$$n = \sqrt{m + o} + c. \quad (13)$$

Finally, the characteristic variables required for the construction of neural network model include location factors and architectural characteristic factors. After quantification, the difference of characteristic variables will affect the training results of the model, so the maximum and minimum method of normalization function is used to process the characteristic variables, so that the distribution trend of variable data is transformed into a conventional function, and its formula is shown in formula (14).

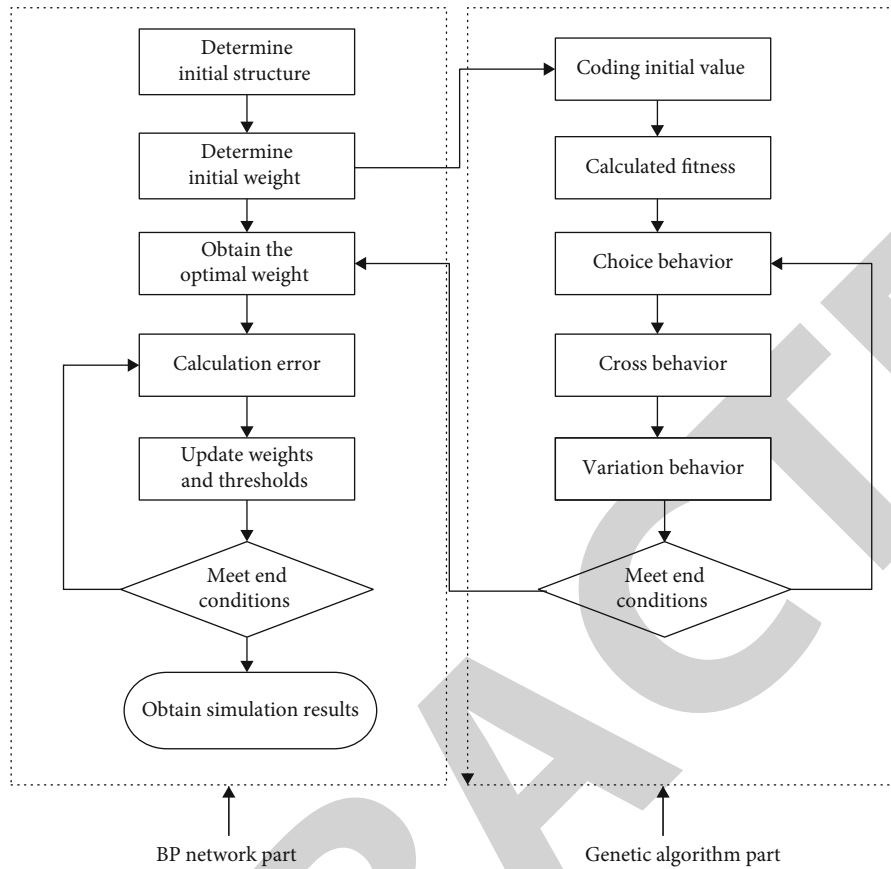


FIGURE 2: Optimization of weight flow of BP network by genetic algorithm.

TABLE 1: Characteristic variables affecting house price and variables' input processing.

Index	Independent variable	Input process	Quantification or unit
Azimuth index	Position	Distance from the city center	Km
	Traffic	Convenience grading evaluation ([1-5])	/
	Greening rate	No process	
	Plot ratio	No process	
	Number of households	No process	(Num)
	Property fee	No process	Yuan/m ² month
	Supporting facilities	Convenience grading evaluation ([1-5])	/
Building index	House type	Type number ([1-5])	/
	Measure of area	No process	m ²
	Orientation	Number by direction ([1, 2])	/
	Residential type	Type number ([1, 2])	/
	Renovation	Type number ([1-3])	/
	Building scale	Number by size ([1-5])	/

$$X_{\text{norm}} = 1 - \frac{X - X_{\min}}{X_{\max} - X_{\min}}. \quad (14)$$

In equation (14), X is the characteristic variable data, and X_{\max}, X_{\min} is the maximum and minimum value of the data. After the model is constructed, the mean square

error is used to evaluate the performance of the model, and its formula is shown in formula (15).

$$\begin{cases} MSE_t = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \\ MSE_p = \frac{\sum_{j=1}^m (y_j - \hat{y}_j)^2}{m} \end{cases} \quad (15)$$

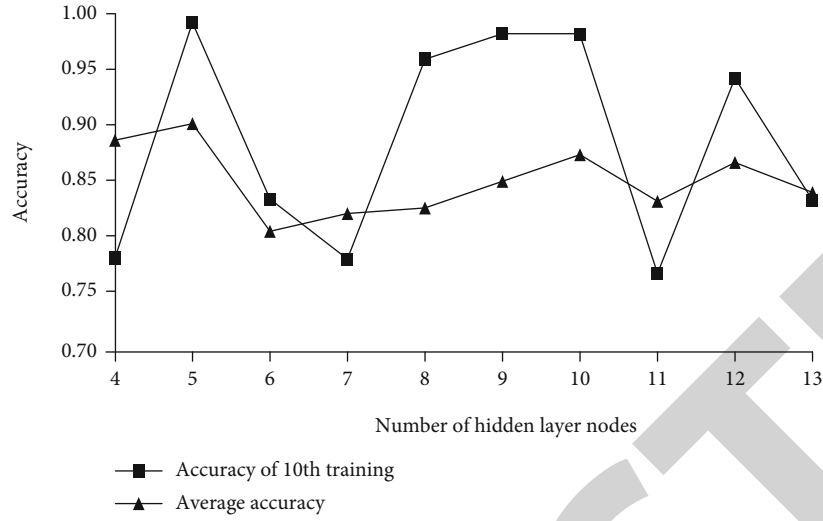


FIGURE 3: Prediction accuracy of nodes in different hidden layers.

TABLE 2: Accuracy of different fitness functions.

Fitness function	1	2	3	4	5	6	7	8	9	10	Average
Genetic algorithm optimization	0.985	0.996	0.975	0.790	0.756	0.706	0.952	0.963	0.728	0.986	0.884
Maximum objective function	0.881	0.742	0.889	0.767	0.985	0.723	0.910	0.870	0.899	0.778	0.844

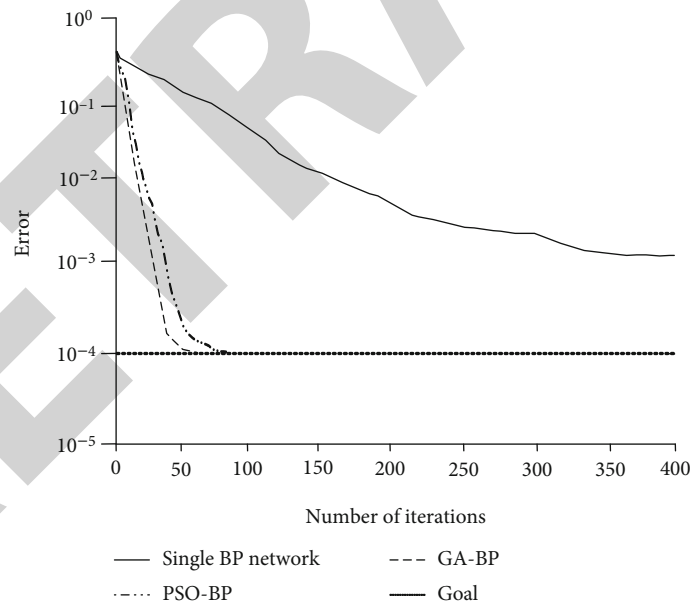


FIGURE 4: Comparison diagram of iteration times of algorithm model.

In equation (15), MSE_t and MSE_p represent the training mean square error and test mean square error of the model, respectively, m, n represents the number of samples tested and the number of samples trained, y_i, \hat{y}_i represents the real value and predicted value of the dependent variable data in the training set, and y_j, \hat{y}_j represents the real value and predicted value of the dependent variable data in the test set.

4. Training and Application Experiment of GA-BP Real Estate Valuation Model

4.1. GA-BP Model Training and Algorithm Optimization Verification. GA-BP algorithm is an iterative algorithm for symmetric diagonally dominant linear equations. It is a probabilistic reasoning algorithm based on recursive update, which has low computational complexity and high parallelism.

Because of these two properties, GA-BP algorithm is very suitable for dealing with large-scale sparse linear equations. GA-BP algorithm is different from classical iterative algorithm and Krylov subspace algorithm. GA-BP algorithm has good convergence for solving symmetric diagonally dominant linear equations. Based on the current market information, this paper makes assumptions on the relevant model market. The study randomly selected some samples and sets up many new samples, of which the number of test samples is 80. At the same time, 13 characteristic variables that have the greatest impact on house prices are extracted. The details of characteristic variables are shown in Table 1.

As shown in Table 1, the house type in the characteristic variable will be numbered from 1 to 5 in turn from one room and one living room, one room and two living rooms, and two living rooms, three rooms and two living rooms, four rooms and two living rooms; The residential types are divided into large flat (1) and ordinary residential (2). The decoration variables number the blank, simple decoration and hard decoration as 1-3 in turn. According to the above 13 characteristic variables, different numbers of hidden layer nodes are brought into the model. After 10 repeated training, the different accuracy of the model is shown in Figure 3.

The accuracy in Figure 3 is the proportion of the number of samples within 10% of the prediction error in the total number of training samples. It can be seen from the figure that after ten training for the same sample, when the number of hidden layer nodes is 5, the accuracy of the model is the highest, and the average accuracy and 10 repeated training accuracy are 0.901 and 0.992, respectively. At the same time, the effects of the maximum objective fitness function and the optimized fun function on the training model are compared, and the results are shown in Table 2.

It can be seen from Table 2 that after taking the traditional maximum objective function as the fitness function of the real estate valuation model, the average accuracy of 10 repeated training is 0.844, while the average accuracy of the optimized fitness function is 0.884. The experiment shows that after the optimization of the fitness function, through the above model, we can combine all the projects of real estate enterprises as one project. The land acquired in that year was also merged into a project. In this way, we can simulate the appropriate sales progress and construct a similar cash flow to estimate their net asset value (NAV) after the year. Then, get the net present value or valuation through an appropriate discount rate.

4.2. Research on Application Performance of GA-BP Real Estate Valuation Model. In order to calculate the performance of the real estate model, 100 samples were selected as the test target. The BP neural network optimized by GA-BP neural network model and particle swarm optimization algorithm is trained and tested. The maximum number of iterations is 400, the learning rate is 0.1, the hidden layer transfer function is logsig function, and the target accuracy is 0.0001. The error accuracy comparison of the three algorithm models is shown in Figure 4.

As shown in Figure 4, when the training samples are the same, the first to achieve the target error is the neural net-

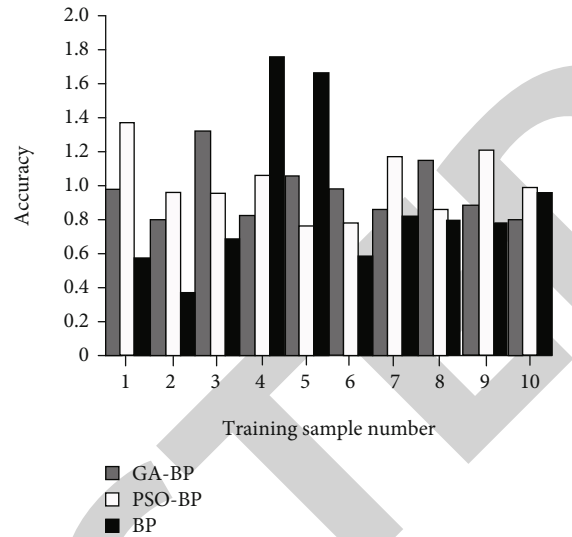


FIGURE 5: Training output error of different models.

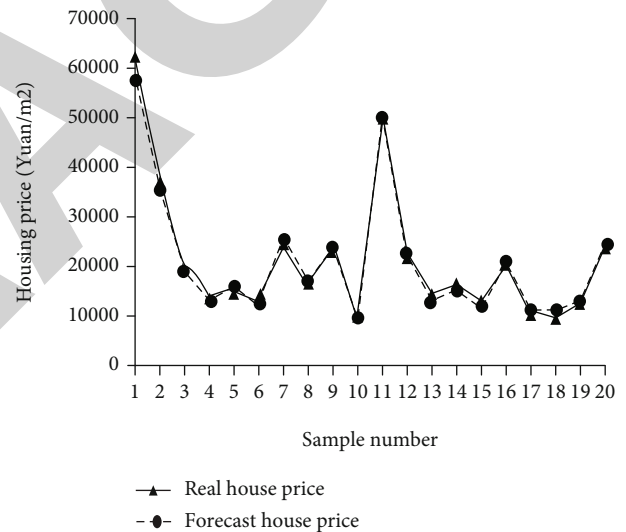


FIGURE 6: The error between the predicted house price and the actual value of the model.

work model after genetic optimization. The GA-BP curve tends to be stable when it is close to 50 iterations, and the convergence is completed only in 56 iterations; at the 84th iteration, it reaches the target value of 0.0001; finally, when the traditional BP neural network completes the maximum number of iterations of 400 times, it still fails to achieve convergence and falls into the result of local extremum. Experiments show that after the optimization, the anti-interference of BP neural network is improved, separated from the local extreme value, and the estimation efficiency of the model is also improved. However, the convergence times of GA-BP model are 28 times less than that of PSO-BP model, which proves that the calculation efficiency of Y genetic algorithm is higher. In the iterative training process of the three algorithms, 10 groups of data are randomly selected for comparison, as shown in Figure 5.

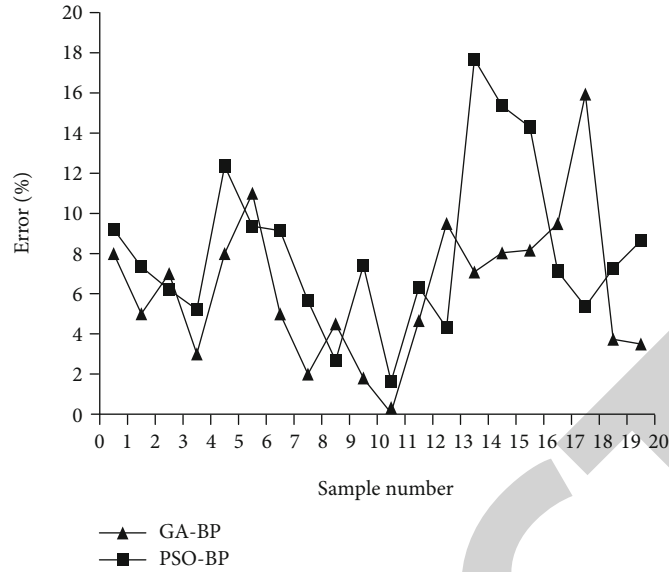


FIGURE 7: House price prediction error of PSO-BP and GA-BP.

The accuracy calculation formula in Figure 5 is the predicted house price/actual house price output by the model. The maximum ratio is 1.75 and the minimum ratio is 0.3; the training result of PSO-BP model is small, with the maximum ratio of 1.36 and the minimum ratio of 0.76. The predicted house price is closest to the actual house price, with the maximum ratio of 1.32 and the minimum ratio of 0.78. Experiments show that the accuracy of genetic algorithm is higher and higher than that of traditional BP model. After completing the training of the model and extracting the optimal initial value and threshold, number 20 test samples bring them into the trained GA-BP neural network system for simulation experiments to compare the actual house price with the error of the estimated house price of the model, as shown in Figure 6.

In Figure 6, the broken line graph of the actual house price is basically consistent with the broken line graph of the predicted house price. Among the 20 test samples, the sample No. 18 has the largest estimation error. The actual house price is 9517 yuan/m² and the estimated house price is 11034 yuan/m². Calculated according to the formula (actual house price - estimated house price)/actual house price, the prediction error is 16%. The actual house price of sample No. 6 is 13988 yuan/m², the model predicts that the house price is 12449 yuan/m², and the prediction error is 11%. The prediction error of other samples is less than 10%. The test sample No. 11 has the smallest error. The actual house price of the sample is 49969 yuan/m², and the model predicts that the house price is 50012 yuan/m², with an error of only 0.08%. If the prediction error of 10% is qualified, the prediction accuracy of the model is 90%. At the same time, the same 20 groups of test sample data are brought into the PSO-BP model, and the prediction errors of the two models are compared. The specific results are shown in Figure 7.

As can be seen from Figure 7, the overall error of particle swarm optimization model is greater than that of genetic

algorithm optimization model. Most of the prediction errors of GA-BP are in the range of 2%-8%, and the overall error of PSO-BP model is in the range of 6%-15%. The average error of GA-BP model is 6.27%. Compared with 8.12% of the average error of PSO-BP model, the relative error optimized by genetic algorithm is reduced by 1.85%. If the prediction error is less than 10% as the qualified accuracy, the prediction accuracy of GA-BP model is 90%, while the prediction error of PSO-BP model is more than 10% in four places, so the prediction accuracy is only 80%. From the experimental data, it can be concluded that the BP neural network model optimized by genetic algorithm is more neutral in real estate valuation. The advantages of the algorithm are not only reflected in higher prediction accuracy and stable prediction qualified rate but also reflected in less iterative training times of the model. Through MATLAB programming, the accuracy of GA-BP neural network, PSO-BP neural network, and traditional BP neural network is compared. In addition, through calculation, it is found that although the prediction accuracy of PSO-BP neural network has been improved, the mean square deviation of multiple runs is lower than that of GA-BP neural network, which can be considered that the model optimization has achieved good results.

5. Conclusion

This paper uses machine learning method to establish the relationship model between characteristic variables and house prices. From the perspective of hidden layer node optimization and fitness function optimization, the calculation results show that when the number of hidden layer nodes is 5, the prediction accuracy of the model is the highest. The model training experiment and model performance test simulation experiment are carried out successively by using MATLAB tools. The former is 4 percentage points higher than the latter, which proves the accuracy and performance of the real estate appraisal model. In the part of

Retraction

Retracted: Optimization of Three-Dimensional Model of Landscape Space Based on Big Data Analysis

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Wang, "Optimization of Three-Dimensional Model of Landscape Space Based on Big Data Analysis," *Journal of Function Spaces*, vol. 2022, Article ID 7002983, 11 pages, 2022.

Research Article

Optimization of Three-Dimensional Model of Landscape Space Based on Big Data Analysis

Cuixia Wang 

School of Education and Modern Art, Shangqiu Institute of Technology, Shangqiu 476000, China

Correspondence should be addressed to Cuixia Wang; 1350008017@sqgxy.edu.cn

Received 2 June 2022; Revised 29 July 2022; Accepted 2 August 2022; Published 17 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Cuixia 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.

Based on virtual reality technology, landscape 3D modeling provides users with the possibility to construct a simulated garden landscape environment design effect online, so it has high requirements for accuracy. With the continuous improvement of precision requirements, the number of people involved in the construction of 3D models is also increasing, which puts forward higher requirements for modeling. Based on this, this paper studies the optimization strategy of landscape space 3D model based on big data analysis. Based on the analysis of the establishment of the 3D model and the related algorithm research, this paper analyzes the optimal design of the 3D model under the background of big data. In the 3D modeling of the edge folded area, it is based on the traditional quadratic error measurement grid simplification algorithm, combined with the vertex error matrix to simplify, so as to shorten the modeling time. Based on an efficient search algorithm, an adaptive nonsearch fractal image compression and decoding method is proposed in the image compression and decoding stage of 3D modeling. The search is performed by specifying the defined area block. Finally, an experiment is designed to analyze the performance of the optimization algorithm. The results show that the improved edge folding region algorithm can reduce errors on the basis of ensuring image quality, and the adaptive search algorithm can shorten the search time and improve the compression rate. This method provides a technical reference for the visualization experience and simulation system of garden landscape design and improves the presentation quality of virtual garden landscape design scenes.

1. Introduction

The construction of 3D landscape model is a comprehensive discipline, which not only covers computer graphics technology, visualization technology, and remote sensing technology but also involves virtual reality technology, spatial data structure technology, etc. [1]. Based on computer technology, virtual reality technology creates a nearly real three-dimensional virtual environment and realizes interactive functions [2, 3]. The current virtual environment has been able to realize the modeling of scenes and models. In this process, large-scale data is involved and the operation speed is slow [4, 5]. Therefore, in the three-dimensional model design of landscape based on big data analysis, it is necessary to optimize the design, focusing on improving the compressed model and improving the operation efficiency.

Based on this background, this paper studies the three-dimensional modeling and optimization of landscape space

based on big data analysis, which is mainly divided into four chapters. The first chapter briefly summarizes the construction of 3D landscape model and introduces and analyzes the chapter arrangement of this study. Chapter 2 introduces the algorithms of 3D modeling at home and abroad and summarizes the shortcomings of the current research. In Chapter 3, the optimization model of landscape 3D modeling based on big data analysis is constructed. The edge folding algorithm is used to optimize the folding area in 3D modeling. In the compression design of 3D modeling, the traditional algorithm is optimized and improved, and the nonsearch fractal algorithm is used to shorten the compression time. Chapter 4 simulates and analyzes the landscape 3D modeling optimization algorithm constructed in this paper and evaluates the performance of the algorithm through indicators such as error rate and compression ratio. The simulation results show that compared with the standard algorithm and quadtree fractal algorithm, the algorithm proposed in

this paper can improve the image compression ratio and reduce the coding error on the basis of ensuring the image quality.

The innovation of this paper is to propose a self-realization nonsearch fractal algorithm. In the image compression of 3D modeling, in addition to the distance factor, the geometric characteristics of sharp edges and corners in the original model are also considered. Taking the product of absolute curvature and quadratic error matrix as the error matrix, more data information is retained on the basis of ensuring the optimization model. The algorithm uses a fast nonsearch matching method between the range block and the defined region to decompose and combine and uses an adaptive decomposition method to improve the image quality. Secondly, in the model optimization design, in order to optimize the folding area, the edge folding algorithm is used. The algorithm is based on quadratic error measurement. The improved edge folding region algorithm can reduce the error on the basis of ensuring the image quality, while the adaptive search algorithm can shorten the search time and improve the compression rate. This method provides a technical reference for the visual experience and simulation system of landscape design and improves the presentation quality of virtual landscape design scenes.

2. State of the Art

In recent years, with the development of computer technology, virtual display technology, graphics, and remote sensing technology have also made a lot of research results. 3D model construction and visualization have been developed in an all-round way. Many scholars improve the fidelity of 3D modeling and shorten the modeling time from the perspective of modeling technology. For example, Shan and Sun and Yang et al. designed a new landscape planning effect simulation system based on virtual reality technology to preprocess 3D landscape images, remove noise information and redundant information, and adopt the parametric description rules of plant spatial layout [6, 7]. Zheng et al. used the spatial simulation of settlement distribution driven by random forest to conduct three-dimensional simulation in order to improve landscape visibility [8]. Tastan et al. compared and studied the availability and constraints of two modeling methods for 3D modeling in immersive virtual reality (IVR). They collected data through the screen capture video of the modeling screen and the video recording of user gestures during the modeling session. Through the analysis of qualitative coding method, they believed that DM modeling can be digitally input through a new keyboard [9, 10]. In their research, Li et al. used real-time computer graphics technology, three-dimensional modeling technology and binocular stereo vision technology to study multivision animated character objects in virtual reality technology, designed binocular stereo vision animation system and 3D graphics algorithm for three-dimensional geometric transformation of computer graphics, and studied image output processes such as basic texture technology and basic lighting model used in fragment processing stage [11]. In the research and analysis, Roy et al. proposed a new method to

calculate the affine parameters of fractal coding to reduce its computational complexity and derive a simple but effective approximate value of scaling parameters [12]. Hernandez-Lopez and Muiz-Pérez proposed the fractal compression of Kelley table in their research and developed a parallel implementation of fractal image compression using quadtree partition [13].

To sum up, up to now, there are many researches on 3D modeling, and many scholars have improved the model from a technical point of view. The image data itself has a certain redundancy. The existence of these data not only wastes modeling time but also affects the image quality. In the research of eliminating redundancy, most of them rely on distance parameters for optimization. However, this kind of algorithm plays a very limited role in optimizing time and model. It only focuses on the simulation of individual geometric features and ignores other geometric features. On the other hand, there is little analysis on the modeling and optimization of overlapping areas. Most of them are optimized for a terrain simulation, which is not scalable. Therefore, the research on the optimization of three-dimensional model of garden space based on big data analysis has important practical significance.

3. Methodology

3.1. 3D Modeling Data of Landscape Space. The three-dimensional modeling of garden landscape space needs to cover all space substances. According to the existing point of view, the space should include terrain, buildings, transportation, vegetation, and public facilities. The specific data sources can come from satellite data, aerial photography data, scanning data, photographing data, manual measurement data, etc.; the content and form of different data sources will be different.

In the three-dimensional modeling of landscape space, many terrains will be involved. In the terrain simulation, underground pipelines, bridges, and buildings are difficult parts. The above ground part only considers independent space in the modeling, which is convenient to determine the spatial location, but the carrier needs to be determined and the specific location needs to be marked in the modeling. Once the spatial position of these terrain is determined, the spatial changes of terrain surface can be analyzed [14]. The terrain surface itself is a continuous space, and the buildings are built on this plane. Therefore, without any treatment, the buildings will overlap with other terrain in space to form gaps, and similar problems will occur in other terrain modeling. At present, the popular GIS modeling methods and terrain simulation are very necessary and basic tools, which play a vital role in other carrier modeling [15]. Therefore, in the three-dimensional modeling and calculation, this paper sets that all vertices are displayed in four-dimensional coordinates. When transforming the space, the coordinate system can be changed by using mathematical transformation methods, such as translation, rotation, and scaling. After the landscape 3D modeling is displayed, the transformed results can be expressed as follows:

$$[x', y', z', 1] = [x, y, z, 1] \times T, \quad (1)$$

where T represents the transformation matrix, which is the result of transformation. At present, there are mainly two kinds of model projection: parallel projection and perspective projection. Among them, parallel projection is mainly positive parallel and oblique parallel, and perspective projection is mainly applied to quadrangular platform. No matter what projection method is adopted, after determining the projection window and size, only the objects in the projection will be displayed, and all other objects will be cut off.

The parallel projection direction refers to the projection vector of the center point of the window, and the oblique parallel projection is that the equation in the projection space is not standardized, so it is necessary to use some staggered transformation to convert the oblique plane projection into orthogonally parallel projection [16]. The process of converting face coordinates into plane coordinates is called projection. The essence of projection coordinate system is plane coordinate system. The projection coordinate system is defined in a two-dimensional plane. Different from geographical coordinate system, the length, angle, and area of projection coordinate system are constant in two-dimensional space. When defining the projection volume, it is necessary to define the projection vector and convert it into the observation coordinate system. The transformation of the projection matrix can be expressed as follows:

$$M = \begin{bmatrix} 1, 0, -p_x/p_z, 0 \\ 0, 1, -p_y/p_z, 0 \\ 0, 0, 1, 0 \\ 0, 0, 0, 1 \end{bmatrix}. \quad (2)$$

In the three-dimensional landscape modeling, the reality of the object surface feels the influence of lighting conditions, so it is necessary to analyze the lighting model in the modeling. Considering the reflection of the object caused by the light source, the local illumination model can be expressed as follows:

$$I = K_a I_{pa} + K_d \sum \frac{J_{pd} N_o L_o}{r + c}, \quad (3)$$

where V represents the observed object, K_a represents the brightness of the ambient reflected light, K_d represents the diffuse reflection intensity, N_o represents the light vector, c represents the constant, and I represents the brightness.

3.2. Optimization of 3D Scene Simulation in Folded Area. In terrain modeling, folding area is a problem that must be faced. If there are relatively small objects in modeling or they are far away from each other, LOD model can be used to draw objects directly [17]. Node LOD model: itself is a resolution structure. Different resolution models are connected by nodes, and the corresponding components are operated by activating the nodes. When all nodes are activated, it is essentially a full resolution structure. Its advantages are

simple structure, convenient operation, and suitable for expressing complex discontinuous volume model objects. If the object is large and close to the eye, the model needs to be refined. In the current detailed description of spatial modeling, LOD model is a common model. This algorithm is based on the triangular network algorithm. First locate different vertices and classify them, then sort the importance, delete the unimportant vertex coordinates, and complete the design [18].

It is inevitable to encounter model construction with boundary in landscape 3D modeling. In simplified design, it is necessary to maintain the clarity and information of the boundary model. Generally, a virtual plane will be formed to maintain the triangular vertical space angle with the edge interface, and the weight of the quadratic error matrix is set through the plane, and then, the matrix at the end point is added to the virtual plane. This method can well maintain the information integrity of the folded area [19]. The specific process is shown in Figure 1. In the analysis, all folding costs need to be calculated first and sorted according to the size; delete the folding operation with the lowest cost, calculate the new vertex folding cost, and reorder until the requirements are met.

In the three-dimensional construction of the model with boundary, it is necessary to keep the boundary information as much as possible. The quadratic error measurement simplification algorithm has great advantages in both quality and complexity in the model simplification information, but this algorithm itself also has some shortcomings [20]. The distance parameter is considered in the design of quadratic error measurement, so the grid distribution is uniform, especially in large-scale model construction, this uniform distribution will lose a lot of typical edge information. During model construction, some unimportant features will be deleted first. For example, in platform area model construction, it will be directly combined and simplified. Therefore, it is necessary to optimize the quadratic error algorithm. In describing the characteristics of the object, the curvature reflects the surface details of the model, which is an important feature. The curvature of the sharp area is large, and the value of the flat area is very small. Therefore, in the simplification, it is considered to establish a new matrix by using the product of quadratic error matrix and absolute curvature. In the quadratic error measurement algorithm, multiplying the quadratic error matrix by the adaptive weight will not affect the basic form of the matrix. The bending degree of adjacent patches is expressed by the average curvature, and the formula can be expressed as follows:

$$H = \frac{(\sum M(e_i))}{A}, \quad (4)$$

where A represents the sum of triangle areas associated with vertices and m represents the included angle of normal vectors of adjacent triangular patches. The included angle of two adjacent triangular patches can be expressed as follows:

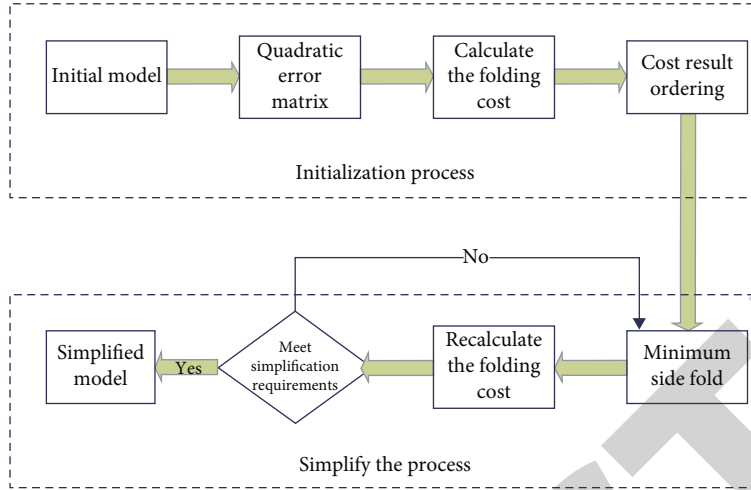


FIGURE 1: The flow of the edge folding algorithm.

$$\cos \varphi = \cos \langle n_1, n_2 \rangle = \frac{n_1 n_2}{|n_1| |n_2|}. \quad (5)$$

The degree of curvature of a surface is Gaussian curvature, and its absolute value is of great significance. Assuming that the small surface on the surface is $\Delta\delta$, translate the normal vectors of all points on the small surface to the origin, and you can get the unit spherical area centered on the origin, with the area unchanged. The absolute value of Gaussian curvature of a surface is the limit value of the shrinkage of a small surface. The bending degree of the grid surface needs to be calculated by discrete differentiation, and the formula is as follows:

$$K = \frac{2\pi - \sum_{i=1}^k \theta_i}{A}, \quad (6)$$

where θ_i represents the vertex angle related to the vertex. Assuming that the edge is folded to a new vertex, the improved folding cost can be expressed as follows:

$$\Delta''(\bar{v}) = \|V_i V_j\| \bar{v} T (K_{iabs} Q'_i + K_{jabs} Q'_j) \bar{v}, \quad (7)$$

where $\|V_i V_j\|$ represents the distance of the vertex, K represents the absolute curvature, and Q represents the error matrix.

3.3. Model Compression Optimization Algorithm. From the gardens of Ming and Qing Dynasties to the modern landscape of natural harmony, great changes have taken place in people's life and recreational space, which are inseparable from the artistic design of landscape architecture. Garden landscape renderings are composed of garden sketches, including rockeries, garden pavilions, flower racks, garden bridges, plant landscaping, and garden sculptures. Because 3DS max modeling is more complex, compared with the SketchUp software, especially in complex garden sketch modeling, such as garden pavilion, although the garden sketch is exquisite, the complex structure makes modeling

extremely difficult. Generally speaking, the software involved in the process of making landscape renderings includes 3DS max. Generally, the size of this file will directly affect the superiority of 3D model. Excessive volume will affect storage and reduce rendering speed. Certain technologies need to be used to reduce complexity and compress the volume of 3D model [21]. Fractal algorithm is a common algorithm for optimizing model at present, such as multiple linear subdivision method and random midpoint iteration method, which can meet the requirements of landscape 3D modeling. This paper improves the fractal algorithm in the research and analysis.

Among fractal image compression algorithms, full search coding algorithm is the most basic algorithm, and other improved algorithms are basically based on this algorithm [22]. This algorithm assumes that there is scale redundancy in different regions of natural image and then uses different redundancy to realize compression. First, image segmentation is an important step that affects the compression ratio and the quality of the model. At present, the commonly used segmentation strategies are quadtree segmentation, fixed block segmentation, and so on [23]. Then, form the codebook, set the step size, and slide the original image according to the step size. This step allows overlap. Each block adopts the average value of four adjacent pixels and compress the word block to obtain the initial codebook. The process of calculating fractal code is to find the best matching block in each codebook, adjust brightness and contrast, minimize regional block differences, and find the best matching block. The formula is as follows:

$$d_i = \frac{1}{n} \sum_{i=1}^n (c_{kj} b_i + h_{k,l} - a_i)^2, \quad (8)$$

when the value range block is mapped to the corresponding defined area block position, it needs to be searched continuously. The error is minimized through continuous isometric transformation. The derivation of this formula is obtained:

$$c_{kj} = \frac{m \sum_{i=1}^m a_i b_i - \sum_{i=1}^m a_i \sum_{i=1}^m b_i}{m \sum_{i=1}^m b_i^2 - (\sum_{i=1}^m b_i)^2}, \quad (9)$$

$$h_{kj} = \frac{1}{m} \left(\sum_{i=1}^m a_i - c_{kj} \sum_{i=1}^m b_i \right), \quad (10)$$

where m represents the number of pixels, h_{ki} represents the contrast coefficient, b_i represents the pixel value of the definition domain, k represents the abscissa of the best matching block, and j represents the ordinate of the best matching block. When the following conditions are met $c_{kj} = 0$,

$$m \sum_{i=1}^m b_i^2 - \left(\sum_{i=1}^m b_i \right)^2 = 0. \quad (11)$$

When the mean square error reaches the minimum, the area block position is stored, and the contrast factor and brightness parameters are obtained at the same time to complete the coding of the value range block. Fractal coding has some disadvantages, that is, the coding speed is slow. For each range block, it is necessary to search the matching range block, which takes time [24]. At present, fractal algorithms generally take reducing search time as an improvement idea, such as classification method and neighborhood search method. These algorithms still need search and matching to be realized [24].

No search fractal image algorithm means that it does not need to search each matching optimal region block, but to specify a specific defined region block, which can reduce a lot of search time. In the research and analysis of this paper, an adaptive search free fractal algorithm is proposed. In addition, the algorithm does not need to record the information of each defined area block, and it can reduce the typing code. There is no need to search, and the search time can be shortened on the basis of ensuring the quality. Compared with the traditional algorithm, the nonsearch algorithm reduces the search process when matching the best fixed block. Directly calculate according to the set threshold, which not only ensures the quality but also compresses the fractal time.

The core problem of no search algorithm is to determine the position relationship between the specially defined area block and the value range block and then carry out the corresponding mapping transformation to calculate the coding error. In the analysis, it is assumed that the value range block size is $B_1 \times B_2$, the coordinate of the center point is (row_R, col_R) , and the specified optimal area block size is $2B_1 \times 2B_2$. When performing search free coding, first divide the original image into nonoverlapping area blocks, expressed as follows:

$$\begin{cases} T = UR_i, \\ R_i \cup R_j = \varphi (i \neq j), \end{cases} \quad (12)$$

where R_i represents a value range block, and a definition area block is specified for each value range block. No search algorithm can greatly reduce the coding time, but this algorithm needs to determine the best matching area block. How to define the best matching block is a key problem, which is also one of the key indicators to determine the performance of no search algorithm. In this paper, the matching error between the value range block and the specified area block is described by mean square deviation. The smaller the error is, the closer the value range block is to the specified area block. The formula is expressed as follows:

$$d_i = \frac{1}{n} \sum_{i=1}^n (stb_i + ot_i - a_i)^2, \quad (13)$$

where n refers to the number of pixels in the value range block, st represents the contrast parameter of the value range block, ot represents the brightness parameter of the defined area block, b represents the pixel gray value of the defined area block, and a represents the pixel gray value of the value range block. From the analysis, we can see that the definition area block of nonsearch fractal image coding is not the cause of coding error, but the error existing in advance. Therefore, there is bound to be a mismatch between the definition area block and the value area block, that is, the mean square error is very large, which shows that they are not similar.

When the nonsearch fractal compression ratio is not high, the adaptive combination method is adopted to solve it. For the value range blocks that meet the error matching conditions, they are combined with the surrounding area blocks that cannot meet the error requirements to form large area blocks, reduce the number of storage and transmission, and improve the compression ratio. This combination method needs to calculate the basic value range blocks that meet the error matching conditions. In this paper, the matrix composed of basic value range blocks is coded in the design, and the abscissa and ordinate of the matrix where the value range block is located are defined. Multiple value range blocks are combined to form a new value range block, and the initial value is the basic value range block. Continue to carry out adaptive combination, as shown in Figure 2.

The optimization of model compression algorithm needs to be evaluated to see the difference between the decoding model and the original image and explain the compression quality of the model. At present, there are two evaluation criteria for model compression. One is subjective evaluation, which evaluates the modeling status through the eyes. The other is the objective standard, which compares and analyzes the data error between images. The objective evaluation is simple and more persuasive. Therefore, in the evaluation model compression algorithm, the objective evaluation method is used. The objective evaluation index is required to reflect the change of gray level as a whole [25]. At present, the coding algorithm is generally evaluated by coding time, compression ratio, and peak signal-to-noise ratio. Model compression algorithm can effectively reduce parameter redundancy, thereby reducing storage occupation, communication bandwidth, and computational complexity. The

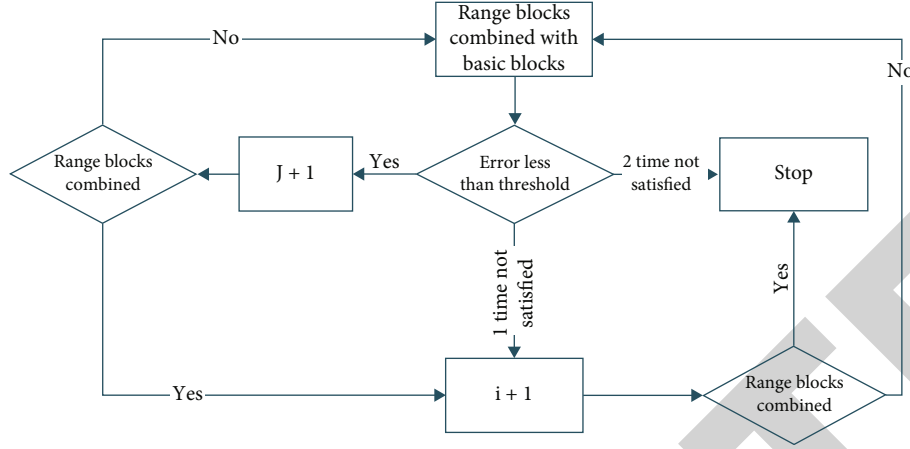


FIGURE 2: Adaptive combination process of range blocks.

more the quantization level is, the richer the image level is, the higher the gray resolution is, the better the image quality is, but the amount of data is large. The less the quantization level is, the less the image level is, the lower the gray resolution is, the false contour phenomenon will appear, and the image quality will become worse, but the amount of data is small.

Peak signal-to-noise ratio refers to the ratio of the maximum possible power of a signal to the noise function affecting accuracy, expressed in logarithm decibels. The formula can be expressed as follows:

$$\text{MSE} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (f(i, j) - \hat{f}(i, j))^2 \quad (14)$$

MSE is the mean square error, $f(i, j)$ represents the mean square of the compressed image, and $\hat{f}(i, j)$ represents the mean square of the original image. The gray image has a precision of 8 bits, the maximum value is 255, and the peak signal-to-noise ratio can be expressed as follows:

$$\text{PSNR} = 10 \times \log \left[\frac{255^2}{\text{MSE}} \right], \quad (15)$$

where PSNR represents the peak signal-to-noise ratio, in decibels. The larger the value, the clearer the image. This index does not need human participation, can be evaluated directly, and has stronger stability. Compression ratio refers to the ratio of the ratio of the data occupied by the image after compression to that before compression. It is generally believed that the larger the compression ratio, the smaller the storage quantity and the higher the application value. The reciprocal of the compression ratio is the compression efficiency. Coding time refers to the time spent from the beginning to the end of coding. The shorter the time, the better the performance of the algorithm.

4. Result Analysis and Discussion

4.1. Simulation of Folding Region Model Optimization Algorithm. In the simulation analysis, Visual Studio development platform and OpenGL are used for graphic rendering, and different simplified models are compared and analyzed, including geometric error, intuitive quality, and grid distribution. In the graphic simulation, the vehicle pictures and animal pictures are used for simulation analysis. The original model covers 10544 and 5804 triangular pictures, respectively. When the model is simplified to 1200 triangular patches, the geometric features of the animal can be saved successfully, which is very close to the original model. When it continues to be simplified to 5%, the overall contour can still be seen. However, the algorithm used in this paper can well retain all kinds of sharp features of the animal, and the original algorithm loses these effective features. The same conclusion can be obtained in the simplification of vehicle simulation. When the simplification is 15%, many feature information in the traditional algorithm disappears. It shows that the algorithm used in this paper can retain more geometric features on the basis of maintaining the shape of the whole model after large-scale simplification; especially in the case of large curvature, the geometric features of the picture will not be lost. The traditional algorithm is easy to lose the characteristics of sharp areas after simplification, because it adopts uniform grid distribution. The algorithm used in this paper is more reasonable in the distribution. It distributes according to the change of curvature. The distribution with large curvature has more meshes. In addition, the time quadratic error measurement algorithm used in the analysis of traditional algorithms only considers the distance factor and has no other surface features.

Compare and analyze the error variation characteristics under different models, as shown in Figures 3 and 4. From the data in the figure, it can be seen that the average error of the algorithm used in this paper is significantly reduced. Although QEM algorithm has certain advantages, the error is almost the same. Moreover, in terms of geometric error, the algorithm proposed in this paper has higher advantages and can meet the needs of large-scale 3D modeling.

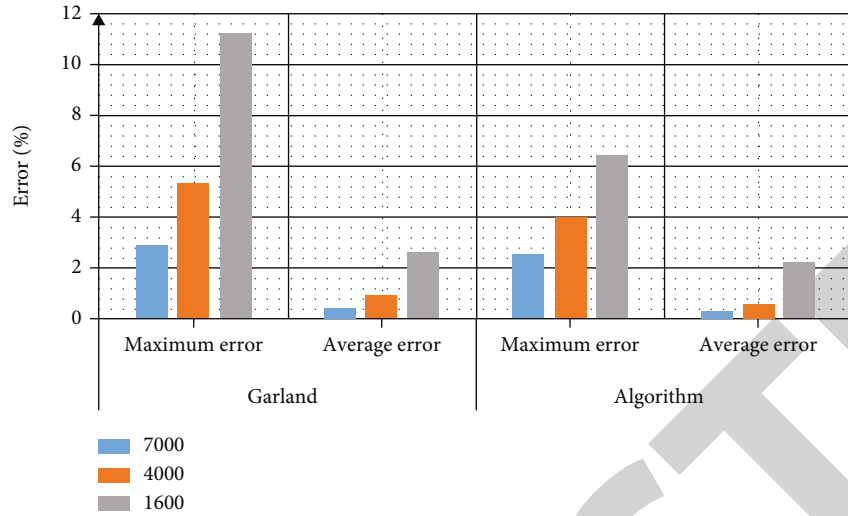


FIGURE 3: Comparative analysis of vehicle model errors.

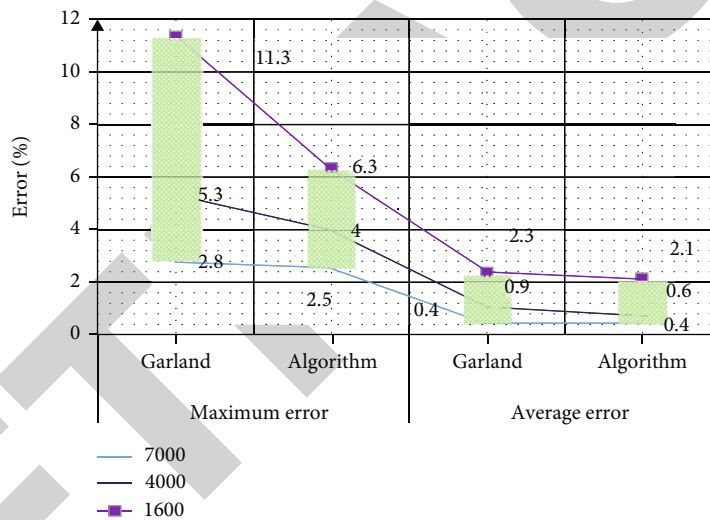


FIGURE 4: Comparative analysis of animal model errors.

4.2. Model Compression Simulation Analysis. In 3D model compression, the coding quality is affected by the coding error. Only by controlling the error can we ensure the similarity of block simulation and the quality of the model. Therefore, the nonsearch fractal compression coding designed in this paper needs to set a threshold. Only when the error value is less than this threshold can the model matching be guaranteed. If the error is greater than the threshold, the model is considered to be mismatched and the image needs to be re divided.

The threshold is not a fixed value and varies within a range. Generally speaking, the smaller the threshold, the more details can be retained. A large threshold is more suitable for flat image compression and improves the compression ratio. Combined with the previous data, the threshold range is set at 80~200 to achieve a balance, which will not reduce the image quality or have a great impact on the com-

pression ratio. In the application, the threshold also needs to be adjusted. If the quality requirements are relatively high, the value can be 80 to ensure that the value range block is decomposed into small modules for rematching and improve the coding quality. If the compression ratio is required, and the value is 200, more value range blocks must be required to realize decomposition and improve the compression ratio. As shown in Figure 5, combined with the actual requirement of threshold 130, the compression ratio is improved on the basis of meeting the image quality.

The threshold value is selected in combination with the demand in the adaptive nonsearch image compression coding of the image. Figure 6 shows the proportion of combined range blocks under different threshold values, taking into account the image quality and compression ratio, and the threshold range is located at 120~150.

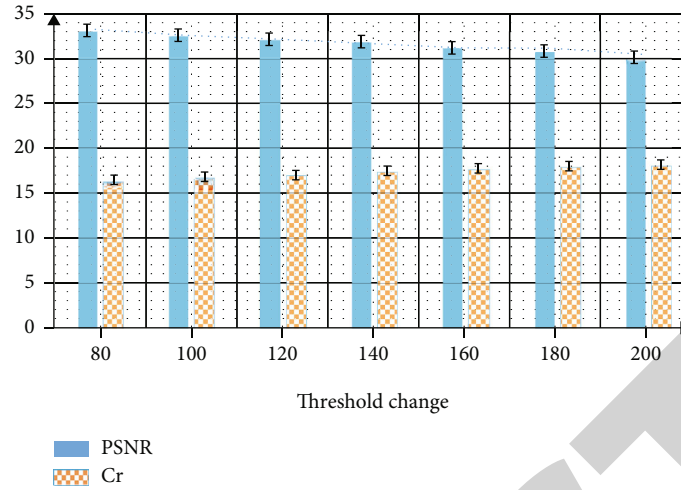


FIGURE 5: Changes of peak signal-to-noise ratio and compression ratio under different thresholds.

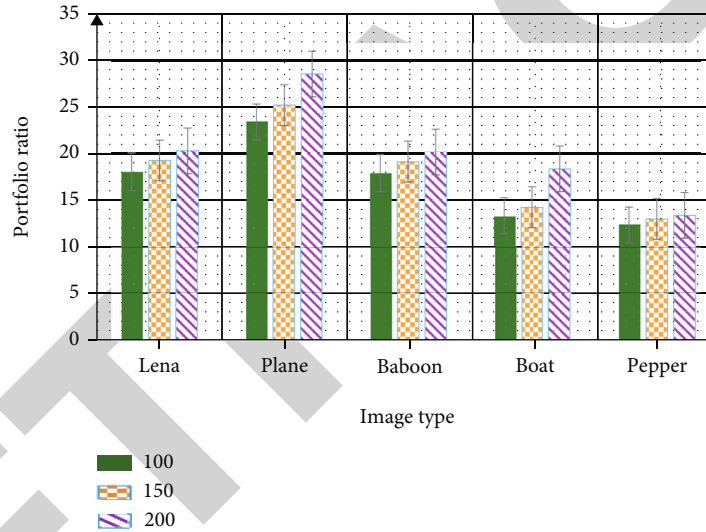


FIGURE 6: Combined range block proportions under different thresholds.

In order to facilitate the experimental analysis, four images are selected for fractal compression analysis. The algorithm, standard algorithm, and quadtree fractal image coding algorithm are used for analysis, respectively. The coding compression ratio results are shown in Figure 7. From the data in the figure, it can be seen that the compression ratio of value domain block of the algorithm used in this paper is greatly improved, because the adaptive decomposition and combination of value domain block reduces the classification code.

Comparing and analyzing the coding time, the measurement results are shown in Figure 8. From the data in the figure, it can be seen that the algorithm used in this paper saves a lot of time, because each value range block does not need to be matched.

The comparative analysis results of coding peak signal-to-noise ratio are shown in Figure 9. From the data in the

figure, it can be seen that the peak signal-to-noise ratio of this algorithm is significantly lower than that of the other two algorithms.

In the optimization and improvement design of the folded region, the product of the quadratic error matrix of the vertex and the absolute curvature of the vertex is used as a new error matrix to improve the modeling quality and retain more geometric features. Based on the basic value range block, the value range block that does not meet the threshold adopts the improved adaptive decomposition and combination algorithm and directly adopts the adaptive combination method to meet the threshold. Simulation results show that compared with standard algorithm and quadtree algorithm, this algorithm can not only improve the compression ratio but also reduce the peak signal-to-noise ratio and save search time.

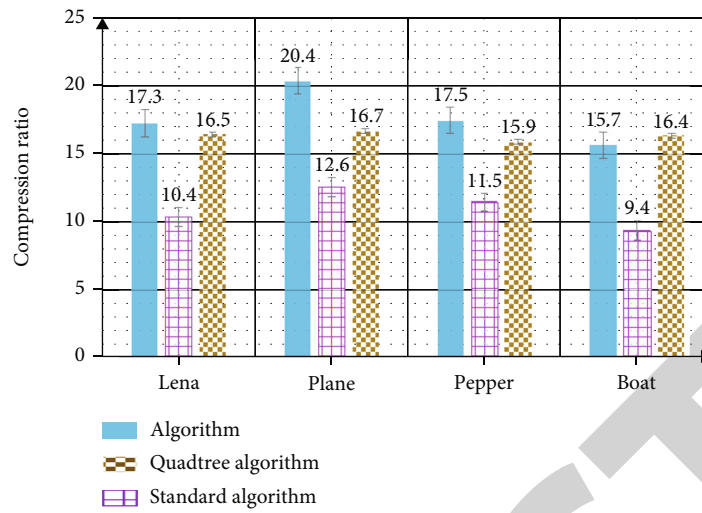


FIGURE 7: Comparative analysis of coding compression ratio.

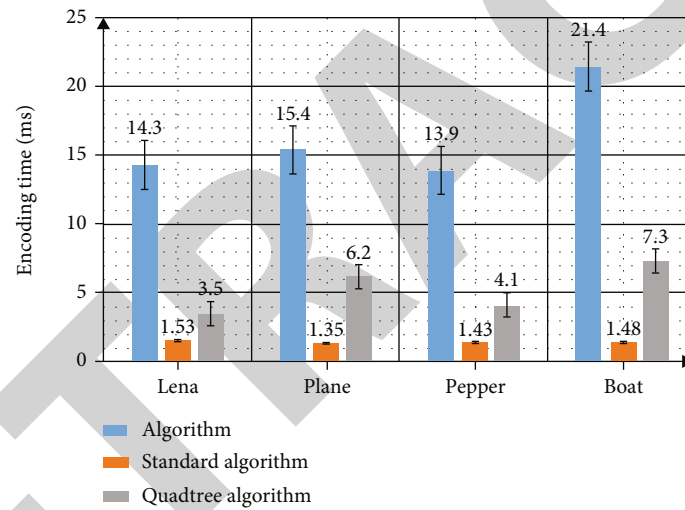


FIGURE 8: Comparative analysis of encoding time.

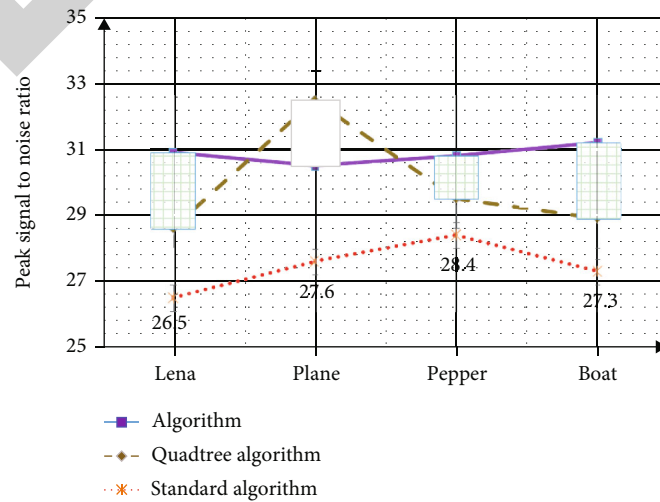


FIGURE 9: Coding peak signal-to-noise ratio comparison.

5. Conclusion

The construction of three-dimensional landscape is the expression form of map. It is one of the important research contents of geographic information to bring three-dimensional feeling to people with the help of computer-aided technology. Big data virtual landscape design can not only display the objects of landscape in a full range but also bring new challenges. Based on this, this paper studies the optimization of landscape 3D model based on big data analysis. In the optimization and improvement design of edge folding area, the quadratic error matrix of vertex and the product of absolute curvature of vertex are used as a new error matrix to improve the modeling quality while retaining more geometric features. Based on the basic value range block, the value range block that does not meet the threshold adopts an improved adaptive decomposition and combination algorithm. The method of adaptive combination is directly used to meet the threshold. The simulation results show that compared with the standard algorithm and quadtree algorithm, the proposed algorithm can not only improve the compression ratio but also reduce the peak signal-to-noise ratio and save the search time. It should be pointed out that although the geometric features are retained in the improvement of the quadratic error algorithm, the factor of viewpoint is not considered in the examination, which needs to be further studied.

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 no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] B. Gan, C. Zhang, Y. Chen, and Y. C. Chen, "Research on role modeling and behavior control of virtual reality animation interactive system in Internet of Things," *Journal of Real-Time Image Processing*, vol. 1, pp. 1–15, 2020.
- [2] Z. Liang, D. Qiao, and T. Sung, "Research on 3D virtual simulation of geology based on GIS," *Arabian Journal of Geosciences*, vol. 14, no. 5, 2021.
- [3] H. Chen, J. Zhao, Z. Wang, J. Dong, and T. Yu, "Modeling virtual abrasive grain based on random ellipsoid tangent plane," *The International Journal of Advanced Manufacturing Technology*, vol. 113, no. 7–8, pp. 2049–2064, 2021.
- [4] Y. Deng, S. Y. Han, J. Li, J. Rong, W. Fan, and T. Sun, "The design of tourism product CAD three-dimensional modeling system using VR technology," *PLoS One*, vol. 15, no. 12, pp. 205–244, 2020.
- [5] L. Luo, X. Yang, and X. Yang, "A 3-D scene management method based on the triangular mesh for large-scale Web3D scenes," *IEEE Multimedia*, vol. 26, no. 3, pp. 69–78, 2019.
- [6] P. Shan and W. Sun, "Research on landscape design system based on 3D virtual reality and image processing technology," *Ecological Informatics*, vol. 63, article 101287, 2021.
- [7] J. Yang, F. Wu, E. Lai, M. Liu, B. Liu, and Y. Zhao, "Analysis of visualization technology of 3D spatial geographic information system," *Mobile Information Systems*, vol. 2021, no. 2, Article ID 9173281, 9 pages, 2021.
- [8] M. Zheng, W. Tang, A. Ogundiran, and J. Yang, "Spatial simulation modeling of settlement distribution driven by random forest: consideration of landscape visibility," *Sustainability*, vol. 12, no. 11, p. 4748, 2020.
- [9] H. Tastan, T. Tong, and C. Toker, "Using handheld user interface and direct manipulation for architectural modeling in immersive virtual reality: an exploratory study," *Computer Applications in Engineering Education*, vol. 30, no. 2, pp. 415–434, 2022.
- [10] L. Guo and P. Wang, "Art product design and VR user experience based on IoT technology and visualization system," *Journal of Sensors*, vol. 2021, no. 5, Article ID 6412703, 10 pages, 2021.
- [11] L. Li, W. Zhu, and H. Hu, "Multivisual animation character 3D model design method based on VR technology," *Complexity*, vol. 2021, no. 4, Article ID 9988803, 12 pages, 2021.
- [12] S. K. Roy, S. Kumar, B. Chanda, B. B. Chaudhuri, and S. Banerjee, "Fractal image compression using upper bound on scaling parameter," *Chaos, Solitons & Fractals*, vol. 106, pp. 16–22, 2018.
- [13] F. J. Hernandez-Lopez and O. Muiz-Pérez, "Parallel fractal image compression using quadtree partition with task and dynamic parallelism," *Journal of Real-Time Image Processing*, vol. 19, no. 2, pp. 391–402, 2022.
- [14] N. Li, X. Song, K. Xiao, S. Li, C. Li, and K. Wang, "Part II: a demonstration of integrating multiple-scale 3D modelling into GIS-based prospectivity analysis: a case study of the Huayuan-Malichang district, China - ScienceDirect," *Ore Geology Reviews*, vol. 95, pp. 292–305, 2018.
- [15] P. C. Lee, L. L. Zheng, T. P. Lo, and D. B. Long, "A risk management system for deep excavation based on BIM-3D GIS framework and optimized grey Verhulst model," *KSCSE Journal of Civil Engineering*, vol. 24, no. 3, pp. 715–726, 2020.
- [16] W. Y. Lin and P. Lin, "Intelligent generation of indoor topology (i-GIT) for human indoor pathfinding based on IFC models and 3D GIS technology," *Automation in Construction*, vol. 94, no. 10, pp. 340–359, 2018.
- [17] D. Černá, "Postprocessing Galerkin method using quadratic spline wavelets and its efficiency," *Computers & Mathematics with Applications*, vol. 75, no. 9, pp. 3186–3200, 2018.
- [18] P. Mutunge and D. Haugland, "Minimizing the tracking error of cardinality constrained portfolios," *Computers & Operations Research*, vol. 90, pp. 33–41, 2018.
- [19] S. H. Huo, Y. S. Li, S. Y. Duan, X. Han, and G. R. Liu, "Novel quadtree algorithm for adaptive analysis based on cell-based smoothed finite element method," *Engineering Analysis with Boundary Elements*, vol. 106, no. 9, pp. 541–554, 2019.
- [20] L. Zhao, X. Zhao, S. Zhu, and R. Fu, *A Multi-Level Adjacent Searching Algorithm of Degenerate Quadtree Grid on Spherical Facet*, vol. 43, no. 4, 2018 Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics and Information Science of Wuhan University, 2018.
- [21] Z. Wang and X. Lü, "Terrain rendering LOD algorithm based on improved restrictive quadtree segmentation and variation

Retraction

Retracted: Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xu and Z. Liang, "Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles," *Journal of Function Spaces*, vol. 2022, Article ID 9295395, 10 pages, 2022.

Research Article

Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles

Jijin Xu  and Zhen Liang 

Fujian Key Laboratory of Agricultural Information Sensing Technology, College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou 350100, China

Correspondence should be addressed to Jijin Xu; michaelxu@fafu.edu.cn

Received 16 May 2022; Revised 21 July 2022; Accepted 26 July 2022; Published 17 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Jijin Xu and Zhen Liang. 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.

Unmanned technology is an important development project of today's cutting-edge science and technology, which has a significant impact on social and economic development, national defense construction, and scientific and technological development. The rapid development of industrial information technology has driven the unmanned intelligent vehicle system to innovate and gradually enter the public's view, and at the same time, the driving safety of unmanned intelligent vehicles is also widely concerned. Target information perception system is the foundation of unmanned system and the fundamental guarantee of safety and intelligence of unmanned vehicles. There are three key problems of target recognition in unmanned driving, namely, target classification, localization, and attitude determination. In the implementation of a networked virtual environment system, a flexible and complete perception model is needed as the guiding model of the system. Since 3D point cloud data can provide more spatial information than 2D RGB image data, it is more beneficial to determine the target category, position, and pose in 3D. In this paper, based on the existing research of unmanned intelligent vehicle perception system, we combine the application of fusion of 3D target information perception model and develop a nighttime unmanned system based on multiview fusion of 3D target information perception model. This system can simultaneously perform the detection of multiple categories of objects and predict the center point, length, width, height, and orientation of the objects, so that the unmanned car can sense the location of the surrounding objects when driving in the actual scene.

1. Introduction

As a class of intelligent robots, the development of unmanned vehicles has a significant impact on social and economic development, national defense construction, and science and technology, and has become a strategic target for research in high-tech fields in various countries around the world [1]. Unmanned driving technology includes key technologies such as target information perception, path planning, vehicle control, and intelligent decision-making [2]. The multiview channels used in unmanned intelligent driving vehicles are visible light cameras, but ordinary visible light cameras cannot accurately obtain specific information in road scenes at night in dim and low-light environments

[3]. In the field of 3D spatial perception research, it is an important capability for robots to acquire 3D information through relevant sensors and to be able to understand the scene and interact with the surrounding environment through this data [4]. Among them, target information perception technology is the underlying module of unmanned technology and the basic guarantee for safe vehicle driving [5]. Through unmanned driving technology based on 3D target information perception model with multiview fusion, we can effectively control many unstable human factors, such as drunk driving and fatigue driving [6].

Unmanned vehicles traveling on the road at night need to sense the surrounding environment of vehicles, pedestrians, etc., and although the infrared camera-based

perception scheme can sense the obstacles in front of the vehicle, it cannot accurately obtain the location information of the obstacles and has limited accuracy [7]. Target information perception for unmanned vehicles is a unique machine-like human-like understanding of the vehicle's surroundings through some multiview channels, such as LIDAR and cameras [8]. 2D target detection is unable to obtain real-world target parameters such as physical dimensions and 3D coordinates due to the use of image information only and to identify target classes and locate targets in the image plane [9]. The basic requirement for 3D target recognition in a typical unmanned scenario is accurate classification, localization, and pose recognition of the target to be detected [10]. According to the differences in the acquired data, the existing 3D target detection algorithms can be divided into two categories: visual and laser point cloud.

Target detection and localization in three-dimensional space is an important basis for people to perceive the environment and has important applications in robotics, autonomous driving, security monitoring, industrial production, intelligent agriculture, and many other fields [11]. And because such sensors as cameras are vulnerable to external factors such as light and weather, and it is too difficult to infer three-dimensional information of the target from two-dimensional images, these algorithms are difficult to achieve satisfactory results in obtaining three-dimensional detection frames of the target [12]. In most of the night vision related research generally uses infrared camera instead of visible camera, the imaging principle of infrared camera is different from visible camera, which is able to transform the temperature distribution of the object in the surrounding scene into an image by detecting the object's own infrared radiation, and then through photoelectric transformation [13]. For safety reasons, autonomous driving is a scenario that requires extra precision as well as stability. In order to improve the performance of the algorithm, most of the current 3D target detection algorithms in autonomous driving scenarios use LiDAR as the main sensor. The chance of traffic accidents is much higher at night than during daytime; therefore, the 3D target information perception model using multiview fusion is important to study how unmanned intelligent vehicles can better perceive the surrounding environment information in the night scenario.

In this paper, a 3D target detection algorithm is proposed by fusing 3D target perception with LIDAR point cloud. The algorithm effectively improves the target detection accuracy of the unmanned intelligent vehicle in the night scene and makes the unmanned intelligent vehicle have better night target information perception ability. The innovative contribution lies in the research of driving behavior planning based on multiview fusion of three-dimensional target information perception model. It enables the driverless intelligent vehicle to make correct driving behavior planning in real time after acquiring the surrounding environment information at night. The basic data format of 3D target recognition algorithm is point cloud, which is relatively stable and not affected by illumination. It can truly reflect the position of objects in the scene and is very suitable for target recognition.

2. Application of Three-Dimensional Target Information Perception Model in Unmanned Intelligent Vehicle at Night

2.1. Construction Method of Three-Dimensional Target Perception Model Based on Multiview Fusion. The 3D target detection model with multiview fusion consists of a feature extraction module, a candidate region generation module, and a channel fusion module. Unmanned vehicle is a comprehensive technology, and if we look at the functions of its parts, we can divide it into 3 parts: target information sensing, decision planning, and vehicle control [14]. However, the point cloud data is very sparse and uneven compared with image data and the data volume is large, which makes the data application and storage calculation face new problems. The 3D target recognition algorithm focuses on the design of algorithms for the above two problems, including various forms of reprocessing of the point cloud and clever design of the network structure to understand the geometric information in the point cloud [15]. The prediction-tracking relationship between the front and back frames of lane line detection is then established to further improve the accuracy and efficiency of lane line detection. The function of general tracking is to predict the location of road features in the next frame image, detect road features in a small range, and improve efficiency. If no road feature is detected within the prediction range, the position of the estimated or previous frame feature is used. If no road feature is detected for several consecutive frames, the full image road feature detection is started. The flow chart of the entire lane line detection algorithm is shown in Figure 1.

First, the feature extraction module acts on the encoded bird's eye view form and the front view form of the infrared image data and the LiDAR point cloud data, respectively. Each LiDAR scan is processed independently instead of creating a two-dimensional point cloud map. This reduces the complexity of the algorithm by analyzing only a single dimension of laser data, thus reducing the complexity of the algorithm. The target information sensing system acquires real-time and accurate information about the road environment through a multiview channel device. The threshold value T of the method is calculated by averaging E , squared difference between pixels P , and root mean square value Q between pixels for this window together, and the parametric equation for performing the threshold calculation is as follows:

$$T = a * E + b * p + c * Q. \quad (1)$$

The multiview channel data fusion discovers the existing and possible risk factors in the environment and reports them to the decision-making module in time. In the urban road environment, the main road types are tarmac, concrete, grass, and gravel roads, and the vehicles will have different characteristics when driving on different road types due to different friction coefficients and side

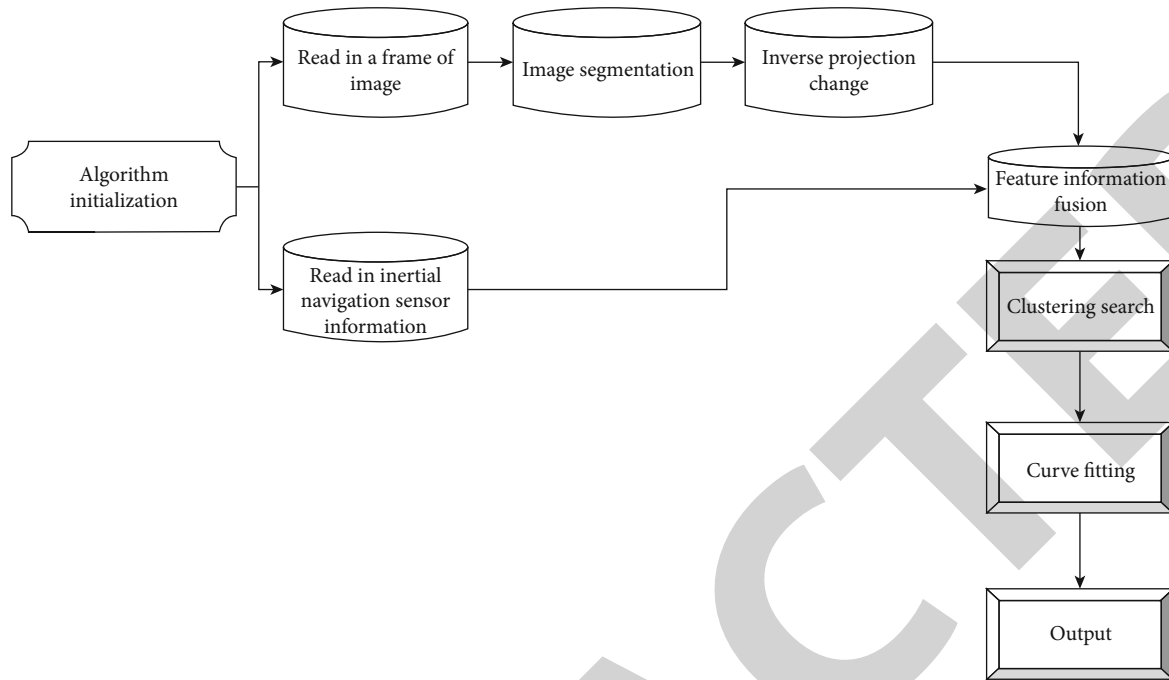


FIGURE 1: Flow chart of lane detection algorithm.

slip angles. The equations to calculate the position translation based on the unmanned vehicle speed and angular velocity are as follows:

$$E = B + V\Delta t \cos(w\Delta t), \quad (2)$$

- E —end position
- B —starting position
- V —speed of a motor vehicle
- w —angular velocity
- Δt —camera sampling period

Therefore, for the linear discrete system, it is required to satisfy the process model as a linear system with superimposed process excitation noise, the measurement model as a linear system with superimposed measurement noise, and the noise both obey normal distribution. Using the end-to-end convolutional neural network, the image obtained from the color camera is used as the input, and the steering wheel angle signal is output directly after the network decision, and the model schematic is shown in Figure 2.

Second, the candidate region generation module is highly effective in 3D target detection algorithms. The network draws on the ideas of the 3D target detection algorithm and uses a network instead of manual extraction to generate candidate regions of objects, using a bird's eye view form of LiDAR data as the input to the candidate region generation module. The input 3D image is given to the 3D detector for target detection. Using preprocessed dense/sparse labels, the nonroad labels in the sparse part are moved from the sparse points to the dense points closest to the vehicle perimeter. If a stochastic process is a Markov decision process, then for that stochastic process there must be a strategy that yields

better results than the other strategies in all cases, which is also called the optimal strategy, then.

$$\pi^*(s) = \arg \max Q(s, a), \quad (3)$$

- s —time
- $\pi^*(s)$ —optimal strategy
- a —action

Then, the optimal value return function can be obtained:

$$Q^*(s, a) = \sum_{x_s \in \mathcal{S}} T(s, a)(R(s, a) + \gamma V^*(s^*)), \quad (4)$$

- γ —influence factor

LIDAR emits tens or hundreds of thousands of laser pulses per second to the environment, and by measuring the return time of the pulses, information such as target distance and orientation can be calculated. Therefore, the perception model design of the network virtual environment must take into account the user requirements and the environment of the system, to be able to dynamically adjust the display and network transmission data according to the implementation conditions and to maintain the user's sense of reality. Instead of sticking to the optimization of the global model, the approach we have chosen is a transition from the local optimum to the global optimum. It can be used to describe, understand, analyze, and evaluate people's perception of information security, and can quantitatively analyze and compare the characteristics of different information security threats, thus establishing a theoretical basis for research in the field of information security and human-computer interaction. It requires loading some binary code into each client and then converting the JDBC calls from the loaded client API to Oracle, Sybase, Informix,

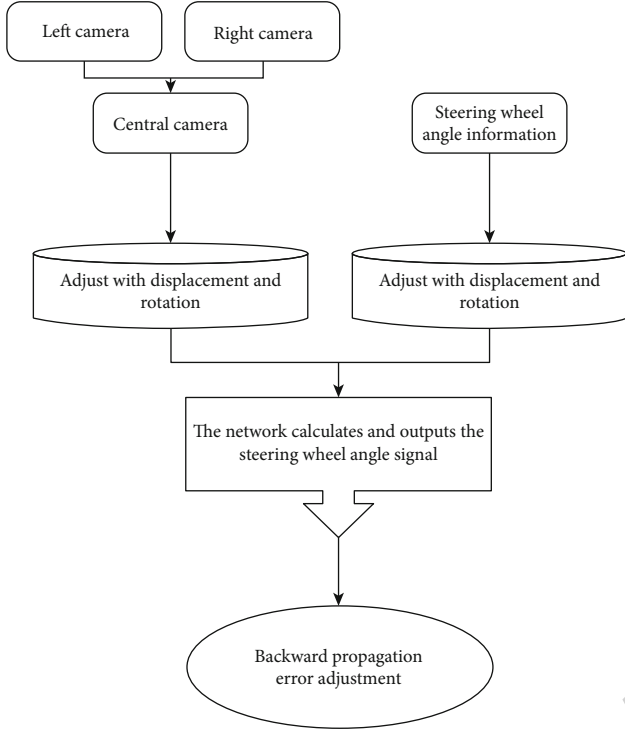


FIGURE 2: 3D target perception model.

DB2, or other DBMS transfers. Thus, given by two point sets:

$$\begin{aligned} A &= \{(A_1, w(A_1)), (A_2, w(A_2)), \dots, (A_m, w(A_m))\}, \\ B &= \{(B_1, w(B_1)), (B_2, w(B_2)), \dots, (B_m, w(B_m))\}, \end{aligned} \quad (5)$$

$w(A_i), w(B_i)$ —weight

Finally, the channel fusion module uses a channel fusion network based on regional unification. This network effectively combines the features of each viewpoint, jointly classifies the targets on each viewpoint, and performs directional regression for the 3D candidate regions. Thus, road edge point detection is performed, and after processing such as filtering and fusion, a travelable area optimal selection algorithm is then designed to obtain the most passable area. Before filtering and fusion, the point cloud in the KITTI dataset is projected onto the RGB image plane to perform a matching verification operation with RGB. The purpose is to ensure the data match and to exclude the effect of bias caused by the data itself in the subsequent algorithm. The view cone of point cloud data is passed to the image-point cloud fusion segmentation network for front and background segmentation and to determine whether there is valid point cloud data, i.e., the scanned point cloud is obtained with the target object returned. Then, sparse points are removed from the classification list, and the final result list represents the location of possible road areas and surrounding geometric features.

2.2. Working Process of 3D Target Perception Model in Unmanned Driving. The application of 3D target perception

model in unmanned driving is divided into two parts: front-end data acquisition and data processing, with data transmission between the two parts via wireless communication. Using feature-centered Euclidean distance for target detection is the first method to use sparse convolutional layers and L1 regularization for processing large-scale 3D data. When the Euclidean distance is greater than a threshold, the point is considered a feature point; otherwise, it is a non-feature point, and the screening formula is as follows:

$$p_i = \begin{cases} \text{True}, & d_m > \sigma_h \\ \text{False}, & d_m < \sigma_h \end{cases}, \quad (6)$$

σ_h —threshold value

P_i —a point in the set of points P

It is very necessary and important to study how to use 3D target perception model to process point cloud data efficiently and obtain accurate 3D target detection results for autonomous driving. The unmanned system generates many conditional branches for the information perceived by the multiview channels, and each conditional branch is divided into modules, making the whole system respond quickly to different information inputs. The unmanned behavior planning system is shown in Figure 3.

First, the data from the network and the terminal are passed through a parameter conversion processor to complete the data separation. The acceleration data is preprocessed by means of a low-pass filter with a frequency of up to 100 Hz to process the raw acceleration data for non-true values and high-frequency noise. The same point clouds undergo certain rigid changes in space (rotation or translation) and their coordinate values change. However, we expect that the network's recognition of the point cloud should remain unchanged regardless of the rotation and translation of the point cloud. Then, the residuals squared at each point are summed as the total residuals for the whole fitted plane. The residuals from a point to a plane $aX + bY + cZ = 1$ are calculated as shown in the following equation.

$$\delta_i = \frac{|ax_i + by_i + cz_i + 1|}{2\sqrt{a^2 + b^2 + c^2}}, \quad i = (1, 2, \dots, n). \quad (7)$$

It is then necessary to pre-set thresholds to determine the properties of the grid, and the size of the threshold selection reflects the ability of obstacle detection. It is assumed that the average distance calculated from all points conforms to the Gaussian distribution, and the mean and standard deviation can be obtained, and the points whose average distance is outside certain criteria are eliminated. A multilayer point deconvolution is built as the main structure of the segmentation network, and feature fusion with the corresponding point convolution layer of the feature network is added. By adding a classification loss function for each point, the final target segmentation network can predict a category for each point, thus achieving the effect of target segmentation. And after transforming into the form of bird's eye view data, the amount of data can be largely reduced, while the information is arranged in pixel form with a uniform and regular pattern. Suppose the

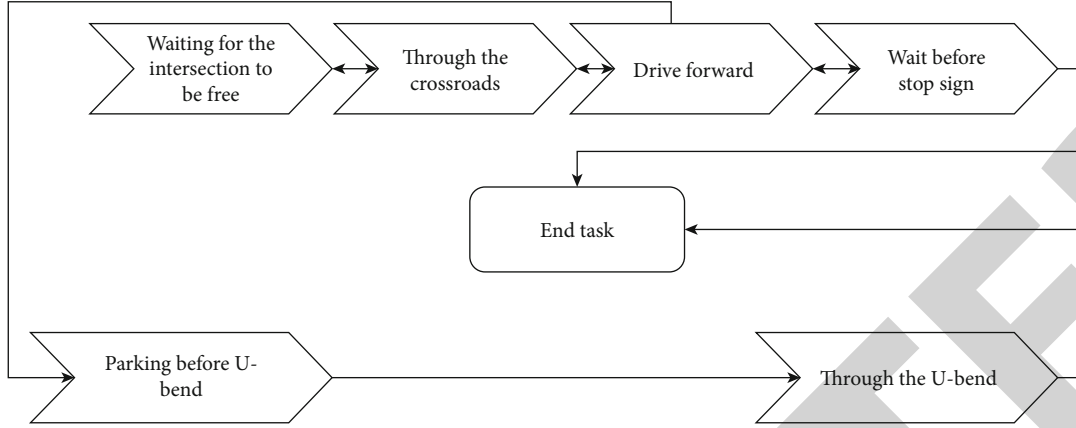


FIGURE 3: Unmanned driving behavior planning system.

system uses n multiview channels to measure a physical quantity, the i th multiview channel measurement is $X_i (i = 1, 2, \dots, n)$, and the weighting factor of each multiview channel is w_i , then the final fusion result is:

$$\bar{X} = \sum_{i=1}^n w_i X_i. \quad (8)$$

Secondly, the separation is based on the data such as model position state obtained from the static model, separated into two parts, one part is used as parameters of the model data converter, and the other part modifies the dynamic data area in the model. The network structure of the auxiliary path is the same as the number of intermediate layers of the original network structure, and each intermediate layer in the auxiliary path shares the weights of the corresponding intermediate layers of the original network. In the backpropagation process, each auxiliary path also performs loss calculation, and the calculated loss on the auxiliary path is called the auxiliary loss, and the weights are also updated on the auxiliary path after calculating the auxiliary loss. Radius filtering is to define the points lacking enough nearest neighbors as isolated noise points. It is necessary to set the search radius and the threshold of the number of nearest neighbors, and if the set conditions are not satisfied, the corresponding points are removed. This part performs a convolution operation on the image block and the view cone point cloud data in parallel, where the image is convolved using a dimensionally invariant multiset convolution and the point cloud is convolved using a point convolution network and the features of the image are fused at each layer of the point convolution network. For this purpose, candidate point sets need to be constructed separately for the points in the point set P , and the candidate point sets in Q are constructed as follows.

$$C = \{c_i | c_i = \langle p_i, q_{i1}, q_{i2}, q_{i3}, \dots, q_{ik}, 1 \leq i \leq 4 \rangle\}, \quad (9)$$

c_i —the i th point

p_i —candidate point set

k —number of candidate points

Finally, after adjustment, getting the adjustment and data acquisition strategy, the model conversion processor obtains the model data of the display from the static model data, then merges the dynamic model data, and finally sends the data to the display or network device. Point cloud noise refers to the existence of a certain probability that a point lies in a threshold radius around its original position. Compared with other similarity measures, the use of rank function is not so sensitive to image noise and luminance differences and is good in real time, and its calculation formula is:

$$\begin{cases} C_{\text{Rank}}(u, v, d) = g_1(u, v) - g_r(u + d, v) \\ g_1(u, v) = \sum_{(i,j)} g_1(u + i, v + i) < g_1(u, v) \end{cases} \quad (10)$$

Also, there are outliers, i.e., random distributions of coordinates, in point clouds. Radar data are in the form of point clouds with varying amounts of data, but they all contain a large amount of noise, so the noise needs to be filtered out by a filtering algorithm. However, the sampling frequency of the horizontal axis of the road contour is no longer uniform due to the changing vehicle speed, and the sample interval of the road contour is small when the driving speed is slow and large when the driving speed is fast. So after data acquisition with noise removed, the system will fuse LiDAR features and image features and a priori knowledge of the road model to perform drivable area detection. For the occluded objects with different attributes, the size and arrangement of the objects described in the high-precision scene environment as well as are different. To make the acceleration data and velocity data sampled at the same frequency, the velocity data are interpolated.

3. Application Analysis of 3D Target Detection Algorithm in Unmanned Driving

3.1. Analysis of 3D Target Detection Algorithm. The 3D target detection algorithm can obtain the classification result of the target object and the coordinate information in 3D space by processing the 3D data. Due to the binary output characteristics, the cross-entropy function is more suitable

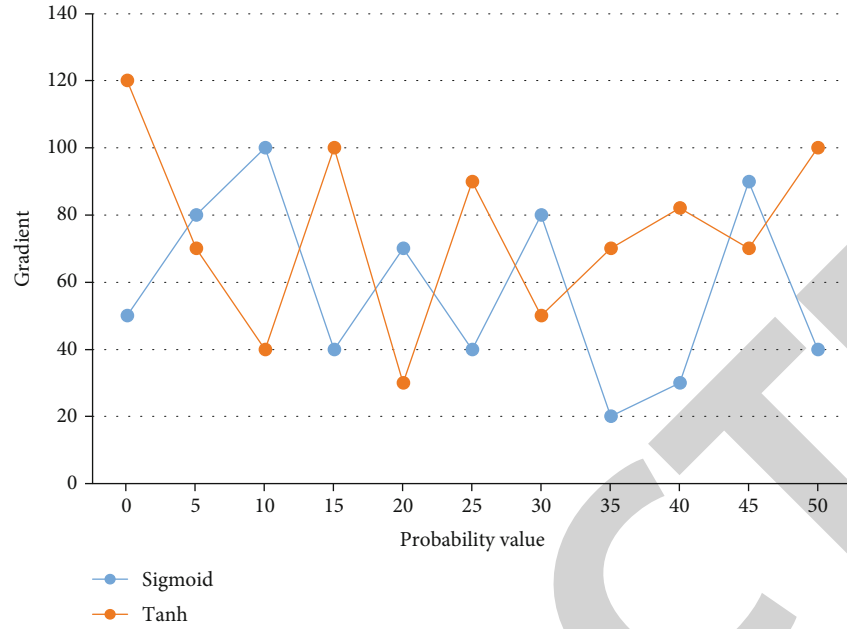


FIGURE 4: Comparison of gradients produced by different activation functions.

for pavement type classification than the squared error cost function, and the cross-entropy function is closer to the criteria of real classification. The 3D target detection algorithm separates the point clouds in space according to the foreground background based on the class prediction of each point by the target segmentation network. And the activation function is used to extract the points of the foreground object category out of them, so as to complete the extraction of the target object. The activation function must be nonlinear, and the activation function must be derivable everywhere to ensure computability in gradient calculation. The comparison of the gradients generated by different activation functions is shown in Figure 4.

First, after the depth information of each pixel on the image is determined by monocular or binocular vision, each coordinate is assigned to the corresponding depth information. The corresponding depth map is obtained and then transformed into a point cloud by geometric principles. The input point cloud data needs to be gridded, and the resolution of the gridding will determine the operation efficiency and detection accuracy of the algorithm later, too large may lead to slow operation and overfitting of the algorithm, too small will not be able to accurately detect the target location. Reasonable grid node distribution plays an important role in the calculation of nonlinear hyperbolic conservation law equations. The numerical solution of hyperbolic conservation law equations has always been one of the important research topics in computational fluid dynamics. In the numerical solution of nonlinear hyperbolic conservation law equations, if the mesh is uniformly divided, in the region where the properties of the solution change little, the sparse mesh distribution can obtain more accurate results. When there is a vehicle signal, the sampled signal will have greater fluctuation, and the median filtering method can effectively remove the impulse interference signal, so as to get a

smooth waveform signal; Figure 5 shows the comparison of the signal after AMR-X and AMR-Y filtering.

Due to the limitation of the working principle of LiDAR, the point clouds collected have the characteristics of near-dense and far-sparse, and the high density of near points easily leads to the slow operation of the clustering algorithm. Therefore, voxel filtering of the point cloud data is needed first to improve the real-time performance of the algorithm. Compared with the basic detection network architecture such as VGG, Darknet53 has a residual module that allows the network to be optimized more easily at the same depth. Therefore, the residual module allows to increase the network depth to improve the network performance without degradation. After the channel fusion network fuses the features of each viewpoint channel, the fusion results are used to perform regression correction on the 3D candidate regions generated by the candidate region module.

Next, the pseudo point cloud data corresponding to these target objects are found based on the pixel mapping of the target objects segmented on the image to the pseudo point cloud. After meshing the point cloud data, the features of each mesh need to be specified, which can be set artificially or extracted using deep learning. The front-end data acquisition part collects data from each lane of the traffic intersection and packages the data for transmission to the upper computer data processing host for calculation and processing via wireless sensor network according to the developed transmission protocol. The upper computer data processing host displays traffic information such as current traffic flow, speed, and time of obtaining vehicle information, and saves the traffic data. The decomposition of image basis functions with ICA is limited to the spatial domain only, and if temporal variations are considered, such as continuous horizontal motion, vertical motion, and rotational motion, image sequences in the time domain can be

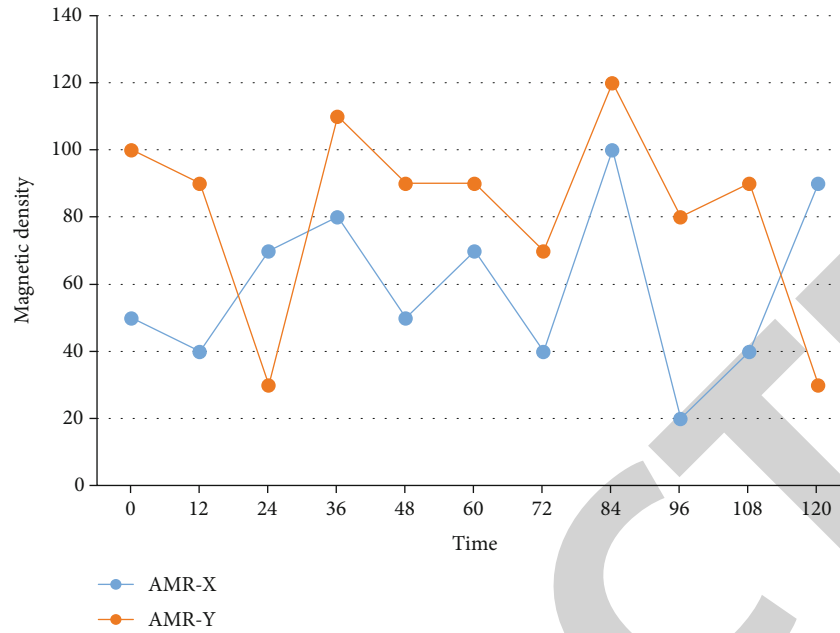


FIGURE 5: Comparison of AMR-X and AMR-Y filtered signals.

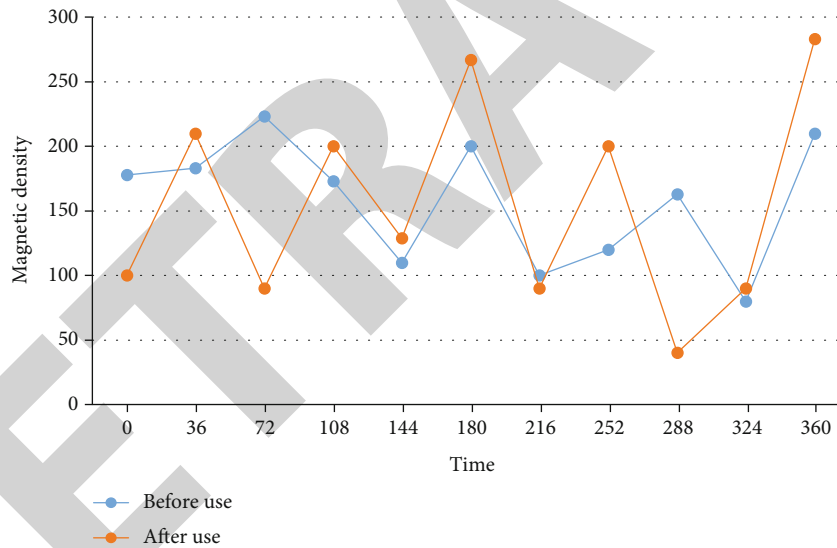


FIGURE 6: Waveform changes of a driving vehicle before and after using 3D target detection algorithm.

composed. The classification and regression of regions are performed by two branches of the region recommendation network, and then feature extraction is performed on the feature map according to the size scaling of the filtered recommended regions and sent to the classification and regression branches to obtain the class and location of the target. In order to detect low-speed vehicles in congested lanes, the waveform change of a moving vehicle before and after using the 3D target detection algorithm based on the multi-state machine vehicle detection algorithm is shown in Figure 6.

Finally, the pseudo point cloud in the candidate frame and the corresponding labeled frame of the candidate frame are fed into the second stage of the Point RCNN network for

candidate frame correction training. The 3D convolution of the gridded data is performed to extract the necessary features. Based on the point cloud data, the laser beam to which the point belongs and the distance from the radar center can be calculated. If the distance is less than the set distance of the laser beam to which it belongs, then the point is a non-ground point; if the distance is not much different from the set distance of the laser beam to which it belongs, then the point is a ground point. The distance between the data points near the vehicle is still much smaller than the distance far away from the vehicle. If we simply use the same radius threshold for point cloud clustering, we will not be able to achieve good clustering effect in both near and far away. If the threshold is too large, the objects in the near area and

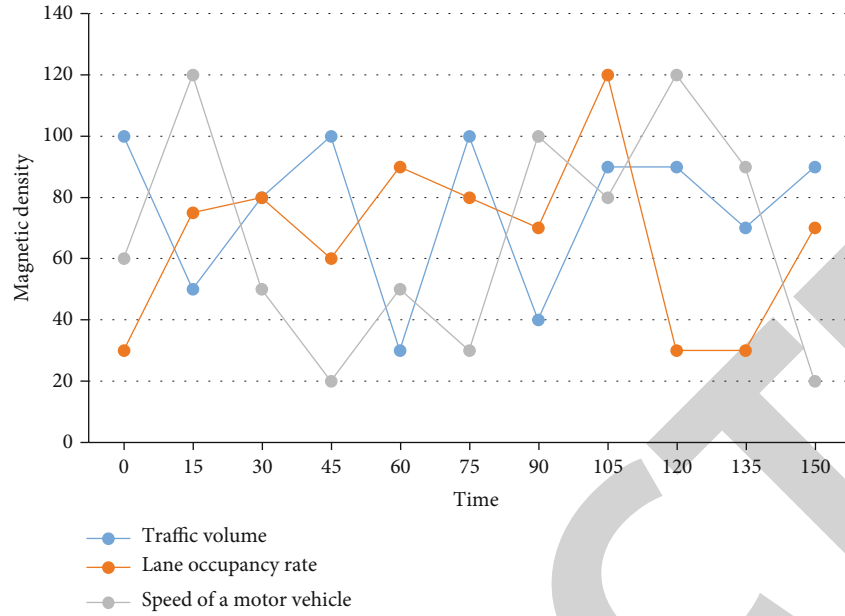


FIGURE 7: Vehicle identification results.

the environment will be grouped into the same cluster, while if the threshold is too small, the targets in the far area of the vehicle will not be clustered correctly. Therefore, one dimension can be reduced by setting the height span of the 3D convolution, fusing the dimension of height into the information of the channel, and then regressing the position and size of the target using the anchor frame and the final feature map as in the image detection algorithm.

3.2. Target Detection and Location Analysis. Different sensors output different forms of data, and the corresponding data processing processes are different. Before radar data can be processed, target detection and localization must be analyzed according to the radar protocol. After multiview channel calibration and data acquisition, the algorithm incorporates laser features, image features, and the necessary a priori knowledge to achieve driving area detection. Since the acquisition frequency of IR camera and LIDAR is not the same, this paper matches the IR image according to the LIDAR frame number, sets the threshold value to 20 frames, and looks for the image with the radar frame number less than the threshold value and closest to it as the corresponding image. After acquiring the parameters, the various data provided by the system implementation layer should be integrated, and the scheduling and adjustment should be carried out. The variance is multiplied as a weight factor by the original signal as the feature signal, so that the characteristics of the magnetic field strength signal can be maintained and the drift of the signal can be suppressed by the characteristics of the variance. The vehicle recognition results are shown in Figure 7.

First, a pseudo point cloud is generated and the data with height more than 2 m above the observation origin is removed, and the point cloud is centered on the ground. Since most of the information of the point cloud target is gathered at its edges, the features at the edges can be

extracted effectively by using deformable convolution to achieve the same effective accuracy, so the setting of the anchor frame is not necessary, and the target frame is re-clustered for each new scene. A multilayer point convolutional network is built using point convolutional layers as the main structure of the localization network, and several branching networks are added after the main body of the network according to the prediction task as the head structure of the prediction network. When the vehicle passes above, the AMR sensor detects the information of the vehicle's perturbation to the earth's magnetic field in real time, and the subsequent processing circuit amplifies and filters this perturbation signal and then converts it into a digital signal by the controller's A/D sampling channel processing to analyze the sampled time series signal. The time synchronization algorithm is used to time synchronize the two sensor nodes, and the vehicle detection algorithm mentioned in this paper can be used to obtain the moment when the vehicle passes the sensor node, and the waveform perturbation schematic of the dual sensor node speed measurement process is shown in Figure 8.

Next, the left-eye images are put through Mask RCNN model and inference is performed. Then, the minimum enclosing box algorithm is used to calculate the minimum enclosing box of the point cloud to get the position, size, and direction of each target in the LIDAR column coordinate system. All the 3D radar data falling into the grid are calculated to get their maximum-minimum quotient difference, and if the difference is greater than the set idle value, the grid is an obstacle point, and vice versa; it is a ground point. The main concern is whether the three-dimensional detector has missed detection, because if it misses detection, the three-dimensional target location network is unable to retrieve the lost object again. By analyzing and processing the geomagnetic vehicle characteristic information collected by the receiving station, the real-time data waveform of each

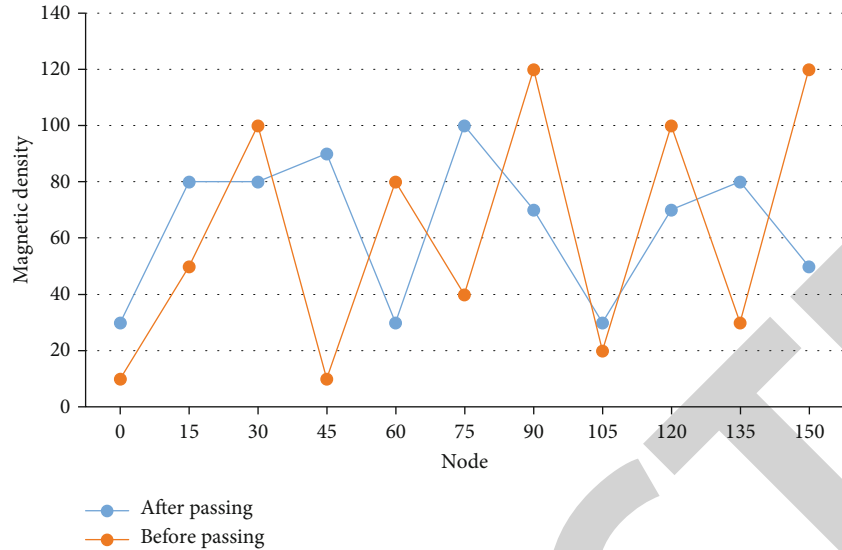


FIGURE 8: Schematic diagram of magnetic field waveform disturbance when a vehicle passes through a dual sensor node.

TABLE 1: Comparison of target test results.

	Fast RCNN	SSD	Mask RCNN
Voc index (mAP)	67.26	73.45	86.71
COCO (AP50) index	34.99	45.62	57.38

TABLE 2: Comparison of three-dimensional target detection results.

	VoxelNet	AVOD	MV3D
Model	LiDAR	RGB + LiDAR	RGB + LiDAR
Difficulty index	52.99	76.17	56.32

lane is displayed. By calculating and obtaining traffic information such as road traffic flow, vehicle speed, single vehicle length, vehicle running direction and vehicle type, reliable data are provided for the traffic information center. And for different scene obstacles, the final multipath effect is different because of their material, shape, and other property differences, and the reflected wave characteristics are also different.

Finally, the mask generated by Mask RCNN is combined with the pseudo point cloud to delineate the point cloud foreground and background and filter out the background information. In the process of updating the network parameters using stochastic gradient descent and chained derivatives, the gradient descent algorithm must accurately calculate the error of each layer before updating the parameters. The comparison of the target detection results of Fast RCNN, Mask RCNN, and SSD is shown in Table 1.

The average precision mean result of Mask RCNN is 13.26 higher than SSD on the PASCAL VOC test set and 19.45 better than Fast RCNN.

If the selected training data does not fully contain all the feature information that the test data has, it can seriously affect the classification results. Therefore, a sampling window of size 5×5 pixels is randomly placed at a pixel location

in a large image. It is important to note that the entire sampling window boundary should be within the boundary of the large image, cropping a small piece of the same image size. The real regions of the three categories of vehicles and pedestrians and cyclists are also marked on each copy of data by manual annotation. A bottom-up candidate region search network is used to extract the candidate regions from the point cloud. The detection results of each 3D target detection algorithm MV3D, AVOD, and VoxelNet are shown in Table 2.

The foreground points are marked in green in the whole point cloud scene to observe the performance of the 3D segmentation network. And the results of the 3D prediction module are represented by 3D boxes to observe whether the network predicts the 3D pose correctly. The purpose is to increase the nonlinear representation capability of the network to obtain more refined image features. Also, the design of filtering the reflected waves by means of antenna polarization, etc., in order to eliminate the negative impact of the localization process due to multipath effect is applied in many receivers.

4. Conclusions

Unmanned self-driving vehicle is also known as intelligent vehicle. It uses on-board multiview channels to sense the vehicle's surroundings and controls the vehicle's steering and speed based on the road, vehicle attitude, and obstacle information obtained from the sensing, thus enabling the vehicle to travel safely and reliably on the road. If RGB data is used, although the 2D recognition network is more mature, it will face the problems of dark light failure and the presence of strong occlusion. In addition, the method of classifying different road types using acceleration multiview channels alone is relatively ineffective for classifying and recognizing tarmac and concrete roads with similar vibration characteristics. In contrast, using a 3D target information perception model with multiview fusion, after

Retraction

Retracted: The Construction of Adaptive Learning for Sports Based on Aerobics Trajectory Recognition Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xi, "The Construction of Adaptive Learning for Sports Based on Aerobics Trajectory Recognition Model," *Journal of Function Spaces*, vol. 2022, Article ID 8339745, 8 pages, 2022.

Research Article

The Construction of Adaptive Learning for Sports Based on Aerobics Trajectory Recognition Model

Chaojie Xi 

Department of Sports, Sanquan College of Xinxiang Medical University, Xinxiang, Henan 453003, China

Correspondence should be addressed to Chaojie Xi; 11572011@sqmc.edu.cn

Received 28 May 2022; Revised 16 July 2022; Accepted 23 July 2022; Published 9 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Chaojie Xi. 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.

Perceiving the movement track of aerobics is a key element of learning aerobics, but the current aerobics movement is not very professional, the ability to identify the movement track is weak, and improper movement in the movement process is easy to cause physical injury. In order to improve the safety of athletes in bodybuilding training, this paper uses Kinect to hold the coach's body contour, determine the standard level of coaches' sports, and combine the characteristics for aerobics training, so as to improve the sports level of coaches, through data acquisition, data processing, and feature extraction to assist sports learning, as well as human posture recognition. The calculation and recognition of human skeleton joints are completed by two algorithms, which improve the human motion recognition algorithm. The aerobics data collected by Kinect device is specified and digitized, which enhances the robustness of the system and improves the performance of the algorithm and the accuracy of the motion data.

1. Introduction

The quality of our national physical fitness is gradually becoming a topic that requires attention. And enhancing the national physical fitness will need a certain degree of physical exercise, and the public learning aerobics program movements, the need for repeated practice of the action, and aerobics training will inevitably encounter high costs, risks, and other difficulties, which inspired us to design a set of techniques or equipment that can simplify the daily physical exercise. Perceiving the track of aerobics is the key element of learning aerobics, but the current way of aerobics is not professional, and the recognition of track movement is weak. Improper actions during exercise are easy to cause physical damage. The cost of employing professional aerobics coaches is higher and the audience is smaller. And the judgment of human motion trajectory is not accurate, which affects the actual motion effect. [1].

At present, there are many devices for virtual simulation experiments. These devices can carry out different virtual somatosensory analysis and simulation for human body.

By analyzing the movement state of the human body, the organization recognizes the wearing sports suits of different movements. So athletes can carry out synchronous data tracking in the process of sports [2]. And virtual training will need an animation that can show the characteristics of hot dance to assist training, and 3D animation technology can solve this problem well; it can simulate human movement [3], natural and smooth display of human posture, due to its high accuracy and operability and other characteristics, which is widely used in all aspects of life, including medical detection of human health status game character model design. It is increasingly accepted and used by the general public. The algorithm applied to 3D animation technology is also born and developed rapidly: from the very first frame animation, such as "matchmaker," to the skinning technology which this system focuses on, combining the advantages of both linear skinning algorithm and quaternion linear skinning algorithm [4], removing the dross and taking the essence, better applying to the present day 3D animation technology. The Kinect depth data stream sensor can provide 3D depth data [5]. This paper analyzes the problems

in aerobics training. An innovative aerobics track recognition system based on device motion capture technology is designed, and the corresponding solutions are proposed.

This paper is divided into five structural parts. The first part explains the research background of this system. The research of aerobics track recognition system based on device motion capture technology is analyzed. The second part compares the current research status and references in related fields. The third part introduces the proposed human motion recognition algorithm using Kinect to optimize bone joints. This paper clusters these features based on static k -means algorithm and analyzes the improved hidden Markov model and artificial neural network algorithm, which focus on extracting joint distance features. The fourth chapter tests and analyzes the scheme proposed in this paper. The experiment analyzed 400 different aerobic posture sequences. Finally, this paper is summarized. The aerobics data collected by Kinect device is designated and digitized, which enhances the robustness of the system and improves the performance of the algorithm and the accuracy of sports data.

2. The Related Works

The launch of Microsoft Kinect device has injected fresh blood into the field of artificial intelligence, and with it, the problem of collecting human skeletal joint positions has been solved, and more and more researchers have set out to develop many systems that can be applied to life. Through Kinect recognition of palm bones, for the study of gestural skeletal movements [6–8], through Kinect device accessing, through the design and analysis of a synchronous motion diagnosis and rehabilitation system, some scholars have explored the intelligent home style by developing some learning tools to simplify the lifestyle. Assist athletes in daily medical diagnosis [9]; Xu et al. designed and researched a somatosensory educational game in response to the trend of the times and combined it with contemporary preferences [10]. Mao et al. used Kinect to get a better application in swimming events to guide athletes in stroke contact [11].

The characteristics of the system are as follows. (1) Kinect is a camera used with XBOX360. It is like a camera, which can be connected to the game console through USB interface. (2) Use infrared positioning: Kinect is more intelligent than ordinary cameras. First of all, it can emit infrared rays, so as to carry out stereo positioning of the whole room, and the camera can recognize the movement of the human body with the help of infrared rays. (3) Multiple additional functions: this product can not only recognize the human body through infrared ray but also recognize the complete RGB color and automatically log in for users with the help of face recognition technology. (4) Equipped with its own interface: when Kinect is installed, users must use an independent menu system instead of the original interface of XBOX360. You can also pause the game directly through voice or put your hand in the air and hold the virtual pause button. (5) Built in chat software videokinect.

Many studies have conducted in-depth data collection on some motion data by using human body markers. Through the use of artificial intelligence and retrograde

analysis of movement characteristics under different sports modes, the movement of athletes under video monitoring is disassembled and marked. Recognize the best barrier free movement mode under the condition of human vision. From it, the movement position tracking simulation is carried out to pave the way for future training. Xue et al. solved the problem of body behavior recognition and description, thus making good use of the device to capture the pose and action with inertia and developed the corresponding system [12].

In the choice of human modeling for Kinect, both HMM and ANN algorithms are widely used in the field of modeling due to the ability of HMM algorithm to optimize the computational process using its own dynamic modeling characteristics and ANN algorithm to classify and integrate the modules and resources of the system with its powerful classification capabilities [13]. As a result, techniques as well as devices designed for action recognition using HMM and ANN algorithms are being developed [14]. If the HMM and ANN algorithms are combined to optimize the system model, then the system's ability to collapse, i.e., its stability, is greatly improved and the performance aspect is superior to that of the HMM algorithm alone [15].

The previously mentioned 3D animation technology in the construction of human models needs to take into account the degree of smoothness of limb movements, that is, the coherence of the movement as well as the variability, the system in the calculation not only to consider the trainer's movement data for smooth improvement, but also the robustness of the system, smoothness cannot be discounted, so the requirements of the modeling algorithm is very high. As the difficult problems of the modeling algorithm are not well solved, the development and application of the human model making are limited to some extent. These difficult problems mainly include mannequin modeling techniques, motion data capture, and bone exclusion skinning [16]. Among them, the algorithm of the virtual human modeling technology is not mature enough to achieve the algorithm changes with the action, and the strain is poor, which leads to the trainer to take into account the smoothness of the algorithm but give up the system resources when making the action [17], and the process is slow [18] or can reflect the data changes in real time while ignoring the filtering effect [19], which has a negative impact on the smoothness of the system and the trainer's. Therefore, designing an algorithm that optimizes the human motion detection technique has become a high priority [20].

Inspired by the above idea, this system combines the first estimated static initial center of mass outperform the random center of mass initially estimated using the K -means method [21].

3. Improved Static Aerobics Movement Recognition Algorithm

3.1. Aerobics Movement Recognition Model. Moving target detection and tracking is one of the core topics of computer vision. It integrates the research results of image processing, pattern recognition, artificial intelligence, automatic control,

and other related fields. For different monitoring scenes, the moving target detection and tracking algorithms are also different. This paper mainly studies the detection of moving objects in static scenes and constructs an aerobics action recognition model. The motion node monitoring and identification diagram is described in Figure 1. This system mainly analyzes human joints through different aerobic exercises.

In this paper, the static mean value is collected by using the analysis results of moving nodes of different bones and joints. Using the recognition and analysis of three-dimensional key points, the distance feature is controlled in the psychological degree. Feature extraction is carried out through the initial state of different positions. Estimation of K centers of mass the performance of human aerobics gesture selection and are always random centers of mass for K centers of mass. The category labeling of each aerobics stance is determined by using ANN. Finally, aerobics moves are identified aerobics moves gestures using HMM. The first is to train each movement. Let us assume that the first movement is trained first. Secondly, we cluster 64 groups of data of each same action together. For discrete measurements, that is, the measurements can be exhaustive. At this time, the emission probability is matrix, but for continuous measurements, it is generally GMM model. At this time, the emission probability is generally the value of Gaussian model parameters. These data can be used for action recognition based on gmm-hmm.

The human aerobics posture in each frame is represented by the position of 20 skeletal joints:

$$H_t = \{p_t^1, p_t^2, \dots, p_t^{20}\}. \quad (1)$$

The transformed joint coordinates are as follows.

$$p_t^{ki} = p_t^i - p_t^{\text{hipcenter}}, \quad 1 \leq i \leq N. \quad (2)$$

The feature vector f for each skeleton frame of the aerobics gesture sequence is defined as follows:

$$\begin{aligned} f &= \{p_t^{k1}, p_t^{k2}, \dots, p_t^{kN}\}, \\ F &= \{f_1, f_2, \dots, f_m\}. \end{aligned} \quad (3)$$

The aerobics gesture selection module by using subframe representations of gestures instead of using all similar aerobics gestures. The aerobics gesture similarity is reduced using a well-known K -means-based clustering algorithm with a squared Euclidean distance metric for aerobics gesture selection techniques.

The conventional (nonstatic) K -means algorithm obtains randomized center of masses in the initial step standalone each time, and these centers of masses are sometimes different. Using these cluster identifiers, all actions can be correctly classified.

After reducing the repetitions of the aerobics gesture sequences using each aerobics gesture separately, the common aerobics action figure is shown in Figure 2.

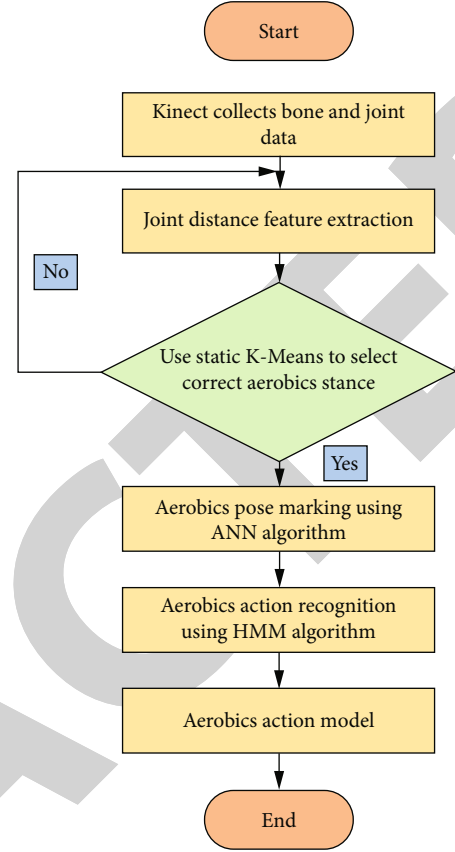


FIGURE 1: Flowchart of aerobics movement recognition system.

In this paper, different neural discrimination of joint points is used for locking. Through the three-dimensional analysis of bones, gesture analysis feature extraction of position matching pattern is carried out. Because the hidden layer of each key node needs and retrograde intelligent matching analysis, it still needs to be further deepened.

The Markov model can correctly identify many instances of aerobics gesture sequences, which are instances of aerobics gestures by using labeled artificial neural networks. Dynamic gesture recognition based on the hidden Markov model is generally based on the temporal characteristics of gestures. A single gesture can be considered as a sequence of different hand shapes, and multiple gestures can be distinguished by hand shapes and their motion trajectories.

$$\lambda = (\pi, A, B),$$

$$\pi_i = P[S_i = q_t], \quad 1 \leq i, t \leq N,$$

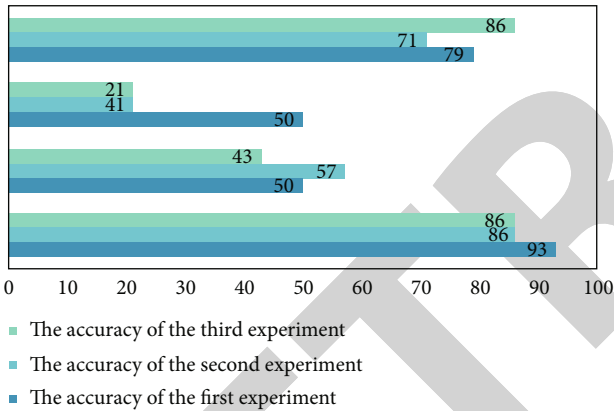
$$a_{ij} = P[S_j = q_{t+1} | S_i = q_t], \quad 1 \leq i, j, t \leq N, \quad (4)$$

$$b_j(k) = P[u_k \text{ at } t | S_j = q_t], \quad 1 \leq j, t \leq N, 1 \leq k \leq R.$$

3.2. Design of Aerobics Track Recognition Method. For the traditional aerobics action recognition, it is necessary to decompose the aerobics action into multiple static actions in advance, because the data computation complexity of multi-image action sequences is high and difficult to implement, so the features of skeletal data are extracted by Kinect



FIGURE 2: Common aerobics movement diagram.

FIGURE 3: Identification process accuracy (%) for the nonstationary K -means training set.

capture, and about 30 frames per second are collected to represent the coherence of aerobics action with continuous human skeletal frame data [22]. The continuous aerobics gesture sequence over a period of time is preset here to represent the change of aerobics. For an aerobics gesture sequence M , G_i denotes the distance feature corresponding to the i -th frame of the aerobics gesture sequence, i.e., the feature quantity, which is N in total, so G_i is the set of distance features extracted from the skeletal data of a human body while performing aerobics gestures.

$$M = (G_1, G_2, \dots, G_i, \dots, G_n). \quad (5)$$

To compare whether two sets of body movements belong to the same aerobics action, which needs to be judged by waveform similarity, and the dynamic time regularization method can stretch the aerobics action sequences of different lengths of the same aerobics action in the time axis accord-

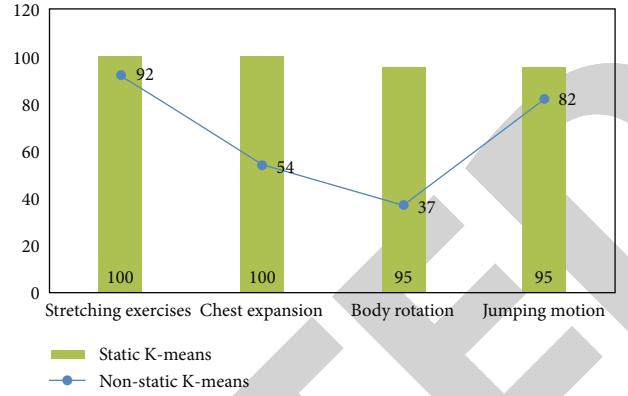


FIGURE 4: Comparison of the UTKinect dataset training set (%).

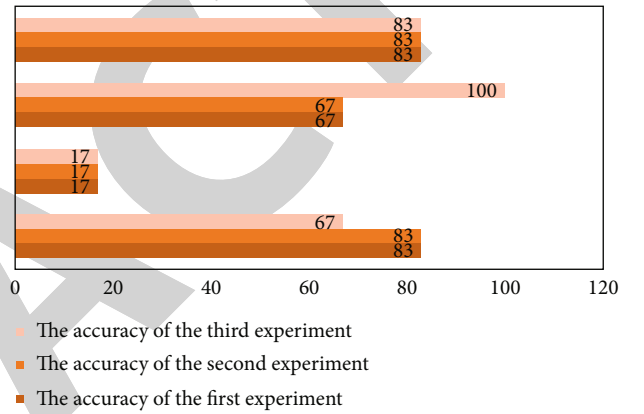
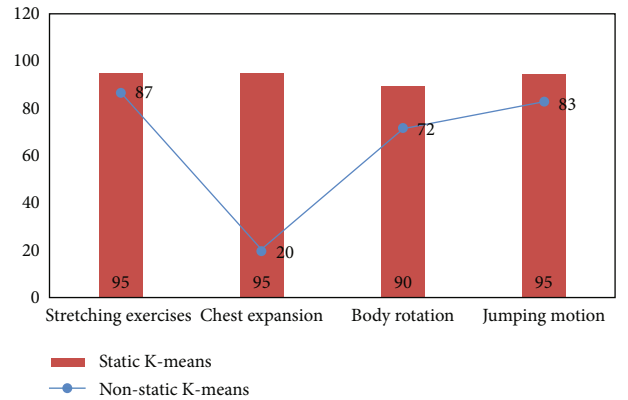
FIGURE 5: Accuracy (%) of the test set identification process using nonstationary K -means.

FIGURE 6: Mean performance of different datasets on the test set.

ingly, so the two aerobics action sequences are of the same length. This means that the two sequences are similar. However, for the processing of time series, the two aerobic movements comparing the length of time series may not be equal, even though the similarity of two aerobic movements is high and the lengths of the sequences are equal, the values of the aerobic movement features at the same time points may be deviated. To solve the above problem, the dynamic time planning idea of DTW algorithm is introduced here to reduce the gap between the action sequences by finding the

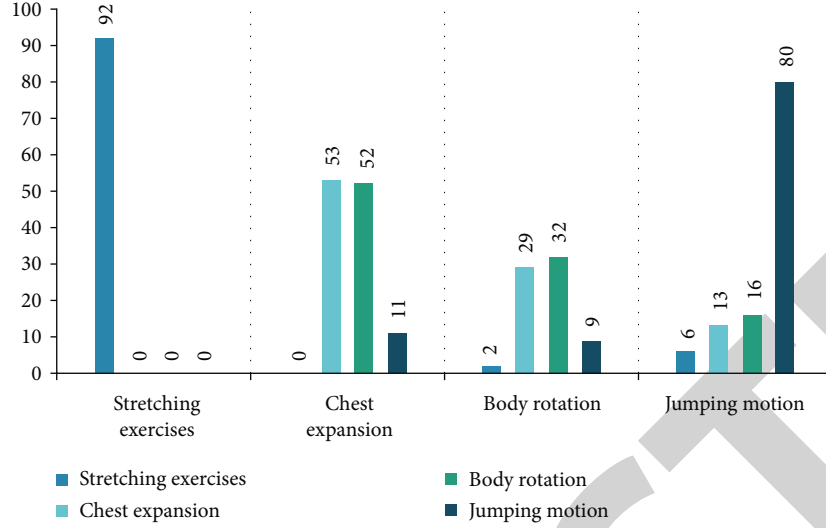


FIGURE 7: Confusion matrix training analysis results.

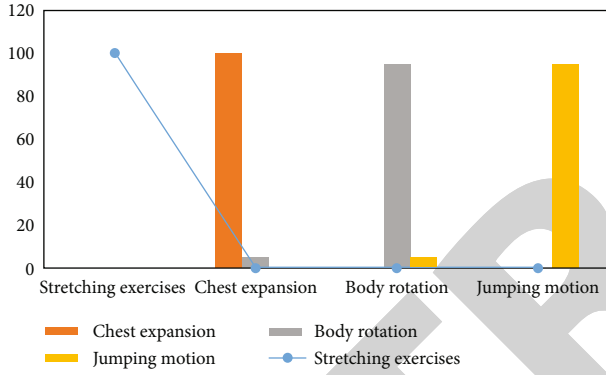


FIGURE 8: Matrix mean analysis of UTKinect dataset.

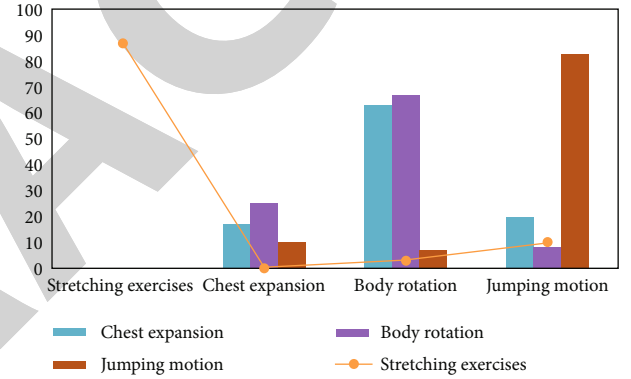


FIGURE 9: Nonstationary K-means confusion matrix.

point-to-point mapping relationship between the two aerobics action sequences, i.e., the matching path with the smallest distance.

The method of extracting motion feature vectors itself may have the problem of time series time point alignment, and DTW can solve this problem very well. The principle of dynamic time regularization is mainly to solve for the minimum distance between two sequences. Suppose T and S are the reference and test aerobics sequences, respectively, and there are two motion sequences of lengths n and m , respectively.

$$\begin{aligned} T &= (T_1, T_2, \dots, T_n), \\ S &= (S_1, S_2, \dots, S_m). \end{aligned} \quad (6)$$

Each sequence contains n and m with different aerobics postures, and their values are the feature vectors of one frame at a time.

However, because linear scaling ignores the possibility that the sequence may be extended or shortened due to the overload of different phases, the recognition efficiency is affected. In order to overcome this effect, this paper proposes

a new dynamic scaling technology. If there are sequences with unequal and, align the two aerobic sequences by linear scaling, shortening the longer sequence or lengthening the shorter sequence. When each sequence contains an equal sum, the minimum distance between the two sequences is obtained by numerically summing the eigenvectors of the corresponding poses of the two aerobic sequences.

In this paper, we construct a matrix grid matrix i and j is expressed as similarity by $d(T_i, S_j)$, and the distance and similarity are inversely proportional.

$$d(T_i, S_j) = \sqrt{\sum_{\omega=1}^N (T_{i\omega} - S_{j\omega})^2}, \quad 1 \leq \omega \leq N, N = 24. \quad (7)$$

Equation (7) represents the Euclidean distance formula for the corresponding points of two different aerobics posture sequences at a 24-dimensional posture feature vector at a point in time, where N denotes the dimension of the distance feature of the aerobic gesture and $T_{i\omega}$ and $S_{j\omega}$ denote the distance feature values corresponding to frame i and frame j of the sequences of different aerobic gestures T and aerobic gestures S . The matrix grid coordinates of (i, j)

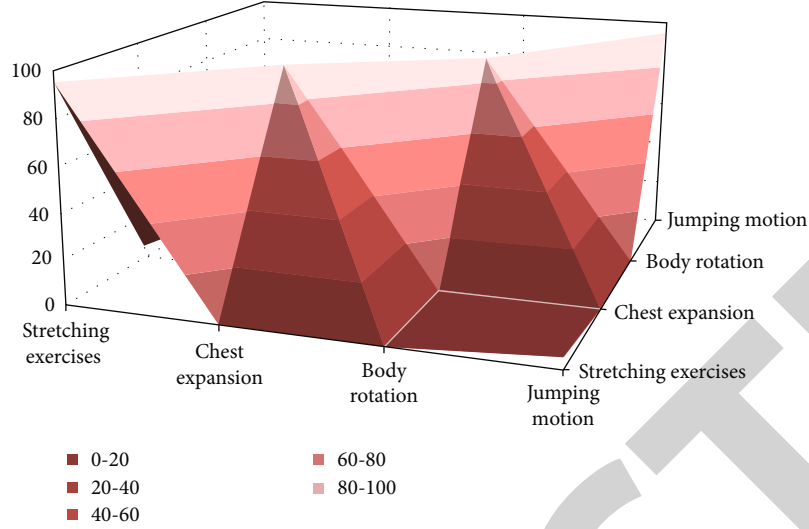


FIGURE 10: Confusion matrix for the UTKinect dataset test set.

represent the correspondence between the points of the aerobic gesture sequence T_i and S_j .

The line through which each point of two different aerobics sequences are aligned is called the planning path, and this line is the optimal $(1, 1)$ to the point (n, m) . This algorithm is called the regularized path algorithm. Equation (8) defines the mapping relationship between different aerobics sequences T and S , where W represents the planning path and k represents the k point in the planning path

$$W = \{w_1, w_2, \dots, w_k, \dots, w_K\} \quad \max(m, n) \leq K < m + n - 1. \quad (8)$$

There are three selection conditions for the planning path, namely, boundary constraint, continuity constraint, and monotonicity constraint:

Boundary constraint: the regularization of two different aerobics posture sequences is always at the two endpoints. To facilitate the study of the skeletal data of aerobic movements, the starting point of the path is $w_1(1, 1)$, the ending point is $w_K(m, n)$, the length of the aerobic movement sequence is set to 30 frames every second, the duration of the aerobic movement is set to 2 seconds, and the output rate of the human skeletal data frames is 30 fps.

Continuity constraint: in order to ensure that the planning path covers each point in the aerobic gesture sequence T and S , the adjacent frames are aligned; assuming that there is a point $w_{k-1}(a', b')$ in the path, the next point a in $w_k(a, b)$ needs to satisfy $a - a' \leq 1$ and b needs to satisfy $b - b' \leq 1$.

Monotonicity constraint: suppose there is a point $w_{k-1}(a', b')$ in the path, the next point $w_k(a, b)$ in a needs to satisfy $a - a' \leq 0$ and b needs to satisfy $b - b' \leq 0$. So the frames in the regularized path are monotonic at the point in time.

After three selection constraints, the point (i, j) can be passed in only three directions, $(i + 1, j)$, $(i, j + 1)$, and $(i + 1, j + 1)$.

Set $Y(i, j)$ as the sum of the Euclidean distances of the points T_i and S_j . The distances of the nearest elements that can reach the point, which is called the cumulative distance. Under the constraint of the selection condition, we find the path that satisfies the condition from the starting point $(1, 1)$ to the end point (n, m) , which is the optimal path to find the point with the minimum cumulative distance corresponding to two different aerobics sequences. The cumulative distance formula is

$$Y(i, j) = d(T_i, S_j) + \min \{Y(i-1, j-1), Y(i-1, j), Y(i, j-1)\}. \quad (9)$$

In order to calculate the similarity of two aerobic gesture sequences, a dynamic time regularization algorithm is used to match them, and the similarity is obtained by inputting the aerobic gesture sequence to be tested and comparing it with the gesture sequence in the standard template. Set the set of action sequences $M = \{S_1, S_2, \dots, S_i, \dots, S_M\}$, and solve the class label(T) of the test aerobic action sequences T by the formula

$$\text{label}(T) = \text{label}(S_c), \quad c = \text{argmin}_i \gamma(T, S_i), \quad i = 1, 2, \dots, M, \quad (10)$$

where i and c are the serial numbers of the action sequence of i in the template database and the sample with the smallest distance in the template database, $\gamma(T, S_i)$ indicates the similarity between T and the action sequence of i , and $\text{label}(S_c)$ is the class of the action sequence corresponding to c .

In τ the process of testing the sample by the above method, the sample to be tested may not be entered into the template database beforehand. To avoid this error, we set a threshold value τ , which represents the similarity of two aerobic sequences, and mark the aerobic sequences outside the template database as nonidentified objects:

$$\tau = \max dtw(S_i, S_j), \quad i = 1, 2, \dots, M, j = 1, 2, \dots, M, i \neq j. \quad (11)$$

4. Analysis of Simulation Results

The experiment was evaluated on 400 different sequences of aerobic gestures (4 movements, 10 objects, different sequences of aerobic gestures) (4 movements, 3 objects, 2 instances, and 5 classes).

First, the experiments are using nonstationary. Based on this method, the training set will be tested and the process will be repeated three times. The average accuracy is shown in Figure 3.

After analysis, the accuracy of all proposed walking actions is high. Each of these measures has high limitations. After the first formal analysis and static simulation, the accuracy of the action studied is high. After the definition analysis, the formal centroid analysis shows a high value. Figure 4 compares and analyzes all training sets. The analysis of different results shows that the method in this paper is higher than other methods in the past.

In this paper, the mean experimental set analysis under different states is carried out. Through the simulation of different action results, it shows that the action accuracy of doing and standing is very high. The significance is strong. Because the analysis repetition of this action is large, the value of stable analysis is high. The accuracy is shown in Figure 5.

The results of the test set show that the nonlinear relationship of the experiment shows good results. The action recognition nodes of each joint are very accurate. Through the static set simulation experiment analysis of the adopted method, the static performance in the mean state is displayed. Figures 6–9 show this process well.

The recognition rates for the case where the training set has a nonstatic K -means confusion matrix are shown in Figure 7.

The recognition rates for the case where the training set has a static K -means confusion matrix are shown in Figure 8.

The recognition rates for the case with a nonstationary K -means confusion matrix on the test set are shown in Figure 9.

The recognition rates on the test set with a static K -means confusion matrix are shown in Figure 10.

From the results of the above simulation comparison experiments, it can be clearly seen that, compared with the nonstatic K -means scheme, the static K -means with static initial centroids has a better effect in correctly identifying the sequence of aerobics motion trajectories.

5. Summary and Outlook

This paper analyzes the problems existing in aerobics training. An aerobics motion trajectory recognition system based on device motion capture technology is designed, and the corresponding solutions are proposed. This scheme improves the performance of traditional human motion rec-

ognition algorithm. Compared with the traditional human motion recognition algorithm, the accuracy of pose selection is improved by using the bone characteristics of Kinect sensor to distinguish motion. Through the simulation and analysis of bone movements in different positions, this paper expounds the posture level of the action model in detail. It not only improves the simulation accuracy of the system but also evaluates it on the public dataset. Compared with the Markov model of neural network, this paper has high reference value. However, the research has certain limitations. In the process of human bone modeling, although the standard data of aerobics items and coach data are compared in real time, each limb movement exceeds a certain range threshold, and the connection of bone joint models will lead to overlapping and unevenness, which is undoubtedly the loss of modeling effect. The next step should focus on the envelope in the process of 3D human modeling. Supporting the envelope will make the action of the model more crash resistant.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] G. Jiamin, Z. Qin, C. Hui, and L. Yibin, "Obtaining lower-body Euler angle time series in an accurate way using depth camera relying on Optimized Kinect CNN," *Measurement*, vol. 188, article 110461, 2022.
- [2] Z. Zhuoyu, H. Ronghua, A. Lin et al., "Automated and accurate assessment for postural abnormalities in patients with Parkinson's disease based on Kinect and machine learning," *Journal of Neuroengineering and Rehabilitation*, vol. 18, no. 1, 2021.
- [3] B. Peter, D. B. Anderson, P. Matthew, and W. R. Walsh, "The reliability of the Microsoft Kinect and ambulatory sensor-based motion tracking devices to measure shoulder range-of-motion: a systematic review and meta-analysis," *Sensors*, vol. 21, no. 24, p. 8186, 2021.
- [4] B. Alireza, K. Shunsuke, T. Teppei, and K. Atsushi, "Improved 3D human motion capture using Kinect skeleton and depth sensor," *Journal of Robotics and Mechatronics*, vol. 33, no. 6, pp. 1408–1422, 2021.
- [5] K. JungYup, P. MinJong, K. Sungjun, and S. Dongjun, "Vision-guided six-legged walking of Little Crabster using a Kinect sensor," *Applied Sciences*, vol. 12, no. 4, p. 2140, 2022.
- [6] B. Wang, H. W. Dong, M. M. Zhang, and Z. G. Pan, "Kinect-based dynamic gesture recognition," *Sensors and Microsystems*, vol. 37, no. 2, pp. 143–146, 2018.
- [7] B. Liu, S. Q. Zhou, and X. Zou, "Smart home system based on Kinect motion control," *Science and Technology Innovations*, vol. 3, pp. 84–85, 2019.

Retraction

Retracted: Adjustment Method of “FAST” Active Reflector Based on Optimal Fitting Strategy

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xia, “Adjustment Method of “FAST” Active Reflector Based on Optimal Fitting Strategy,” *Journal of Function Spaces*, vol. 2022, Article ID 6994858, 15 pages, 2022.

Research Article

Adjustment Method of “FAST” Active Reflector Based on Optimal Fitting Strategy

Yu-Chen Xia 

School of Mathematics, Hefei University of Technology, Hefei, Anhui 230000, China

Correspondence should be addressed to Yu-Chen Xia; 2019213092@mail.hfut.edu.cn

Received 28 June 2022; Accepted 22 July 2022; Published 8 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Yu-Chen 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.

The adjustment method of the fast active reflector of “Chinese heavenly eye” is studied. Firstly, this paper studies the offset characteristics of fast main cable node position, combined with the position of the celestial body to be measured, the follow-up coordinate system is established, and the optimal fitting model of active reflector adjustment based on discrete main cable node is established by using the idea of optimal fitting of the ideal surface. Secondly, the genetic algorithm is used to solve the minimum root mean square error (RMSE). The offset of the triangular reflector is driven by the offset of the main cable node, so that the working area is closer to the ideal paraboloid, and the maximum electromagnetic wave signal receiving ratio is achieved. Finally, rotate the coordinate data of each main cable node according to the known celestial body angle and use the Lagrange operator method to obtain the shortest distance from the actual offset point to the ideal paraboloid as the objective function, with the radial expansion range of the actuator and the variation range of the distance between adjacent points as constraints, and the size of the electromagnetic wave contact area is the basis for judging the amount of information, and a reflector adjustment model for the optimal displacement strategy of discrete main cable nodes is established. Through the test, concluded that the optimal fitting strategy is ideal, the model can accurately obtain the fast active reflector adjustment method under the constraint conditions. Through the test, the conclusion that the optimal fitting strategy is ideal is obtained, and the fast active reflector adjustment method that the model can accurately obtain under the constraints is verified.

1. Introduction

In the early 1990s, since the existing radio telescopes in the world can no longer meet the needs of existing scientific and technological development, the international astronomical community proposed the construction of large-scale radio telescopes to promote further human exploration of the universe (Wu, 1995) [1]. The five-hundred-meter Aperture Spherical radio Telescope (FAST) is a major scientific and technological innovation achievement with independent intellectual property rights in China. As one of the countries with the earliest start and fastest development of celestial literature in the world, many scientists have focused on many cutting-edge and hot topics of astrophysics and creatively proposed to independently develop a new type of large aperture antenna, namely, China’s celestial eye (hereinafter referred to as “fast”), which is a major scientific and technological innovation achievement with independent intellec-

tual property rights in China [2]. The “FAST” telescope is a 500 m-aperture spherical reflector radio telescope with extremely high sensitivity in the world, covering a wide range of astronomical contents, and is engaged in the initial turbidity of the universe, dark matter, dark energy and large-scale structure, the evolution of galaxies and the Milky Way, stellar objects, and even observational research on solar system planets and adjacent space events (Nan and Jiang, 2017) [3]. The “FAST” telescope mainly observes celestial objects through the adjustment of its active reflector, and the active reflector adjustment technology is an important technical means to improve the observation accuracy of radio telescopes (Wang et al., 2022) [4]. Therefore, the active surface adjustment method used in this study drives the offset of the triangular reflection panel and the offset of the main cable node, so that the working area is closer to the ideal paraboloid, so as to achieve the maximum electromagnetic wave signal reception ratio and ensure the safe and efficient

operation of the fast telescope, which is of great significance to improve the level of scientific research in China, and also provides a good reference for the research in the same field of the international community.

At present, there are many studies on the “FAST” active reflector [5–8]. As early as 2006, some scholars theoretically and empirically studied the motion of the paraboloid of each unit spherical block during the telescope tracking process [9]. After that, some scholars have analyzed the surface shape accuracy of FAST based on the structure and working principle of FAST active reflector and on the basis of reflector unit dynamics. Also, the fitting accuracy of FAST instantaneous paraboloid was studied [10–12]. Wu and Zhang (2008) fit a paraboloid based on the least squares method and calculated the fitting accuracy of the paraboloid in the FAST reflection surface and found that the surface accuracy is the highest when the paraboloid vertex is at the center of the circle [13]. There are also some scholars who have verified, analyzed, and optimized the FAST project’s active reflector measurement control scheme on the scaled-down radio astronomical telescope model [14–17]. Luo et al. (2011) proposed an integrated astronomical trajectory planning model and carried out the corresponding derivation and analysis to obtain the relationship between the observation trajectory of FAST and the corresponding coordinate of the center point of the cable-net paraboloid with time [18]. Li and Zhu (2012) analyzed the pros and cons of the parabolic deformation strategy based on the premise of the minimum difference between the parabolic arc length and the spherical arc length during the deformation process of the reflecting surface and through the smooth transition between the parabolic edge and the spherical surface [19]. Wang et al. (2022) used the azimuth and elevation angles to determine the positional relationship between the observed object, the origin, the curved surface where the feed cabin is located, and the reference sphere, and based on the limitation of the adjustment range of the reflective panel, the particle swarm algorithm was used to determine the equation of the paraboloid [20]. Based on the research on the “FAST” active reflector, some scholars have established a model to adjust the paraboloid to determine the receiving ratio of the feed cabin to improve the observation accuracy of radio telescopes [21–24]. Li et al. (2020) combined with the joint simulation of kinematics and control, hardware in the loop simulation test and other methods, realized the accurate positioning of the feed source in the feed cabin in the 100 m large-scale space and overcome the effects of high-altitude wind disturbance, steel cable vibration, and its own motion coupling [25]. Li et al. (2022) combined Fermat’s principle to simulate the electromagnetic wave reflection in three-dimensional space, calculate the receiving ratio of the feed cabin, and test the data results, indicating that FAST can freely adjust the autonomous surface in a certain radio frequency band and obtain a higher acceptance ratio [26].

To sum up, the existing extensive research on the “FAST” active reflector, the fitting accuracy of fast instantaneous paraboloid, adjusting paraboloid to determine the receiving ratio of feed cabin, and other aspects has laid a

solid theoretical and practical foundation for this paper. However, the existing research also needs to be further deepened, such as the distance between the actual offset point of the main cable node and the ideal paraboloid can still be further reduced, and the electromagnetic wave receiving ratio of the paraboloid reflector still needs to be further improved [27]. By studying the offset characteristics of FAST’s main cable node position, establishing the follow-up coordinate system combined with the position of the celestial body to be measured, and adopting the idea of optimal fitting of the main cable node to the ideal surface, this paper establishes the optimal fitting model of active reflector adjustment based on discrete main cable nodes [28, 29]. On this basis, the genetic algorithm is used to solve the minimum root mean square error, and the offset of the main cable node is used to drive the offset of the triangular reflector. Thus, the working area is better close to the ideal paraboloid to achieve the maximum electromagnetic wave signal reception ratio.

In the following innovation of this paper, the following coordinate system is introduced, which effectively simplifies the geometric operation process, reduces the calculation and model complexity, and adopts the idea of overall deviation distribution, which can make the main cable nodes approach the ideal paraboloid more accurately. At the same time, the optimization model is established by using the known constraints and solved by the genetic algorithm. In the process of solving, the global control variables are added to screen out the bad values.

Firstly, the following coordinate system with the direction of the observed celestial body as the reference is introduced to find the optimal approximation model for the distribution of the mean square deviation of the radial direction. In order to make the expansion distance of the actuator within the adjustable range, take h as the variable to investigate the overall deviation distribution vector and find the h value with the minimum overall offset, that is, the minimum overall RMSE, so as to determine the trajectory equation of the ideal paraboloid. Using the traversal algorithm to search the h value under the restriction of radial expansion, it is obtained that the h value with the smallest RMSE is 300.817 5, and the trajectory equation of the ideal paraboloid is

$$x^2 + y^2 = 562.47(z + 300.8175). \quad (1)$$

The global optimal solution of the offset of each main cable node in the working area is obtained by a genetic algorithm, and the final optimal solution is $h = 300.8169$, $f = 140.6192$. The paraboloid trajectory equation (1) of the actual offset in the follow-up coordinate system and the coordinates of the main cable node after the offset are obtained. The coordinates of the main cable node after the offset in the reference coordinate system are obtained through the inverse transformation of the rotation matrix. The mean value of the distance change between adjacent nodes is 0.0633%, and the average length of the radial shrinkage of the main cable node is 0.20343.

2. Basic Assumptions and Index Selection

2.1. Basic Assumptions. In order to facilitate the study of the problem, the following hypotheses are put forward: (i) electromagnetic wave signal and reflected signal are regarded as linear propagation. (ii) The reflecting panel is smooth, and the electromagnetic beam is completely reflected after passing through the reflecting surface. (iii) The electromagnetic beam from the measured celestial body is parallel incidence. (iv) All components of the active reflecting surface are rigid bodies. (v) All data sources involved in this study are true, accurate and scientific.

2.2. Symbol Description. In combination with modeling theory and method, there are 6 variables involved, as shown in Table 1.

3. Trajectory Equation of Ideal Paraboloid under Nonworking State of Active Reflector of Radio Telescope Based on Ergodic Algorithm

3.1. Research Approach. According to the measurement requirements of the radio telescope, the focus of the ideal paraboloid is in the feed cabin; that is, the coordinates are $(0, 0, -0.534R)$. When the object is directly above the reference sphere, the main cable node of the ideal paraboloid vertex is the easiest to study the offset direction of all nodes; so, the paraboloid trajectory equation can be transformed into a function related to the radial expansion distance h of the vertex. For the convenience of calculation, the following coordinate system with the direction of the observed celestial body as the reference is introduced, and the unit vector outward along the spherical normal direction at the main cable node is introduced to study the radial displacement deviation d . The component of the deviation in the coordinate system forms a functional relationship with the original coordinates of the main cable node, and the objective function only related to the radial expansion distance h can be obtained by substituting the assumed ideal paraboloid equation. The method of RMSE of overall deviation distribution is used to find the most ideal radial displacement of all displaced main cable nodes in the working area. Combined with the constraints of radial shrinkage range, the optimization model of the ideal paraboloid trajectory equation is established. Using the traversal algorithm to calculate the deviation of each node, the h of the optimal solution can be found, and the trajectory equation can be solved [30].

3.2. Displacement Model of Active Reflector. When the active reflector of the radio telescope is in the nonworking state, that is, in the reference state, each main cable node falls on the reference sphere with a radius of 300 meters and a diameter of 500 meters. When observing the celestial body, it is in the working state, and the area of the illumination part of the feed is adjusted to an approximate ideal paraboloid with an aperture of 300 meters. To determine the ideal paraboloid, the benchmark space rectangular coordinate system $OXYZ$ is established with the spherical center of the reference

sphere as the origin and the vertical direction as the z -axis, and the spherical trajectory equation of the feed cabin can be obtained as follows:

$$x^2 + y^2 + z^2 = (R - 0.466R)^2 = (0.534R)^2. \quad (2)$$

Benchmark spherical trajectory equation is as follows:

$$x^2 + y^2 + z^2 = R^2. \quad (3)$$

At the same time, taking the spherical center of the reference sphere as the origin and the direction pointing to the observed celestial body as the z -axis, we establish the spatial follow-up coordinate system $OX'Y'Z'$, as shown in Figure 1. The models in the subsequent articles are established and solved based on this follow-up coordinate system $OX'Y'Z'$.

Assumed that the trajectory equation of the ideal paraboloid in the $OX'Y'Z'$ coordinate system is as follows:

$$x^2 + y^2 = 4f(z + h), \quad (4)$$

Where variable f is the focal length of the ideal paraboloid, and h is the offset of the paraboloid vertex along the Z' axis relative to the reference sphere. Obviously, under the working state, the point with the maximum offset of each main cable node in the working area is the vertex.

Since only each main cable node can be used for paraboloid fitting, the optimal approximation model is found by fitting the distribution of the mean square deviation from the reference sphere center to each point, that is, the radial direction deviation. To make the expansion distance of the actuator within the adjustable range, we take h as the variable to find the h value with the smallest overall offset, that is, the smallest overall RMSE, so as to determine the ideal paraboloid.

3.3. Establishment and Solution of Overall Offset Model. Take a point $A(x_0, y_0, z_0)$ in the above follow-up coordinate system and set it as any main cable node in the lighting area on the ball.

After being pulled by the actuator, point A is offset to point A' on the paraboloid. The radial expansion of the actuator is known; so, we can think that the displacement of main cable node A only moves along the normal direction of the sphere, while the displacement along the tangential direction of the sphere is ignored. According to the actual situation, the paraboloid is close to the spherical surface; so, here we think that the vector AA' is approximately along the normal direction of the spherical surface, then the line segment AA' is parallel to the Z' axis A , and its value is set as d . When the point is located on the inner side of the ideal paraboloid, the deviation value d is positive; otherwise, it is negative.

Introduce the unit vector $q = (q_x, q_y, q_z)$ outward along the spherical normal direction at the main cable node A , where A are their projections on the OX' , OY' , and OZ'

TABLE 1: Symbol description.

Serial number	Symbol	Description
1	R	Reference spherical radius
2	f	Focal length of ideal paraboloid
3	h	Vertex offset
4	d	Deviation value of main cable node
5	D	Overall deviation distribution vector of working area
6	q	The unit vector outward along the normal direction of the sphere at the main cable node

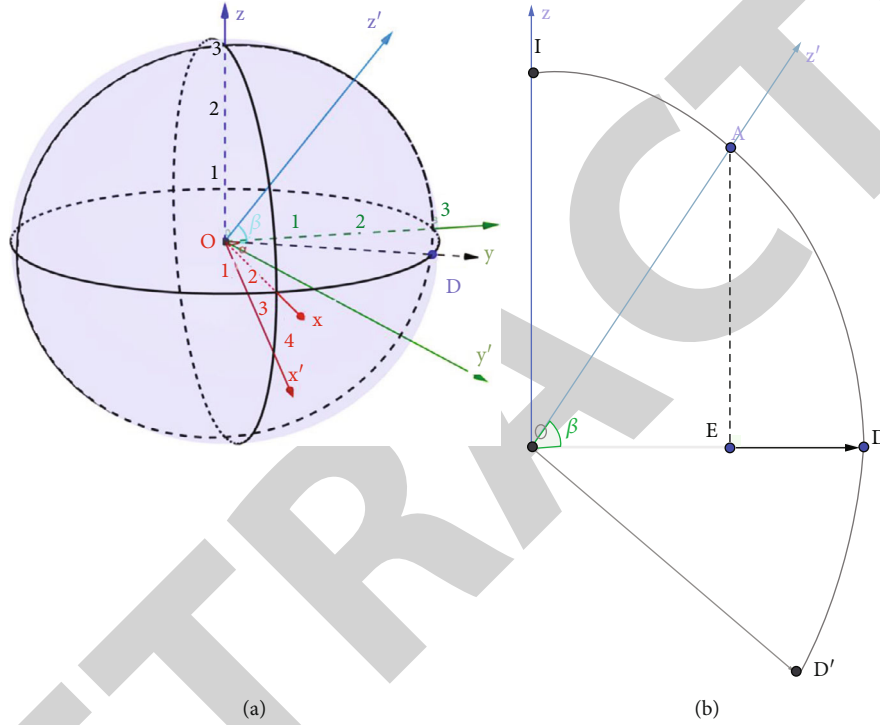


FIGURE 1: Rotation diagram of spatial follow-up coordinate system (a) and surface rotation diagram (b).

axes, respectively, and then point A' about point a has parametric equation as follows:

$$\begin{cases} x = x_0 + q_x d, \\ y = y_0 + q_y d, \\ z = z_0 + q_z d. \end{cases} \quad (5)$$

Then, the coordinates of point A' can be obtained as follows:

$$(x_0 + q_x d, y_0 + q_y d, z_0 + q_z d). \quad (6)$$

Then, we can set the direction for the deviation d of point A . at this time, d is transformed into a deviation

vector:

$$\vec{d} = (d_x, d_y, d_z)^T = (q_x d, q_y d, q_z d)^T. \quad (7)$$

Known conditions indicated that point A' is a point on the ideal paraboloid and satisfies the paraboloid equation. Therefore, substitute it into the coordinate solution equation to obtain the function of deviation d with respect to h :

$$d = -\frac{x^2 + y^2 - 4fz}{2q_x x + 2q_y y - 4q_z f} + h \cdot \frac{4f}{2q_x x + 2q_y y - 4q_z f}. \quad (8)$$

Remember $\gamma = -x^2 + y^2 - 4fz / 2q_x x + 2q_y y - 4q_z f$, $\eta = 4f / 2q_x x + 2q_y y - 4q_z f$. Then, \vec{d} can be simplified as

$$\vec{d} = (\gamma + h\eta) \cdot \vec{q}. \quad (9)$$

Let $D = (\vec{d}_1^T, \vec{d}_2^T, \dots, \vec{d}_n^T)^T$ be the overall deviation distribution vector of each main cable node in the working area, where \vec{d}_i^T is the deviation vector of the i -th main cable node. Transformed into a vector formula about h :

$$D = \left(\vec{q}_1\gamma_1, \vec{q}_2\gamma_2, \dots, \vec{q}_n\gamma_n \right)^T + \left(\vec{q}_1\eta_1, \vec{q}_2\eta_2, \dots, \vec{q}_n\eta_n \right)^T h. \quad (10)$$

Then, we can get the RMSE calculation formula of the overall deviation distribution D with respect to h :

$$\delta = \frac{\sqrt{D^T D}}{n}. \quad (11)$$

where n is the total number of offset main cable nodes. Then, the objective function of the overall offset optimization model is

$$\min \delta = \frac{\sqrt{D^T D}}{n}. \quad (12)$$

If the celestial body to be observed is directly above the reference sphere, the follow-up coordinate system $OX'Y'Z'$ coincides with the reference coordinate system, and the deviation vector is as follows:

$$\vec{D} = (d_x, d_y, d_z)^T = (0, 0, q_z d)^T = (0, 0, d)^T, \quad (13)$$

which is downward along the Z -axis. Then, the trajectory equation of the ideal paraboloid in the reference coordinate system is

$$x^2 + y^2 = 4f(z + h), \quad (14)$$

and the functional expression of the deviation d with respect to h is obtained:

$$d = \frac{x^2 + y^2 - 4fz}{4q_z f} + h \cdot \frac{f}{q_z f}. \quad (15)$$

Then, the optimization model is

$$\min \delta = \frac{\sqrt{D^T D}}{n},$$

$$\text{s.t.} \begin{cases} -0.6 \leq d \leq 0.6, \\ \gamma = \frac{x^2 + y^2 - 4fz}{4q_z f}, \\ \eta = \frac{fz}{q_z f}, \\ D = \left(\vec{q}_1\gamma_1, \vec{q}_2\gamma_2, \dots, \vec{q}_n\gamma_n \right)^T + \left(\vec{q}_1\eta_1, \vec{q}_2\eta_2, \dots, \vec{q}_n\eta_n \right)^T h. \end{cases} \quad (16)$$

3.4. Result Analysis. In the case of an ideal paraboloid, the electromagnetic waves reflected by the paraboloid can converge on the feed cabin, then the coordinates of the feed cabin are the focal coordinates of the paraboloid, and the relationship between f and h can be determined according to the formula as follows:

$$f - h = -(1 - 0.466)R. \quad (17)$$

Since the relationship formula between f and h is determined, the objective function of the optimization model is the univariate variable function about h . Based on the idea of piecewise step size and using the ergodic algorithm, the final optimal solution of h is 300.8175, and the RMSE is 0.19791. Substitute h to obtain the trajectory equation of the ideal paraboloid as follows:

$$\begin{aligned} x^2 + y^2 &= 2 \times 281.235 \times (z + 300.8175) \Rightarrow x^2 + y^2 \\ &= 562.47(z + 300.8175). \end{aligned} \quad (18)$$

4. Problem Two Model Establishment and Solution

4.1. Research Approach. A follow-up coordinate system has been established according to the direction of the measured celestial body. When the angle of the celestial body is known, the ideal paraboloid in the reference coordinate system can be transformed into an ideal paraboloid with the vertex on the z -axis in the follow-up coordinate system. To study the coordinate changes of each main cable node, the Lagrange operator method is used to solve the shortest distance from the point to the parabola, which is used as the objective function and added constraints to make the actual offset point approach the ideal paraboloid. The constraint conditions include the expansion amount of radial actuator and the variation range of node spacing. According to this condition, the optimization model is established, and the global variables are defined for the variation range of node spacing. 5.1 coordinate rotation and solution of ideal paraboloid trajectory equation.

4.2. Theoretical Derivation. If the object to be observed is located at $\alpha = 36.795^\circ, \beta = 78.169^\circ$, then in the rotating follow-up coordinate system, the observed object rotates directly above the spherical center, that is, $\alpha = 0^\circ, \beta = 790^\circ$. We need to discuss the rotation of the coordinate system. The trajectory equation of the ideal paraboloid in the follow-up coordinate system $OX'Y'Z'$ is solved by MATLAB software as follows:

$$\begin{aligned} x^2 + y^2 &= 2 \times 281.235 \times (z + 300.8175) \Rightarrow x^2 + y^2 \\ &= 562.47(z + 300.8175). \end{aligned} \quad (19)$$

The rotation transformation between the reference coordinate system and the follow-up coordinate system is discussed below.

We also choose the follow-up coordinate system to discuss the coordinate offset of each main cable node. Process

the coordinates of the main cable node transform it from the reference coordinate system OXYZ to the follow-up coordinate system OX'Y'Z'.

Let the base vector under the reference coordinate system be $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ and the base vector under the follow-up coordinate system be $(\vec{e}'_1, \vec{e}'_2, \vec{e}'_3)$. Let the original coordinate of \vec{e}'_1 in the reference coordinate system be $\vec{e}'_1 = (x_1, y_1, z_1)$, and there are

$$x_1^2 + y_1^2 + z_1^2 = 0, \quad (20)$$

where \vec{e}'_3 can be expressed as $\vec{e}'_3 = (\cos \beta \cos \alpha, \cos \beta \sin \alpha, \sin \beta)$ by the original basis vector. Since \vec{e}'_1 is perpendicular to \vec{e}'_3 , there is an angular relationship:

$$x_1 \cos \beta \cos \alpha + y_1 \cos \beta \sin \alpha + z_1 \sin \beta = 0. \quad (21)$$

Take the projection line of the OZ' axis on the xOy plane and its unit vector \vec{OD} and use the above formula to obtain the vector $\vec{OD} = (\sin \beta \cos \alpha, \sin \beta \sin \alpha, -\cos \beta)$ after the rotation of point D. According to the geometric angle relationship, it is known that the rotation angle is the same, that is, $\cos \angle \vec{e}'_1, \vec{OD} \geq \cos \angle (\vec{e}'_1, \vec{OD}')$; so, the angle relationship is obtained:

$$x_1 \sin \beta \cos \alpha + y_1 \sin \beta \sin \alpha - z_1 \cos \beta = \cos \alpha. \quad (22)$$

According to equations (15), (20), and (21), replace the known angle $\alpha = 36.795^\circ, \beta = 78.169^\circ$ to obtain the solution

$$x_1 = 0.9864, y_1 = -0.01019, z_1 = 0.1642. \quad (23)$$

The variation of the basis vector is obtained:

$$\begin{cases} \vec{e}'_1 = 0.9864\vec{e}_1 - 0.01019\vec{e}_2 - 0.1642\vec{e}_3, \\ \vec{e}'_2 = -0.01019\vec{e}_1 + 0.9924\vec{e}_2 - 0.1228\vec{e}_3, \\ \vec{e}'_3 = 0.1642\vec{e}_1 + 0.1228\vec{e}_2 + 0.9788\vec{e}_3. \end{cases} \quad (24)$$

Then, the final coordinate transformation formula is as follows:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 0.9864 & -0.01019 & -0.1642 \\ -0.01019 & 0.9924 & -0.1228 \\ 0.1642 & 0.1228 & 0.9788 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad (25)$$

The processed part of the coordinate data is shown in Table 2.

Use the rotation coordinate transformation formula of formula (25) to rotate the coordinates of formula (4), that is, the equation of the ideal paraboloid in the follow-up coordinate system, and get the trajectory equation of the ideal paraboloid when the celestial body is at the position of angle

$\alpha = 36.795^\circ, \beta = 78.169^\circ$ in the reference coordinate system:

$$\begin{aligned} &0.9738302955942003 \cdot x^2 + 0.9849199553123014 \cdot y^2 \\ &+ 0.039679480346832406 \cdot z^2 - 0.0403312962919332 \cdot xy \\ &- 0.32151881303998514 \cdot xz - 0.32151881303998514 \cdot yz \\ &- 92.34698831460001 \cdot x - 63.447065976000005 \cdot y \\ &- 550.521224802 \cdot z = 169200.819225. \end{aligned} \quad (26)$$

4.3. Reflector Pane Adjustment Model Based on Discrete Main Cable Node Displacement. To make the main cable node approach the ideal paraboloid as much as possible, a reflector panel adjustment model based on discrete main cable node displacement is established. The first mock exam is the best one in the problem, and the ideal paraboloid is also for the main cable node. The actual paraboloid with triangular panels should have a larger radius of curvature, as shown in Figure 2. Therefore, discrete points to approximate the ideal paraboloid are adopted [31].

In the follow-up coordinate system, the distance between the actual offset point of the main cable node and the ideal paraboloid is discussed, and the coordinate of the actual offset point closest to the ideal paraboloid under the limiting conditions (i.e., the radial expansion is -0.6 and 0.6, and the change of the distance between adjacent nodes is no more than 0.07%) is found. Therefore, the optimization model is established to find out the adjustment model of the reflection panel based on the displacement of discrete main cable nodes.

In practice, when the actuator pulls the triangular reflector for displacement, there will be a small change in the distance between adjacent nodes, which deviates from the theoretical offset point calculated above, so that the main cable node is not on the ground anchor point and the connecting line of the ball center; that is, the three points cannot be strictly collinear. Therefore, the main cable node and ground anchor point do not move completely along the radial direction in the actual deformation [32, 33].

We take any main cable node $A(x_0, y_0, z_0)$ on the working area of the reference sphere, set the ideal offset point of A point as $A'(x', y', z')$ point, and the A' point is located on the ideal paraboloid $x^2 + y^2 = 4f(z + h)^2$. Due to the existence of deviation, we set the actual offset point of A point as $A''(x'', y'', z'')$ point. Figure 3 shows the geometric relationship between theoretical displacement and actual displacement of main cable node.

The Euclidean distance between point A' and point A'' is

$$d_{A'A''} = \sqrt{(x' - x'')^2 + (y' - y'')^2 + (z' - z'')^2}. \quad (27)$$

To establish the adjustment model of the main cable node approaching the ideal paraboloid, we use the shortest distance from the actual offset point to the ideal paraboloid as the objective function to establish the optimization model.

TABLE 2: Coordinates of each main cable node after rotation.

Node	X	Y	Z	Thea X	Thea Y	Thea Z
A0	49.320 026 47	36.889 381 35	-294.018 482 6	-0.164 181 18	-0.122 800 87	0.978 756 6
B1	55.229 444 59	45.147 999 76	-291.807 329 5	-0.183 874 204	-0.150 286 726	0.971 392 903
C1	59.071 297 35	33.580 077 66	-292.614 262 4	-0.196 623 253	-0.111 766 87	0.974 088 014
D1	49.396 380 8	26.555 491 25	-295.118 526	-0.164 435 398	-0.088 394 798	0.982 419 134
E1	39.575 149 93	33.781 467 71	-295.859 369 1	-0.131 760 932	-0.112 436 88	0.984 884 257
A1	43.180 250 62	45.272 464 75	-293.812 901 1	-0.143 721 369	-0.150 701 494	0.978 076 287
A3	49.071 321 78	53.521 360 38	-291.492 125 4	-0.163 353 742	-0.178 141 503	0.970 351 048
.....
E428	-152.060 690 4	187.429 294 9	-178.852 558 8	0.506 172 8	-0.623 929 324	0.595 400 116
E429	-146.547 628 5	195.355 187 4	-174.930 084 6	0.487 849 82	-0.650 329 634	0.582 300 544
E430	-140.771 871 8	203.086 310 4	-170.820 626 6	0.468 615 102	-0.676 066 534	0.568 624 596

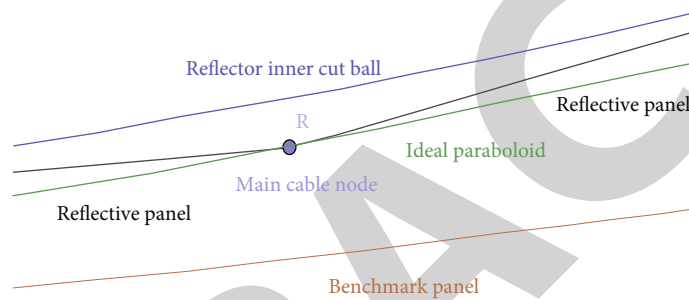


FIGURE 2: Schematic diagram of ideal paraboloid, benchmark sphere, and inscribed sphere of reflector.

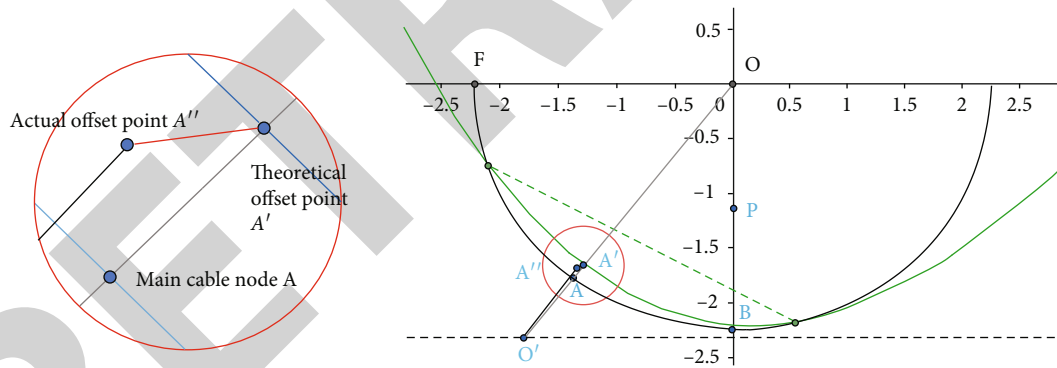


FIGURE 3: Displacement of main cable node.

The Lagrange operator method is used to calculate the minimum distance between the actual offset point and the ideal paraboloid. For the actual offset point $A'(x', y', z')$, take the square of Euclidean distance as the objective function, i.e.,

$$\begin{aligned} \min D &= (x - x'')^2 + (x - y'')^2 + (x - z'')^2, \\ \text{s.t. } x^2 + y^2 - 4f(z + h) &= 0. \end{aligned} \quad (28)$$

Lagrangian operator λ with $\lambda \geq 0$ is introduced to construct Lagrangian as follows:

$$Q = D + \lambda [x^2 + y^2 - 4f(z + h)] = 0. \quad (29)$$

Calculate the partial derivative of the Lagrangian on x, y, z, λ , respectively, and make it zero, including

$$\begin{cases} \frac{\partial Q}{\partial x} = 2(x - x'') + 2\lambda x = 0, \\ \frac{\partial Q}{\partial y} = 2(y - y'') + 2\lambda y = 0, \\ \frac{\partial Q}{\partial z} = 2(z - z'') - 4\lambda f_z = 0, \\ \frac{\partial Q}{\partial \lambda} = x^2 + y^2 - 4f(z + h) = 0. \end{cases} \Rightarrow \begin{cases} x = \frac{x''}{1 + \lambda}, \\ y = \frac{y''}{1 + \lambda}, \\ z = z'' + 2\lambda f, \\ \frac{x''^2 + y''^2}{(1 + \lambda)^2} - 4f(z'' + 2f + h) = 0. \end{cases} \quad (30)$$

The solution of the equations is obtained

$$\lambda = -\frac{z'' + h}{2f}, x = -\frac{2fx''}{z'' + h - 2f}, y = -\frac{2fy''}{z'' + h - 2f}, z = -h, \quad (31)$$

Replace the x, y, z solved into the original objective function D to obtain

$$\begin{aligned} D_{\min} &= (x - x'')^2 + (y - y'')^2 + (z - z'')^2 \\ &= \left[\frac{(z'' + h)x''}{z'' + h - 2f} \right]^2 + \left[\frac{(z'' + h)y''}{z'' + h - 2f} \right]^2 + (z'' + h)^2. \end{aligned} \quad (32)$$

The objective function gives the minimum value of the distance between any point in space (take the actual offset point here) and the ideal paraboloid. In order to make the objective function reach the minimum value, that is, the actual offset point can best approach the ideal paraboloid, it is necessary to add constraints to it. Obviously, in order to make the actual offset point closer to the ideal paraboloid, the actuator, the main cable node, and the transformed point should be on the same straight line and perpendicular to the reference sphere. The direction vector of the straight line is obtained by using the coordinates:

$$\frac{x_{o'} - x'}{x'' - x_0} = \frac{y_{o'} - y'}{y'' - y_0} = \frac{z_{o'} - z'}{z'' - z_0}. \quad (33)$$

The expansion and contraction amount of the actuator tends to the direction of the center of the ball along the radial direction of the reference sphere, and its radial expansion and contraction range are $-0.6 \sim +0.6$ m. Then, we can solve the radial deviation according to this condition:

$$(x' - x'')^2 + (y' - y'')^2 + (z' - z'')^2 \leq 0.6^2 = 0.36. \quad (34)$$

According to the above constraints, we establish the adjustment optimization model of reflective panel based on the displacement of discrete main cable nodes:

$$\begin{aligned} \min D &= \left[\frac{(z'' + h)x''}{z'' + h - 2f} \right]^2 + \left[\frac{(z'' + h)y''}{z'' + h - 2f} \right]^2 + (z'' + h)^2, \\ \text{s.t.} \begin{cases} \frac{x_{o'} - x'}{x'' - x_0} = \frac{y_{o'} - y'}{y'' - y_0} = \frac{z_{o'} - z'}{z'' - z_0}, \\ (x' - x'')^2 + (y' - y'')^2 + (z' - z'')^2 \leq 0.36. \end{cases} \end{aligned} \quad (35)$$

At the same time, the topic shows that the small change range of the spacing between adjacent nodes does not

exceed 0.07%. Therefore, we add constraints in the process of solving the optimization model, judge in the global situation, and discard the coordinate sequence with the change range exceeding 0.07%.

4.4. Result Analysis. By changing h and f , the objective function value is minimized. Since this problem is a multivariable optimization function, the ordinary traversal method requires a large amount of computation and is affected by the accuracy of the traversal step size; so, the local optimal solution is often obtained. Therefore, the genetic algorithm is used to find the global optimal solution in this problem, and the specific flow chart is shown in Figure 4.

The genetic algorithm steps are roughly as follows:

Step 1. Firstly, the variables are binarized to generate a random initial population, and the fitness value of each individual in the population is calculated.

Step 2. Sort the population individuals according to their fitness and then perform gene recombination and gene mutation in turn to obtain a new population of progeny.

Step 3. Insert the offspring group into the parent group and screen out individuals with excellent fitness values.

Step 4. Enter the next evolution and stop and output the optimal solution when the evolution reaches the formulation of algebra.

Using the above genetic algorithm and substituting the data in the attachment to solve the optimal offset coordinates of each main cable node, the final optimal solution is $h = 300.8169$ and $f = 140.6192$. Substitute h and f into the calculation formula (4) of the deviation d . Since the length does not change with the change of the coordinate system, the value of d is the same as the value in the reference coordinate system. Therefore, it can be obtained that the parabolic trajectory equation of the actual migration is $x^2 + y^2 = 562.4768(z + 300.8169)$, the average length of the adjacent nodes is 0.0633%, and the average length of the radial shrinkage of the main cable node is 0.20343. According to the solved main cable node coordinates, perform inverse transformation according to the rotation matrix, solve the transformation coordinates in the reference coordinate system, and adjust the main cable node number, position coordinates, and the expansion and contraction amount of each actuator within the 300-meter diameter of the rear reflection surface. The resulting genetic iterative evolution diagram is shown in Figure 5.

5. Receiving Ratio of Feed Cabin Based on Reflector Adjustment Scheme

5.1. Research Approach. According to the active reflector mediation scheme obtained, in order to calculate the effective amount of information received by the feed cabin, the area contacted by the electromagnetic wave can be used as the basis for judging its amount of information. For each triangular reflector, the projection triangle formed when the electromagnetic

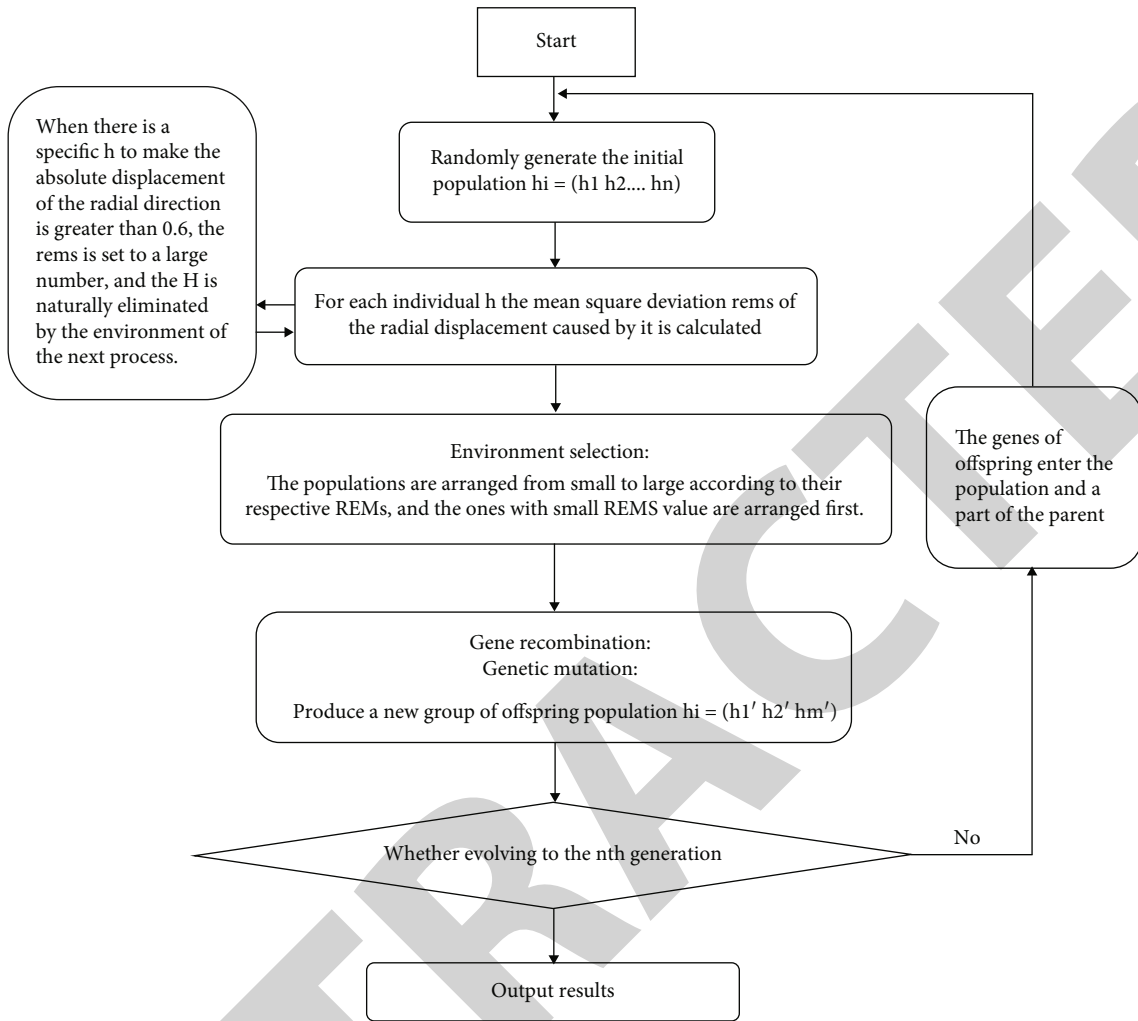


FIGURE 4: Schematic diagram of the rotation of the space follow-up coordinate system.

wave propagates to the plane of the feed cabin through reflection is considered. Since the effective receiving area of the feed cabin is small, about $0.7854m^2$, it is necessary to judge the area of the projected area. By meshing the plane where the feed cabin is located and decomposing it into a set of points, the effective coverage area of the projection to the feed cabin can be judged according to the distribution of points [34].

5.2. Determination of Receiving Ratio of Feed Cabin after Adjusting Paraboloid. The topic has given that electromagnetic wave signals and reflected signals are regarded as linear propagation; so, their propagation can be considered to obey the optical principle. Since we have rotated the coordinate system, the electromagnetic wave from the observed celestial body can be regarded as a parallel electromagnetic beam of vertical incidence, and its direction vector is set as $\vec{e} = (0, 0, 1)$. We use the size of the area touched by the electromagnetic wave as

the basis for judging the amount of information. For each triangular reflecting surface, consider the projection on its plane when the electromagnetic wave received and reflected by its three vertices propagates to the feed cabin located at the ideal parabolic focus. The projection triangle formed by these three projections is the projection area of the electromagnetic wave reflected by the triangular reflective panel, as shown in Figure 6.

Among them, \vec{n} is the unit normal vector of the triangular panel, and \vec{p} is the direction vector of the outgoing light after the incident electromagnetic wave is reflected by the reflective panel. Let $\vec{p}(x, y, z)$ be used to calculate the two side length vectors $\vec{AB}(x_{AB}, y_{AB}, z_{AB})$ and $\vec{AC}(x_{AC}, y_{AC}, z_{AC})$ of the triangular panel using the known vertex coordinates of the triangular panel, and then the coordinate expression of the normal vector can be calculated as follows:

$$\vec{n} = \vec{AB} \times \vec{AC} = (y_{AB}z_{AC} - z_{AB}y_{AC}, x_{AB}z_{AC} - z_{AB}x_{AC}, x_{AB}y_{AC} - y_{AB}x_{AC}). \quad (36)$$

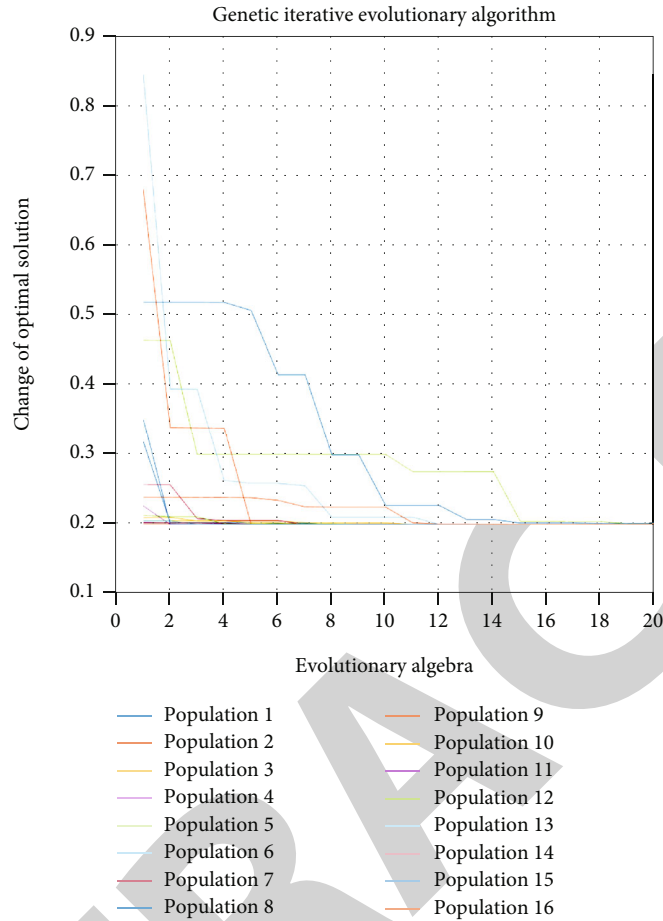


FIGURE 5: Iterative evolution diagram of the genetic algorithm.

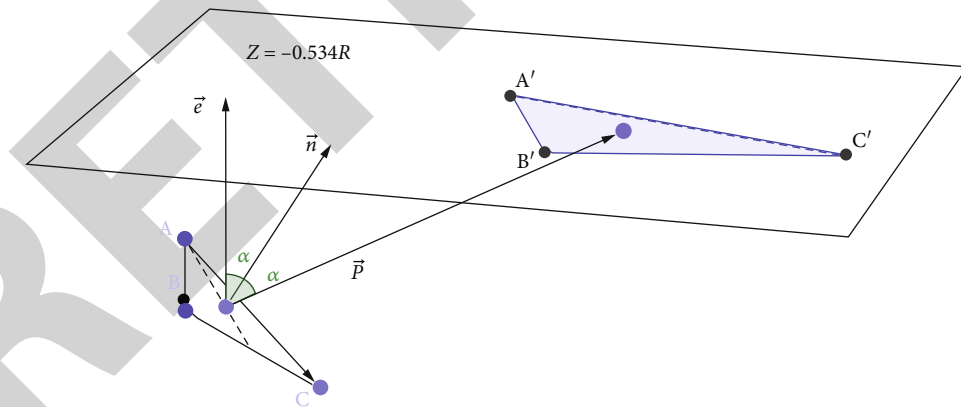


FIGURE 6: Projection diagram of the triangular reflection panel.

According to the optical principle, the incident ray, the normal line, and the reflected ray are in the same plane; that is, the reflected ray direction vector \vec{p} satisfies the condition:

$$\vec{p} \cdot (\vec{n} \times \vec{e}) = 0. \tag{37}$$

At the same time, the normal is the angle bisector of the incident ray and the reflected ray; that is, the angle between

\vec{n} and \vec{p} is equal to the angle between \vec{n} and \vec{e} . The specific formula is as follows:

$$\cos \langle \vec{p}, \vec{n} \rangle = \cos \langle \vec{n}, \vec{e} \rangle, \tag{38}$$

$$\vec{e} \Rightarrow \frac{\vec{p} \cdot \vec{n}}{|\vec{p}| \cdot |\vec{n}|} = \frac{\vec{n} \cdot \vec{e}}{|\vec{n}| \cdot |\vec{e}|}.$$

Since \vec{p} is only used to represent the direction of the reflected light, for the convenience of calculation, we set its component in the x direction to be 1. Combining the above conditions, we get the system of equations:

$$\begin{cases} \vec{p} \cdot (\vec{n} \times \vec{e}), \\ \frac{\vec{p} \cdot \vec{n}}{|\vec{p}| \cdot |\vec{n}|} = \frac{\vec{n} \cdot \vec{e}}{|\vec{n}| \cdot |\vec{e}|}. \end{cases} \quad (39)$$

If the normal vector of the triangular unit panel is \vec{n} (a, b, c), then the coordinate expression of \vec{p} is obtained by solving the system of equations as $(1, (b/a), (a^4 + 2a^2b^2 + b^4 - a^2c^2 - b^2c^2/2a^3c + 2ab^2c))$.

Combining the reflected light direction vector \vec{p} and the vertex coordinates of the reflective panel, we can obtain the trajectory equation of the reflected electromagnetic wave projected to the reflective panel as

$$\frac{x - x_A}{p_x} = \frac{y - y_A}{p_y} = \frac{z - z_A}{p_z}. \quad (40)$$

Among them, $p_x = 1, p_y = b/a, p_z = a^4 + 2a^2b^2 + b^4 - a^2c^2 - b^2c^2/2a^3c + 2ab^2c$. The intersection of the trajectory and the electromagnetic wave receiving plane where the feed cabin is located is the projection point, and the triangle formed by the projection points of the three vertices of the reflective panel is the projected triangle on the plane of the feed cabin. In the follow-up coordinate system, the x -coordinate of the plane where the feed cabin is located is always $-0.534R$, and using this conditional formula (39), the coordinates of the projection point are solved as follows:

$$\left(-\frac{0.534R + z_A}{p_z} + x_A, -\frac{b(0.534R + z_A)}{a p_z} + y_A, -0.534R \right). \quad (41)$$

Similarly, the coordinates of B and C can also be obtained similarly.

Based on the adjustment scheme of the main cable nodes obtained in the second 4, we can find out the projection of the main cable nodes on the plane of the feed cabin according to the position coordinates of each main cable node in the working area, so that the amount of electromagnetic wave information received by the feed cabin can be calculated, as shown in Figure 7.

5.3. Result Analysis

5.3.1. Determination of the Effective Receiving Area of the Circular Surface of the Feed Cabin. It is known that the effective receiving surface of the feed cabin is a central disk with a diameter of 1 m; that is, the effective receiving area of the feed cabin is small, about 0.7854 m^2 ; so, we need to judge the area of the projected area. If the projected area completely covers the central disk, the effective receiving area is 0.7854 m^2 ; if the projected area does not completely

cover the central disk, the effective receiving area is the overlapped area; if the projected area does not cover the center disk, the effective receiving area is 0.

Using the finite difference method to decompose the plane where the feed cabin is located into a set of points, for the discretization of the $\{-1 < x < 1; -1 < y < 1\}$ region on the plane where the feed cabin is located, taking $\Delta = 0.001$ as the step length, use two sets of straight line families parallel to the OX' axis and the OY' axis to mesh the circular surface and divide it into a square grid, then any point on the circular surface falls on the node of the square grid, and its coordinates are as follows:

$$\begin{cases} x = i\Delta, i = 0, 1, 2, \dots, m, \\ z = k\Delta, k = 0, 1, 2, \dots, n. \end{cases} \quad (42)$$

For each point on the circular surface, it is discussed whether it falls inside the projected triangle or on the boundary, and then the problem turns into the problem of whether the point falls inside the triangle. Take any discrete point P as the following discussion [35].

Whether the point P is inside the triangle $A'B'C'$, the positional relationship between the triangle and the disk can be judged by the barycentric method. The two sides of the projected triangle are used as the direction of a set of two-dimensional basis vectors, and then any point on the plane can be represented by this pair of basis vectors, that is, $\vec{A'P} = u\vec{A'B'} + v\vec{A'C'}$. Among them, u and v are real numbers, and the positional relationship between point P and triangle $A'B'C'$ is judged by the values of u and v . The result is shown in Figure 8.

When the value of u and v falls in the shaded part, that is, when the condition is satisfied, the point P falls inside the triangle $A'B'C'$; otherwise, it is not inside the triangle.

When the point P is inside the triangle $A'B'C'$, then we think that the point P receives the electromagnetic wave of the panel triangle ABC . It can be seen that N ($N = 1290$) panels participate in the reflection on the paraboloid with a diameter of 300 meters. Traverse the N panels. Whenever a transmitting panel is mapped to a triangle containing a point on the plane where the feed cabin is located, the number of electromagnetic waves Q received by it increases by 1. When the traversal is completed, the electromagnetic wave reception ratio of this point is $\lambda = Q/N$. Then, traverse other discrete points according to the above process. Each discrete point stores the electromagnetic wave reception ratio and draws an image, as shown in Figures 9 and 10.

Then, the total amount of electromagnetic wave radiation received by the feed cabin is λ :

$$\lambda = \sum_i^N \lambda_i \cdot \Delta^2. \quad (43)$$

Among them, λ_i is the reception ratio of discrete points in the circular surface of the feed cabin, and $\Delta = 0.001$ is the step size. According to this principle, the feed cabin

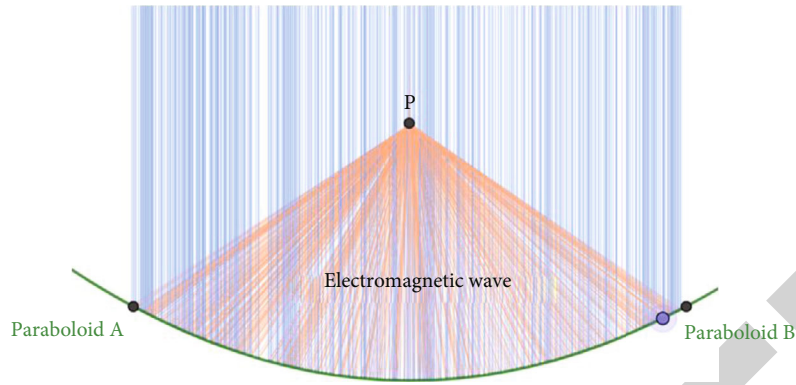


FIGURE 7: Schematic diagram of the projection of parabolic reflected electromagnetic wave.

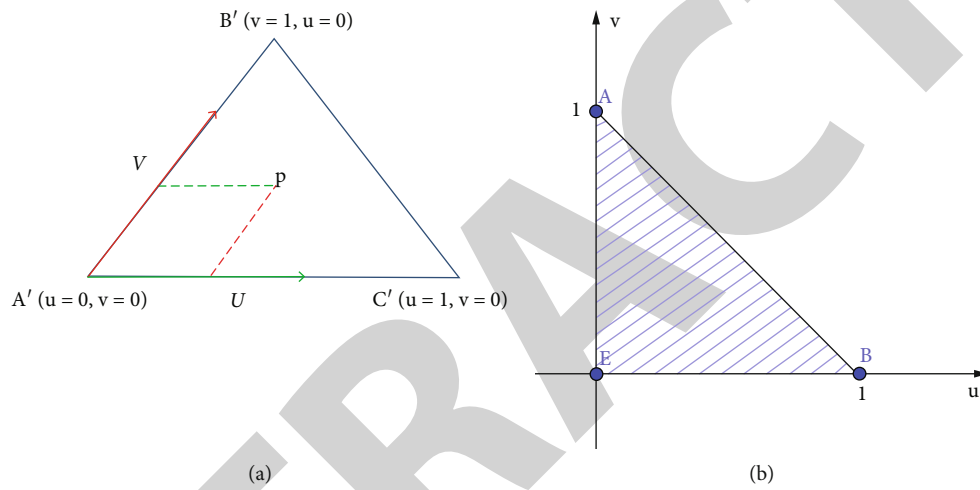


FIGURE 8: Schematic diagram of geometric relationship of landing point judged by the gravity center method.

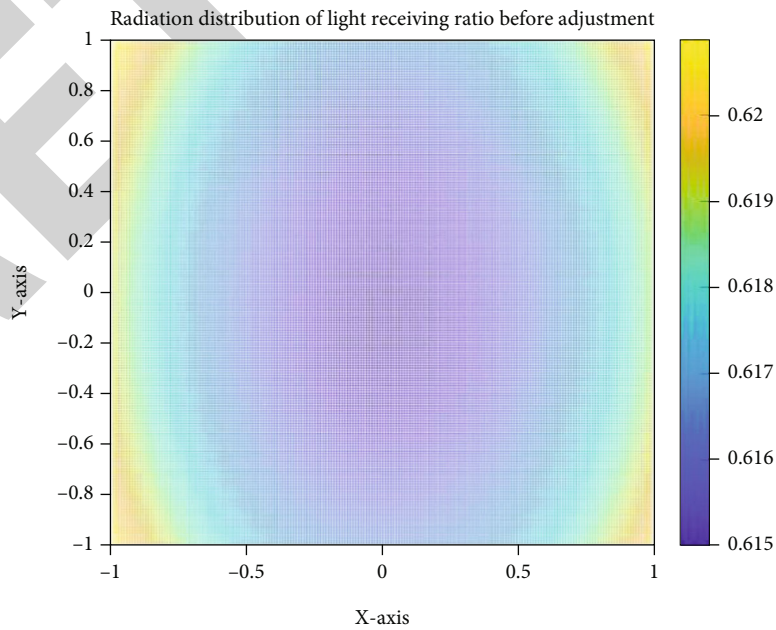


FIGURE 9: Distribution of electromagnetic wave reception ratio of parabolic reflector.

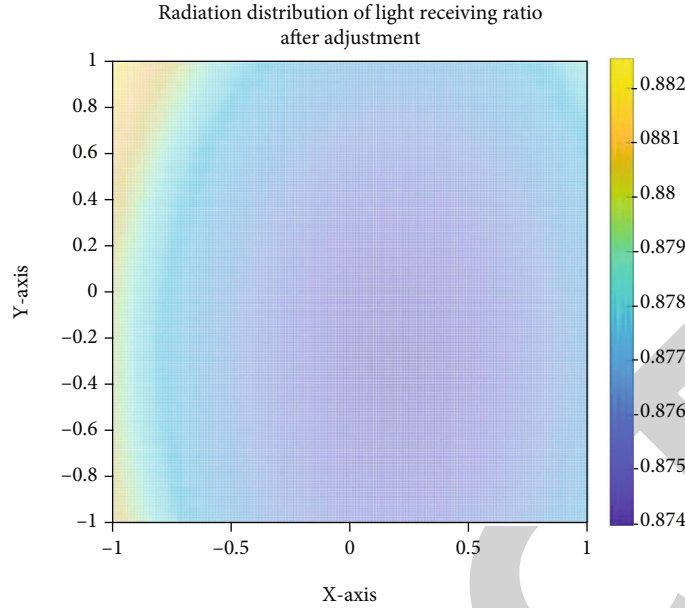


FIGURE 10: Distribution of electromagnetic wave reception ratio of parabolic reflector.

TABLE 3: Area of each reflective panel participating in reflection within 300 m aperture.

Main index node 1	Main index node 2	Main index node 3	Reflector area
A0	B1	C1	51.359 744 99
A0	B1	A1	51.360 016 64
A0	C1	D1	51.357 310 45
A0	D1	E1	51.357 310 45
.....
D245	D267	D268	65.874 906 28
D246	D247	D269	62.299 899 98
D246	D268	D269	65.594 604 13

can receive the information of 61.628% of the total number of panels before adjustment and 87.752% of the total number of panels after adjustment.

5.3.2. *Calculation of Receiving Area Ratio of Feed Cabin.* It is known that the effective receiving surface of the feed cabin is a central disk with a diameter of 1 m; that is, the effective receiving area of electromagnetic wave is small, about 0.7854m². From the idea of discrete instead of continuous, the formula of the average total electromagnetic wave radiation area of the feed cabin is as follows:

$$S_k = \lambda N \pi r^2. \tag{44}$$

The formula of the total electromagnetic radiation area provided by the reflection panel is as follows:

$$S_{\text{total}} = \sum_i^N s_i - s_{\text{cabin}}. \tag{45}$$

Among them, s_{cabin} is the shielding area of the feed cabin, and s_i is the area of each reflection panel participating in the

reflection within a diameter of 300 meters. The complete data can be found in the supporting material, and some data are shown in Table 3.

When the reference sphere is not adjusted, the feed cabin can receive the electromagnetic wave signal reflected by about 61.628% of the total panels. The ratio of the area of electromagnetic waves received by the feed cabin to the area of electromagnetic waves received and reflected by the reflector is about 0.636%. After the base sphere is adjusted to a paraboloid, the feed cabin can receive the electromagnetic wave signals reflected by about 87.75% of the total panels. The ratio of the area of electromagnetic waves received by the feed cabin to the area of electromagnetic waves received and reflected by the reflector is about 0.906%.

6. Conclusion and Prospect

In this paper, the following coordinate system is introduced, which effectively simplifies the geometric operation process, reduces the calculation and model complexity, and adopts the idea of overall deviation distribution, which can make the main cable nodes approach the ideal paraboloid more accurately. The optimization model is established with

known constraints and solved by genetic algorithm. In the process of solving, global control variables are added to screen out bad values. However, due to the complexity of the optimization algorithm, the large amount of data, the impact of various data, and the long optimization time, the specific details of the change of the distance between the main cable nodes cannot be solved and can only be screened according to the global control variables. The displacement model of the main cable node in the follow-up coordinate system established in this paper can be applied to any angle of the observed celestial body. The active reflector adjustment optimization model can add constraints to it; so, it can be extended to more conventional cases, such as considering the length change after the pull-down force. The specific research conclusions are as follows:

- (1) A follow-up coordinate system with the direction of the observed celestial body as the reference is introduced to find the optimal approximation model for the distribution of the mean square deviation of the radial direction. In order to make the telescopic distance of the actuator within the adjustable range, the overall deviation distribution vector is investigated with h as a variable to find the H value with the minimum overall offset, that is, the minimum overall root mean square error, and the trajectory equation of the ideal paraboloid is determined by the ergodic algorithm
- (2) The coordinate data of each main cable node are rotated according to the known angle of the celestial body, and the shortest distance from the actual offset point to the ideal paraboloid is obtained by using the Lagrange operator method. Taking this as the objective function, the reflective panel adjustment model of the optimal displacement strategy of discrete main cable nodes is established. The global optimal solution of the offset of each main cable node in the working area is obtained by using genetic algorithm. The mean value of the distance variation between adjacent nodes is 0.0633%, and the average length of the radial contraction of the main cable node is 0.20343
- (3) The area contacted by the electromagnetic wave is taken as the basis for judging its information content. For each triangular reflecting surface, considering the projection triangle formed when the electromagnetic wave is reflected and propagated to the plane where the feed cabin is located, the effective coverage area of the projection to the feed cabin is judged by using the finite difference method and the center of gravity method. When the reference sphere is not adjusted, the feed cabin can receive the electromagnetic wave signal reflected by about 61.628% of the total number of panels, and the ratio of the electromagnetic wave area received by the feed cabin to the electromagnetic wave surface product received and reflected by the reflector is about

0.636%. After being adjusted to a paraboloid, the feed cabin can receive about 87.75% of the total number of panels, and the ratio of the electromagnetic wave area received by the feed cabin to the electromagnetic wave area received and reflected by the reflector is about 0.906%

However, there are still some shortcomings, mainly including the following: first, the optimization algorithm is complex, the amount of data is large, there are impacts between various data, and the optimization time is long. Secondly, the change of the distance between the main cable nodes cannot be solved in detail and can only be filtered out according to the global control variables. In order to obtain more accurate results, future research will focus on increasing the constraints of the dynamic reflector adjustment optimization model, so as to extend it to more conventional cases.

Data Availability

The data in this paper come from question A of the 2021 Contemporary Undergraduate Mathematical Contest in Modeling.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was funded by the Natural Science Foundation of the Higher Education Institutions of Anhui Province (KJ2021A0486).

References

- [1] S. Y. Wu, "A project on the threshold of the 21 century: next generation of radio telescope," *Progress in Astronomy*, vol. 13, no. 2, pp. 164–172, 1995.
- [2] R. D. Nan, "Five hundred meter aperture spherical radio telescope (FAST)," *Mechanics & Astronomy*, vol. 49, no. 2, pp. 129–148, 2006.
- [3] J. Han, D. W. Fan, Y. H. Tao et al., "The project management information system for FAST scientific observation," *Frontiers of Data & Computing*, vol. 4, no. 1, pp. 20–29, 2022.
- [4] Y. Wang, Q. Deng, B. Q. Niu, H. B. Wei, and X. Guo, "Design of Active Reflector Adjustment Scheme for 'FAST' based on search optimization," *Guizhou Science*, vol. 40, no. 1, pp. 93–98, 2022.
- [5] J. X. Xue, Q. M. Wang, X. D. Gu, Q. Zhao, and H. Q. Gan, "Estimation and improvement for fitting accuracy of instantaneous parabolic reflector in FAST," *Optics and Precision Engineering*, vol. 23, no. 7, pp. 2051–2059, 2015.
- [6] Q. W. Li, P. Jiang, and R. D. Nan, "Initial gap between panels of the reflector of FAST," *Journal of Mechanical Engineering*, vol. 53, no. 6, pp. 119–121, 2022.
- [7] Z. X. Luo, X. H. Qin, and X. Chen, "Research of reflector adjustment of FAST: based on OLS model," *Technology Innovation and Application*, vol. 12, no. 5, pp. 57–59, 2022.

Retraction

Retracted: Participation in Targeted Poverty Alleviation and Enterprise Innovation Investment: Analysis of the Mediating Effect Test Model Based on Financing Constraints

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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] T. Gao and H. Wang, "Participation in Targeted Poverty Alleviation and Enterprise Innovation Investment: Analysis of the Mediating Effect Test Model Based on Financing Constraints," *Journal of Function Spaces*, vol. 2022, Article ID 7060462, 16 pages, 2022.

Research Article

Participation in Targeted Poverty Alleviation and Enterprise Innovation Investment: Analysis of the Mediating Effect Test Model Based on Financing Constraints

Ting Gao  and Huaiming Wang

College of Finance, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China

Correspondence should be addressed to Ting Gao; gaoting@nau.edu.cn

Received 9 June 2022; Revised 14 July 2022; Accepted 16 July 2022; Published 8 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Ting Gao and Huaiming 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.

In China, enterprises participating in targeted poverty alleviation can be provided with a strategic resource to enhance their ultimate value creation ability and directly win reputation and obtain sustainable competitive advantages, thus enhancing their innovation ability. In this study, the influence and mechanism of the participation in targeted poverty alleviation on enterprise innovation investment are empirically examined based on financing constraints with the data of nonfinancial listed companies in Shanghai and Shenzhen from 2012 to 2019 as the object. The research suggests that enterprises taking the initiative to participate in targeted poverty alleviation can first have increased innovation investment and especially enhanced exploratory innovation ability. Second, enterprises' participation in targeted poverty alleviation can reduce financing constraints of enterprises' innovation, improve their resource allocation efficiency, and stimulate their innovation vitality primarily through the influence of resource, reputation, and information. This study enriches the theoretical research on the economic consequences of targeted poverty alleviation and provides a reference for guiding enterprises in practice to fulfill social responsibilities to promote enterprises' innovation vigor and boost their high-quality development.

1. The Introduction

Poverty has always been a major obstacle hindering the progress of human society. Poverty eradication has been a strategic event in governance since ancient times, a basic mission of the CPC, and an important task for China to build a moderately prosperous society in an all-round way and realize the great rejuvenation of the Chinese nation. "The world of chaos is not in the rise and fall of a surname, but in the people's happiness." In 2013, President Xi Jinping proposed the idea of targeted poverty alleviation and sounded the charge of the battle against poverty. In 2014, the targeted poverty alleviation policy was officially implemented and made aid accessible to all households. In 2015, the policy of targeted poverty alleviation was promoted as a national strategy. In the above context, the whole Party and the people of all ethnic groups in China have taken out the spirit of "dare to teach the sun and the moon to

change the sky," summoned up the momentum of "Our men wouldn't come back till they beat the Loulans," and launched a general attack on poverty. After five years of poverty alleviation and seven years of targeted poverty alleviation, China has achieved accomplishments in extreme poverty eradication and achieved decisive progress in building up a well-off society in an all-round manner, creating another human miracle that will go down in history.

In China, enterprises serve as the critical important market players and the direct undertakers of social production and circulation, showing unique advantages in achieving targeted poverty alleviation. Notably, since the targeted poverty alleviation action of "ten thousand enterprises helping ten thousand villages" was carried out nationwide, numerous enterprises have made an active response to the call of the central government, fulfilled social responsibility, participated in actions relating to poverty alleviation, and contributed to winning the battle against poverty. The participation

in targeted poverty alleviation takes on a strategic significance for enterprises in fully exploiting the advantages of the industrial chain. To be specific, enterprises can leverage and integrate social resources and their capital strength for alleviating poverty [1], assume actively social responsibilities, and return their corporate value [2, 3]. Despite the possible consumption of human, material, and financial resources, enterprises participating in targeted poverty alleviation will also gain strategic resources to optimize their value chain, industrial chain, and supply chain; build a good social image and brand image; and take advantage of the government's tax relief and other support policies, which is considered a win-win strategic project [4, 5]. On the one hand, enterprises in poor areas that participate in targeted poverty alleviation can be directly beneficial to develop industries in the poor areas, expand markets and build brands, and help industrial projects relating to poverty alleviation connect to the market and generate benefits quickly via their sales channels, procurement channels, and ecosystems [6]. As a result, the coordinated development of local economy will be stimulated; the employment of poor people will be promoted; the conditions of education, medical care, and public facilities will be improved; and the living standards of poor areas will be improved through the overall allocation of resources [7]. On the other hand, enterprises taking the initiative to participate in targeted poverty alleviation can convey the spirit of dedication to the outside world, help them build a good social image, improve their corporate reputation, develop solid social relations with stakeholders, and further acquire more strategic resources (e.g., capital, talent, and technology) [8], thus contributing to the profit-making of enterprises and the combination of economic and social benefits [9].

Moreover, innovation continues to be the primary driver of China's economic and social development and strategically supports the building of a modernized economy. With the stress on the need to unswervingly advance supply-side structural reform, foster new growth points, and achieve high-quality development in a down-to-earth manner, the 19th National Congress of the Communist Party of China (CPC) clearly declared that China will become one of the leading innovative countries by 2035. China's science and technology has been leaping forward, the innovation system has been progressively optimized, the innovation environment has been constantly optimized, and the innovation capacity has been significantly enhanced during the 13th Five-Year Plan period. China's innovation governance has formed a novel pattern, whose scientific and technological strength has also reached a new level. China's R&D investment has increased from 1.42 trillion yuan in 2015 to 2.21 trillion yuan in 2019, with the intensity of R&D investment reaching 2.23 percent in 2019. China has moved up the GII ranks consistently and steadily since 2013, as suggested by the Global Innovation Index released by the World Intellectual Property Organization (WIPO), which has led to continuous progress from ranking 14th in 2020 to 12th in 2021. Furthermore, China has been a global innovation leader, knocking at the door of the GII top 10. China has become the world's second largest R&D investment country and second largest knowledge output country.

Innovation has become an enduring research topic in economic management and finance, giving birth to several theoretical and practical achievements on innovation. In the context that China has committed to becoming an innovative country and the system design of innovation has driven its development, innovation has become an enduring research topic in economic management and finance, giving birth to several theoretical and practical achievements on innovation. Moreover, based on the continuous improvement of talent training level and the rapid change of higher education, "enterprise innovation" has served as a solid force at intermediate level to optimize the functions of the capital market and adjust and change the industrial structure, which is also a significant cornerstone for the healthy development and competitiveness improvement of microenterprises. Despite the existing literature, comprehensive studies have been conducted on the factors for enterprise innovation, including industrial policy at the macrolevel [10], such as governance system [11, 12] and financial environment [13], and at the microlevel, corporate governance [14], internal control [15], and accounting conservatism [16, 17]. How to play the role of corporate social responsibility in corporate innovation through targeted poverty alleviation has been rarely investigated.

In China's emerging and transforming market, the main force to facilitate innovation and development in China is R&D; although R&D is of critical significance for enterprises in improving their core competitiveness and achieving sustainable development [18], enterprises' innovation projects are commonly faced with a high degree of information asymmetry and financing constraints due to high risk, substantial capital requirements, and long cycle of innovation, thus restricting enterprises' R&D innovation activities. With the continuous promotion of "the rule of China," China's enterprises have been more extensively involved in social governance activities and play an increasingly vital role in the process of improving national governance capacity. Subsequently, a question is raised whether the participation of enterprises in targeted poverty alleviation will help ease financing constraints and thus affect the innovation input of enterprises.

Under the above policy background and the current enterprises' financing difficulties and lack of innovation, this study is aimed at exploring whether enterprises' participation in industrial targeted poverty alleviation, besides creating value for the society, will reduce their financing constraints, increase their innovation investment, and achieve a win-win situation between social value and enterprise value in terms of corporate social responsibility under the policy guidance. The marginal contribution is embodied in two aspects. First, the innovative behavior of enterprises through poverty governance is investigated in this study based on China's actual situation. Thus, this study enriches the relevant literature on the development of targeted poverty alleviation based on China's social governance while providing more insights for enterprises to play the governance role in consolidating the achievements of poverty alleviation and implementing the Rural Revitalization Strategy. Second, in terms of financing constraints, the theoretical analysis and

empirical test of the influence of targeted poverty alleviation on enterprises' innovation investment, and financing constraints in deep insight into the relationship between the function mechanism, the study can promote the motivation of enterprises to fulfill social responsibility and lead to the economic consequences of a rich China's enterprises to be involved in targeted poverty alleviation. As a result, empirical evidence can be provided for enterprises to carry out targeted poverty alleviation and reduce financing constraints, help enterprises actively launch and disclose targeted poverty alleviation projects, and provide novel ways for enterprises to reduce financing constraints. Third, enterprises' participation in targeted poverty alleviation is a specific behavior to roll out national strategic policies. The study on the economic consequences of enterprises' participation in targeted poverty alleviation is beneficial to provide more insights into the influence of national macropolicies on corporate behavior, and it can guide enterprises to promote their high-quality development by actively engaging in public welfare undertakings, thus helping regulators further optimize the social responsibility disclosure system and regulatory mechanism and actively encouraging enterprises to launch investment projects in underdeveloped areas to contribute to rural revitalization.

2. Theoretical Analysis and Hypothesis

2.1. Participation in Targeted Poverty Alleviation and Innovation Investment of Enterprises. Innovation is a vital factor for an enterprise to enhance its value creation ability and maintain its competitive edge. Innovation largely determines the success or failure of an enterprise, especially in the increasingly competitive international market environment at present [19]. The participation in targeted poverty alleviation is a vital measure for enterprises to assume social responsibility and practice socialist core values, and it takes on a strategic significance for enterprises in enhancing their innovation ability. On the one hand, the participation in targeted poverty alleviation can facilitate product optimization and upgrading of enterprises and increase innovation efforts. Accuracy is essential to the essence of poverty alleviation "accurate," requiring accurate objects for poverty alleviation and placing more emphasis on project scheduling precision and poverty results accuracy. Accordingly, enterprises participating in targeted poverty alleviation are no longer simply making donations for charity, but to "foster" object for poverty alleviation, help poor areas develop characteristic industry, stimulate the endogenous power of poverty alleviation, and fundamentally achieve poverty alleviation and prosperity. In accordance with the poverty area, human resources, environment, and others, an enterprise can combine its development strategy, market position, product process, and others to retrofit existing technologies, develop novel products, optimize the industrial structure, and break through the geographical position, the consumption structure, the limitation of the natural environment, etc. The above process requires the enterprise to enhance its innovation ability and increase the intensity of innovation, so the enterprise can develop emerging industries adapting to poor

areas under local conditions, form novel business models, and design advanced production technologies capable of facilitating the common development of enterprises and poor areas. As a result, the participation of enterprises in poverty alleviation activities refers to a process of continuous innovation and development [20, 21].

On the other hand, the participation in targeted poverty alleviation can improve employees' sense of identity with the organization and stimulate enterprise innovation vitality. Human capital, the carrier of knowledge, is the vital factor of productivity to boost the development of enterprises and the basic unit of enterprise innovation activities, thus taking on a critical significance in enterprise innovation. The enhancement of human capital quality can enhance the ability of enterprises to absorb new knowledge and develop novel products and help enterprises acquire and share knowledge, form innovation culture, and transform innovation achievements [22]. Enterprises participating in targeted poverty alleviation, assuming their social responsibilities with high quality, and gaining considerable government recognition and high social recognition will increase the company's visibility and external governance efficiency. Moreover, employees will enhance their sense of ownership and sense of pride and belonging to the organization through the external honor of the enterprise [23]. In accordance with social exchange theory, the stakeholders of an enterprise allocate resources in the balance between the reward and the cost. If one gets a large reward through the transaction, it will voluntarily repay others through material and emotional behaviors. Employees' sense of identity to the organization will improve their work enthusiasm, strengthen the cohesion of the enterprise, deepen communication and team cooperation, further integrate innovative ideas, stimulate the innovation vitality of the enterprise, and enhance the innovation ability [24]. Furthermore, enterprises can establish a good image, enhance social reputation, transmit favorable signals, make stakeholders have full confidence in the business status and development potential of the enterprise, and improve their investment desire and strength by actively assuming their social responsibilities to reduce the financing constraints of enterprises, increase innovation investment, and contribute to R&D and innovation activities.

Exploratory innovation and exploitative innovation are two types of innovation ways with different logics and tracks. To achieve breakthrough change and high-quality development, enhance the competitive ability, and achieve the transformation and upgrading in accordance with the development of new problems, new economic mode, and new consumption demand, enterprises should investigate novel technologies, develop novel products, create new markets, develop new industry innovation behavior, which is high-risk, large-scale, high-return, and radical exploratory innovation [25]. Exploitative innovation is more related to the updating, transformation, and improvement of products and technologies based on the original and the reuse of original knowledge and resources [26, 27]. Compared with exploitative innovation, exploratory innovation will acquire, integrate, digest, and absorb new knowledge, resources, and

technologies. With participating in targeted poverty alleviation, enterprises often develop emerging industries based on the resource advantages of poor areas and their own development needs, which requires considerable capital to update technology, research, and development of novel products and innovation of business models and stimulate the economic vitality of poor areas. Thus, targeted poverty alleviation itself is an exploratory innovation, and enterprises' participation in targeted poverty alleviation will be more conducive to promoting their investment in exploratory innovation.

Based on the above analysis, this study proposes the following hypotheses:

- (i) *H1a*: the participation in targeted poverty alleviation is conducive to promoting enterprise innovation investment
- (ii) *H1b*: compared with development innovation, the participation in targeted poverty alleviation is more conducive to increasing investment in exploratory innovation

2.2. Participation in Targeted Poverty Alleviation and Corporate Financing Constraints. In accordance with the MM theory, the market is a perfect market, where there is no difference in the cost of obtaining funds from internal and external sources. However, the information asymmetry theory suggests the existence of agency problems and the fact that there is no perfect world market, causing friction between the supply and demand sides of funds in transactions. Moreover, internal and external financing cannot be completely replaced by each other due to the problems of risk and quantity. As a result, the cost of external financing is significantly higher than the cost of internal financing and finally affects the relevant decisions of investment and financing. Due to the limited self-owned funds, enterprises will always face financing activities from external markets or financial institutions in the process of operation and financial management. The R&D and innovation activities of enterprises require considerable continuous capital investment as support. Internal funds are often difficult to meet the needs of R&D and innovation activities, so enterprises consider external sources. When obtaining funds from the outside, the fund supplier should often spend more costs to acquire the information of the fund demander and prevent moral hazard and other problems. Enterprises should pay the above costs when obtaining funds from the external market, so the cost of external financing is significantly higher than the cost of internal financing, thus limiting the financing channels of listed companies.

The development of enterprises is inseparable from R&D and innovation activities. R&D and innovation require considerable continuous capital investment as a guarantee. The lack of financing and capital support hinders the investment of numerous enterprises in innovation. When constrained by financing, R&D innovation is generally forced to be interrupted or stopped due to insufficient

funds. Due to the imperfection of China's capital market, financing constraints (e.g., financing difficulty and high cost of capital) have been a main factor hindering the development of China's enterprises. If enterprises can assume social responsibility actively, it will reduce financing constraints through the influence of resource, reputation, and information [5].

First, the participation in targeted poverty alleviation can obtain strategic resources and promote resource integration. Resource-based theory holds that an enterprise is an aggregate of various resources, and the enterprise with unique resources has "resource positioning barrier" and can obtain unique competitive advantages. The participation in targeted poverty alleviation can help enterprises improve their social influence and bring positive economic benefits, which are differentiated resources for enterprises to participate in market competition [28]. On the one hand, although poor areas lack social resources, they are often rich in natural and economic resources, thus leading to a good public opinion environment for enterprises. Enterprises participating in targeted poverty alleviation can effectively exploit the resource endowment advantages of poor areas; obtain the support of local land, labor, and other characteristic resource factors; organically combine enterprise capital, technology, and management; and boost the efficient allocation of a variety of resources, industries, and production factors. As a result, a material basis is laid for enterprises to develop local characteristic industries in poor areas; superior brands are built; the added value of products is increased; local characteristic industry chains are established accurately and effectively; the innovation and transformation of the original technology, products, and business model are achieved; and innovation activities are promoted. On the other hand, under the background that philanthropy, a new "soft competitiveness" of enterprises, will play a radiating role in the economy, targeted poverty alleviation is a national strategic deployment, and enterprises' participation in it can help them be more recognized from the government and obtain core technologies and resources required for innovation with the help of the government [1]. The government will provide policy, information, and financing support to encourage and support enterprises to carry out targeted poverty alleviation. Accordingly, the participation in targeted poverty alleviation will ease corporate financing constraints through the acquisition and integration of strategic resources.

Second, the participation in targeted poverty alleviation can respond to stakeholder expectations and improve social reputation. Stakeholder theory suggests that enterprise operation is aimed at maximizing the value of the company on the premise of protecting the interests of stakeholders. Enterprises and society are a type of symbiotic relationship. The development of enterprises should be balanced with economy, environment, and society. Enterprises should assume corresponding social responsibilities and create certain social value by realizing self-value [29]. As a social and economic organization, enterprises should not exist for the purpose of shareholder value, whereas they should effectively assume corresponding social responsibilities to the society, community, relevant individuals, and other

stakeholders. An enterprise refers to a collection of contracts of stakeholders who are the source of creating value for the enterprise. The business activities of the enterprise will have an effect on the interests of stakeholders, and stakeholders also take on a critical significance in the formulation and implementation of the enterprise strategy [30].

Participating in targeted poverty alleviation refers to the natural responsibility of enterprises and the need to create a responsible image. Besides, it is a strategic social responsibility of shared value [31]. Enterprises' participation in targeted poverty alleviation can improve their social image; accumulate good reputation capital; gain recognition, support, and trust from stakeholders; and bring strategic competitive advantages to enterprises. Social reputation is a critical intangible asset for the survival and development of enterprises. It is the basic standard for enterprises to enhance business communication, expand market share, and obtain credit support. First, the enterprise participating in targeted poverty alleviation sets up a good image, which will help the upstream and downstream enterprises of the supply chain develop a stable relation of cooperation and get more market share and investors to attract more capital. The enterprise can also improve the core position in the value chain of enterprises and market share, enhance the value creation ability and profitability of the enterprise, and increase the cash flow and capital accumulation of business activities, so as to reduce its internal financing constraints. Second, in China's current economic system, banks are still the main channel of enterprise financing, and social reputation lays a basis for banks and other financial institutions to measure whether enterprises can obtain credit. The good reputation capital of enterprises makes it easier for them to obtain more loans and credit support, which solves the problem of difficult and expensive financing to some extent. Samet and Jarboui suggested that enterprises that actively fulfill their social responsibilities are more likely to obtain bank loans with lower interest rates and longer terms [32].

Third, participating in targeted poverty alleviation can reduce the degree of internal and external information asymmetry of enterprises and enhance their financing ability. Information asymmetry is the root cause of enterprise financing constraint. The key to reduce the degree of enterprise financing constraint is to decrease the degree of internal and external information asymmetry. Enterprises' participation in targeted poverty alleviation information has incremental information, which may increase the confidence of information demand of capital providers. Enterprises consume considerable resources to participate in targeted poverty alleviation activities and assume the social responsibility of poverty alleviation and prosperity, suggesting that enterprises are in good operating condition and have certain financial capacity to launch public welfare activities. Enterprises that disclose accurate information for poverty alleviation can improve enterprise information transparency; reduce the degree of information asymmetry; make potential investors or creditors more objectively forecast the future of the enterprise to the financial situation, to enhance the accuracy of its assessment of the value of

the company; eliminate the fear of the market; and reduce the transaction costs in the corporate finance activities, thus easing financing constraints [33]. Gao et al. [34] confirmed that CSR fulfillment by enterprises can reduce financing constraints; the higher the quality of CSR information disclosure, the stronger the mitigation effect will be.

Based on the above analysis, this study proposes the following hypothesis:

- (i) *H2*: the participation in targeted poverty alleviation will help ease financing constraints of enterprises

2.3. Targeted Poverty Alleviation, Financing Constraints, and Investment in Innovation. An enterprise's innovation project is characterized a long cycle, a great capital demand, and high risks. The research and development of project success are directly related to the increase of enterprise competitiveness, so enterprises will adopt a series of measures to strictly protect an important commercial secret. As a result, the enterprise innovation activities often face serious information asymmetry and high degree of innovative financing constraints, thus threatening the implementation of innovation strategy. Enterprises facing higher financing constraints expect increased operational risks and increased investment difficulty, as well as reduced returns. Thus, enterprises are more inclined to adopt sound investment strategies, reduce innovation input with greater uncertainty, and reduce the possibility of failure of investment projects [35]. When enterprises can raise the funds needed for investment more easily and at a lower cost, they will have higher profitability, and they often have sufficient cash flow and working capital, thus weakening the information asymmetry between enterprises and investors. In addition, the enterprise financing environment can be optimized, which can further alleviate financing constraints, enhance enterprise innovation confidence, and improve enterprise innovation ability. Thus, the participation in targeted poverty alleviation will promote innovation investment by alleviating corporate financing constraints.

In addition, exploratory innovation has a longer return cycle and higher uncertainty than development innovation, which is more related to the improvement of the competitiveness level of enterprises, and its information asymmetry is more serious, and underinvestment is more common. Accordingly, this type of innovation is more restricted by financing constraints. Moreover, participation in targeted poverty alleviation is more conducive to the strategic development of enterprises, exploratory innovation is also more focused on the realization of long-term interests of enterprises, and enterprises often carry out the two activities for the same purpose. Thus, strategic resources brought by participating in targeted poverty alleviation will play a greater role in facilitating exploratory innovation.

Based on the above analysis, this study proposes the following hypotheses:

- (i) *H3a*: the participation in targeted poverty alleviation will promote the increase of enterprise innovation input through the alleviation of financing

constraints; that is, financing constraints play a partial intermediary role between the participation in targeted poverty alleviation and innovation input

- (ii) *H3b*: compared with developmental innovation, the participation in targeted poverty alleviation can promote the increase of exploratory innovation input by alleviating financing constraints; that is, financing constraints play a greater mediating role between the participation in targeted poverty alleviation and exploratory innovation input than between developmental innovation input

3. Research Design

3.1. Sample Design and Data Sources. Due to the disclosure of the listed company accurate data for poverty alleviation, which began in 2016, to compare the company involved in precise changes before and after the innovation investment for poverty alleviation, the symmetry of guarantee during the study period, the Shanghai and Shenzhen A-share nonfinancial listed companies between 2012 and 2019 are selected as the research samples, and the observation and related variables for ST and *ST company samples after missing value are excluded in each year. All continuous variables in 1% and 99% quartile are Winsorized. A total of 15,464 observations of 3084 listed companies are obtained.

Excel 2007 and Stata15 software are adopted to process and verify the data. Targeted poverty alleviation data are collected and classified manually from the annual corporate social responsibility report, and other data are collected from the CSMAR database.

3.2. Model Design and Variable Definition. Listed companies' participation in targeted poverty alleviation is an active charity behavior, and there may be significant differences between participating enterprises and nonparticipating enterprises in advance. In this study, a multiperiod differential model is adopted by referring to the study by Zhang and Dong [8] to test the net impact of participating in targeted poverty alleviation on the level of innovation investment of enterprises, avoid the interference of prior differences, and reduce the estimation error. The model design is expressed as

$$RD_{i,t} = \beta_0 + \beta_1 TPA_{i,t} + Controls_{i,t} + Year + Industry + Firm + \varepsilon_{i,t}. \quad (1)$$

The explained variable RD in model 1 represents the company's innovation investment level, which is measured by the ratio of R&D expenditure to operating income. Furthermore, in accordance with the different nature of innovation, the measurement indexes of exploratory innovation (RD1) and developmental innovation (RD2) are constructed, which are measured by the ratio of expenditure to operating revenue in the research stage and development stage to test hypotheses H1a and H1b.

The explanatory variable TPA is the dummy variable of listed companies' participation in targeted poverty allevia-

tion. Targeted poverty alleviation information disclosed by listed companies comprises general information and itemized input information, whereas the two are inconsistent or lack of information. Considering the information disclosure of listed companies of the inconsistency problem, to increase the accuracy of the measure, this study refers to Deng et al. [5], only to have a combination of the overall situation and the samples as to participate in the precision of the sectional input information samples for poverty alleviation. The TPA value is set to 1; otherwise, it is considered to be not involved in an accurate sample for poverty alleviation, with TPA value of 0. To control other factors that affect enterprise innovation input, this study also selects a series of indicators including financial characteristics and corporate governance characteristics and controls annual, industry, and firm individual fixed effects. Specific variable definitions are listed in Table 1.

To test whether financing constraints play an intermediary role in the relationship between targeted poverty alleviation and innovation input, the following model is constructed to test hypothesis H2 and hypotheses H3a and H3b:

$$KZ_{i,t} = \beta_0 + \beta_1 TPA_{i,t} + Controls_{i,t} + Year + Industry + Firm + \varepsilon_{i,t}, \quad (2)$$

$$RD_{i,t} = \beta_0 + \beta_1 TPA_{i,t} + \beta_2 KZ_{i,t} + Controls_{i,t} + Year + Industry + Firm + \varepsilon_{i,t}. \quad (3)$$

The explained variable KZ index in model 2 represents the financing constraints of the company. The larger the value of KZ index is, the more serious the financing constraints the company faces. The calculation method refers to Kaplan and Zingales [36] and Wei et al. [37] in accordance with five indicators (including operating cash flow, cash dividends, cash holdings, asset-liability ratio, and Tobin's Q). The KZ index, measuring the financing constraint degree of listed companies, is fitted and estimated. In model 3, the KZ index is introduced into the test model of innovation input and combined with model 2; the mediating effect of enterprises participating in targeted poverty alleviation improving innovation input by alleviating financing constraints is verified.

4. Empirical Analysis

4.1. Descriptive Statistical Results and Analysis. The descriptive statistical results of the main variables are listed in Table 2. To improve the visibility of regression coefficient, the innovation input indicators applied in this study are all percentages. The average value of (RD) is 4.707, suggesting that the average proportion of R&D expenditure to revenue of the sample enterprises is nearly 4.707%, in which the level of exploratory innovation investment (RD1) is about 4.222%, and the level of development innovation investment (RD2) is 0.410%, suggesting that the overall innovation ability of China's enterprises is not high, and most enterprises' innovation is concentrated in the exploratory activities

TABLE 1: Variable definition table.

Variable categories	The variable name	Variable code	Variable definitions
Explained variable	The innovation	RD	T period R&D expenses/ T period operating revenue
	Exploratory innovation	RD1	T phase research phase expenditure/ T phase operating income
	Development innovation	RD2	T phase development phase expenditure/ T phase operating income
Explanatory variables	Targeted poverty alleviation	TPA	Dummy variable: when the company discloses the investment amount of at least one specific targeted poverty alleviation project in the current year, the value is 1; otherwise, the value is 0.
Intervening variable	Financing constraints	KZ	It is estimated by reference to relevant research
	The company size	Size	\ln (total assets in phase T)
	Establishment year of company	Age	\ln (company establishment year of period T)
	Return on assets	ROA	Net profit of period T /total assets at the end of period T
	Company growth	Growth	Growth rate of operating income in period T
	Financial leverage	Lev	Total liabilities at the end of period T /total assets at the end of period T
	Debt paying ability	Pay	Ebit for period T /total liabilities at the end of period T
Control variables	The cash flow	CFO	Net cash flow from operating activities/total assets of the company
	Proportion of fixed assets	PPE	Net fixed assets/total assets of the company
	Ownership concentration	Top1	The shareholding ratio of the largest shareholder
	The joining together of two jobs	Dual	The value is 1 if the chairman and the general manager are concurrent; otherwise, the value is 0
	Board size	Bsize	The number of board members
	Proportion of independent directors	Indep	Ratio of number of independent directors to number of directors
	The annual	Year	Annual dummy variable
	Industry	Ind	Industry dummy variable

TABLE 2: Descriptive statistical results.

Variable	Observations	Mean	Standard	Minimum	The median	Maximum
RD	15464	4.707	4.544	0.030	3.650	26.100
RD1	15464	4.222	3.903	0.014	3.439	22.519
RD2	15464	0.410	1.247	0.000	0.000	7.827
TPA	15464	0.216	0.418	0.000	0.000	1.000
KZ	15464	0.418	2.043	-9.667	0.649	9.167
Size	15464	22.113	1.221	19.847	21.941	25.984
Age	15464	2.835	0.333	1.792	2.890	3.466
ROA	15464	0.040	0.069	-0.290	0.039	0.214
Growth	15464	0.169	0.379	-0.518	0.110	2.445
Lev	15464	0.404	0.200	0.050	0.391	0.907
Pay	15464	0.216	0.349	-0.646	0.129	2.053
CFO	15464	0.046	0.065	-0.148	0.044	0.230
PPE	15464	0.208	0.143	0.004	0.182	0.649
Top1	15464	0.337	0.143	0.088	0.318	0.736
Dual	15464	0.301	0.459	0.000	0.000	1.000
Bsize	15464	8.473	1.598	5.000	9.000	14.000
Indep	15464	0.376	0.053	0.333	0.353	0.571

rather than in the development stage. In the research sample, the proportion of enterprises participating in targeted poverty alleviation (TPA) is nearly 21.6%, suggesting that the intensity of China's enterprises' participation in poverty alleviation needs to be improved. The mean value and standard deviation of financing constraint (KZ) are 0.418 and 2.043, respectively, suggesting that all listed companies have financing constraint to a certain extent, and there are significant differences in financing constraint degree among different enterprises. The descriptive statistics of other control variables are basically consistent with the general statistical results and roughly consistent with the actual situation.

4.2. Correlation Analysis. To preliminarily test the influence of the participation in targeted poverty alleviation on enterprise innovation input and judge whether the proposed research hypothesis is reasonable, this study conducted Pearson correlation test on the participation in targeted poverty alleviation and innovation input, and the results are listed in Table 3. The correlation coefficient between innovation input (RD) and the participation in targeted poverty alleviation (TPA) is 0.09, but is not significant, which cannot verify H1a. The possible reason is that enterprises participating in targeted poverty alleviation have large individual heterogeneity, and the net impact of targeted poverty alleviation on company innovation needs to be observed based on controlling various factors and company characteristics. The correlation between other explanatory variables and innovation input remains significant, suggesting that the explanatory variables selected in this study are reasonable. The correlation coefficients between all variables are less than 0.5, and the VIF test value is far less than 10, suggesting that there is no multidiscipline in the research model.

4.3. Multiple Regression Results and Analysis. Table 4 lists the basic regression results of the influence of targeted poverty alleviation on the company's innovation investment. In column 1, considering the fixed effects in control of the year, the industry, company, and other factors which affect innovation input levels, the regression coefficient of TPA was 0.138, which is significant at the 5% level, suggesting that businesses participate in precision while consuming a certain amount of resources for poverty alleviation, but will ensure accurate poverty alleviation policy implementation to promote the development of the enterprise innovation activities, and will increase investment in innovation by enterprises. Further, to test the heterogeneous influence of the participation in targeted poverty alleviation on enterprise innovation activities, columns 2 and 3 also test the influence of the participation in targeted poverty alleviation on exploratory innovation and developmental innovation. There is a significant positive correlation between TPA and RD1 at the 1% level, but the correlation between TPA and RD1 is not significant. Accordingly, the significant positive influence of the participation in targeted poverty alleviation on innovation input only exists in exploratory innovation, and there is no significant influence on development innovation. This indicates that compared with development-oriented innovation, the participation in targeted poverty alleviation is more

conducive to improving the investment level of exploratory innovation, thus verifying hypothesis 1B.

Because the viscosity of the firm's innovation investment exists for a certain period, the innovation investment may be the heterogeneity of the early decision or individual difference. Thus, if the absolute value of innovation input index is adopted, it is not reliable, objective, and scientific to accurately reflect the influence of poverty alleviation on the real influence of innovation activities. Accordingly, this study further uses the variable increment index of innovation input (ΔRD) to reexamine model 1. The regression results of columns 4 to 6 show that TPA is significantly positive with ΔRD and $\Delta RD1$ at the level of 1% and 5%, respectively, but the correlation coefficient with $\Delta RD2$ is not significant, which verifies H1a and H1b again, suggesting that the participation in targeted poverty alleviation can not only promote the level of innovation input but also improve the growth level of innovation input. Meanwhile, enterprises participating in targeted poverty alleviation pay more attention to long-term development benefits. To deepen industrial transformation and achieve high-quality development, enterprises will actively absorb new knowledge and skills, develop novel products, expand new markets, and carry out exploratory innovation activities according to the needs of market development.

Corporate social responsibility can be considered a strategy to obtain resources required for development, and strategic resources are the vital mechanism to affect corporate financial decisions. Table 5 lists the test results of financing constraint as an influencing mechanism. Columns 1 and 2 represent the regression results for the model. The regression coefficient of the TPA with KZ -0.953, significant at 1% level, suggests that through the integration of resources and superior resources, the poor areas receive recognition and financial support from the government to help enterprises to participate in targeted poverty alleviation. These enterprises can also gain social recognition through reputation effect and improve profitability. In addition, these enterprises can reduce the degree of information asymmetry between investors and bondholders and obtain the required credit resources at a lower cost, thus easing the level of financing constraints of enterprises.

Columns 2–7 represent the regression results of model 3. After joining the financing constraint index, TPA and innovation investment (RD, delta RD) still show a significant positive relationship. TPA and KZ are negatively related at the 1% significance level. Based on model 1, the regression result suggests that as the enterprises participating in targeted poverty alleviation have obtained more funds for poverty alleviation resources, they have reduced the financing constraints on innovation and become more innovative. The level of innovation input has significantly increased. Thus, financing constraints play a partial intermediary role in the relationship between the participation in targeted poverty alleviation and enterprise innovation input. The participation in targeted poverty alleviation can significantly facilitate the R&D and innovation investment intensity of enterprises by alleviating financing constraints, consistent with the theoretical analysis results above and supporting

TABLE 3: Table of phase relation.

	RD	TPA	Size	Age	ROA	Growth	Lev	Pay	CFO	PPE	Top1	Dual	Bsize	Indep
RD	1													
TPA	0.09	1												
Size	-0.27***	0.27***	1											
Age	-0.12***	0.18***	0.19***	1										
ROA	-0.02**	0.04***	-0.01	-0.06***	1									
Growth	-0.03***	0.01	0.04***	-0.06***	0.28***	1								
Lev	-0.29***	0.10***	0.53***	0.16***	-0.37***	0.00	1							
Pay	0.07***	-0.01*	-0.17***	-0.08***	0.33***	0.12***	-0.23***	1						
CFO	-0.04***	0.07***	0.06***	0.03***	0.39***	0.01	-0.17***	0.37***	1					
PPE	-0.24***	0.07***	0.11***	0.04***	-0.10***	-0.10***	0.13***	-0.09***	0.21***	1				
Top1	-0.17***	0.05***	0.17***	-0.06***	0.12***	-0.02***	0.05***	0.06***	0.10***	0.10***	1			
Dual	0.13***	-0.06***	-0.19***	-0.10***	0.04***	0.03***	-0.13***	0.06***	-0.00	-0.08***	-0.02***	1		
Bsize	-0.13***	0.09***	0.29***	0.09***	0.01	-0.02**	0.16***	-0.05***	0.03***	0.14***	0.01*	-0.18***	1	
Indep	0.07***	-0.00	-0.01	-0.04***	-0.02**	-0.01	-0.01	-0.00	0.00	-0.04***	0.05***	0.12***	-0.53***	1

Note: ***, **, and * are significant at the level of 1%, 5%, and 10%, respectively.

TABLE 4: Participation in targeted poverty alleviation and company innovation investment.

Variable	(1) RD	(2) RD1	(3) RD2	(4) Δ RD	(5) Δ RD1	(6) Δ RD2
TPA	0.138** (2.10)	0.158*** (2.67)	0.004 (0.13)	0.211*** (3.05)	0.156** (2.39)	0.045 (1.44)
Size	0.046 (0.84)	-0.124** (-2.54)	0.170*** (7.15)	0.367*** (6.41)	0.312*** (5.78)	0.029 (1.11)
Age	-2.087*** (-5.48)	-1.346*** (-3.93)	-0.297* (-1.79)	0.390 (0.97)	0.601 (1.59)	-0.322* (-1.78)
ROA	-4.580*** (-9.72)	-4.494*** (-10.60)	0.239 (1.16)	-3.172*** (-6.39)	-3.756*** (-8.03)	1.063*** (4.74)
Growth	-0.809*** (-16.67)	-0.610*** (-13.99)	-0.164*** (-7.74)	-1.594*** (-31.21)	-1.323*** (-27.50)	-0.230*** (-9.95)
Lev	-2.500*** (-12.12)	-2.271*** (-12.25)	-0.262*** (-2.91)	-0.805*** (-3.71)	-0.821*** (-4.02)	-0.037 (-0.37)
Pay	-0.568*** (-5.28)	-0.425*** (-4.39)	-0.155*** (-3.28)	0.150 (1.33)	0.028 (0.26)	0.003 (0.06)
CFO	-0.226 (-0.67)	-0.092 (-0.30)	-0.085 (-0.58)	0.348 (0.98)	0.454 (1.35)	-0.089 (-0.55)
PPE	0.375 (1.29)	0.130 (0.50)	0.271** (2.13)	-0.825*** (-2.69)	-0.609** (-2.11)	-0.078 (-0.56)
Top1	-0.157 (-0.45)	0.006 (0.02)	-0.260* (-1.70)	0.270 (0.74)	0.204 (0.59)	-0.016 (-0.09)
Dual	0.135** (2.21)	0.101* (1.84)	0.048* (1.78)	0.211*** (3.28)	0.131** (2.16)	0.047 (1.62)
Bsize	0.002 (0.07)	-0.034 (-1.50)	0.019* (1.70)	0.022 (0.83)	-0.008 (-0.30)	0.017 (1.43)
Indep	-1.079* (-1.70)	-1.409** (-2.47)	0.015 (0.06)	-0.600 (-0.90)	-0.660 (-1.05)	-0.263 (-0.87)
Constant	12.237*** (6.85)	15.589*** (9.71)	-2.100*** (-2.69)	-9.140*** (-4.86)	-5.955*** (-3.36)	0.346 (0.41)
Year	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes	Yes	Yes
N	15464	15464	15464	15464	15464	15464
R ²	0.103	0.098	0.043	0.101	0.093	0.022

Note: ***, **, and * represent significant at the level of 1%, 5%, and 10%, respectively, and are reported as *T* values in brackets, the same as in the following table.

the hypothesis H3a. Furthermore, TPA has a significant positive relationship to exploratory innovation (RD1, Δ RD1), while TPA is not significantly related to developmental innovation (RD2, Δ RD2), suggesting that the participation in targeted poverty alleviation can promote exploratory development investment by alleviating financing constraints, but has little impact on developmental innovation investment. The results also confirmed H3b.

4.4. Endogenous Test. There may be an endogenous problem of reverse causality between targeted poverty alleviation and innovation poverty alleviation, i.e., the innovation ability of listed companies. Listed companies with the innovation abil-

ity may be more competitive, have more ability and resources to undertake social responsibility, or have strong incentives to obtain the required resources for innovation activities. These companies can use targeted poverty alleviation to release positive signals from the capital market and build a positive image of the companies. To alleviate the endogenous problem of reverse causality, the proportion of listed companies participating in targeted poverty alleviation in the same industry and region (IndReg_TPA) is adopted as an instrumental variable to conduct an endogenous test in accordance with the study of Deng et al. [5]. There are significant copies following effects and geographical differences in companies' social responsibility assuming behaviors [38].

TABLE 5: Test of mediating effect based on financing constraints.

Variable	(1) KZ	(2) RD	(3) RD1	(4) RD2	(5) Δ RD	(6) Δ RD1	(7) Δ RD2
TPA	-0.953*** (-2.62)	0.136** (2.08)	0.156*** (2.64)	0.004 (0.15)	0.208*** (3.02)	0.154** (2.36)	0.044 (1.43)
KZ		-0.059*** (-3.15)	-0.068*** (-4.06)	0.021*** (2.59)	-0.079*** (-4.02)	-0.063*** (-3.39)	-0.011 (-1.20)
Size	-2.946*** (-9.75)	0.010 (0.17)	-0.167*** (-3.33)	0.183*** (7.53)	0.319*** (5.44)	0.273*** (4.95)	0.022 (0.84)
Age	2.462 (1.17)	-2.083*** (-5.47)	-1.341*** (-3.92)	-0.299* (-1.80)	0.395 (0.99)	0.605 (1.60)	-0.322* (-1.78)
ROA	-7.802*** (-2.98)	-4.706*** (-9.95)	-4.639*** (-10.92)	0.284 (1.37)	-3.341*** (-6.71)	-3.890*** (-8.30)	1.040*** (4.62)
Growth	-1.459*** (-5.42)	-0.841*** (-16.96)	-0.648*** (-14.54)	-0.153*** (-7.05)	-1.637*** (-31.39)	-1.358*** (-27.62)	-0.236*** (-9.98)
Lev	9.361*** (8.18)	-2.194*** (-9.63)	-1.917*** (-9.36)	-0.372*** (-3.73)	-0.394 (-1.64)	-0.495** (-2.19)	0.018 (0.17)
Pay	-4.510*** (-7.53)	-0.571*** (-5.31)	-0.428*** (-4.43)	-0.154*** (-3.26)	0.147 (1.30)	0.025 (0.24)	0.003 (0.05)
CFO	-22.834*** (-12.19)	-1.036** (-2.44)	-1.030*** (-2.70)	0.206 (1.11)	-0.742* (-1.66)	-0.411 (-0.98)	-0.235 (-1.16)
PPE	20.016*** (12.37)	0.509* (1.73)	0.286 (1.08)	0.222* (1.73)	-0.644** (-2.08)	-0.466 (-1.60)	-0.054 (-0.38)
Top1	0.005 (0.00)	-0.170 (-0.49)	-0.009 (-0.03)	-0.255* (-1.67)	0.252 (0.69)	0.190 (0.55)	-0.018 (-0.11)
Dual	-0.526 (-1.55)	0.132** (2.17)	0.098* (1.79)	0.048* (1.81)	0.208*** (3.24)	0.128** (2.12)	0.047 (1.60)
Bsize	0.048 (0.34)	0.004 (0.17)	-0.031 (-1.37)	0.018 (1.62)	0.026 (0.95)	-0.005 (-0.20)	0.018 (1.47)
Indep	-1.165 (-0.33)	-1.018 (-1.60)	-1.339** (-2.34)	-0.007 (-0.02)	-0.518 (-0.77)	-0.595 (-0.94)	-0.252 (-0.83)
Constant	45.084*** (4.55)	12.914*** (7.18)	16.373*** (10.13)	-2.344*** (-2.98)	-8.229*** (-4.35)	-5.232*** (-2.93)	0.469 (0.55)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	15,464	15,464	15,464	15,464	15,464	15,464	15,464
R ²	0.112	0.104	0.099	0.044	0.102	0.094	0.022

Whether a company participates in targeted poverty alleviation may be affected by enterprises in the same industry and region, while the participation of enterprises in targeted poverty alleviation in the same industry and region does not directly affect the company's innovation investment. Accordingly, the two-stage instrumental variable method is adopted in this study to retest the main research questions, and the regression results are listed in Table 6.

The first-stage regression results suggest that the regression coefficient of IndReg_TPA is significantly positive, and the partial F weak instrumental variable test can reject the

original hypothesis, consistent with the theoretical expectation. At the second stage, after the endogenous problems are controlled, TPA still shows a significant positive relationship to RD and Δ RD, thus again verifying the basic conclusion of this study. Thus, it is indicated that the promotion effect of the participation in targeted poverty alleviation on enterprise innovation is not affected by reverse causality problems.

4.5. Robustness Test. The positive relationship between targeted poverty alleviation and innovation input may be the

TABLE 6: Endogeneity test.

Variable	The first stage	The second stage	
	(1) TPA	(2) RD	(3) Δ RD
IndReg_TPA	0.911*** (39.81)		
TPA		1.807** (2.03)	0.211** (1.96)
Size	0.053*** (19.87)	-0.180*** (-4.75)	0.026 (1.56)
Age	0.039*** (4.74)	-1.016*** (-10.03)	0.007 (0.15)
ROA	0.379*** (6.66)	-7.003*** (-9.81)	-2.168*** (-3.77)
Growth	-0.010 (-1.51)	-0.531*** (-6.31)	-1.398*** (-16.14)
Lev	-0.063*** (-3.62)	-4.915*** (-22.79)	-0.293** (-2.24)
Pay	-0.040*** (-3.51)	0.223 (1.59)	0.231** (2.55)
CFO	0.038 (0.87)	1.452*** (2.71)	0.099 (0.37)
PPE	0.110*** (5.38)	-3.614*** (-14.25)	-0.741*** (-5.92)
Top1	0.050*** (2.79)	-2.239*** (-10.21)	-0.087 (-0.99)
Dual	-0.001 (-0.15)	0.505*** (7.51)	0.109*** (3.43)
Bsize	0.005** (2.57)	0.018 (0.76)	0.008 (0.83)
Indep	0.026 (0.47)	3.438*** (5.11)	0.052 (0.17)
Constant	-1.483*** (-13.56)	12.794*** (13.36)	-0.198 (-0.51)
Year	Yes	Yes	Yes
Industry	Yes	Yes	Yes
Firm	Yes	Yes	Yes
N	15,464	15,464	15,464
R ²	0.255	0.340	0.086
Partial F-test of IVs	439.117 [0.000]		

result of mechanical correlation rather than causality. Compared with nonparticipating enterprises, enterprises participating in targeted poverty alleviation may show significant advantages in terms of resource base, business capability, and corporate governance, while having better innovation foundation and stronger innovation capability. As a result, the conclusions of this study are affected by the endogenous influence of the omission of important explanatory variables.

To eliminate the individual differences between the two types of enterprises, this study uses the propensity score matching (PSM) method to match the samples that have participated in targeted poverty alleviation according to the nearest neighbor 1:1. The regression results after matching are listed in Table 7. The mediating effect of targeted poverty alleviation on enterprise innovation investment and financing constraint still exists significantly, suggesting that the conclusion of this study is not affected by the omission of important explanatory variables and remains robust after eliminating individual differences of enterprises.

Since there may be measurement bias in the research conclusions of this study, this study builds a new financing constraint measurement index (DKZ) by referring to the research of Wang [39]. Moreover, the SA index is adopted as a substitute variable of financing constraints for the robustness test in accordance with the study by Wu and Huang [40]. Columns 1–4 of Table 8 present the regression results of alternative financing constraint measures. Since the information disclosure of targeted poverty alleviation began in 2016, the test of samples from 2016 to 2019 can be adopted to directly compare the difference in innovation input between enterprises participating in targeted poverty alleviation and those that do not. Regression results of replacement samples are shown in columns 5–7 of Table 8. Robustness test results also support the hypothesis of this study.

5. Research Conclusions and Implications

The participation in targeted poverty alleviation is the main performance of enterprises in rolling out national poverty alleviation policies, feeding back poverty, actively assuming social responsibilities, and realizing corporate value regression. It is also a strategy for enterprises to obtain social and economic resources and improve social identity. In the context of “mass entrepreneurship and innovation” and targeted poverty alleviation, the influence of the participation in targeted poverty alleviation on enterprise innovation is discussed with China’s Shenzhen and Shanghai A-share listed companies from 2016 to 2019 as the research sample. The KZ index is constructed to measure financing constraints, and it is suggested that the promotion of targeted poverty alleviation on the R&D investment intensity of enterprises is through the channel of alleviating financing constraints.

This study suggests that first enterprises’ participation in targeted poverty alleviation can increase the efficiency of social resource allocation, achieve Pareto optimization, reduce the financing constraints of enterprises, and thus increase the investment in innovation of enterprises and especially enhance exploratory innovation ability of the enterprise. Compared with development innovation, exploratory innovation is characterized by greater uncertainty, greater exploitation, and higher returns. It requires acquiring, integrating, digesting, and absorbing new knowledge, new resources, and new technologies, and it is more dependent on the integration of strategic resources. Thus, the participation in targeted poverty alleviation more significantly

TABLE 7: Robustness test of model transformation.

Variable	(1) RD	(2) RD1	(3) RD2	(4) RD	(5) RD1	(6) RD2
TPA	0.166*** (2.58)	0.147*** (2.70)	0.029 (0.99)	0.165** (2.57)	0.146*** (2.68)	0.029 (1.00)
KZ				-0.065*** (-2.67)	-0.081*** (-3.93)	0.023** (2.04)
Size	0.036 (0.50)	0.000 (0.00)	0.126*** (3.81)	-0.000 (-0.00)	-0.045 (-0.73)	0.138*** (4.12)
Age	-1.438*** (-2.83)	-0.992** (-2.31)	-0.087 (-0.37)	-1.459*** (-2.88)	-1.018** (-2.37)	-0.080 (-0.34)
ROA	-4.818*** (-6.97)	-4.464*** (-7.62)	0.055 (0.17)	-5.015*** (-7.21)	-4.709*** (-8.00)	0.124 (0.39)
Growth	-0.638*** (-10.31)	-0.477*** (-9.10)	-0.141*** (-4.95)	-0.680*** (-10.66)	-0.529*** (-9.79)	-0.126*** (-4.30)
Lev	-2.871*** (-10.34)	-2.803*** (-11.91)	-0.284** (-2.22)	-2.537*** (-8.33)	-2.386*** (-9.26)	-0.401*** (-2.87)
Pay	-0.640*** (-4.05)	-0.467*** (-3.49)	-0.267*** (-3.69)	-0.648*** (-4.11)	-0.477*** (-3.57)	-0.264*** (-3.65)
CFO	-0.073 (-0.17)	0.161 (0.43)	-0.182 (-0.90)	-0.954* (-1.73)	-0.937** (-2.01)	0.127 (0.50)
PPE	1.143*** (3.04)	0.604* (1.90)	0.470*** (2.72)	1.284*** (3.38)	0.780** (2.43)	0.420** (2.41)
Top1	0.432 (1.00)	0.588 (1.60)	-0.150 (-0.75)	0.415 (0.96)	0.567 (1.55)	-0.144 (-0.72)
Dual	0.074 (0.91)	0.070 (1.01)	0.018 (0.49)	0.072 (0.88)	0.067 (0.97)	0.019 (0.51)
Bsize	-0.047 (-1.56)	-0.076*** (-3.00)	0.024* (1.77)	-0.044 (-1.46)	-0.073*** (-2.85)	0.023* (1.69)
Indep	-1.379* (-1.80)	-1.706*** (-2.63)	0.076 (0.21)	-1.305* (-1.70)	-1.614** (-2.49)	0.050 (0.14)
Constant	10.032*** (4.25)	9.310*** (4.66)	-2.210** (-2.04)	10.743*** (4.53)	10.197*** (5.07)	-2.460** (-2.26)
Year	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes	Yes	Yes
N	8500	8500	8500	8500	8500	8500
R ²	0.110	0.112	0.043	0.111	0.115	0.044

affects exploratory innovation. Second, the participation in targeted poverty alleviation can reduce financing constraints of innovation, improve resource allocation efficiency, and stimulate innovation vitality of enterprises mainly through resource effect, reputation effect, and information effect. Accordingly, the participation in targeted poverty alleviation more significantly affects exploratory innovation. The above analysis suggests that enterprises' participation in targeted poverty alleviation is not a simple social welfare payment but a mutually beneficial behavior that facilitates high-quality development of enterprises while assuming their social responsibilities.

Based on the above conclusions, this study draws the following points of enlightenment:

First, from the perspective of the Chinese government, the goals of the Party and the country consist of targeted poverty alleviation, building a moderately prosperous society in all respects, and fulfilling the "Chinese Dream" of the great rejuvenation of the Chinese nation, which requires the joint efforts of the people of all ethnic groups in China. Enterprises as the backbone of the development of high-quality fuel economy in the market, technology, information organization, capital, human resources, and management show a unique advantage and a strong ability of resource

TABLE 8: Robustness test of substitution variables.

Variable	Substitution variables				Replace the sample		
	(1) DKZ	(2) SA	(3) RD	(4) RD	(5) Δ RD	(6) Δ RD1	(7) Δ RD2
TPA	-0.162** (-2.00)	-0.019*** (-14.61)	0.139** (2.12)	0.140** (2.11)	0.205** (2.13)	0.158* (1.84)	0.027 (0.75)
DKZ			-0.026** (-2.46)				
SA				-0.098** (-2.21)			
Size	-0.420*** (-14.22)	0.019*** (17.90)	0.048 (0.88)	0.044 (0.80)	0.628*** (5.61)	0.552*** (5.50)	0.054 (1.26)
Age	-0.034 (-0.39)	0.007 (0.97)	-2.086*** (-5.48)	-2.088*** (-5.48)	0.974 (0.95)	0.039 (0.04)	1.014*** (2.61)
ROA	-17.186*** (-18.09)	0.056*** (6.11)	-4.549*** (-9.55)	-4.586*** (-9.71)	-2.843*** (-4.06)	-4.170*** (-6.64)	1.279*** (4.80)
Growth	-0.328*** (-4.29)	0.004*** (3.83)	-0.807*** (-16.60)	-0.809*** (-16.66)	-1.671*** (-23.62)	-1.450*** (-22.89)	-0.190*** (-7.05)
Lev	7.497*** (35.16)	0.021*** (5.17)	-2.521*** (-11.93)	-2.502*** (-12.11)	-0.746** (-2.03)	-0.944*** (-2.88)	0.188 (1.35)
Pay	0.616** (2.48)	-0.018*** (-8.62)	-0.571*** (-5.30)	-0.567*** (-5.25)	-0.332* (-1.77)	-0.256 (-1.52)	-0.040 (-0.56)
CFO	-14.693*** (-28.08)	0.016** (2.48)	-0.195 (-0.57)	-0.228 (-0.67)	0.651 (1.31)	0.535 (1.21)	-0.027 (-0.14)
PPE	11.579*** (44.07)	-0.024*** (-4.20)	0.353 (1.20)	0.377 (1.29)	-0.205 (-0.39)	-0.374 (-0.79)	0.195 (0.97)
Top1	-0.202 (-1.06)	-0.083*** (-12.19)	-0.157 (-0.45)	-0.149 (-0.42)	0.790 (1.10)	0.704 (1.10)	0.287 (1.06)
Dual	-0.073 (-1.22)	-0.002 (-1.51)	0.135** (2.21)	0.135** (2.21)	0.231** (2.34)	0.153* (1.73)	0.019 (0.49)
Bsize	-0.044** (-2.13)	0.002*** (3.09)	0.002 (0.07)	0.002 (0.06)	0.025 (0.56)	-0.001 (-0.03)	0.035** (2.05)
Indep	0.045 (0.08)	0.006 (0.50)	-1.080* (-1.70)	-1.079* (-1.70)	0.648 (0.59)	0.189 (0.19)	0.413 (1.00)
Constant	4.455*** (5.91)	3.437*** (98.29)	12.189*** (6.81)	11.899*** (4.99)	-17.720*** (-4.37)	-11.960*** (-3.29)	-4.820*** (-3.13)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	15,462	15,464	15,464	15,464	9725	9725	9725
R ²	0.490	0.869	0.103	0.103	0.115	0.122	0.019

integration, which can be consistent with the specific environment in poor areas, adjust measures to local conditions, and vary from person to person to create production and employment opportunities for poor people. Thus, to guide enterprises to actively participate in targeted poverty alleviation, the government should give targeted support to enterprises participating in industrial poverty alleviation. First, the position of enterprises in targeted poverty alleviation

should be clarified, and the huge role of enterprises in targeted poverty alleviation should be confirmed. Second, the government should fully exploit the advantages of the platform, formulate preferential policies for enterprise poverty alleviation, provide more financial support for poverty alleviation enterprises, do a good job in supporting services for poverty alleviation enterprises, and call on enterprises to think about the source of wealth, think about progress when

becoming rich, and assume social responsibilities on their initiative. Lastly, an enterprise incentive mechanism for poverty alleviation should be established, to give enterprise resource, technology, and human help, and to promote the integration of enterprises and take advantage of all aspects of the resources and scientific strengths for the reasonable selection and implementation of poverty alleviation projects, stimulating entrepreneurs as a crucial contributor to targeted poverty alleviation of practitioners and the leader of the social fashion, so as to win the battle against poverty for the enthusiasm of enterprises to mobilize the great potential of targeted poverty alleviation.

Second, enterprises, an important force in the capital market, should actively participate in targeted poverty alleviation projects and fulfill their social responsibilities under the call of the state. Targeted poverty alleviation exhibits a unique policy nature. Enterprises' participation in targeted poverty alleviation can obtain corresponding preferential and favorable development conditions, while improving their social image, establishing and maintaining ties with the government, and conveying confidence to the outside world. It needs multilateral cooperation and complementary advantages to win the battle against poverty. Enterprises should be aware of their important role in people's livelihood projects, and their participation in targeted poverty alleviation will bring economic and social benefits to them. Therefore, enterprises should seize the opportunity and use their own advantages. To be specific, they should set up poverty alleviation projects in poor areas, develop resources, cultivate industries, and revitalize local natural and human resources. Enterprises should optimize their own industrial chain, strengthen the construction of enterprise brands, enhance the visibility and competitiveness of enterprise products, and achieve sustainable development of enterprises and society.

Third, the financing environment of enterprises should be improved in depth. The trend of China's economic development has been very good over the past few years, and the mixed system reform is also in full swing. However, the state-owned economy still takes on a critical significance, and the development space for private enterprises remains insufficient. As revealed by the existing data statistics and research, private enterprises have a higher level of innovation investment and innovation power, thus contributing to the improvement of the innovation ability of the whole society, and the performance of social responsibility by private enterprises is more conducive to enterprise innovation than state-owned enterprises. However, affected by the policy system and financing capacity, the private economy and development lag behind the state-owned economy, which indirectly has an adverse effect on the technological innovation and development of the whole society. To vigorously facilitate R&D innovation in China and fully use the positive role of enterprises, it is particularly important to improve the financing environment of enterprises, not only the operation of enterprises but also the competitiveness and long-term development. To solve the above problem, the assessment and supervision of financial institutions should be improved, a multilevel capital market should be developed, and a favorable environment should be created.

This study also has some limitations. Impacted by the single form of data disclosure on R&D innovation, only quantitative indicators of R&D investment and patent volume, and the short observation year of targeted poverty alleviation data, this study only selects R&D investment to measure enterprise innovation in a single manner, which is insufficiently comprehensive to reflect the reality. It is expected that future research can have increasingly appropriate ways to measure innovation. Moreover, this study only studies the influence of the participation in targeted poverty alleviation on enterprise innovation and does not involve other aspects of corporate governance and enterprise development. Furthermore, a question is raised whether there is any way other than financing constraints for the influence of the participation in targeted poverty alleviation on enterprise investment activities, which will be investigated in depth 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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this study.

Acknowledgments

This work was funded by the Humanities and Social Science Fund of Ministry of Education of China: Financialization of Non-Financial Enterprises: Research on Economic Consequences and Factors (No.: 19yja630077).

References

- [1] D. Shifeng, S. Henggui, and Z. Yiqun, "Research on the impact factors of targeted poverty alleviation of Chinese listed companies – based on the perspective of social responsibility," *Fiscal Research*, vol. 2, pp. 104–115, 2019.
- [2] W. Jiabin, C. Jin, and L. Zhiyu, "Precision poverty alleviation, policy benefits to enterprises and enterprise value," *Modern Finance & Economics*, vol. 390, no. 7, pp. 74–93, 2022.
- [3] S. Begum, M. Ashfaq, E. Xia, U. Awan, and R. Welford, "Does green transformational leadership lead to green innovation? The role of green thinking and creative process engagement," *Business Strategy and the Environment*, vol. 31, no. 1, pp. 580–597, 2022.
- [4] M. T. Azim, L. Fan, M. A. Uddin, M. Jilani, and S. Begum, "Linking transformational leadership with employees' engagement in the creative process," *Management Research Review*, vol. 42, no. 7, pp. 837–858, 2019.
- [5] B. Deng, T. Cunjie, and J. Li, "Firms' targeted poverty alleviation and the ease of financial constraints," *Chinese Journal Of Finance And Economics*, vol. 46, no. 12, pp. 138–151, 2020.
- [6] K. V. Lins, H. Servaes, and A. Tamayo, "Social capital, trust, and firm performance: the value of corporate social

Retraction

Retracted: Mathematical Modeling and Simulation of Online Teaching Effect Evaluation Based on Decision Tree Algorithm

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Sun, "Mathematical Modeling and Simulation of Online Teaching Effect Evaluation Based on Decision Tree Algorithm," *Journal of Function Spaces*, vol. 2022, Article ID 7425196, 11 pages, 2022.

Research Article

Mathematical Modeling and Simulation of Online Teaching Effect Evaluation Based on Decision Tree Algorithm

Pingli Sun 

Department of Education and Teaching, Zhengzhou Preschool Education College, Zhengzhou, Henan 450000, China

Correspondence should be addressed to Pingli Sun; sunpingli@zzpec.edu.cn

Received 27 May 2022; Revised 18 July 2022; Accepted 25 July 2022; Published 8 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Pingli 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 support of big data technology, the field of education is also facing new problems and opportunities. Network teaching has become the mainstream means of higher education. In order to explore the changes of students' learning effect in the process of online teaching, this paper proposes to build an online teaching effect evaluation model with the support of data mining technology and decision tree algorithm. This paper records the factors and objects that reflect the teaching effect in network teaching and traditional teaching, respectively. A decision tree algorithm is used to divide the attributes of influencing factors from relevant rules. Using the Kirschner model to build the evaluation system, add two attribute elements: students' teaching evaluation and teachers' self-evaluation. Data mining technology is used to preprocess and clean up the sample set, which improves the accuracy of the calculation results. In the evaluation model, the association rule algorithm is also constructed to classify the data of the same element type and delete the data of different elements after marking. Through this evaluation model, teachers can accurately judge students' learning interests and improve students' academic performance. The results show that compared with the traditional data mining algorithm, the decision tree algorithm has obvious advantages in computing speed and accuracy.

1. Introduction

There is an obvious gap between online teaching and traditional teaching methods. It has the characteristics of diversity, flexibility, and interaction [1]. According to different disciplines, choosing various forms of teaching can achieve the purpose of solving students' after-school problems anytime and anywhere. Teachers' choice of teaching methods also needs to refer to teaching needs, teaching objectives, online platform stability, and other factors [2]. Compared with traditional classroom, online classroom also includes check-in, in-class test, live broadcast time, homework release, and other forms. The main elements of network teaching mode are teachers, students, teaching content, teaching media, supporting platform for teaching and learning, etc. Teachers choose the teaching content according to the requirements of the syllabus, the characteristics of students, and the cognitive basis of students. Then, design and compile into multimedia teaching software and network courseware or design and develop into network courses and

publish them to the course center [3]. The corresponding data generated during students' online learning can also be collected through the teaching platform. How teachers use behavior data to form accurate and effective effect evaluation is the main content of our exploration [4]. At present, online teaching modes include Tencent conference, Mu class live broadcast, and nail class; this kind of video and audio live broadcast can enable students to acquire and share knowledge in the family environment. Due to the full implementation of the network education model in colleges and universities, teaching resources also have richer choices [5]. The popularization and development of online teaching mode is one of the ways for colleges and universities to display modern educational technology means. It also brings new ideas for the improvement of classroom environment and teaching quality management. Teaching effect evaluation is one of the links of teaching quality management and supervision, which can ensure the effectiveness of teaching methods [6]. Whether the teaching effect is perfect needs to be further reflected in the evaluation indicators [7].

How to evaluate and test the effect of online teaching objectively and fairly is the main concern of college educators all over the world [8]. The various attributes of evaluation indicators can directly affect teachers' planning of teaching mode and the way of carrying out teaching activities. Although a relatively complete teaching effect evaluation system has been established in the traditional teaching classroom, due to the unique form of online teaching, it is necessary to reconstruct an online teaching effect evaluation system [9]. According to the survey, the acquisition process of online teaching effect evaluation has the following problems: first, there are many online teaching modes, resulting in different judgment standards for teaching effect [10]. Many colleges and universities do not have a deep understanding of the online teaching mode and still use the traditional classroom evaluation indicators, resulting in the evaluation coefficient which cannot be objective and fair. Second, online teaching has distinct characteristics, and the evaluation of teaching effect must meet the elements of online teaching [11]. Third, the concept of teaching effect evaluation is vague, the evaluation angle is single, and the evaluation standard is single. There are still one-sided levels of teachers' evaluation views, ignoring the main function of students in teaching effect evaluation [12]. According to the above situation, this paper studies the online teaching effect by using the mathematical methods such as decision tree algorithm and Coriolis model in data mining technology, so as to realize the construction of teaching effect evaluation model.

2. Application Status of Data Mining Technology at Home and Abroad

With the development of educational modernization, the field of education has updated teaching aids and teaching means in combination with computer intelligence technology [13]. The development speed of artificial intelligence technology in the field of education will be significantly improved in the future. I believe that with the landing and application of 5g communication technology, the application scenarios of artificial intelligence technology will be gradually improved, and a large number of products based on artificial intelligence technology will be gradually implemented, such as the application of AR and VR products, which will open up a huge value space for artificial intelligence technology. People's understanding of the way of education has gradually changed. Modern information technologies such as Internet, big data, and artificial intelligence are changing people's living habits and learning styles [14]. The data generated in the teaching process of colleges and universities will be stored in the information database. These data can provide teachers with students' learning status and performance change trend [15]. However, how to quickly query information and calculate and count data sources is the primary problem perplexing educators. As a multidisciplinary research technology, data mining was officially proposed at the first academic conference in Canada [16]. Subsequently, this technology has been widely used in the fields of medicine, pedagogy, and economics and

achieved good results. They began to pay attention to the application of data mining technology in the field of e-learning. According to data mining technology, the real-time path of computer software in application is analyzed to mine system vulnerabilities and execution records [17].

Japan's modern industry needs to deal with massive data information in the big data environment [18]. In order to improve the retrieval speed and accuracy, they use data mining technology to improve the analysis quality in the economic field and make a scientific transformation of enterprise finance [19]. The application of this technology in financial management is highly feasible, and some research results have been obtained. The German construction team is characterized by complexity in its work. This kind of specialized work has great difficulty and error in practical operation [20]. In order to improve the quality of construction projects, they use big data mining technology to detect the construction quality. Data collection and pre-treatment shall be carried out in the project development stage, and a certain logical process shall be adopted to analyze whether the construction quality meets the standard. Finally, the purpose of optimizing building quality and building process structure is realized [21].

As a common method in data mining, a decision tree algorithm has been widely used in the field of education in China [22]. The decision tree algorithm has been used in student grade point analysis, behavior data analysis, and so on. Through the data feedback information, we can judge the enrollment proportion and predict the employment situation of students. The decision tree algorithm can also get the following results in the process of analyzing the changes of students' grades: students' age, gender, learning time, and discipline basic ability [23]. According to the main factors affecting students' performance, the corresponding performance prediction model is generated. Through the improvement from generation to generation, the decision tree algorithm can also realize the early warning processing of students' performance from the specific factors affecting students' learning behavior in the relationship rules and classification rules [24]. In the field of education, foreign countries have also taken a series of research on decision tree algorithm. They not only applied this algorithm to student performance monitoring but also analyzed the investment of education funds and student information management [25]. Obtain students' favorite subject data from the big data environment, and then cultivate students' learning habits. Using the association rules in the decision tree algorithm can also accurately judge the enrollment rate, enrollment rate, and transfer rate of students in different regions, so as to realize the allocation and planning of educational resources. Based on the above research status, this paper also uses the decision tree algorithm to build and study the model of online teaching effect evaluation.

3. Analysis of Evaluation Model Construction

3.1. Research on the Construction of Online Teaching Effect Simulation Model Based on Decision Tree Algorithm. In today's college education process, students' learning effect

and teachers' teaching effect have become important indicators affecting students' performance. One can directly reflect teachers' teaching level, and the other can objectively reflect students' learning ability. It has a guiding function for teachers to choose teaching methods and schools to plan teaching plans. At present, online education platforms emerge one after another, and online teaching has become a widely used way in teaching mode. With the support of data mining technology, this paper uses decision tree association rule algorithm to build its simulation model and form an evaluation system for online teaching. So as to realize the detection and improvement of students' learning effect. According to various influencing factors, this paper analyzes the change trend of students' performance and then puts forward reasonable targeted suggestions to improve the teaching quality. As a widely used classification algorithm in data mining technology, the decision tree algorithm can get the hidden information contained in the data through data analysis and preprocessing. These information can be divided into effective and ineffective functions. Further processing can summarize the factors affecting the calculation results. For the teaching effect, the use of data mining and classification is helpful to understand the shortcomings of the teaching model, improve their own advantages, and clarify the defects in the teaching methods. For college teachers, it can help teachers adjust teaching methods, change teaching plans, and formulate educational plans for students' learning effects. At the same time, it can also improve their own teaching level and optimize teachers' teaching skills in the teaching effect. The focus of higher education is to improve teaching quality and cultivate compound talents with high quality and strong comprehensive ability for the society. As an important basis to measure the quality of teaching, students' performance is also an important symbol to evaluate students' mastery of knowledge and learning attitude. Therefore, the prediction and analysis of students' scores can provide an important basis for teaching managers to deepen teaching reform, reasonably arrange teaching plans, and improve teaching quality. With the rapid growth of student performance data, it is difficult for teachers to find rules directly according to the distribution of student performance and make decisions according to this rule. The traditional analysis and treatment of students' performance by teachers are generally only to count the number of students whose performance is at the level of excellent, good, average, and poor. However, it is impossible to understand the reasons for the distribution of students' achievements. If teachers can fully understand the reasons for students' achievements, they will be able to better "suit the remedy to the case" and improve the teaching quality. For the massive data generated in the teaching process, in the past, only the primary data backup, query, and simple statistical stages were used, so these data were not fully utilized. Now, we can use data mining classification technology to transform a large amount of data into classification rules, so as to better analyze these data and obtain useful information.

Firstly, this paper refers to a large number of literature and analyzes the difference between traditional teaching effect and online teaching effect. It is found that the final

score and process detection of students' learning can reflect the effect of online teaching. Because the student achievement information belongs to discrete data, the classification rule algorithm in the decision tree algorithm is selected as the technical support. This paper analyzes the main factors affecting teaching effect and students' achievement from various angles. The whole research idea is shown in Figure 1.

It can be seen from Figure 1 that after determining the research object and method, it is also necessary to sort out and collect the data. We mainly collect objective factors such as online teaching platform, students' age, gender, teachers' teaching level, and teachers' rank. Among the research objects, a university was randomly selected for online teaching, and a questionnaire survey was conducted on students and teachers. A total of 8000 valid survey results were recovered. According to the survey results, there are obvious differences in teachers' choice of online platforms and teaching methods. We show the way of online teaching in terms of data and proportion, as shown in Figure 2.

As can be seen from Figure 2, more than half of college teachers have overlapping and mixed behaviors in choosing online teaching methods. More than 20% of teachers use the three teaching methods. Most teachers make their own teaching contents and resources, and a small number of teachers choose the teaching resources provided by the Mu class platform. According to the above survey results, we will calculate the collected data by decision tree algorithm to judge the impact of the above factors on the effect of online teaching. The decision tree algorithm can classify and sort the irregular events to form the corresponding mathematical model and use the classification rules to predict and analyze the hidden data. The decision tree algorithm can solve both classification problems (the corresponding target values are classified data) and regression problems (the output results can also be continuous values). Compared with other algorithms, decision tree has a very obvious advantage; that is, it can be visualized intuitively. Each leaf node under the tree structure is given a flag to record the tree conditions contained therein separately. This arrangement from top to bottom can achieve the purpose of classification after data preprocessing in practical research. Three classes of modern education technology are randomly selected as the representative objects of this paper. There are one professor and three lecturers among the teachers, and a total of 5423 data are generated. Establish a data summary table for the basic information of students from the perspectives of name, gender, student number, and major. In the judgment of teaching effect, the change of students' academic performance is selected as a reference. Compare the performance changes of students in the three classes, as shown in Figure 3.

As can be seen from Figure 3, the number of students in the first class is large, and the average academic performance is relatively uniform. The number of students in the third class is small, of which the number of male students accounts for two-thirds, and the academic performance is high. Before establishing the decision tree model, the data needs to be discretized and split. Discrete processing is for data induction to form a small range of characterization attributes. Remove the data irrelevant to the influencing

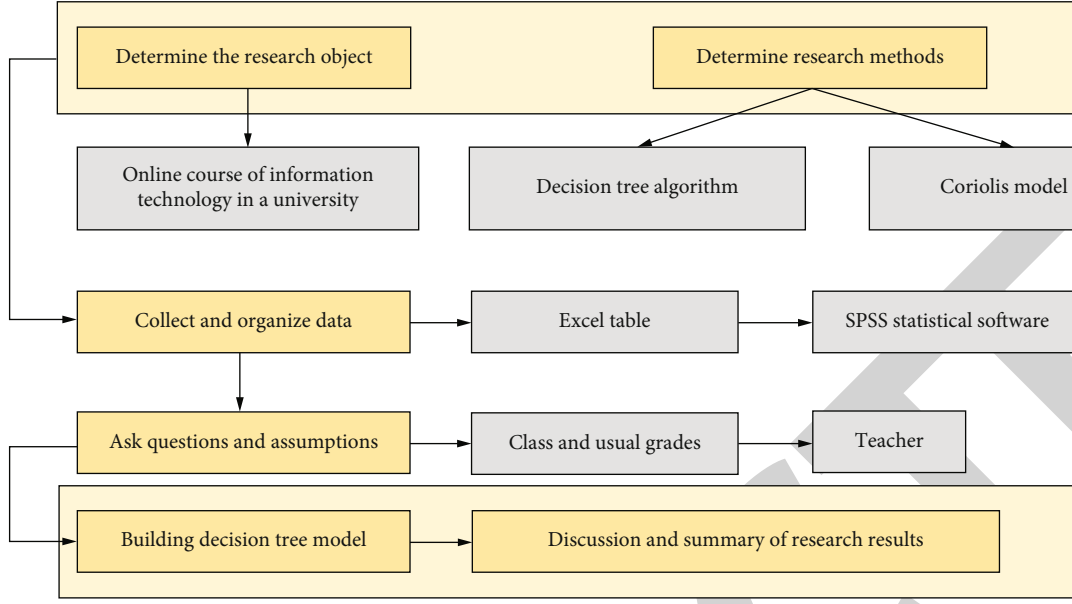


FIGURE 1: Flow chart of research ideas.

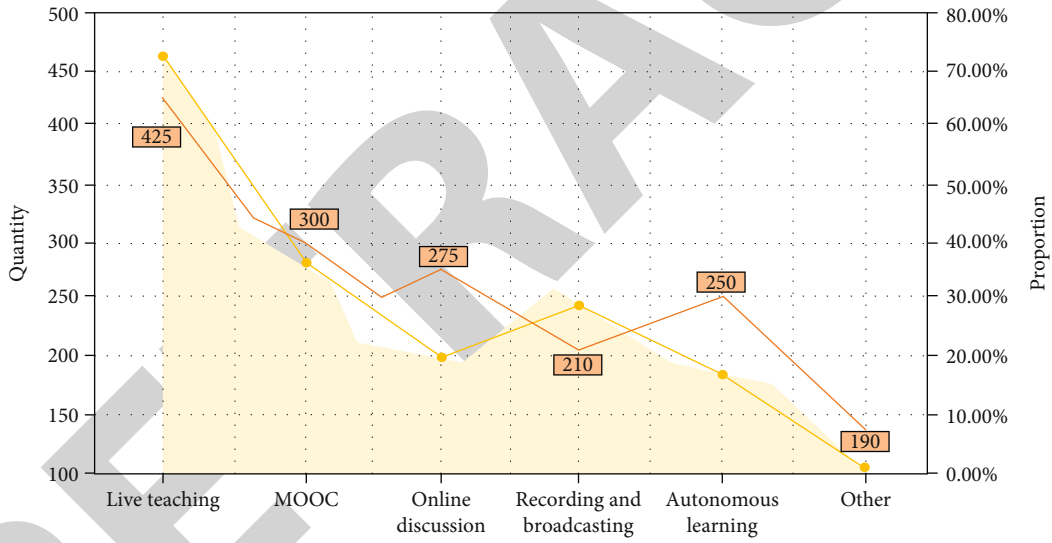


FIGURE 2: Display diagram of data change and proportion change of online teaching mode.

factors, and retain the effective information. Suppose a sample set has different representation values in training, which can be distinguished according to discrete values and ordinary values. The outputs corresponding to these calculation results can form multiple subsets. At this time, the data quotient measurement is calculated according to the subset information:

$$\text{InfoA}(D) = \sum_{j=1}^v \frac{|D_j|}{|D|} \times \text{Info}(D_j). \quad (1)$$

The weight changes in each partition can be divided according to the category. When the effective expected data in the tuple is smaller and smaller, the accuracy of the whole partition is higher. The definition of data quotient value

should also be combined with the difference of information demand:

$$\text{Gain}(A) = \text{Info}(D) - \text{InfoA}(D). \quad (2)$$

The classification data used in the decision tree algorithm is similar to the variable $\text{Info}(D)$, and its operation formula is

$$\text{SplitInfoA}(D) = - \sum_{j=1}^v \frac{|D_j|}{|D|} \times \log_2 \left(\frac{|D_j|}{|D|} \right). \quad (3)$$

The calculated data values are divided into corresponding attribute areas by the set. At the same time, the output

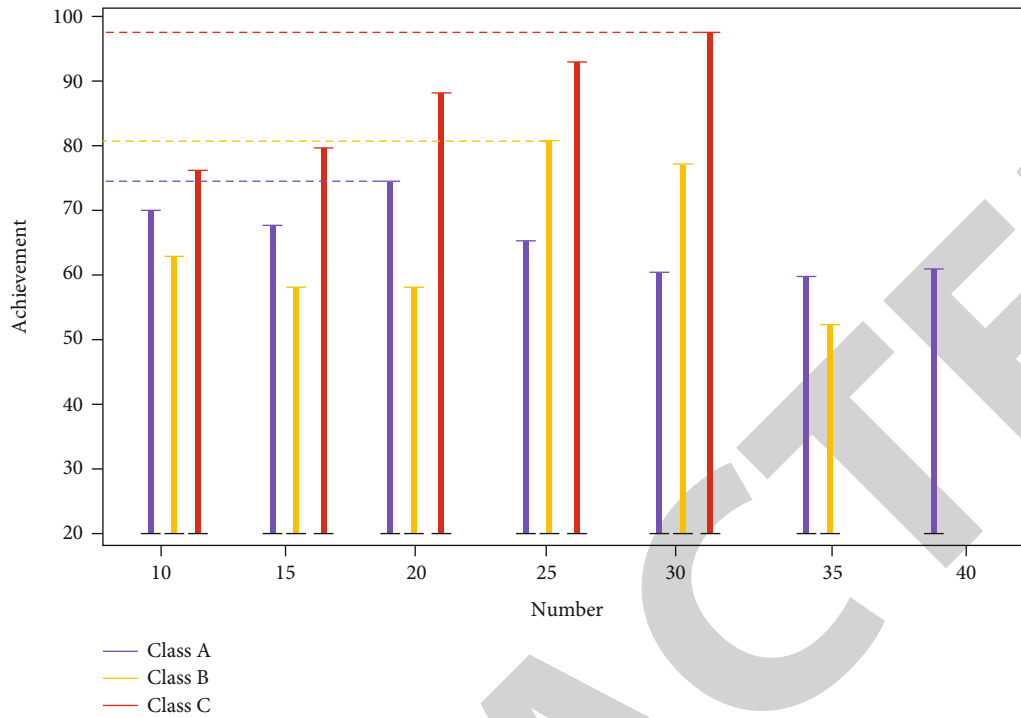


FIGURE 3: Student achievement change chart of three classes.

variables of the calculation results are compared. The whole comparison formula is as follows:

$$\text{Gainratio}(A) = \frac{\text{Gain}A}{\text{SplitInfo}A(D)}. \quad (4)$$

According to the above formula, students' usual test scores and final test scores are constructed. To determine whether the students are qualified or not, they need to be sorted into two grades, and the data quotient value is calculated by using the classification formula:

$$\text{Info}(D) = -\frac{534}{1096} \times \log_2 \frac{534}{1096} - \frac{562}{1096} \times \log_2 \frac{562}{1096} = 0.9995. \quad (5)$$

The above formula includes sample quantity and effective data quantity. Students' grades are divided into three grades: excellent, good, and average. Among the three grades, there are 527 excellent samples above the average coefficient, 58 good samples, and only 3 general samples. We define the usual score formula:

$$\begin{aligned} \text{info}(p) &= \frac{858}{1096} \times \left(-\frac{527}{858} \times \log_2 \frac{527}{858} - \frac{282}{858} \times \log_2 \frac{282}{858} \right) \\ &= 0.9623. \end{aligned} \quad (6)$$

Classify the data of usual performance:

$$\begin{aligned} \text{Split info}(p) &= -\left(\frac{858}{1096} \times \log_2 \frac{858}{1096} + \frac{282}{1096} \times \log_2 \frac{282}{1096} \right) \\ &= 0.8631. \end{aligned} \quad (7)$$

We take the test results of students' online teaching as test data and maximize the selection of each attribute according to the decision tree model. The test score is the root node of the decision tree model, so as to further determine each child node. The final decision tree model is as follows:

As can be seen from Figure 4, three branches are generated under the root node, which are the summary results of usual grades in online teaching. Excellent, good, and general are used as the data sources of the three branch nodes. Finally, the decision tree model needs to be further simplified. This pruning method adopts the operation of deleting after a long time. Eliminate the data without research value to make the model calculation more simple and convenient. From the above research, it is concluded that the online teaching effect has obvious advantages over the traditional teaching effect. With the increase of the number of students, the range of excellent students is gradually expanded, and ordinary students account for the vast majority of the sample set. Next, we need to build an evaluation model for the effect of online teaching.

3.2. Application Analysis of Network Teaching Effect Evaluation Model. With the more and more frequent

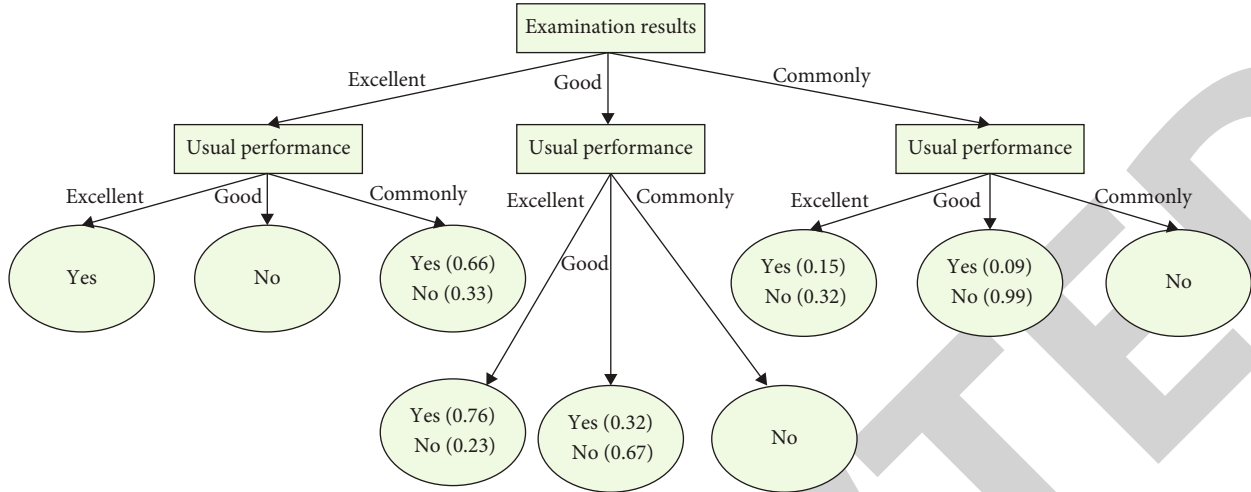


FIGURE 4: Decision tree model.

application of data mining technology in college teaching activities, we should rejudge the impact of teaching effect on students' learning activities in the face of online teaching mode. The new teaching mode brings new challenges to the evaluation of teaching effect. Teachers' analysis of teaching effect is the main means to ensure teaching quality and improve teaching efficiency. Whether the evaluation mechanism is reasonable and the evaluation index is accurate is one of the various factors affecting the effect evaluation. At present, most colleges and universities do not fully understand the online teaching methods, and the evaluation angle is relatively single, which is quite different from the actual teaching process. Data mining technology records the learning situation and learning trajectory of different users in the database. The sequential pattern mining method of data mining is used to classify documents. Improve the speed of students' information retrieval. Teachers can conduct mining and analysis based on the data accessed by students. Cluster analysis of the access data, understand the contents of students' interest and the relationship between access sites, and analyze students' access habits and interests. Actively push some interesting content to them, and set up hyperlinks between related pages to improve the structure of the website. Make the links between pages more in line with students' visiting habits. Facing the shortcomings of online teaching effect evaluation, this paper constructs and studies the online teaching effect evaluation index by using the decision tree algorithm and Coriolis model structure. The whole evaluation process is shown in Figure 5.

As can be seen from Figure 5, the evaluation system has two parts: process and summary. After many online teaching environment simulations, record the generated student data. Classify and summarize the data set to reflect the teaching effect and operation. It can reflect the actual learning effect of students from the aspects of class rate, learning time, and performance change. Traditional evaluation methods generally use numerical calculation, including equal difference calculation method, similar estimation method, and weighted average method. However, in the face of a large

amount of data generated by the online platform, the above methods cannot guarantee the accuracy of the calculation results. Before building the decision tree model, this paper needs to traverse the two variables of quotient value. The feature of numerical maximization is selected as the best tree node. After importing the data, conduct digital conversion to convert the characteristic information into dictionary form. The code formula is as follows:

$$\begin{aligned}
 x &= x.to_dict(\text{orient} = \text{"records"}), \\
 \text{transfer} &= \text{DictVectorizer}(), \\
 x &= \text{transfer.fit_transform}(x).
 \end{aligned} \tag{8}$$

When designing the decision tree model, first define the number of node layers:

$$\text{estimator} = \text{DecisionTreeClass} = \text{max_depth} = 3. \tag{9}$$

Call the compound function to the training model, input the characteristic variables and mark them as the best child nodes. Finally, it is combined with the output characteristics to form tree parameters:

$$\text{estimator.fit}(x, y). \tag{10}$$

There are no regular and repeated paths for the above operations, and users cannot find the regular characteristics of data processing. This can greatly ensure the security and effectiveness of data. It can be seen from the decision tree node that the data on the left side of the object represents the samples with high trust and the data on the right side represents the samples with low trust. Starting from the second level, the subnode with a value of indicates the state of low teaching evaluation. We apply the same data to the traditional data processing methods and compare the results from the accuracy, as shown in Figure 6:

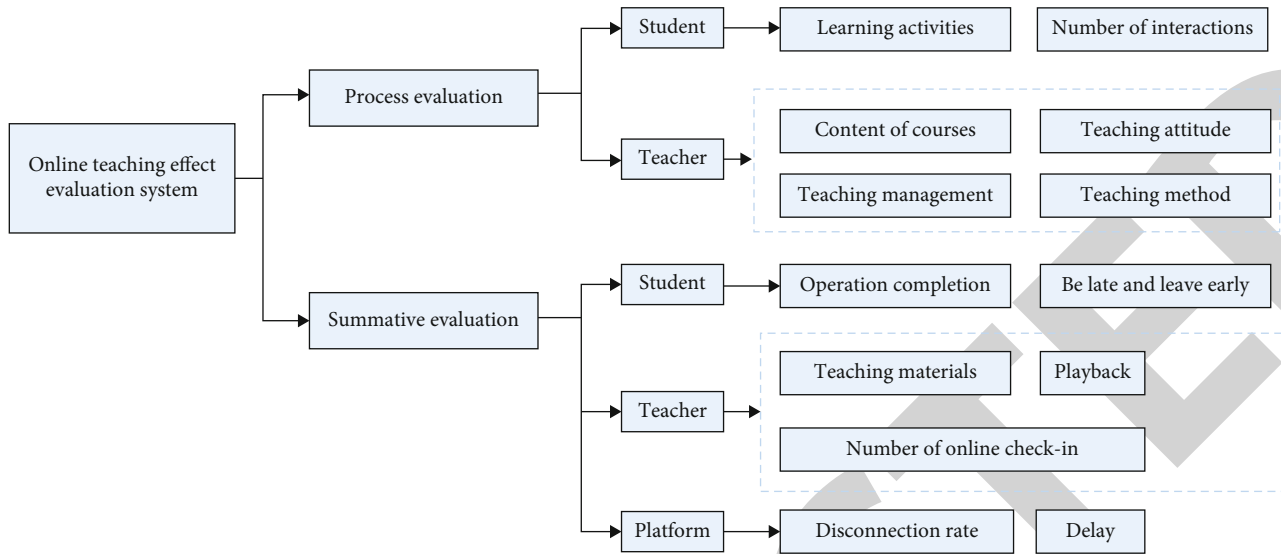


FIGURE 5: Flow chart of the whole evaluation.

It can be seen from Figure 6 that there is little difference between the traditional method and the decision tree algorithm in the initial stage. With the gradual increase of the number of samples, the accuracy of the traditional method shows a cliff decline, while the decision tree algorithm can still maintain high accuracy. Finally, we use Coriolis model to evaluate the prediction effect. This model is mainly composed of reaction group, behavior group, learning group, and result group. The response group includes students' feelings about online teaching, which is mainly used to collect students' feedback on the teaching mode. The behavior group is to judge the status and behavior characteristics of students' online learning, as well as the mastery of each subject's knowledge and skills. The learning group is a comprehensive collection of learning effects, which can represent the effect of online teaching to a certain extent. The final result group can judge whether students fully master knowledge and skills and whether they can use online course content to solve problems in daily life. Teachers need to strengthen interaction. Online teaching must focus on interaction. In the process of interaction, teachers can also have a general understanding of students' learning, so as to determine the next teaching plan. At the same time, pay attention to online assessment. Online assessment is often one of the most important ways for teachers to grasp students' learning. Investigate the student data, respectively, to form the teaching effect evaluation index, and use the analysis and calculation ability to obtain the proportion of each index. Assuming that the coefficient gain of each index can be determined, the attribute of maximum coefficient increment can be obtained by comparing their size in the overall structure. Assuming that the sample set has unknown values, define similar sets as expected information:

$$I(S_1, S_2, \dots, S_m) = - \sum_{i=1}^m p_i \log_2(p_i) \quad (11)$$

Among them, the similar set is $C(i = 1, 2, \dots, m)$. The probability formula of any sample belonging to the set is

$$p_i = \frac{s_i}{s}. \quad (12)$$

It is assumed that subsets with different values of attributes can be represented by similar variables. The data in S_j is divided into output variables by the total set:

$$E(A) = \sum_{j=1}^v \frac{(s_{1j} + s_{2j} + \dots + s_{mj})}{s \times I(s_{1j}, s_{2j}, \dots, s_{mj})}. \quad (13)$$

The smaller the calculated value, the higher the information purity in the representative model. This highly accurate subset data can reflect that the evaluation index of teaching effect is dynamic and effective. According to the calculation results, the online teaching mode is evaluated from the perspective of students. The whole student group has a high acceptance of online teaching methods, and the mastery of learning knowledge in the survey has also achieved the expected results. In order to more comprehensively evaluate the effect of online teaching, we need to supplement the evaluation and suggestions obtained from the perspective of teachers.

4. Research Results of Evaluation Model and Analysis of Influencing Factors

This paper randomly selects a university to conduct a simulation experiment on its online teaching courses. A large number of sample data were obtained in the study, including the main characteristics of students' gender, age, teacher education, and teacher age. In order to explore the effectiveness of online teaching effect evaluation model, we need to analyze the influencing factors from multiple angles. The collected student evaluation forms basically reflecting that

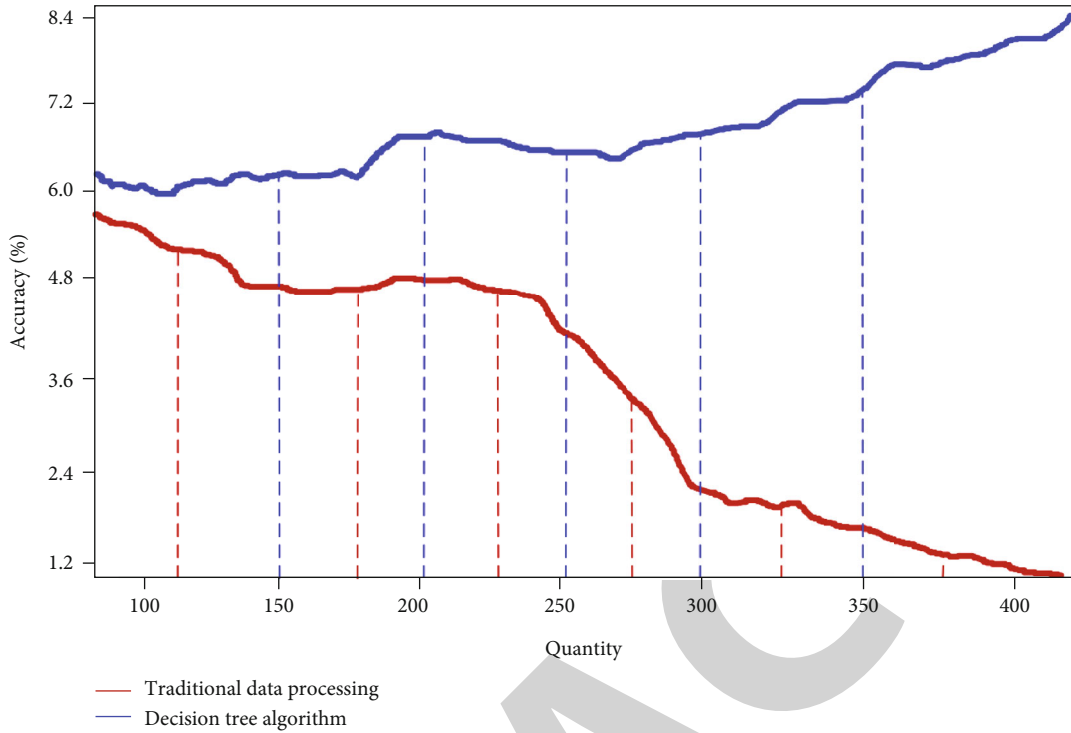


FIGURE 6: Accuracy comparison between traditional method and decision tree algorithm.

TABLE 1: Student evaluation information form.

Full name	Number of participants	Average	Total score	Teacher age	Teacher education	Teacher gender
Baizhuohua	197	4.03	795.6	24	Undergraduate	Male
Wanghai Feng	165	5.42	895.9	26	Undergraduate	Female
Wangqianqian	35	2.92	102.3	28	Master	Male
Dongdaming	95	1.64	156.3	30	Master	Female
Zhangguoqiang	78	7.20	562.2	35	Doctor	Male
Lihua	40	12.12	485.6	37	Doctor	Female

the teaching methods are generally welcomed. In the follow-up, we should improve the whole effect evaluation model in combination with teachers' teaching evaluation and suggestions. The decision tree algorithm is used to preprocess the sample data. In the face of incomplete data and invalid data, the reference index needs to be set in advance. This way of setting indicators can improve the quality of data processing and finally achieve the purpose of improving the evaluation quality. The data we selected is from the 2020 student evaluation information table, which is stored in the database. After sorting the information table, it is expressed as follows.

As can be seen from Table 1, each student's evaluation of teachers is summarized. The table also includes teachers' names, ages, professional titles, and grades. Cleaning the above data can fill in the missing attributes and eliminate useless information. In the summary table, some teachers' professional title grade attribute is empty, so we use cleaning to delete the record. Finally, the online teaching effect evaluation system is constructed. The main functional blocks and system pages are shown in Figure 7:

As can be seen from Figure 7, the system functions include query, management of student files, management of teacher files, class opening, course schedule viewing, after-school homework management, course evaluation, and other modules. And randomly display a student's evaluation page of the course. After the simulation forms the evaluation system, it is also necessary to analyze the factors affecting the teaching effect.

The discrete data attributes processed by decision tree algorithm include peacetime score, final score, and class. In addition to combining students' evaluation of teaching, we also make a statistical analysis of students' learning effect. The lowest score is 50, and the highest score is 100. We plan 15 points into one stage, and the number of middle school students in each stage is not fixed. From 50 to 65 points are the general stage, from 65 to 80 points are the good stage, and above 80 points are the excellent stage. Combined with the research object, the students' usual performance data includes two characteristics: continuous and discrete. In the calculation, due to the difference of data types, it is easy

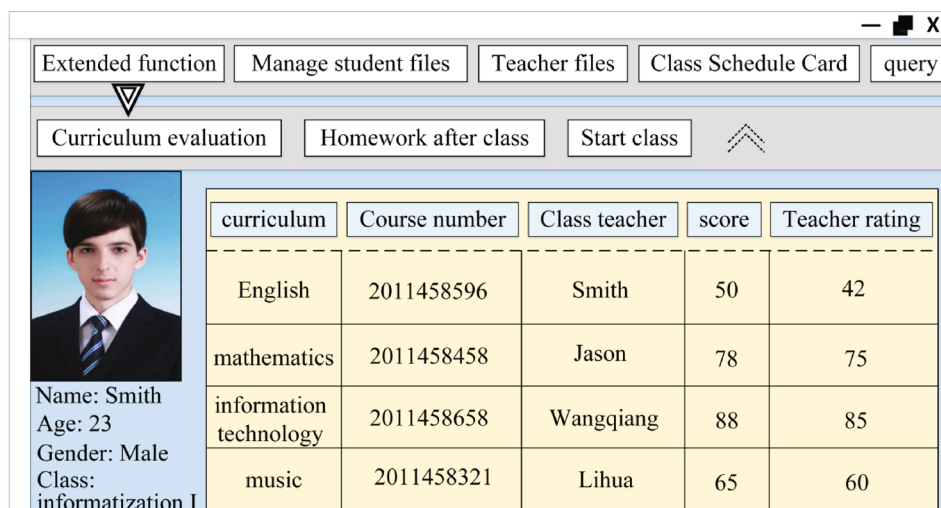


FIGURE 7: Main functional blocks and system page diagram.

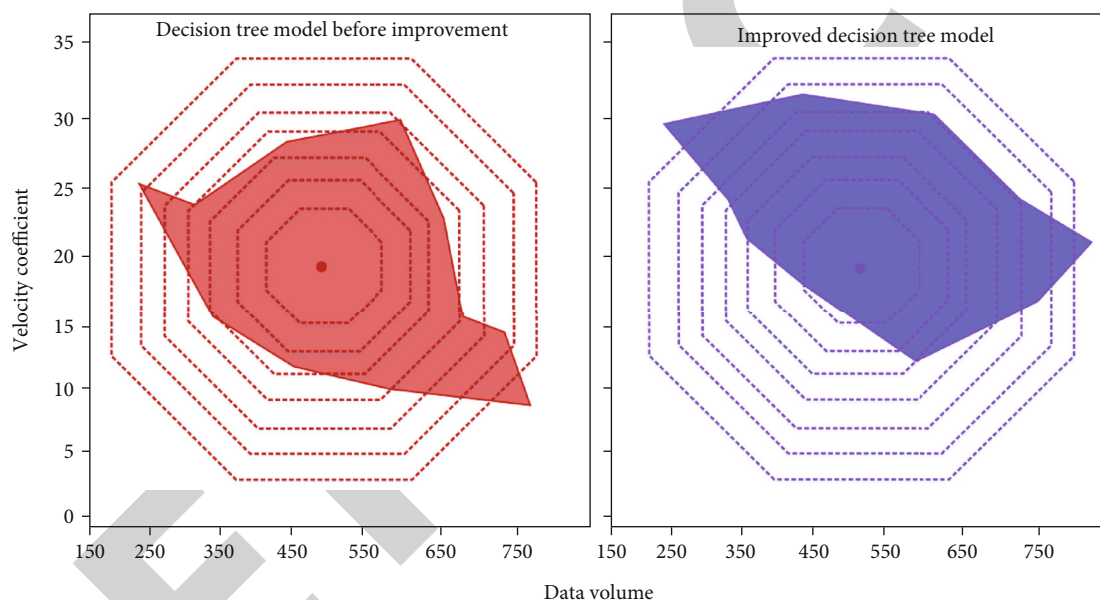


FIGURE 8: The change of speed range of decision tree model in data processing before and after improvement.

to take a higher value. Based on this situation, we choose to use the decision tree algorithm to improve the calculation. In the improved algorithm, the function of association rules is used to find out the object relationship in the data set, and the prediction model can be verified when using student data to predict the evaluation of teaching effect. Compare the speed range changes of the decision tree model before and after improvement in data processing, as shown in Figure 8.

As can be seen from Figure 8, the data set traversed by the correlation algorithm can eliminate the interference of miscellaneous data, and each predictable sequence can be automatically classified according to the set threshold in the search for simultaneous cases. Therefore, the improved decision tree algorithm improves the speed and processing range of the prediction model.

Finally, in the teaching evaluation and prediction system, we automatically generate each student's usual score report,

and the score index is displayed as a normal distribution according to the data uploaded in the above three stages. Among them, the length of learning, online teaching platform, and teachers' education are the main influencing factors. Students in three stages are investigated, as shown in Figure 9.

As can be seen from Figure 9, most of the students who have studied for a long time are in the excellent stage, and a few are in the general stage. When the online teaching platform is mu class network and learning link platform, students' learning effect is good. When the teacher's degree is a master's degree, the student's performance rises rapidly. According to the analysis of the above research results, the online teaching effect is finally presented as the usual score and the ultimate test score. The professional grade, educational background, age, and other factors of subject teachers have the second largest impact on students' learning effect.

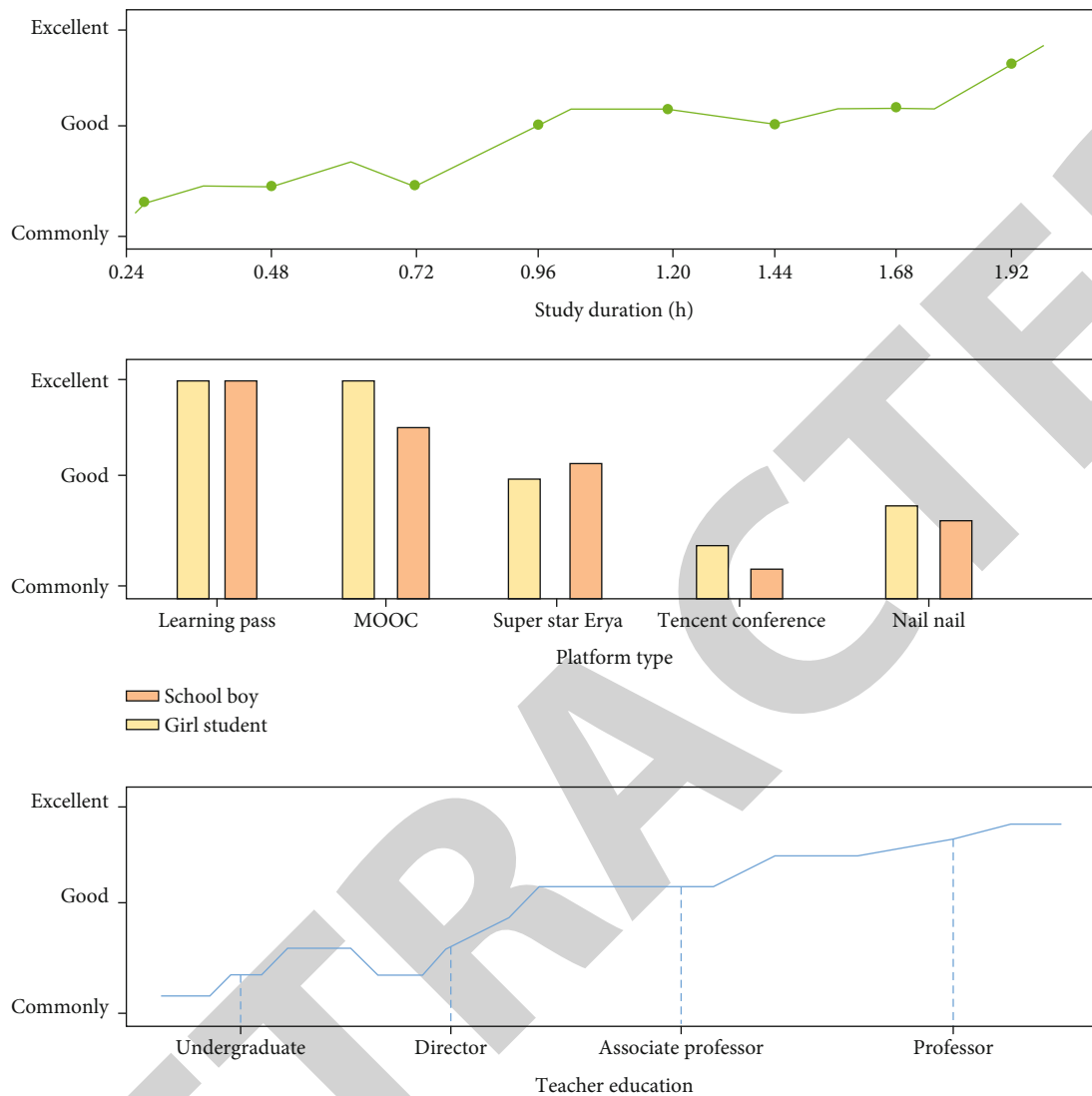


FIGURE 9: The impact of learning time, online teaching platform, and teachers' academic qualifications on students' evaluation.

When the professional grade of teachers is higher, the effect of students' online learning is better. At this time, the more important influencing factors of students' attributes are sorted according to age, discipline nature, and class collective. Students in the excellent stage provide more data in the teaching effect evaluation form. They actively cooperate with teachers in teaching research according to their own learning state. Teachers need to understand the actual needs of teaching and students' learning needs, improve teaching interaction, and promote the transformation of curriculum knowledge. Colleagues who strengthen teaching skills should also enrich their professional abilities, timely cater to modern teaching ideas and improve teaching methods.

5. Conclusion

Under the background of big data environment and computer, modern technology has a greater and greater impact on people's life. Big data technology has achieved effective combination and application in various fields. Among them,

the field of education is also facing new problems and opportunities with the support of big data technology. Online teaching mode is one of the topics frequently studied and investigated by scholars all over the world in recent years. Compared with traditional teaching, online teaching has the advantages of convenience, simplicity, and high interactivity. In order to explore the changes of students' learning effect in the process of online teaching, this paper proposes to build a model for online teaching effect evaluation with the support of data mining technology and decision tree algorithm. Firstly, the decision tree algorithm is used to divide the attributes of influencing factors from relevant rules. Record the factors and objects that reflect the teaching effect in online teaching and traditional teaching, respectively. Secondly, data mining technology is used to preprocess and clear the sample set to improve the accuracy of calculation results. Finally, the Kirschner model is used to construct the evaluation system, adding two attribute elements: students' evaluation of teaching and teachers' self-evaluation. In the evaluation model, the association rule

Retraction

Retracted: Application of Optimized BP Neural Network In Financial Alert System

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Gao, "Application of Optimized BP Neural Network In Financial Alert System," *Journal of Function Spaces*, vol. 2022, Article ID 1816315, 10 pages, 2022.

Research Article

Application of Optimized BP Neural Network In Financial Alert System

Qiaoyi Gao 

Business School, Zhejiang Wanli University, Ningbo 315100, China

Correspondence should be addressed to Qiaoyi Gao; gaoqiaoyi@zwu.edu.cn

Received 7 June 2022; Revised 19 July 2022; Accepted 23 July 2022; Published 8 August 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Qiaoyi 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.

The development of economic globalization not only brings development opportunities for enterprise development but also causes the rapid changes of economic environment. Global enterprises are facing many development challenges. It is very significant to implement financial alert. However, most of the existing financial alert systems are based on traditional quantitative analysis and cannot face the demand of big data growth. Based on this, main body of a book studies and analyzes the request of optimized BP neural network in financial alert system. Based on the financial early warning and a brief analysis of the development requirements of BP neural network, this paper establishes a financial early warning model for BP neural network. Considering the shortcomings of BP neural network, genetic algorithm is used to optimize the specification. Finally, a new reference scheme is provided for economic early warning. And use the following steps to create a list, taking into account the volume analysis in the conceptual design. The emulation results indicates that the optimized BP neural network can accelerate the convergence, has strong stability, and has higher accuracy in financial alert.

1. Introduction

The concept of financial alert first appeared in western countries. At present, the theory of financial crisis is constantly improving, and there are many research results in the research and analysis of financial distress [1]. From the perspective of fiscal forewarning, early research mainly depends on personal development experience, and there will always be some deviation in financial alert [2]. With the development of financial alert theory and the introduction of statistical theory, development of big data technology provides new research methods for risk warning in various fields [3, 4]. Although there are many researches on enterprise financial alert at home and abroad at this stage, there are still no good migration request research results. Many methods immediately refer to a certain early warning model for analysis, which is not reliable [5]. Therefore, it is required to optimize the request of BP neural network in financial alert system.

Under this background, the main body of a book studies the request of BP neural network in financial alert system. The first chapter briefly introduces the situation of financial alert, the commonly used pattern, and the chapter arrange-

ment of this study. Chapter 2 introduces the request and improved algorithm of financial alert algorithm and BP neural network at home and abroad and summarizes the shortcomings of the current research. In Chapter 3, the disadvantages of BP neural network research are combined, and a BP neural network model is established. A new network structure is proposed to optimize the stability and learning of BP neural network by using real coding method on initial gravities and thresholds. In Chapter 4, financial alert model is in view of BP neural network. The performance of optimized BP neural network is judged by analyzing the convergence speed, stability, and early warning accuracy under different algorithms. The experimental results indicate that compared with the traditional BP neural network, the optimized BP neural network proposed in main body of a book has faster convergence speed and shorter running time and has better request attributes in financial alert.

The innovation of this paper is to establish the maximum number of hidden layer nodes in BP neural network optimization. The application requirements of optimized BP neural network in financial early warning system are studied and analyzed. Considering the shortcomings of BP

neural network, genetic algorithm is used to optimize the specification and finally provides a new reference for economic early warning, using the nonuniform characteristics of the initial population space randomly generated by GA. Using equidistant sampling method, it is easier to find individuals and accelerate the convergence speed. The simulation results show that the optimized BP neural network can accelerate the convergence speed, has strong stability, and has high accuracy in financial early warning.

2. State of the Art

With the development of enterprises, the harm caused by financial risks is also deepening, which has a significant impact on the survival and development of enterprises. Considering the prevention of the losses caused by financial risks, it is very indispensable to take financial alert analysis. At present, there are many researches in financial alert analysis, which have completely shifted from qualitative research to quantitative research. For example, Huang and Guo established a new financial risk evaluation system in the research. In the research, fuzzy analysis method is adopted. For noise data and contour graph, virtual set processing method is adopted to analyze the unstable factors in the market and form fuzzy matrix, which helps fuzzy method to obtain more valuable fuzzy membership degree and further improve the request performance of risk assessment system [6]. In their research, in financial risk assessment, Tang et al. measured report data, introduced decision tree C4.0 algorithm for research and mining data rules, and built report detection system to find fraud. Through this model, the error information in finance could either be found in time and the potential knowledge could either be mined [7]. Wang and Xie used BP neural network to implement early warning analysis on the financial affairs of enterprises in coastal areas [8]. In the design of the early warning system, Sun and Lei chose textile enterprises being the study subject to sort out the data characteristics of textile enterprises. Leading textile companies in the A-share textile sector have listed mining companies being their research targets [9]. Deng et al. proposed the kpca-mpso-bp model in the financial alert analysis of e-commerce, constructed it using KPCA, IPSO, and BP neural network and Credit Risk Index of cross-border e-commerce in view of KCA, used MPPO to search the inertia gravity and threshold of BP neural network, and used BP neural network to train the credit risk data of 13 cross-border e-commerce enterprises [10]. Zhao used ARMA model, two types of neural networks (back propagation), and MLP to quantitatively analyze the portfolio in his research [11]. In gas detection, Zhou et al. used BP neural network to optimize the specifications of KF, introduced differences reimbursement to track the system state, and eliminate the changes of dynamic system specifications. The do-akf algorithm indicated the best performance [12]. On the basis of the 2π periodicity of angle measurement under static temperature and the complexity of the influence of ambient temperature change, Jia et al. proposed a method to improve the angle measurement accuracy of rotary encoder in view of Fourier expansion BP neu-

ral network (fe-gabpnn) optimized by GA, which has good fitting performance [13].

To sum up, it could either be seen that in the relevant research of financial alert, in addition to qualitative research, quantitative research is widely used, but there are many factors causing enterprises to get caught up in financial crisis, and some indicators are difficult to be analyzed quantitatively by traditional quantitative research methods. In the request of data mining algorithms in financial alert, most of them are analyzed on the basis of foreign pattern. There are many requests of BP neural network, but they are not localized and lack of practical request. In the request research of BP neural network, in addition to traditional algorithms, BP neural network optimization algorithms continue to emerge, but they are rarely applied to financial alert analysis. Therefore, it is of great practical significance to implement the request research of optimized BP neural network in financial alert.

3. Methodology

3.1. Design of Financial Alert Model. Mature enterprises are relatively mature in the management and development system, especially in the financial alert system. The daily operation and management of enterprises must be disturbed by many factors, but these factors could either be divided into external factors and internal factors in essence. Enterprise operation management is the key element of enterprise survival and profit and the logical relationship between the elements. It determines the market operation results of an enterprise. In the long run, whether an enterprise can find an enterprise operation mode suitable for its business needs and continuously improve it determines whether an enterprise can have a future. The causes of problems in management are nothing more than organizational structure problems; for example, the structure is chaotic. The enterprise goal is not clear, and the enterprise must first have clear goals and plans to make steady progress. We must make clear what the problem is and then find a solution. Through the calculation and analysis of financial alert data, we can realize the supervision of finance and find problems in time. From the perspective of investment, financial alert can provide more information on enterprise operation and management and realize the rationalization of investment. For the regulatory authorities, the use of financial alert model can analyze the potential risks in the market and implement risk control in time. Considering that financial early warning needs to adopt more scientific and accurate evaluation methods. Because of its self-learning, recognition, and judgment ability, BP neural network can process large-scale data in parallel and has the ability of logical thinking emulation. Compared with the early system, BP neural network has great advantages [14].

BP neural network has great advantages in financial alert and has the characteristics of information processing parallelism. Each neuron can calculate on the basis of the obtained information and then output the results, which improves the information processing ability of financial alert [15]. BP neural network has the characteristics of high fault

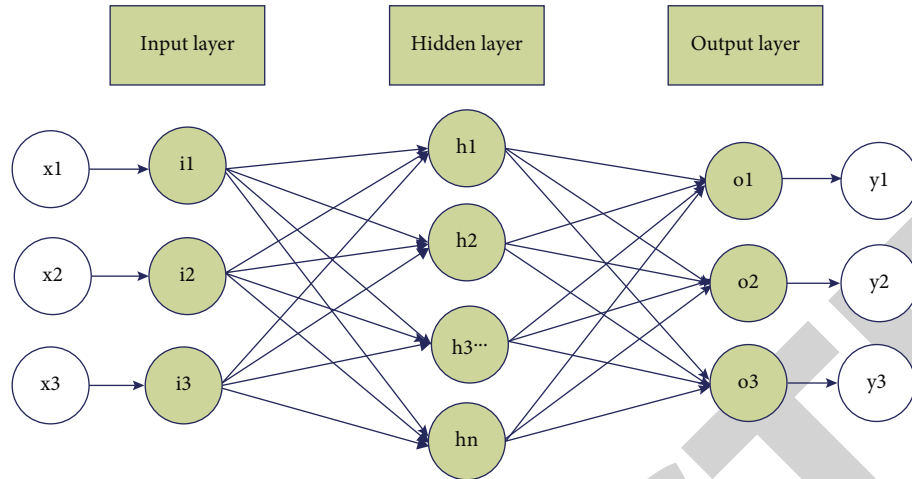


FIGURE 1: Basic structure of BP neural network.

tolerance and inaccurate calculation. In the financial alert indicators, it does not strictly require the enterprise sample data, nor does it need to assume the data situation independently. It can deal with discontinuous and incomplete information and improve the ability of using financial alert data [16]. New model has strong adaptive ability. It can not rely on external coercive forces but also achieve balance on the basis of information exchange and feedback. It can improve the time variability of financial alert and meet the needs of different enterprises [17]. With the accumulation of sample size, financial alert can continuously update knowledge and realize dynamic financial alert. In view of the advantages of new model in financial alert, new model is used to realize financial alert. The basic structure is indicated in Figure 1.

In the financial alert analysis, the selection of financial indicators is very significant. Financial indicators should be sensitive, have timely warning information, and can reflect the real situation of the enterprise. The financial indicators should be comparable and able to conduct comparative analysis among industries and enterprises. Financial indicators should be easy to obtain and highly operable, which can systematically reflect the financial situation of the enterprise and avoid one-sided amplification of the partial financial situation of the enterprise. Therefore, in the financial alert analysis, the financial indicators select the data closely related to finance, including market sales rate, net interest rate, cash sales rate, average return on net assets, amount of accounts receivable, asset cash rate, asset return, current ratio, quick ratio, asset liability ratio, asset turnover rate, accounts receivable collection and transfer rate, current asset turnover rate, operating income growth rate, tangible asset liability ratio Growth rate of total assets, price to book ratio, cash flow liability ratio, and inventory turnover rate.

As an artificial neural network model, new model simulates human neural network, which is divided into input, output, and implicit layers. It has strong nonlinear mapping ability, but new model also has its own shortcomings in request [18]. BP neural network requires little learning effi-

ciency, but it is not fast enough in practical request, so it can only use incremental exercise. The emulation of any function is close to and can deal with all kinds of network problems, but in fact, BP neural network does not always have a solution [19]. In nonlinear systems, it is difficult to choose a good learning rate, and the lack of parameter attributes impacts the exercise performance. Therefore, BP neural network needs to be optimized.

3.2. BP Neural Network Optimized by Genetic Algorithm. In this paper, GA is used to optimize BP neural network. Suppose that the input layer of new model is represented by M and the output layer is represented by P . The linking gravity between the implicit layer and the input layer is W_{ij} , and the linking gravity between the implicit layer and the output layer is W_{jp} . The sigmoid function is selected as the excitation function, and the input sample is represented by X , the output sample is represented by Y , and the expected attributes is represented by d . Forward propagation could either be expressed as

$$u_j^I = \sum_{m=1}^m w_{mj} x_m, \quad (1)$$

$$v_j^J = f \sum_{m=1}^m w_{mj} x_m, \quad j = 1, 2, \dots, J.$$

The total error of the network output layer could either be expressed as

$$E = \frac{1}{2} \sum_{p=1}^P e_p^2(n), \quad (2)$$

where $e_p(n)$ represents the single error of the node. When the output layer propagates forward, the error signal will be transmitted. After modifying the linking gravity layer by layer, the error back propagation will change. New model

uses the gradient descent method to modify the network gravity. The selection of excitation function is

$$f(x) = \frac{1}{1 + \exp(-ax)}. \quad (3)$$

Then, the n gravity adjustment amount could either be expressed as

$$\Delta w_{jp}(n) = \eta \left[y_p(n) (1 - y_p(n)) \right] \left(d_p(n) - y_p(n) \right). \quad (4)$$

Using the same reasoning and assuming that the gradient is expressed as $\delta_j^l(n) = -\delta E(n) / \delta u_j^l(n)$, we can get

$$\Delta w_{mj}(n) = \eta \delta_j^l(n) x_m^M(n). \quad (5)$$

Determine the quantities of network input layer nodes, implicit layer nodes and output nodes on the basis of the set input and output sequences, and initialize the linking gravities of the input layer and implicit layer, and a is for initial threshold of implicit layer, and b is for the threshold attributes of the output layer. Calculate the output of the implicit layer on the basis of the input vector, and the given threshold of the initial gravity, the formula is

$$H_j = f \left(\sum_{i=1}^n w_{ij} x_i - a_j \right), \quad (6)$$

where H is for calculation result of implicit layer, and the attribute range of j is $1 \sim l$. The expected results are calculated on the basis of the actual output of the new model. The formula is

$$O_k = f \left(\sum_{i=1}^l w_{jk} H_i - b_k \right), \quad (7)$$

where k is a natural quantities, with an attribute of $1 \sim m$. Calculate the error result on the basis of the formula.

BP neural network is widely used, in which the quantities of implicit layer nodes need to be reasonably selected. In the selection, the main body of a book takes an implicit layer as the study subject, determines the input layer and output layer in combination with the actual demand, and selects the optimal quantities of implicit layer nodes on the basis of the formula

$$l = \log_2^n \leq \sqrt{m + n} + a, \quad (8)$$

where l represents the quantities of implicit layers, m represents the quantities of access nodes, n represents the quantities of output nodes, and a ranges from 0 to 10.

The introduction of GA can give full play to the advantages of the two algorithms. The exercise of new model is in view of the principle of gradient descent, and the threshold attributes of gravity are constantly revised to avoid falling into local minimum attributes. GA can realize global

search [20]. Through global search, the gravity and threshold of new model could either be redetermined without impacting the nodes, implicit layers, and other specifications of new model [21]. In the optimization, the GA encodes the linking gravities to form the initial population and then calculates and screens each individual through crossover and mutation until the optimal gravity and threshold are obtained [22]. Then, adjust the adaptability on the basis of the output error results to ensure the minimum error attributes. The specific process is indicated in Figure 2.

When using GA, we first need to solve the problem of coding. The encoding operation does not impact the convergence performance. In the operation, the individual is encoded, and the individual is composed of input layer, implicit layer, linking gravity, and output threshold [23]. In new model, the most significant feature is the sum of squares of the error between the output attributes and the expected attributes. The more the error is, the smaller the performance of new model is. In the process of sample propagation, if the error attributes are large, the BP network needs to be corrected [24]. The introduction of GA can search the network gravity, introduce the gravity and threshold of chromosome into it, and reduce the error function. The objective function could either be expressed as

$$\min E(\xi) = \frac{1}{2} \sum_{i=1}^U (y^* - y)^2, \quad (9)$$

where E represents the error. Considering the improvement of the adaptability of chromosome, adaptability function is introduced. Adaptability function is not only a significant standard to distinguish individuals but also a standard of natural selection. Use the degree function to check whether the individual has reached the optimal solution in the calculation [25]. If the adaptability function is not selected properly, abnormal individuals may be produced in the early stage of genetic evolution. These individual experiences impact the performance of the algorithm because they are too prominent. In the later stage of genetic evolution, when the algorithm converges, due to the small difference between individuals, there may be a local optimal solution. Even improper adaptability function will lead to genetic stop. In the selection of adaptability function, the function cannot be negative, and it should be as simple as possible to reduce the computational complexity. The adaptability should have strong universality, and there is no need to revise the specifications of the adaptability function. Fitness of an individual refers to the measure of the degree of dominance of an individual in the survival of a population, which is used to distinguish the "good" and "bad" of an individual. Fitness is calculated using fitness function. Fitness function is also called evaluation function, which mainly judges individual fitness through individual characteristics. The general process of evaluating the fitness of an individual. After decoding the individual coding string, the individual phenotype 2 can be obtained. The objective function value of the corresponding individual can be calculated from the individual's phenotype. According to the type of optimization problem, the

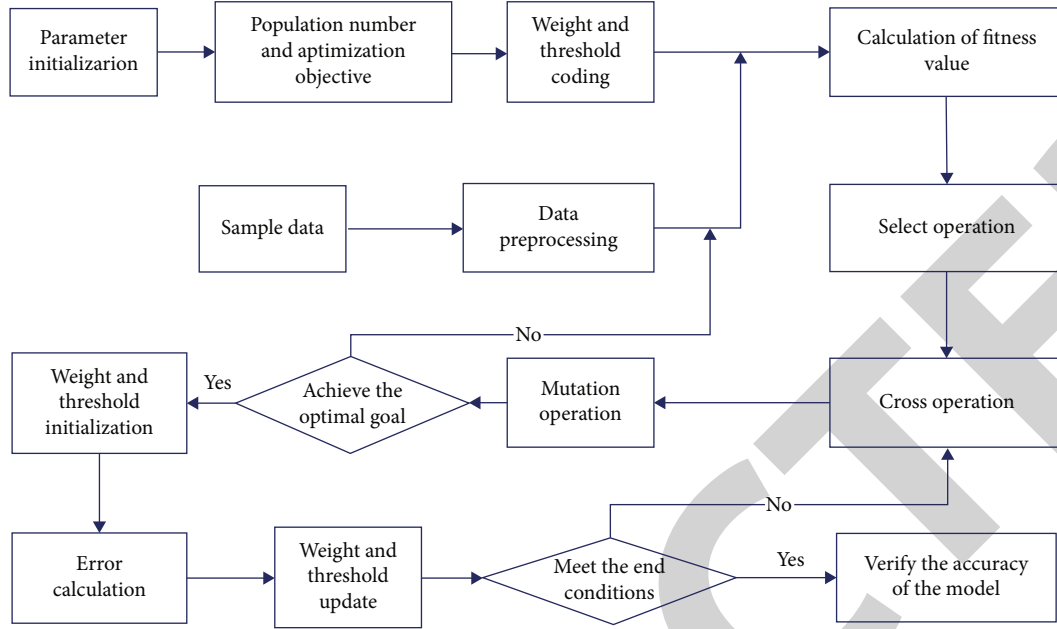


FIGURE 2: GA optimization neural network.

individual fitness is calculated from the value of the objective function according to certain conversion rules. In this paper, the selection of adaptability function is in view of the total error of new model, so the selection of individual adaptability function is

$$f = \frac{2}{2 + \sum_{k=1}^k \sum_{j=1}^n (T_j^k - Y_j^k)^2}, \quad (10)$$

where k represents the quantities of sample sets, T represents the ideal output, and Y represents the real output result.

GA is used to optimize BP neural network. When selecting, the best individual is saved, and the proportion of adaptability is used at the same time. Select the individual with the highest adaptability attributes in the population and immediately enter the next generation without crossover and mutation. This method can ensure that the optimal solution is not impacted by crossover and mutation, but the quantities of optimal individuals may increase sharply and get caught up in local optimal solution. Therefore, it is also indispensable to combine the adaptability ratio method to save the best individual. Assuming that the population size is n , the likelihood of individual selection is expressed as

$$P_{s_i} = \frac{f_i}{\sum_{j=1}^n f_j}, \quad (11)$$

where f_i represents the adaptability function attributes and P_{s_i} represents the likelihood of being selected, which reflects the proportion of individual adaptability in the sum of individual adaptability. The combination of these two methods can select individuals with large adaptability function attri-

butes to enter the next generation with the greatest likelihood and also provide a certain likelihood for individuals with small adaptability attributes to prevent the problem of local optimal solution.

In GA, compared with other algorithms, the most obvious feature is crossover operation, which is the most common method to generate new individuals. Using GA to optimize BP neural network also needs to go through chromosome cross mutation. Different ways are used in coding connecting genes and parameter genes, and the two parts need to be cross operated, respectively. Among them, the coding of connecting genes adopts one-point crossover operation, and the parameter genes adopt arithmetic crossover. The linear combination of two individuals belongs to arithmetic crossing. Assuming that the individual is represented by x and the crossed individual is represented by x' , the formula could either be expressed as

$$\begin{cases} x'_1 = ax_1 + (1-a)x_2, \\ x'_2 = ax_2 + (1-a)x_1. \end{cases} \quad (12)$$

In the formula, the attributes of a ranges from 0 to 1. The mutation operation of the algorithm adopts different compilation methods to adjust the connecting genes and specifications. The basic mutation method is used to change the gravity que gene of the connecting gene, and the nonuniform mutation method is used to change the rate gene. Specifically, the likelihood of the individual is specified as the variation point, and then, the genetic attributes of the variation point is inversely calculated to generate a new individual. The mean variation operation can ensure the free movement of individuals, but it is difficult to search the key areas. Therefore, in the nonuniform variation, the random attributes with uniform distribution are not used for

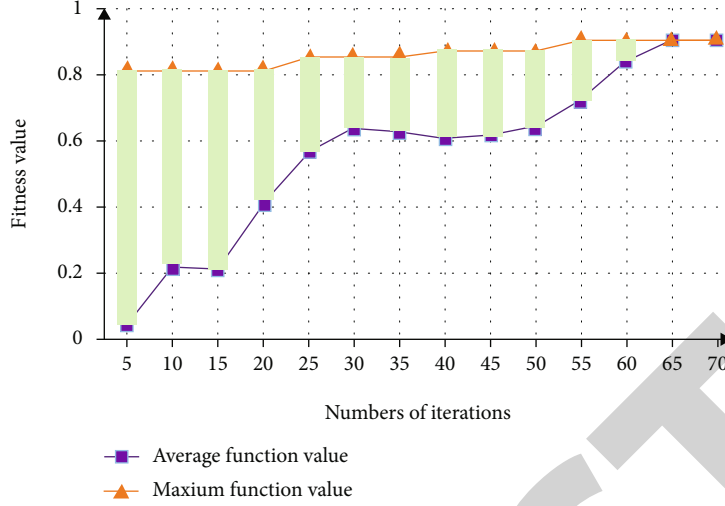


FIGURE 3: Relationship between test function 1 and evolution times.

attributes selection, and the random disturbance results are used as new gene attributes. Each gene is mutated with the same likelihood, and the whole solution vector will change slightly.

$$\begin{aligned}
 P_c &= \begin{cases} \frac{f_{\max} - f}{f_{\max} - f_v}, & (f' \geq f_v), \\ 1, 0, & (f' < f_v), \end{cases} \\
 P_m &= \begin{cases} 0.5 \times \left(\frac{f_{\max} - f}{f_{\max} - f_v} \right), & (f \geq f_v) \\ \frac{f_v - f}{f_v - f_{\min}}, & (f < f_v), \end{cases}
 \end{aligned} \quad (13)$$

where P_c is for represents the intersection likelihood, P_m is for represents the sudden change likelihood, f_{\max} and f_{\min} represent the maximum and minimum adaptability attributes, respectively, f_v represents the average adaptability attributes, and f' represents the larger adaptability attributes of the two crossed chromosomes. When the individual adaptability tends to be the same, the intersection likelihood and sudden change likelihood could either be the largest, and the adaptability is relatively scattered; the likelihood is the smallest. For individuals whose adaptability attributes are higher than the average attributes, they can enter the next generation, and individuals whose adaptability attributes are lower than the average attributes will be eliminated. The adaptive likelihood can provide the individual optimal attributes and ensure the convergence of the algorithm on the basis of ensuring the diversity of the population.

4. Result Analysis and Discussion

4.1. Emulation Analysis of Optimized BP Neural Network. Establish BP neural network, including 5 input neurons, 4 output neurons, and 6 implicit layer neurons. Optimize the neural network model in view of GA for network learning,

store the result gravities, and establish the neural subnetwork knowledge base. The algorithm can only be realized by using MATLAB changes.

Using the real quantity coding method, the population size is set to 100, the initial attribute of variation likelihood is set to 0.05, the initial attribute of intersection likelihood is set to 0.6, and the initial attribute of B is set to 0.01. Then, the test function is optimized, and the same transformation rules are used to translate the test function upward to obtain the evolutionary adaptability function. The measurement results are indicated in Figures 3 and 4.

The data in figure indicates that, compared with the unimproved algorithm, in terms of optimizing the test function, the new model can achieve faster convergence and perform well, can reduce the quantities of iterations, and has high accuracy. The average running time of the traditional algorithm is 0.704 s, and the running time of the optimized new model algorithm is 0.604 s, which indicates that the optimized new model algorithm can shorten the running time at the same time. The improved algorithm can implement genetic optimization on many points at the same time and then search on the basis of the direction of negative gradient, so as to avoid the problems of local minimum points and slow convergence speed.

The algorithm of optimizing new model is used to optimize the maximum upper bound network. Considering the reduction of computational complexity, the amount of network exercise is 30, and the mean square error function is used as the adaptability function. The optimal quantities of implicit layer nodes are 12. The change of adaptability function in the optimization process is indicated in Figure 5. It could either be seen from the figure that after 60 iterations, the optimal individual adaptability will not change, and the average adaptability is still changing.

After determining the optimal quantities of implicit layers, the input layer node is 24 and the output layer node is 4. GA is used to optimize the gravity and threshold of new model. The quantities of network exercise are still 30. The mean square error function is used as the adaptability

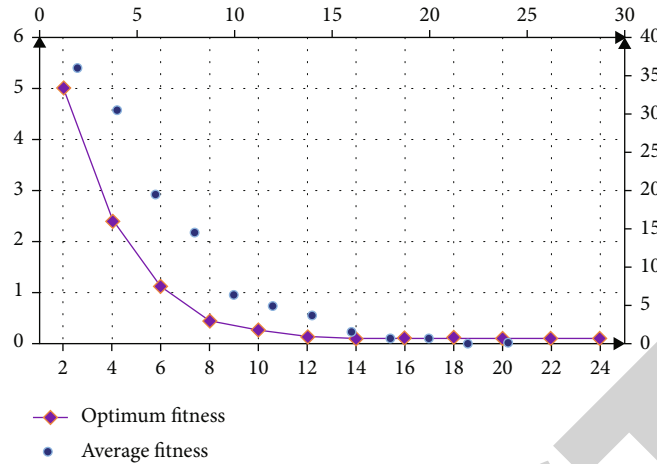


FIGURE 4: Relationship between adaptability of test function 2 and evolutionary algebra.

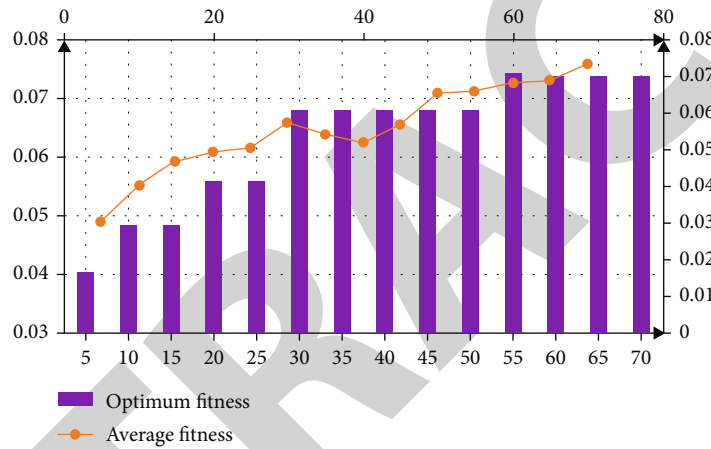


FIGURE 5: Adaptability function optimization process.

function. After 65 times of evolution, the optimal individual adaptability is gradually stable.

The GA is used to find the maximum attributes of the nonlinear function globally, as indicated in Figure 6. After the maximum attributes is obtained, the objective function and binary transcoding are calculated through initialization coding, and then, the maximum adaptability attributes and the optimal individual are found through genetic variation. The results are indicated in Figure 7. The learning error rate of nonlinear function in view of GA is within 3%.

4.2. Financial Alert Analysis. The optimized BP neural network model is used to analyze the financial alert of enterprises. Under the environment of MATLAB, the programming language and the collected data samples are input into the model for analysis. Although the optimized BP neural network has no high requirements for the original data, the amount of data selected is too large. SPSS software is used to analyze the sample factors, and all sample data are input to obtain the significance likelihood, which is 0~1. If the ratio reaches 0.9, it is considered that the factor has great influence, and if it is lower than 0.6, it is considered that it is not indispensable to analyze. Implement correlation analysis on the selected financial index

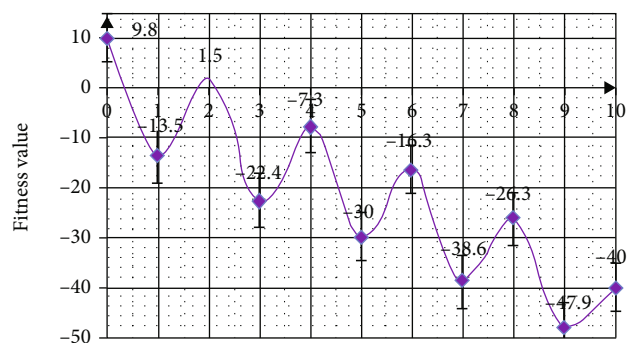


FIGURE 6: Global optimization of nonlinear function.

data to select the factors with great influence. The selected indicators include quick ratio, turnover rate of current assets, asset cash ratio, growth rate of total assets, return on net assets, sales expense ratio, and price to book ratio. The current ratio, the turnover rate of current assets, and the cash ratio of assets all reflect the solvency of the enterprise. The growth rate of total assets reflects the growth ability of the enterprise, the return on net assets reflects the profitability, the proportion

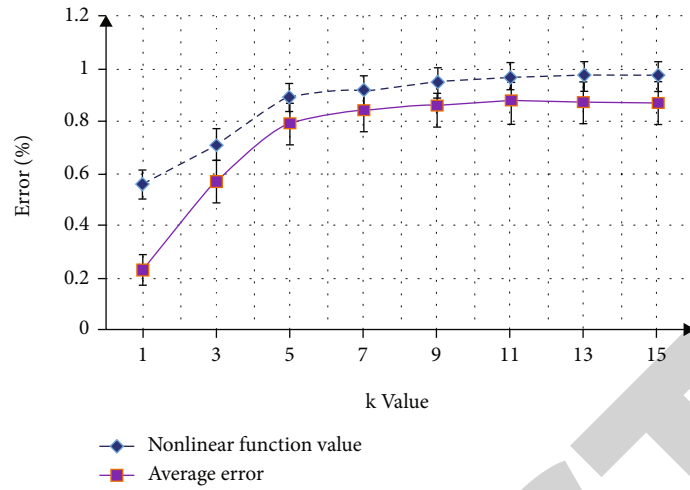


FIGURE 7: Error analysis of nonlinear function emulation.

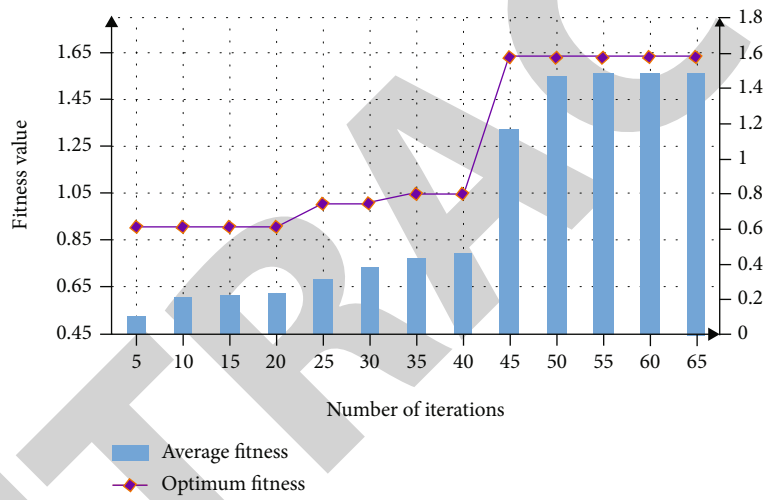


FIGURE 8: Adaptability evolution curve of experimental data.

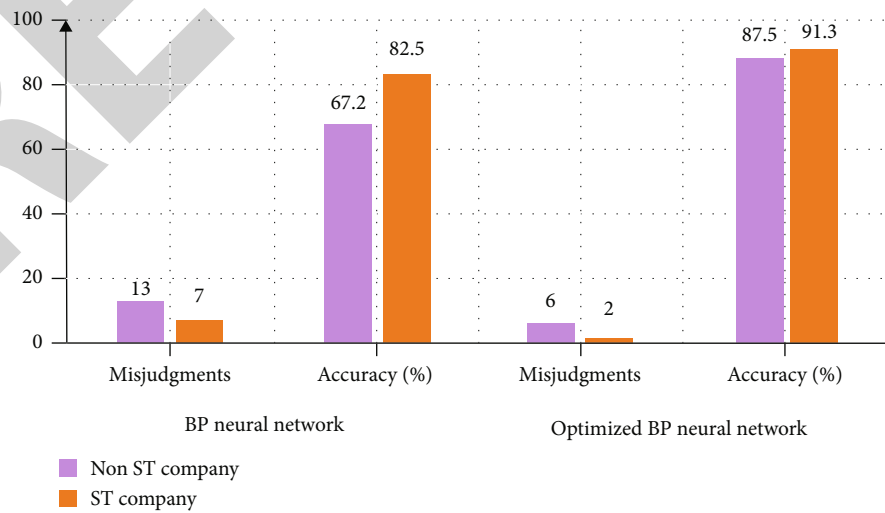


FIGURE 9: Accuracy of financial alert before and after new model optimization.

of sales expenses reflects the market sales ability, and the price to book ratio reflects the stock attributes of the enterprise.

The input layer reflects the neurons of the data, and the financial indicators included are enterprises. Therefore, the input layer sets seven nodes, and the data of nearly five years is brought into the optimized new model for experimental analysis. 60% of the data is used as sample data and 40% of the data is used as test data, which is compared with the real crisis situation of the enterprise to calculate the accuracy. In the optimization design of new model, GA is used to search through adaptability. Therefore, the best adaptability and adaptability function will be inconsistent in request. Considering the analysis results, continuous iteration is required until the requirements are met. Figure 8 is the adaptability evolution curve of experimental data. St refers to special treatment. The stock represented by ST Company is st stock, which has relatively large investment risks.

The accuracy of financial alert before and after new model optimization is indicated in Figure 9. From the figure, it could either be seen that the early warning accuracy of optimized new model has been significantly improved, which proves the superiority of the algorithm.

5. Conclusion

This paper studies the requirements of optimizing BP neural network in financial early warning and constructs a financial early warning analysis model. The stability of the algorithm is improved, and the maximum upper bound model of the number of hidden layer nodes is proposed. GA is used to optimize the design of the new model, and roulette selection is carried out on the basis of population grouping to maintain population diversity. The running time of the algorithm is mainly used to calculate the adaptability function. The simulation results show that compared with the traditional algorithm, the optimized model can give full play to the mapping ability of the new model. It should be pointed out that when optimizing the new network, the reference of GA can only improve the prediction accuracy. It has better learning ability and faster convergence speed and improves the accuracy of financial early warning. In the case of small sample size or uneven distribution, it may affect the prediction results, which needs further research. In this paper, the use of financial early warning can better monitor and diagnose finance, timely discover crises, and better promote the development of enterprises. In the business activities of enterprises, finance is at the core stage. This paper provides reference value for good financial operation and enterprise development.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Zhejiang Public Projects “Development and Integration of Internal Control Information System in Administrative Institutions” (No. LGF20G030003).

References

- [1] M. Qu and Y. Li, “Financial risk early-warning model based on kernel principal component analysis in public hospitals,” *Mathematical Problems in Engineering*, vol. 2021, Article ID 5525354, 7 pages, 2021.
- [2] L. Zhu, M. Li, and N. Metawa, “Financial risk evaluation Z-score model for intelligent IoT-based enterprises,” *Information Processing & Management*, vol. 58, no. 6, article 102692, 2021.
- [3] S. Li, W. Shi, J. Wang, and H. Zhou, “A deep learning-based approach to constructing a domain sentiment lexicon: a case study in financial distress Prediction,” *Information Processing & Management*, vol. 58, no. 5, article 102673, 2021.
- [4] L. Sharafi, K. Zarafshani, M. Keshavarz, H. Azadi, and S. Van Passel, “Farmers’ decision to use drought early warning system in developing countries,” *Science of the Total Environment*, vol. 758, article 142761, 2021.
- [5] Z. Zhu and N. Liu, “Early warning of financial risk based on K-means clustering algorithm,” *Complexity*, vol. 2021, Article ID 5571683, 12 pages, 2021.
- [6] X. Huang and F. Guo, “A kernel fuzzy twin SVM model for early warning systems of extreme financial risks[J],” *International Journal of Finance and Economics*, vol. 26, no. 1, pp. 1459–1468, 2021.
- [7] X. B. Tang, G. C. Liu, J. Yang, and W. Wei, “Knowledge-based financial statement fraud detection system: based on an ontology and a decision Tree,” *Knowledge Organization*, vol. 45, no. 3, pp. 205–219, 2018.
- [8] J. Wang and S. Xie, “Application of BP neural network in early-warning analysis of investment financial risk in coastal areas,” *Journal of Coastal Research*, vol. 106, no. sp1, p. 259, 2020.
- [9] X. Sun and Y. Lei, “Research on financial early warning of mining listed companies based on BP neural network model,” *Resources Policy*, vol. 73, no. 2, article 102223, 2021.
- [10] Z. Deng, M. Yan, and X. Xiao, “An early risk warning of cross-border E-commerce using BP neural network,” *Mobile Information Systems*, vol. 2021, Article ID 5518424, 8 pages, 2021.
- [11] P. Zhao, “RETRACTED: Quantitative analysis of portfolio based on optimized BP neural network,” *Cognitive Systems Research*, vol. 52, pp. 709–714, 2018.
- [12] S. Zhou, C. Y. Shen, L. Zhang et al., “Dual-optimized adaptive Kalman filtering algorithm based on BP neural network and variance compensation for laser absorption spectroscopy,” *Optics Express*, vol. 27, no. 22, pp. 31874–31888, 2019.
- [13] H. K. Jia, L. D. Yu, Y. Z. Jiang, H. N. Zhao, and J. M. Cao, “Compensation of rotary encoders using Fourier expansion-back propagation neural network optimized by genetic algorithm,” *Sensors*, vol. 20, no. 9, p. 2603, 2020.
- [14] D. Li, G. Wu, J. Zhao, W. Niu, and Q. Liu, “Wireless channel identification algorithm based on feature extraction and BP neural network,” *Journal of Information Processing Systems*, vol. 13, no. 1, pp. 141–151, 2017.

Retraction

Retracted: Exploration of Energy-Saving Chilling Landscape Design Based on Algo for Group Intelligence

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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, "Exploration of Energy-Saving Chilling Landscape Design Based on Algo for Group Intelligence," *Journal of Function Spaces*, vol. 2022, Article ID 1851623, 10 pages, 2022.

Research Article

Exploration of Energy-Saving Chilling Landscape Design Based on Algo for Group Intelligence

Zhuo Li 

Art College, Heilongjiang University, Heilongjiang, Harbin 150001, China

Correspondence should be addressed to Zhuo Li; 2005050@hlju.edu.cn

Received 28 May 2022; Revised 12 July 2022; Accepted 15 July 2022; Published 4 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Zhuo 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.

Due to the influence of regional climatic conditions, the development of cold ecological tourism agricultural park in the north is slow, its landscape planning and design research lags behind, and there are still some problems. With the development of electronic computers, the previously unachievable optimization of engineering structures is now possible, while the traditional optimization methods cannot meet the requirements of modern computing due to the implementation constraints. The emergence of algo for group intelligences has made some applications of engineering structural optimization possible. Group intelligence is the characteristic of a system in which the collective behavior of low-level individuals interacting locally with their environment forms an overall pattern of consistent functions. This paper presents a study of energy-efficient cold landscape design based on algo for group intelligences and analyzes the problems that arise in the ecological landscape design of cold urban settlements. Then countermeasures are proposed to solve the problems. The linear model is replaced by the determination of the weight of the evaluation index of the energy-saving design of frozen landscape. And take it as a new migration model. Then, based on the weight of the evaluation index, the group algorithm analysis in landscape design is proposed, which is a generalized definition of the standard model. The experimental results show that the algo for group intelligence outperforms another algorithm in terms of solving ability to the extent that the average optimization ability is improved by 13.45%, so the algo for group intelligence demonstrates its superior ability to take into account both local and global search. Therefore, the use of algo for group intelligences to build an energy-efficient society is a necessary path for China's development, and the construction of energy-efficient gardens is an important part of building an energy-efficient society, and the application of energy-efficient concepts to modern landscape design is an inevitable trend in the development of urban landscaping.

1. Introduction

Energy consumption in buildings has become one of the three major energy consumers in China, together with industrial energy consumption and transportation energy consumption [1]. The uncontrolled squandering of the main resources, the reckless pollution of the environment, and the shortage of water, land, minerals, and other resources, as well as the low development and utilization of some new energy sources, make the contradiction between human and energy resources increasingly prominent [2]. The spatial creation of the landscape of cold urban settlements must “naturally” use the natural environment, adapt to local conditions, use the regional advantages to turn its disadvantages

into favorable factors, and create a landscape environment with regional characteristics [3]. From the perspective of the composition of urban green space, cold landscape is an important part of urban development space, which plays an important role in improving the quality of urban environment, enhancing the overall image of the city and accelerating urban economic development [4]. In cold regions, many residential buildings use active means to meet the humanization of the indoor environment, which in turn increases building energy consumption, which is a huge waste for residential buildings in cold regions with high energy-saving appeals [5]. Natural resources have long been unable to withstand the endless exploitation of human beings, and the environment on which human beings

depend for survival has become full of holes [6]. Therefore, applying ecological and environmental principles to the design and planning of urban public landscapes and proposing public landscape planning and design methods can improve the practical role of public landscapes in energy conservation and environmental protection.

Energy-efficient gardening, as an important part of building an energy-efficient society, plays an extremely important role in implementing the scientific concept of development and creating resource-efficient and sustainable development in the landscaping industry [7]. The core content should be studied in terms of building site selection, zoning, building and road layout orientation, building orientation, building body shape, building spacing, dominant direction of winter and summer monsoons, solar radiation, etc. We will analyze the factors that determine the climate and make full use of them through the planning and layout of the residential area, so as to form good living conditions and a microclimate environment that is conducive to energy saving. However, the development and construction of domestic cold landscapes are still in the initial stage, and blind development, imitation of countries, and quick success are inevitable, such as the planting of seedlings in the opposite season, planting against the environment, introducing a large number of species, and even filling lakes to create gardens are common [8]. At present, there are not many domestic and theses on the study of ecological landscape design in cold urban settlements. In the design of cold urban landscape, design should be combined with ecological design from a holistic perspective to enhance the livability of cold urban areas and promote respect for nature and protection of nature.

As part of the public green space, urban parks are a platform for people to relax and get close to nature, and they are an indispensable green life facility in the urban infrastructure [9]. Under the boom of creating energy-saving green landscape in cities, the concept of energy-saving greening, i.e., “choosing the greening mode with the least amount of land, the least amount of water, the least amount of financial allocation, and the least disturbance to the surrounding ecological environment,” is a landscape construction concept in line with the theory of sustainable development [10]. In its specific design process, we should not only emphasize its artistry, but also pay more attention to its energy-saving and environmental protection role, so as to create a warm, healthy, and sustainable urban living environment for human beings. When the inevitability of sustainable development becomes the common understanding of the whole society, how to apply energy-saving technology in the design of cold-land residential buildings becomes an important research topic.

The innovation of the research is that through the overview of the landscape planning and design system and methods of the cold botanical garden, this paper puts forward an energy-saving cold landscape design method based on swarm intelligence algorithm and studies the evaluation index weight of energy-saving cold landscape design. Then, based on the weight of the evaluation index, the group algorithm analysis in landscape design is proposed, which is a

generalized definition of the standard model. The planning and design of cold ecological sightseeing agricultural park landscape adopts modern algorithms for swarm intelligence and mapping. While traditional gardens are nutritious, they also pay attention to the combination with local culture.

2. Design Ideas of Energy-Saving Cold Landscape Based on Group Intelligence Algorithm

2.1. Energy-Saving Cold Landscape Design Method Based on Group Intelligence Algorithm. With the continuous acceleration of urbanization, the concept of energy conservation has been deeply rooted in urban public landscape design, and energy conservation and environmental protection have become the core concept of public landscape design. This is a complex and systematic work based on the theory of circular economy, the theory of sustainable development, landscape ecology, and other multilevel theories. As urbanization continues to accelerate, the concept of energy conservation has been deeply rooted in the design of urban public landscapes, and energy conservation and environmental protection have become the core design concept of public landscapes. It is a complex and systematic work based on the theory of circular economy, sustainable development theory, landscape ecology, and other multilevel theories. The algo for group intelligence is a multipoint random search algorithm based on population, which is a collection of multiple individuals with similarity, such as bird swarms, ant swarms, fish swarms, and other biotic swarms and abiotic swarms such as quantum swarms. In both versions, ant density and ant population, ants update pheromones directly after moving from one city to a neighboring city, and the former pheromone update is independent of distance. This performance evaluation can be used to measure the variation of the distance between neighboring solutions on the resulting Pareto frontier, which is defined as:

$$S = \frac{1}{n-1} \sum_{i=1}^n (\bar{d} - d_i)^2. \quad (1)$$

Among them:

$$d_i = \min_j (f_1^i(x) - f_1^j(x) + f_2^i(x) - f_2^j(x)), \quad (2)$$

where n is the number of vectors, \bar{d} is the distance average, and S is the distance.

Therefore, we analyze the environmental characteristics of ecological landscape design in cold urban residential areas and propose a method for energy-saving cold landscape design using algo for group intelligences. In this regard, an environmentally friendly and energy-saving landscape design method is proposed, which focuses on the use of resources and energy. The route of energy-saving and environment-friendly public landscape design is shown in Figure 1.

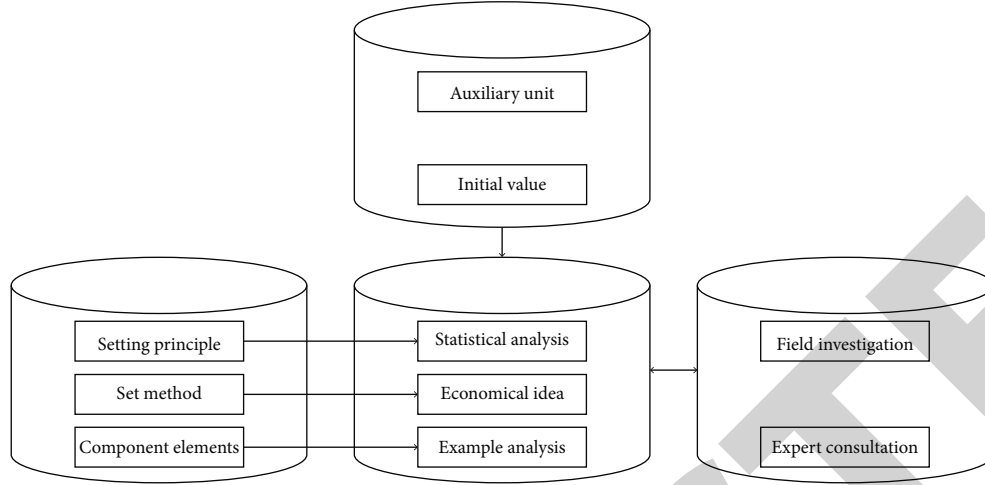


FIGURE 1: Road map of environmental protection and energy conservation public landscape design.

The first is the energy-saving design of landscape power resources. The design adopts the way of multiple platforms working together, splitting the traditional large landscape power supply platform to ensure the intensification of energy supply. This requires the building layout to be reasonably zoned for the thermal environment requirements, i.e., the thermal environment quality requirements are similar to the centralized arrangement. Traditional optimization methods can achieve high speed and accuracy when solving small- and medium-scale optimization problems. However, most of the optimization problems in the actual production process are complete problems, which are too complex to be solved by traditional optimization algorithms. Each individual has its own entry and exit rates. The in-migration and out-migration rates are functions of the population size, and they can be calculated by the following equations:

$$\begin{aligned} \lambda_k &= I \left(1 - \frac{k}{n} \right), \\ \mu_k &= E \left(\frac{k}{n} \right), \end{aligned} \quad (3)$$

where λ is the migration rate, μ is the emigration rate, I is the maximum possible immigration rate, E is the maximum emigration rate, K is the number of individuals, and n is the maximum number of individuals.

Therefore, the basic encoding method to convert the solution into a form that is convenient for GA operation is the binary encoding method. That is, the solution of the general problem is represented as a binary string, and the space in which the solution is encoded becomes the encoding space. Landscape base electricity needs to ensure the effective transmission efficiency of the power system. Its overall landscape electric energy environmental protection mode diagram is shown in Figure 2.

The second is the energy-saving design of landscape land resources, which can save a lot of land resources by improving the utilization of land resources in the landscape design process. It includes the soil and the topographic environ-

ment composed of soil; in landscape ecology, any landscape exists with an optimal spatial pattern to meet the sustainability of a landscape or area. This design proposes an ant colony optimization algorithm to optimize the allocation of land resources in traditional public landscapes. In each step of path construction, ants decide which city to move to next according to a probabilistic behavioral selection rule called stochastic scaling rule. The probability that an ant located in the current city will choose as the next city is

$$\begin{cases} p_{ij}^k = \frac{[\tau_{ij}(t)^\alpha] [\eta_{ij}]^\beta}{\sum_{t \in N_i^k} [\tau_{ij}(t)^\alpha] [\eta_{ij}]^\beta}, j \in N_i^k, \\ 0, j \notin N_i^k, \end{cases} \quad (4)$$

where k is the ant, i is the current city, j is the next city, and N_i^k is the feasible neighborhood that can be directly reached.

Direct analysis methods can stop and give a solution at any time because they deal with complete solutions. However, branch-and-bound, dynamic programming, and separation and conquest all construct solutions during the search process, so they cannot stop and give a complete solution until the entire search process is completed. Therefore, when using the ACO to solve a specific problem, the first step is to solve how to map the movement path of ants in the ACO to the solution of the problem. Ant colony algorithm can be used to solve the path planning problem, and the algorithm is implemented in JAA language. Although this method has not been studied for a long time, preliminary studies have shown that it is a promising algorithm. It simplifies the binary discrete particle swarm algorithm into basic components and provides a base point for further exploration of the nature of discrete particle swarm algorithm in the future. In addition, sunrooms are a useful addition to residential buildings to reduce energy consumption, and the rooms connected to them not only reduce a large amount of heat loss, but also reduce the energy required for cooling.

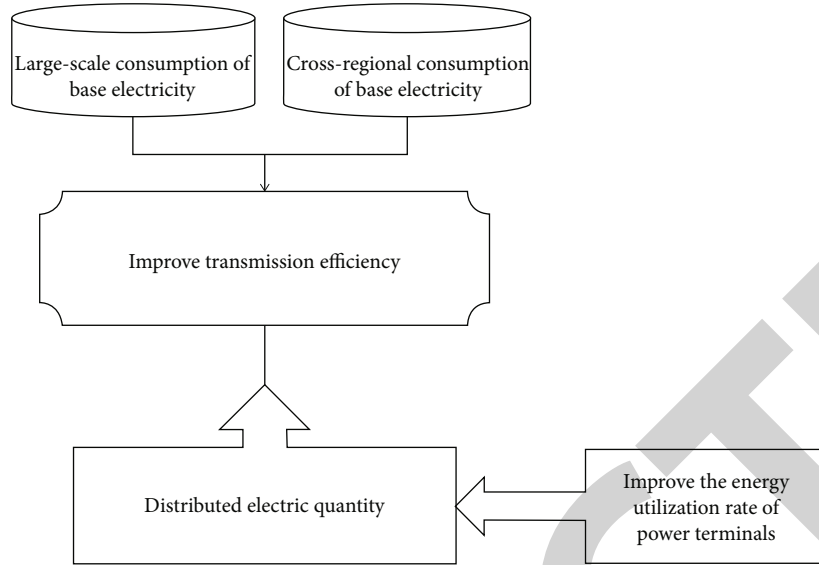


FIGURE 2: Energy-saving realization of electric power.

Finally, the use of energy walls, from the perspective of the climate characteristics of cold regions and building construction, and the use of external wall insulation construction in cold regions are relatively more reasonable. Because the external insulation is smaller than the internal insulation wall heat bridge, heat loss is less, and thermal comfort is high. In order to reduce the indoor heat loss in the main rooms, a “transition zone” can be set up between the living rooms with high temperature requirements and the outdoor spaces with very low temperatures, taking into account the usage conditions. In the evolution of the ant colony algorithm, the ants construct the solution heuristically by using pheromones and heuristic information. The insertion position is decided by the one that saves the most, and the selection and insertion steps are repeated, and when the path cannot be expanded, another route is built. The algorithm generally uses the error sum of squares criterion function as the clustering criterion function, and the error sum of squares criterion function is defined as

$$J_c = \sum_{i=1}^k \sum_{p \in C_i} \|p - M_i\|^2, \quad (5)$$

where p is the sample point and M_i is the average deviation of data.

Therefore, the good or bad construction of the pheromone and heuristic information model directly affects the effectiveness of the ant colony algorithm. In addition, a leader particle is introduced into the algorithm. This particle uses a separate update rule similar to the ACO pheromone and provides information to all other particles. Based on the “self-regulating and self-designing” ecological properties of the ecosystem, the algorithm restores some or all of the natural properties of the cold landscape shorelines and sections, reconstructs the natural ecosystem of the cold landscape, and creates a natural environment rich in biodiversity.

The core is to carry out scientific planning and design guided by the overall development plan of the city, with reasonable allocation of flowers and trees, three-dimensional layout of trees, irrigation, and grass, so that while improving the ecological environment of the city, urban residents can enjoy a beautiful ecological living environment and cultivate their morale.

2.2. Determination of Evaluation Index Weight of Energy-Saving Landscape Design in Cold Region. In fuzzy decision, weight is very important. It reflects the position and role of various factors in the process of comprehensive decision-making and directly affects the results of decision-making. The weight is usually given based on experience, which undeniably reflects the actual situation to a certain extent. However, the weights given by experience sometimes cannot objectively reflect the actual situation, resulting in the “distortion” of the evaluation results. Objective weight determination methods are as follows: The statistical method to determine the weight is expert estimation. The technical design of energy-efficient walls in cold regions needs to be carried out on two levels: first, the material properties of the materials used; second, the construction practices of the insulation structure. Energy-efficient landscaping involves many links and is a systematic whole. Due to its special geographical conditions, cold regions reflect unique landscape characteristics in terms of natural environmental features and human characteristics. Therefore, when selecting the index system, it should be established on a scientific and reasonable basis, and the selected indexes must truthfully reflect the evaluation object and take into account each other. Evaluation from different perspectives forms a complete and unified, hierarchical evaluation system, so that the evaluation results are complete, available, and accurate. The complexity of the model of migration is significantly higher than that of the linear model. Sinusoidal migration models are superior to other models. They can be expressed by the following equation:

$$\begin{aligned}\lambda_k &= \frac{1}{2} \left(-\cos \left(\frac{k\pi}{n} \right) + 1 \right), \\ \mu_k &= \frac{E}{2} \left(-\cos \left(\frac{k\pi}{n} \right) + 1 \right).\end{aligned}\quad (6)$$

First is the hierarchical single ranking. It calculates the weights of indicators in a certain level relative to the indicators in the previous level to which they belong. The main idea is to construct a judgment matrix and then use computer software to calculate and check the weights of the single-level indicators. As the depth of the residential building has limited variation, when the length of the residential building is reduced, the floor area is reduced by different proportions. By modulating the existing public landscape environment data, the grid data of the existing public landscape pattern is taken as the initial data. All the pheromones on the edges are reduced by a fixed size value, and then each ant adds pheromones to the edge through which it passes. This fixed size pheromone is

$$\tau_{ij} \leftarrow (1 - \rho)\tau_{ij}, \forall (i, j) \notin L, \quad (7)$$

where ρ is the evaporation rate of pheromone.

Since garden construction is inextricably linked with urban electricity and water energy, we should maximize energy conservation of water and electricity and other resources, improve the utilization rate of resources, and reduce energy consumption in all aspects of garden design, construction, maintenance, and postmanagement. On the basis of favorable plant growth and maintenance of biodiversity, reduce the large-scale transformation of site topography and use site topsoil as planting soil as much as possible, so as to maintain biodiversity and reduce the economic costs required for exotic planting soil.

Next is the hierarchical total ranking and assignment. The hierarchical total ranking is the weight value of each indicator for the total target (energy-efficient landscape design for urban parks), which is the weight of each indicator that is ultimately needed. As the length of the residential building increases, the curve flattens out, i.e., when the length of the building grows to a certain limit, the growth rate of the resulting energy efficiency decreases to a nonsignificant rate. Let there be Y data ants in the matrix space, i.e., with Y species solutions, and then combine these solutions together to form the initial ant data population, i.e., the ant data genes, and locate the ordinal numbers. In addition, the pheromone and heuristic information model compatible with the solution will be defined and initialized in the initialization of the ant colony intelligence algorithm. Pheromone importance factor refers to the influence of pheromones produced by species on ants during their movement. Heuristic function importance factor reflects the relative importance of heuristic information in guiding ant colony in path search. The heuristic information of the inner loop ant colony algorithm for solving the model represents the heuristic information for the decision-making level to select the process schedule. Different from pheromone information, heu-

ristic information depends on the characteristics of the problem itself. It has nothing to do with the information of the solution generated in the iterative process.

The framework of the ant colony optimization algorithm is shown in Figure 3 below.

Select the most appropriate technical principles according to the characteristics of site conditions. Use renewable resources such as wind energy, solar energy and bioenergy to create landscapes. It can well solve the power demand of site landscape lighting and other services, and energy-saving landscape can not only save the economic cost of municipal pipe network. However, if the proposed square is demolished and trees are cut down to create a square without determining the functional needs of the square, it will inevitably cause unnecessary waste and permanent damage to the original environment and ecosystem. Therefore, if the mean value of the optimal value of an algorithm after normalization is positive, it indicates that the performance of the algorithm is higher than the average. A negative value indicates that the performance of the algorithm is lower than the average. After completing the pheromone evaporation step, all ants release pheromones on the edge through which it passes:

$$\tau_{ij} \leftarrow \tau_{ij} + \sum_{k=1}^m \nabla \tau_{ij}^k, \forall (i, j) \notin L, \quad (8)$$

where $\nabla \tau_{ij}^k$ is the amount of pheromone released by the k first ant to its passing edge.

Finally, there is the analysis of weighting results. The weight value of each indicator can reflect the degree of influence of each indicator on energy-saving landscape design, from which we can see which indicators should be emphasized and which indicators should be focused on, providing a basis for identifying problems and proposing measures. In the case of a certain depth of a residential building in a row, it is beneficial to increase the length of the residential building in an appropriate amount. After the matrix space is established, the adaptation function needs to be established. The fitness function can control the optimization direction of the landscape land allocation and is the basis for judging the merit of each ant in the above population matrix. Then a new variation operation (Gaussian variation) is used. The formula of the probability distribution can be defined as

$$f_{\mu, \sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-((x-\mu)^2/2\sigma^2)}, \quad (9)$$

where μ is the mean value and σ^2 is the variance.

Management intensive landscape refers to the minimum investment of human, material, and financial resources in all phases of garden construction, especially in the postmanagement and maintenance phase. Due to insufficient field research in the early stage, the designer cannot reasonably lay out the design lot. And the design cannot comprehensively consider the functional needs of the lot, the location of landscape nodes, the layout of viewpoints, and other

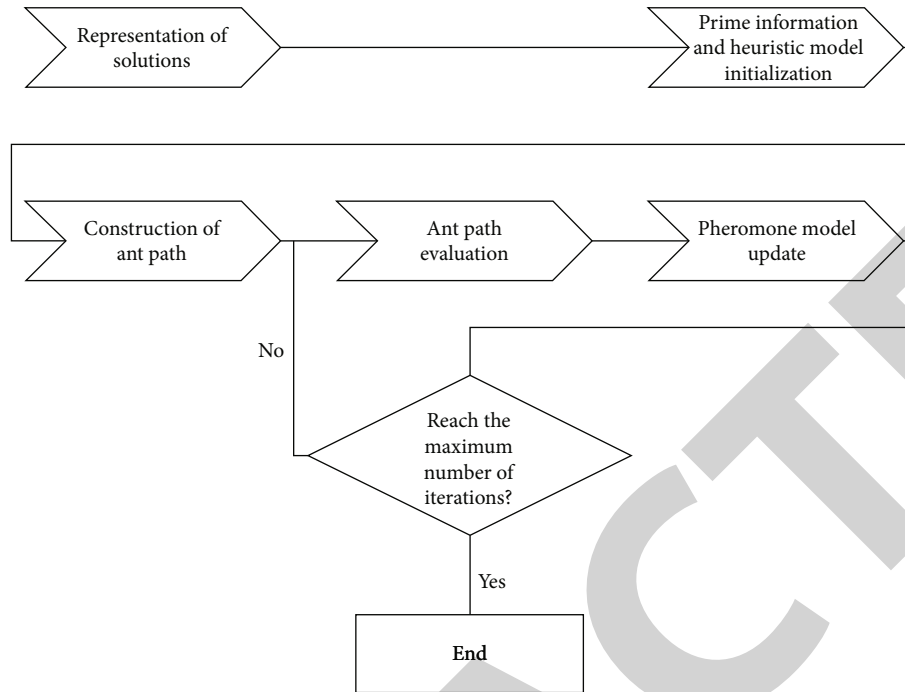


FIGURE 3: Ant colony optimization algorithm framework.

issues. Thus, it is not possible to project a reasonable proportion of hardscape and green landscape. Thus, in the indirect representation of the real number encoding scheme, the real value stored in the particle only indicates the tendency of the element in the particle to take a certain value, and according to these tendencies, the solution arrangement is obtained by “decoding.”

3. Application Analysis of Group Intelligence Algorithm in Energy-Saving Cold Landscape Design

3.1. Analysis of Group Algorithm in Landscape Design. The population algorithm provides a general framework for solving optimization problems of complex systems, which does not depend on the domain and type of the problem. At this time, there is a complex coupling between the decisive variables for solving the problem and the objective function, to find the influencing factors according to the problem analysis, the influence of these factors on the objective function by some combination to find a set of optimal coefficients, that is, the solution of the problem. The trend of the experimental acceleration ratio of the algo for group intelligence is shown in Figure 4.

First, the decision variables and their various constraints are determined, i.e., the individual manifest X and the solution space of the problem are determined. The initial population is generated randomly: for the binary-encoded population, the bit strings are generated randomly according to the encoding length of the solution. Sometimes it is not necessary to search all the solutions; for example, the characteristics of the fitness function (linear) and the search space (convex region) in linear programming only need to search

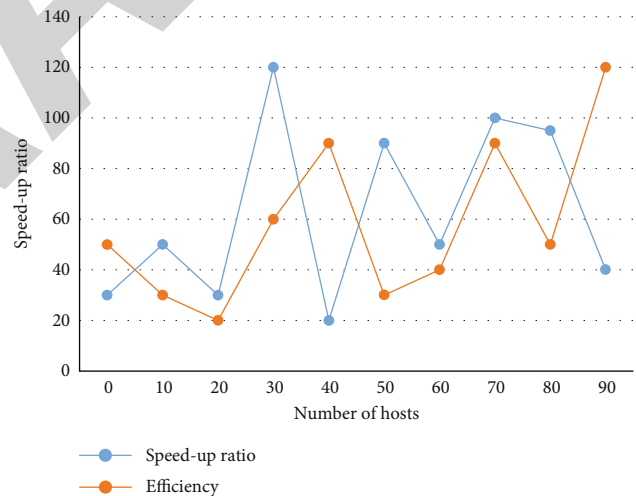


FIGURE 4: Trend diagram of experimental acceleration ratio of group intelligence algorithm.

along the boundary of the search space. There are also methods such as dynamic programming and branch-and-bound, which can cut off part of the solutions in the search space and only need to deal with part of the solutions. Ecological landscape is different from landscape design that simply aims at visual enjoyment or meeting functional needs but elevates the landscape to the height of protecting the ecological balance of the city and improving the whole human ecosystem and living environment. Use integers to store the position of each of its elements in its solution arrangement rather than the arrangement itself. The velocity of each particle is formed by the left-right movement of its element positions, thus forming a judgment matrix for the velocity

TABLE 1: Criteria level index judgment matrix and weight.

Evaluation of conservation-oriented landscape design	B_1 water saving	B_2 energy saving
B_1 water saving	0.549	0.482
B_2 energy saving	0.784	0.652
Weight	0.128	0.336

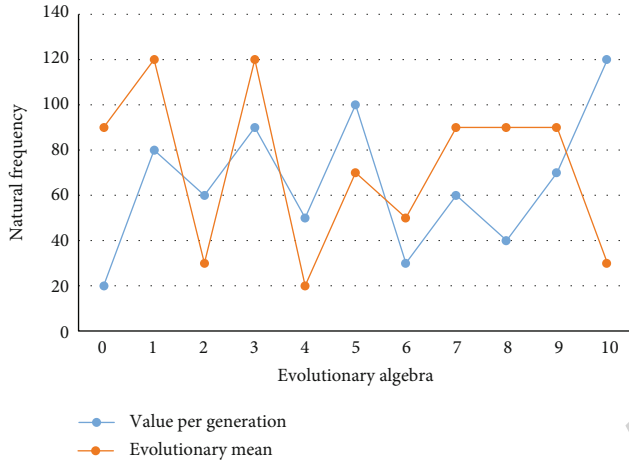


FIGURE 5: Evolution curve.

of the particle. Based on the obtained judgment matrix, the values were entered into the computer software Yaahp, and after the consistency test $CR < 0.05$, the weight values of each element for the elements of the previous layer were obtained, and the results are shown in Table 1.

Second, the individual encoding method representing the feasible solution is determined, i.e., the individual encoding type B and the search space of the population algorithm are determined. Due to the limitation of population size and the length of bit strings, each bit string has an intrinsic probability of being generated. The larger the population, the higher the probability of possible patterns and bit strings appearing, and vice versa will be smaller. The evolutionary process of the algo for group intelligence is shown in Figure 5, where the evolutionary mean is 2109.1 Hz, and Figure 6 shows the deviation curve of the algorithm from the SVM in this chapter.

Even when the defined fitness function is discontinuous, nonregular, or noisy, the population algorithm can find the global optimal solution with a high probability. The design approach incorporating ecological theory creates landscapes that focus on the combination of ecology and art, adapt to modern society, have high artistic standards, and incorporate ecological ideas and techniques. The particle swarm algorithm of continuous type is then used to update the position of each element in its arrangement, which forms a new position of the particle expressed in real values so that the selection is focused on searching in promising regions of the search space. The genetic material in the population is then changed by crossover or mutational recombination in order to develop new points in the search space.

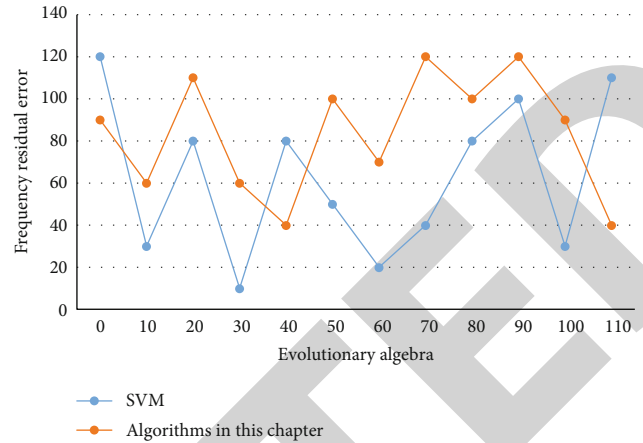


FIGURE 6: Deviation curve.

Finally, the decoding method is determined, i.e., the correspondence or conversion method from individual encoding type B to individual expression type X is determined. Then, the intermediate results of these real values are sorted by using the principle that the element with the smallest value is the first to be scheduled, and the relative position of each element in the arrangement is determined to obtain the updated solution arrangement. When the current position of a particle is greater than its best position, the velocity of the particle decreases (current velocity - random number multiplied by a system parameter); and vice versa increases (current velocity + random number multiplied by a system parameter). In this rule, the largest real number will be selected first as the first position of the new scheduling scheme, and the slightly smaller real number will be selected as the second position of the new scheduling scheme. There is no any special requirement for the search space (such as connectivity and convexity), and only adaptive information is utilized without other auxiliary information such as derivatives, which allows a wider range of adaptation. In ecological conservation design, the positive significance of ecological conservation design is that it closely connects ecological research with landscape design and establishes a scientific view of design theory. The speed of convergence of an evolutionary algorithm as an optimization application is basically determined by the selection operator or, more precisely, by the selection pressure exerted on the search process by the selection operator. Different coding methods and different operations of the population operator are required for different optimization problems. They are closely related to the specific problem being solved, and thus, the degree of understanding of the problem being solved is crucial to the success of the population algorithm application.

3.2. Analysis of Swarm Operators in Group Intelligence Algorithm. The evaluation function of the algo for group intelligence is a function that reflects the degree of merit of the solution expressed by the individuals in the group, which is equivalent to the fitness function of the evolutionary algorithm. The population operator is the operator that operates on the population in the population algorithm, which reflects some activity process of the population. In the algo

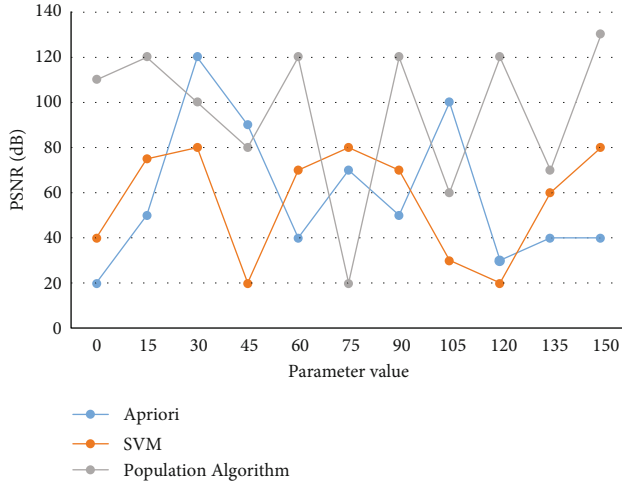


FIGURE 7: Simulation comparison of three different algorithms.

for group intelligence, the components of position and velocity are no longer a single number but a fuzzy set of possible values of this component. Since the update of the pheromone directly affects the selection probability, if the update of the pheromone is made more diverse, the selection of the path will be diversified. Therefore, the simulation results of the Apriori algorithm, SVM algorithm, and population algorithm are compared, as shown in Figure 7.

First of all, for the objective function and the evaluation function, a characteristic of the genetic algorithm is that it uses only the objective function value of the problem to obtain the relevant search information for the next step. Genetic algorithm processes multiple individuals in the population at the same time, that is, it evaluates multiple solutions in the search space, which reduces the risk of falling into the local optimal solution. At the same time, the algorithm itself is easy to realize parallelism. Genetic algorithm basically does not need to search the knowledge of space or other auxiliary information, but only uses the fitness function value to evaluate individuals; on this basis, genetic operation is carried out. The use of the objective function value is reflected by evaluating the fitness of the individual. An appropriate fitness function is constructed to transform the chaotic system parameter problem into a multidimensional optimization problem, and then the global search capability of the fuzzy discrete particle swarm algorithm is used to solve this optimization problem. One advantage of the fuzzy discrete particle swarm algorithm is that it can maintain the continuous nature of the particle swarm algorithm, since the particle swarm algorithm has excellent performance in solving continuous domain problems. By introducing fuzzy sets, it is hoped that discrete problems can be solved in a similar way to continuous problems without losing the efficiency of the particle swarm algorithm. For the slab type building, which is more commonly used in residential building design, when the building length is determined to be less than the optimal side length corresponding to a certain floor area, the building length should be reduced as much as possible, while the depth of the building should be increased. This will greatly reduce the build-

TABLE 2: Group intelligence algorithm.

	f_1	f_2	Average optimization percentage
Algo for group intelligence	27.19%	46.22%	36.71%
Apriori	18.37%	28.15%	23.26%

ing's volume factor, which is very beneficial to the energy-saving design of residential buildings. For the algorithm in this chapter and Apriori algorithm, $f_{ij}(A)$ refers to the optimal fitness value of the algorithm in the case of particle number i and space dimension j . The optimization percentage results of the two algorithms are shown in Table 2.

From the optimization percentages of the two algorithms compared in Table 2, it can be seen that the algo for group intelligence outperformed the other algorithm in terms of solving power by 13.45% in terms of average optimization power; thus, the algo for group intelligence demonstrated its superior ability to take into account both local and global search.

The second is the fitness scale transformation, in which the probability of each individual being inherited into the next generation population is determined by the fitness of that individual in the genetic algorithm. In high-rise residential buildings, the energy consumption index of slab-type residential buildings with north-oriented enclosed traffic corridors is about 7% lower than that of multistory slab-type residential buildings. Under the condition of similar building area, the heat consumption index of high-rise tower residential buildings is about 11%-17% higher than that of high-rise slab residential buildings, and the tower residential buildings with complex body shape and too many concave and convex surfaces are not good for energy saving. The score of underground space utilization is the lowest, because in terms of urban parks, the landscape is usually above ground, and the demand for underground space is less, and it is often used as a parking lot. However, the use of underground space should not be neglected. For parks with limited space and carrying capacity, underground space can be used to create more diverse places. In order to make the group algorithm can directly solve the scheduling problem, a maximum ranking rule is proposed, which uses each location component value between objects that has a size order relationship. And it is combined with a random key encoding method generated to transform the successive positions of objects into a feasible scheduling scheme. The process of the scheduling scheme is different from that of the normal particle swarm algorithm in that two operations are added to the algorithm, one is to use the normalization operation after the position update to constrain the position values to the interval $[0, 2]$, and the other is to call the defuzzification operation to obtain an integer vector before calculating the fitness in order to solve the fitness function later. The results are shown in Figure 8 for single-score samples with nominal values of $1.20 \mu\text{m}$ and $2.40 \mu\text{m}$ for different set thresholds of the population operator.

Finally, the method of dealing with constraints, when dealing with constraints, only for specific application

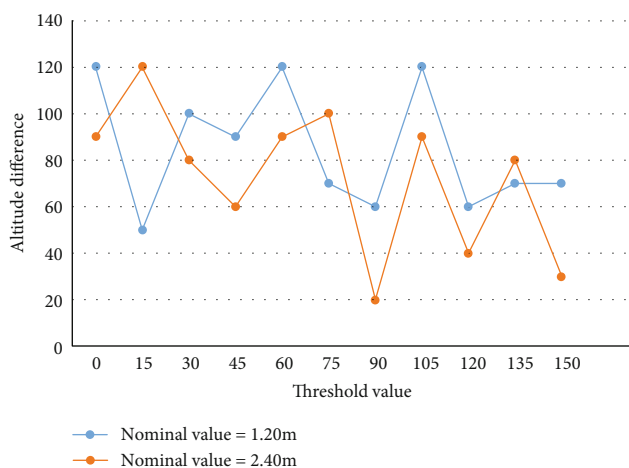


FIGURE 8: Relationship curves of error thresholds of group operators with different nominal values.

problems and the characteristics of constraints. Then, considering the computing power of genetic operators, different processing methods are selected in genetic algorithm. The energy-saving chilling landscape design uses a domain search algorithm. The key idea of this algorithm is to exchange the positions of two adjacent operations, and when the current exchange produces a solution that is better than the original one, we accept this exchange and update the value of the solution. This method does not change any part of the particle swarm algorithm calculation, but only converts the particle positions from real space to integer space when computing the fitness function. The conversion is very simple, either by taking the integer part of the particle positions directly or by rounding the function. The weight size is related to the reasonable traffic organization, and the smooth road planning and space division can maximize the space utilization and avoid unnecessary land waste, so the reasonable traffic organization is crucial.

4. Conclusions

Urban public landscape design has become an important design area of concern for urban builders at this stage. With the booming development of cold cities and the influence of ecological construction and low-carbon economy in the world, the sustainable development of cold cities has attracted much attention. Energy saving is the objective requirement of modern ecological civilization. Modern landscape design has emerged as a result of the dual destruction of nature and human body and mind by industrialization. The most fundamental difference with previous gardening is that its main creation object is the overall human ecosystem. The structure of the algo for group intelligence is simpler, the parameters of control are less, the robustness of the algorithm is stronger, etc. Facing the current situation of resource scarcity in China, landscape greening should implement the scientific concept of development, follow the requirements of the construction of an energy-saving society, and put energy conservation into concrete implementation in landscape design. This paper summarizes the

relevant energy-saving strategies based on the climate of cold regions and the thermal characteristics of residential buildings. The first part of this paper introduces the background and significance of the research and then introduces the main work of this paper. In the second part, the method of energy-efficient cold landscape design based on algo for group intelligence and the determination of the weight of energy-efficient cold landscape design evaluation indexes are reviewed. The third part is the core of the paper, from the analysis of group algorithm in landscape design and the analysis of group operator of group intelligent algorithm, completing the description of the application analysis of group intelligent algorithm in energy-saving cold landscape design. The last part of the paper is the summary of the whole work.

However, this algorithm has some limitations. When the algorithm is used in the overall architectural scheme design, it lacks the unified consideration of architecture and landscape. The unification of architecture and landscape is the ultimate goal of current urban planning. Whether from architectural design or landscape design, we can not only consider a single factor one-sidedly. It needs to be analyzed in future ide 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 author declares no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the 2020 Basic Scientific Research Business Fee Scientific Research Project of Heilongjiang Provincial Colleges and Universities: construction of intangible cultural inheritance and training base for college students' ice lantern and ice sculpture making skills (high-level key project, Project No. 2020-kyywf-0908).

References

- [1] Q. Wenjin and C. Weineng, "Design and application of group intelligence algorithm for smart city," *Artificial Intelligence*, vol. 5, p. 15, 2021.
- [2] D. Mingqing, "Research on environmental protection and energy saving design method based on modern urban public landscape," *Environmental Science and Management*, vol. 44, no. 12, p. 5, 2019.
- [3] D. Xuechao, "The development trend of group intelligence algorithm from the perspective of chicken swarm algorithm," *Digital World*, vol. 5, p. 1, 2017.
- [4] Z. Lei, *Discussion on application measures of energy-saving walls in villages and towns in cold regions*, Housing and Real Estate, 2019.

Retraction

Retracted: Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Liu and Q. Zhang, "Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology," *Journal of Function Spaces*, vol. 2022, Article ID 4292063, 10 pages, 2022.

Research Article

Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology

Yutang Liu ¹ and Qin Zhang ²

¹School of Science, Henan Institute of Technology, Xinxiang, Henan 453003, China

²School of Mathematics and Statistics, Xinxiang University, Xinxiang, Henan 453003, China

Correspondence should be addressed to Qin Zhang; zhangqin@xxu.edu.cn

Received 12 May 2022; Revised 8 July 2022; Accepted 15 July 2022; Published 2 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Yutang Liu and Qin 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.

Attribute encryption is an effective one to many network communication technologies, which supports flexible access control strategies and is very suitable for fine-grained access control in large-scale information systems. In order to improve the attributes of static data, encryption technology can provide a reliable technical guarantee for network security. This paper presents a mathematical modeling method of static data attribute encryption based on big data technology. The big data redundancy elimination algorithm based on similarity calculation is analyzed. By using static data attribute encryption based on big data technology, the length of encrypted data packets will not increase, and partial redundancy of fragments can be eliminated, which can greatly improve the efficiency of the system. The attribute-based encryption mechanism uses attributes as public keys, and the decryption user is a group; so, the encryption efficiency is very high. It can realize efficient encryption and decryption, as well as flexible access control based on user attributes. This scheme can reflect the importance of attributes; so, it is more practical.

1. Introduction

Along with the continuous reform and opening up, China's economy has significantly improved, Internet technology has been widely popularized, and the increasing level of mobile device technology has promoted the further expansion and exponential growth of data volume [1]. At the same time, information technology attacks have emerged continuously, and users' private files cannot be protected by traditional password methods alone, which is a great security risk [2]. Lingfeng proposes a new big data encryption algorithm based on data deduplication technology. This method eliminates the redundancy in big data processing and fully reflects the advantages of elliptic curve encryption algorithm, such as high security and short key. At the same time, combined with CTR working mode, it fully shows the advantages of good parallelism and fast speed [3]. Big data are data sets with huge amount of data and many types of

structures, which are difficult to be processed using existing relational databases or data processing tools [4]. Encrypting data on the basis of big data properties is one of the effective methods to control access to data users in a data outsourcing environment [5]. The characteristics of big data are mostly obvious, such as large volume, diversity, and complex sources, and some of them contain a large amount of information content with certain value [6]. Traditional encryption techniques convert private data information into unrecognizable ciphertext data, losing many semantic features of the data and rendering conventional data processing methods ineffective [7]. Thus, it severely hinders the data computing and processing services provided by cloud service providers to their tenants and is not applicable to the protection of private data in cloud platforms.

Therefore, the impact of factors such as security level and number of attributes on the algorithm should be analyzed in depth through the in-depth study of attribute

encryption, combined with the actual needs of engineering applications using programming simulation in terms of time overhead, communication overhead, CPU, and memory overhead [8]. Database system as the core component of information system and database files as the aggregation of information, obviously, its security will be an important indicator of information system security. This reinforces the need to protect not only economic information in business communications but also personal information in communications [9]. Databases must provide data protection measures and cannot leak information to unauthorized parties. For the present, there are various ways to encrypt network data attributes, including link encryption, node encryption, and end-to-end encryption. However, in the current distributed and network-based application environment, users access the system in various ways, and the database system faces various security threats [10]. So, these several network data encryption methods rely on traditional data encryption, but traditional data encryption and decryption techniques rely on application and individual operations. It is important to go through manual operations, which inevitably leads to leakage of important data through the network by human.

The explosive growth of data volume has brought us into the era of large-scale data processing, which is characterized by high computing intensity and the need for efficient concurrent computing and storage capacity. By using the attribute encryption system, the authentication and access rights of users in the system are no longer described by a single identity or certificate, but each user has a set of attributes and a set of keys corresponding to the group of attributes. However, the widely used encryption methods have different performance, and they need to be located according to the actual encrypted data object. If this method is constantly updated, it will affect the overall algorithm execution efficiency and increase the time overhead. As an extension of the identity-based cryptosystem, attribute-based cryptosystem has attracted more and more attention because of its special application significance and a wide range of usage scenarios.

The innovation points of this paper are as follows.

- (1) The adopted redundant data detection technology can transparently encrypt and decrypt the network data in the passing intermediate layer, allowing users to customize filtering encryption rules and specify encryption algorithms
- (2) This paper constructs a secure and scalable mathematical model of static data attribute encryption around attribute encryption technology, combining specific data types and realistic scenarios
- (3) In In-depth study of database encryption technology and extended stored procedure technology, the paper proposes a redundancy elimination algorithm for big data based on similarity calculation, which enables application developers to have "transparent access" to ciphertext data, thus greatly reducing the difficulty of application development

2. Mathematical Modeling Idea of Static Data Attribute Encryption Based on Big Data Technology

2.1. Encryption Model Classification Method. Mathematical modeling is the process of describing actual phenomena with mathematical language. The actual phenomena here include both concrete natural phenomena such as free fall and abstract phenomena. For example, the value tendency of customers to a certain commodity. The description here includes not only the description of external forms and internal mechanisms but also prediction, experiment, and interpretation of actual phenomena. Through teaching, students can understand the whole process of using mathematical theories and methods to analyze and solve problems and improve their ability to analyze and solve problems. Improve their interest in learning mathematics and their awareness and ability to apply mathematics, so that they can often think of using mathematics to solve problems in their future work. Improve their awareness of making full use of computer software and contemporary high-tech achievements and be able to organically combine mathematics and computer to solve practical problems. Encryption systems are usually divided into three categories: symmetric encryption systems, asymmetric encryption systems, and hybrid cryptosystems. Static data attribute encryption is characterized by the fact that the sender of the message only needs to encrypt the message according to the attribute and does not need to know the number and identity of the user, which reduces the cost of encryption and protects the privacy of the user. The two subjects involved in data transmission can be summarized as follows: the transmitter (sender) and the receiver of the information. Network database encryption can be broadly divided into two ways: external encryption and internal encryption, as shown in Figure 1.

When both users need to communicate, the sender encrypts the plaintext packets with an encryption algorithm to form complex unreadable encrypted data and then sends it to the recipient through the medium. If A sends a message to B , the verification process of its signature is shown in Figure 2.

The first is the symmetric encryption system, in which the key used for encryption is exactly the same as the key used for decryption. The encryption algorithm is the process of generating an unreadable ciphertext from the original readable plaintext through a series of operations and transformations with the key. For the unsigned person j , a randomly selected value x_j on Z_p generates the corresponding signed private key for the attribute i as

$$T_{i,j} = g_1^{x_j/(y+t_i)}. \quad (1)$$

If any of the commands in the transaction fails to execute, then the entire transaction needs to be undone, including the commands that were executed successfully. The state of the database then reverts back to the state before the transaction was executed. When used for general encryption functions, the public key is used to encrypt the plaintext and

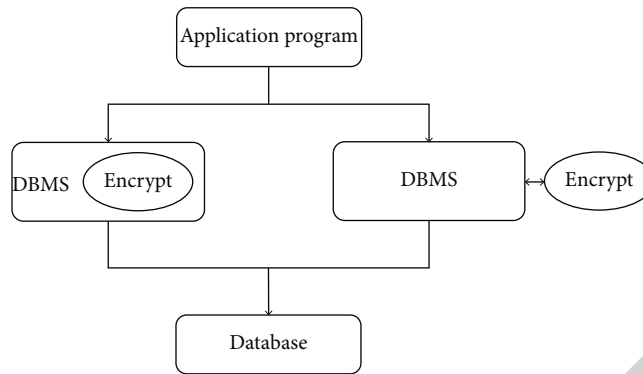


FIGURE 1: Database encryption mode.

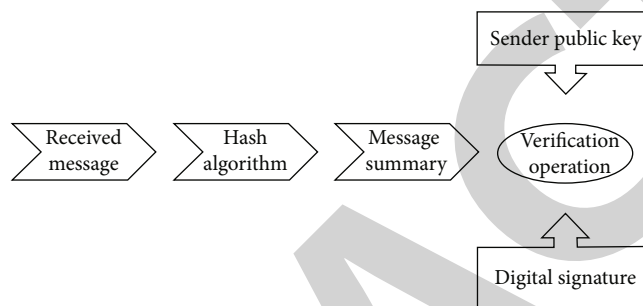


FIGURE 2: Verification process of signature.

the private key to decrypt the ciphertext. Authentication is the security portal of the network database system and the basis for access control. In general, the available measurement information will depend on the current moment. For example, if there is complete data from 0 to τ moment, then the measurement information can be written as follows:

$$P_{\text{data}} = \left\{ (r, y, u) \in R \times U \times Y \mid P_r \begin{bmatrix} u - u_{\text{data}} \\ y - y_{\text{data}} \end{bmatrix} = 0 \right\}, \quad (2)$$

where P_r is the time truncation operator, i.e.,

$$P_r x(t) = \begin{cases} x(t), & 0 \leq t \leq \tau \\ 0, & \text{other} \end{cases}. \quad (3)$$

When the plaintext is encrypted with the private key and the ciphertext is decrypted with the public key, it can be used as a digital signature. The encryption step is called signature, and the decryption step is called verification. Therefore, the network database encryption system must provide a way to identify the user, and the user must provide a username, password, or other relevant security credentials, such as terminal key and user USB Key, in accordance with the system security requirements. If it is not determined behavior information system, it means that the environment has changed; otherwise, it means that the environment has not changed

and has defined the environment trigger function:

$$h(s) = \begin{cases} 0, & (a(x) = a(y)) \wedge (d(x) = d(y)), \\ 1, & (a(x) = a(y)) \wedge (d(x) \neq d(y)). \end{cases} \quad (4)$$

$a(x)$ is the action selection of behavior x , and $d(x)$ is the behavior x acceptable or unacceptable.

The next is the asymmetric encryption system, which is different from the symmetric encryption algorithm, and the asymmetric encryption algorithm needs two keys: public key and private key. At any point of time, the data in the database accessed by the user should be unique and up-to-date, and the integrity of the relational data and the consistency of the business logic cannot be destroyed by the transactions executed in the database. In addition to this, there is the escrow policy for passwords, and when the encryption algorithm is handed over to a third party for management, it must be strictly enforced in a controlled manner. The duplicate detection method is adopted to detect the relevant data that need to be encrypted, and all duplicate data among all data are deleted. That is, different records and different fields of each record in the database are encrypted with different keys. This is complemented by checksum measures to ensure the confidentiality and integrity of the database data storage and to prevent unauthorized access and modification of the data. Verification technology or password technology shall be adopted to ensure the integrity of important data during transmission, including but not limited to identification data, important business data, important audit data, important configuration data, important video data, and

important personal information. Then, in the hash table, the characteristic information of the repetitive data is stored. Let G_1 be a bilinear group of order prime p and g be its generating element, and the Lagrangian parameter is expressed as

$$\Delta_{i,s}(X) = \prod_{j \in S, j \neq i} \frac{x-j}{i-j} \quad (5)$$

S represents Z_p .

Finally, there is the hybrid cryptosystem, which is a combination of symmetric and asymmetric encryption. When a transaction is successfully executed, the resultant changes to the database are permanently preserved, and even if the system crashes, the database is still able to recover to the state it was in after the transaction was successfully executed. The main point of the process is to use symmetric encryption to actually encrypt the plaintext and then encrypt the symmetric encryption key using the public key. Since the key length is shorter than the message, the defect that the processing speed of public key encryption is lower than that of symmetric encryption can be solved. In the database, not all data need to be encrypted, and only the sensitive data of users can be encrypted to improve the database access speed. Nonrepudiation means that in the process of information interaction between the two parties of communication, it must be confirmed that the information of both parties has the true unity. That is, it is impossible for all participants to deny or repudiate their true identity, as well as the originality of the information provided and the operations and commitments completed. In the attribute-based encryption mechanism, the sender controls the access policy, and the more complicated the access policy is designed, the more massive the public key is, and the more difficult it is to prove the security of the algorithm. Therefore, the network database encryption system must provide encryption setting function, so that users can set the database, table, field, etc. that need to be encrypted, which is conducive to the user's autonomy to balance between efficiency and security.

2.2. Mathematical Model Structure of Static Data Attribute Encryption. Big data refers to the technical architecture that extracts its value by capturing, computing, and analyzing a large amount, type, and source of complex data at high speed using economical methods. The database encryption system includes three functional parts: data reader, password manager, and encryption manager. Its structure is shown in Figure 3.

First is the data reader, receives the request and the data from the application, carries on the syntax analysis, and carries on the corresponding operation according to the user's request. After receiving the encrypted information, the data receiver can decrypt the encrypted information only by the same key and encryption algorithm and then read the data to make use of it. For a set of attributes, the t_i on Z_p is randomly selected as the private key of the attributes, and the public key corresponding

to the attributes is published as

$$\{T_1 = g^{t_1}, \dots, T_{|U|} = g^{t_{|U|}}\}. \quad (6)$$

Using the full file detection algorithm, the hash table is initialized with an operation that sets a single file to granularity, detects all data, and finds duplicate data. The data must be encrypted before it is transmitted by the data transmission node and then decrypted after it is received by the data receiving node. Then, it needs to be encrypted again before it can be transmitted out, and here, it is important to note that the key used for reencryption is not for this link but for the next one. The physical NIC driver communicates directly with the transport layer driver, and the network packets are passed between them through the NDIS interface without any processing. When encryption is implemented at the application level, the encryption of data can be done by the application first, and then the encrypted data is transferred to the DBMS and stored in this encrypted form. The control law of the DBMS is

$$u(x) = k_p e(t) + k_i \int_0^t e(t) dt + k_d \frac{de(t)}{t} \quad (7)$$

or in the form of transfer function:

$$G(s) = k_p + \frac{k_i}{s} + k_d s. \quad (8)$$

k_p is the scale factor, k_i is the integral coefficient, and k_d is the differential coefficient.

Next is the password manager, which is the core part of the database encryption system and is responsible for completing the encryption and decryption of the data in the background, transparent to the user and the application. For securing the data, access control (i.e., authentication and authorization) has proved useful, as long as the data can be accessed using a predefined system interface. Therefore, it is required that both the decrypting party, i.e., the recipient and the sender, must have the encryption key and the encryption algorithm. Comparing the stored data in the hash table with the obtained hash function value, if an exact match is obtained, the file will be replaced by a pointer; if it is difficult to match correctly, the file can be stored. Byte substitution, row shifting, and column mixing operations are simple and reversible, and the corresponding inverse functions are used in the decryption algorithm. The round key plus step will reverse the operation on the same round key heterogeneous data group when encrypted. Under this model, the security of a cipher depends only on the security assumption on which the scheme is based, and unless that security assumption is breached, the security of the cipher can be proven to be unbreakable. If the variance of the cost function is negative, the original center of mass is replaced with a noncentral center of mass at the current location; otherwise, the center of mass remains unchanged. The fitness is suitably extended by the simulated annealing algorithm, and

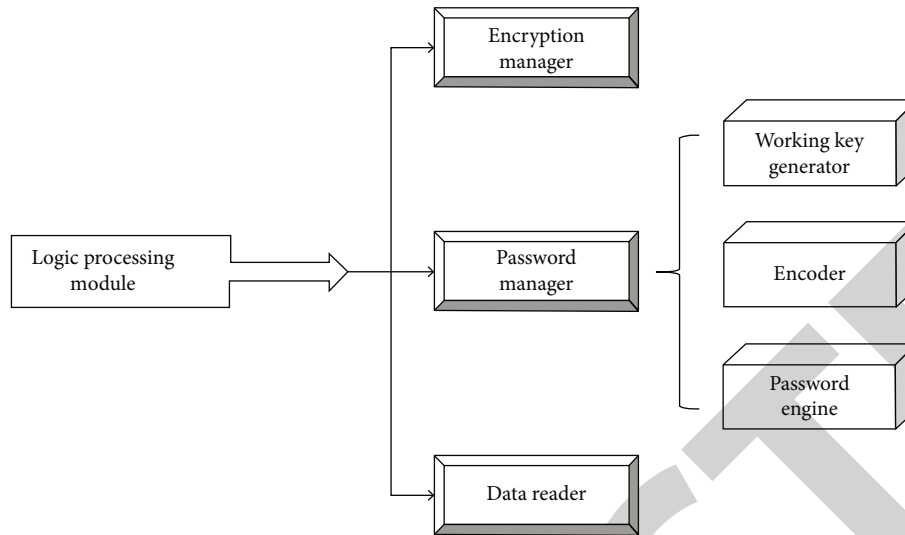


FIGURE 3: Structure diagram of the encryption system.

the fitness stretching is done as follows:

$$f_i = \frac{e^{f_i/T}}{\sum_{i=1}^M e^{f_i/T}}. \quad (9)$$

f_i is the fitness of the i individual.

Finally, the Encryption Manager is the module used by the system administrator to implement the encryption dictionary management functions and to interact with the encryption dictionary. It is responsible for the definition, modification, and deletion of the encryption dictionary. Both the user's private steel and cipher text are associated with attributes. The solution supports attribute-based threshold access policy, i.e., the user can only decrypt the cipher text when the number of attributes intersecting the set of attributes in the user's cipher text reaches the threshold value set by the system. In the case that the fingerprint matches with the matching condition, the block boundary can be set at the window, as well. For the set ω if $|\omega \cap \omega'| \geq d$, then select any d elements belonging to the intersection of the two sets and use Lagrange's difference theorem to get the following:

$$\frac{E'}{\sum_{i \in S} (e(D_i, E_i))^{\Delta_i, S(0)}} = M \quad (10)$$

The provider of the data resource must use the public key of each user in the receiving group when encrypting the message. Then, the ciphertext is sent to the corresponding users separately, which leads to problems such as high processing overhead and high bandwidth consumption. Therefore, use mandatory access control, a mechanism that restricts access to an object based on the sensitivity of the information contained in the object and whether the subject has formal access authorization to the object of that sensitivity. At this point, you can first publish the public key encryption key on a reliable third-party website and let your friends

obtain this public key. The receiver decrypts the digital signature with the sender's public key to obtain the decrypted information summary. And the hash algorithm used by the sender is used for the obtained message, so as to obtain the summary of the information. Compare the two information to determine whether the information has been modified. If the information has been modified, the two summaries must be different. Then, use this public key to encrypt the information and transmit it to you and then use the decryption key private key to recover the plaintext of the information for reading, in which the private key decryption key will not be transmitted in any form. When you write data, you do not need to maintain the relationship between data and data, and you do not need to fix the format of data in a table; in addition, you do not need to maintain the characteristics of ACID, which makes the performance greatly improved.

3. Analysis of Big Data Redundancy Algorithm Based on Similarity Calculation

3.1. Redundant Data Detection and Analysis. In the space of data structure, Bloom filter has the advantage of high data compression efficiency, and the feature values of this algorithm consist of Bloom filter data structure representation. Compared with the traditional data redundancy elimination algorithm, the Bloom filter algorithm has more advantages in query time and space efficiency and is more suitable for handling large data. The similarity calculation-based redundancy algorithm and CBC algorithm are shown in Figures 4 and 5 below to compare the time consumption of each encryption algorithm when the key length is 128 bit and 256 bit.

Firstly, the data in a data set with two or more duplicates are deleted to ensure that only the same data is retained in the final data set, so that the deleted redundant data are replaced by data pointers. The database files as a whole are encrypted with encryption keys and encryption algorithms to ensure the authenticity and integrity of all user data tables,

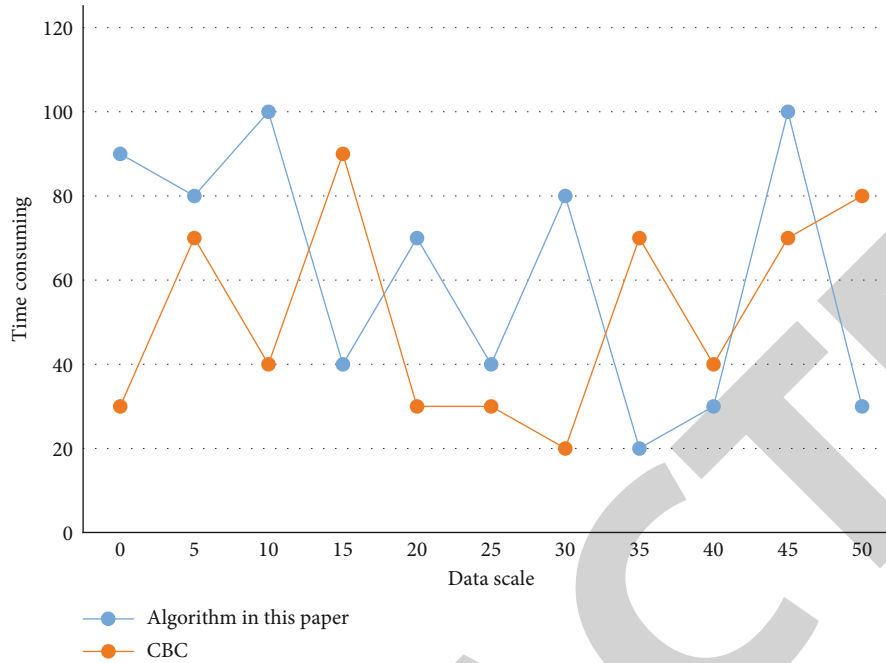


FIGURE 4: Comparison of time consumption of each algorithm when the key length is 128 bit.

system data tables, indexes, views, and stored procedures in the entire database. Data sharing is achieved by decrypting the entire database file with the decryption key. The information sent to others is encrypted with e , and as long as others can decrypt it with d , the information is proved to be sent by you, which constitutes a signature mechanism. It can be encrypted by AES advanced encryption standard using deduplication technology first. Since CTR mode encryption and decryption is fast and both the key and group length can be varied, it can be used as the working method of the group cipher algorithm. Moreover, with such an encryption method, the encryption process in database applications is transparent and can be used directly. Proxy reencryption is very suitable for use in scenarios where encrypted data is to be shared with multiple parties. There is no need to share private keys with recipients, and there is no need to encrypt the entire message for each recipient before proxy. This encryption allows the user to encrypt only once and then authorize the recipient according to his public key. In the proxy re encryption scheme, there is a semi-trusted proxy with encryption. It is able to convert the ciphertext encrypted with the public language to the ciphertext encrypted with the public key, but the proxy does not obtain any information about the plaintext during the whole process. To achieve strong backward security requires updating the keys of the underlying encryption. A comparative analysis of the additional communication burden and computational burden required when the underlying encryption is stream cipher, group cipher, and public key cipher, respectively. The comparison of the extra burden required to update the underlying key for different underlying encryption schemes is shown in Table 1 below.

Secondly, the ratio of the number of bytes before redundant data deletion to the number of bytes processed is

mainly used in determining the reduction rate of data to achieve. The group length and key length are both variable, and only to meet the requirements are the processed group size limited to 128 bits, while the key length is 128 bits or 256 bits various options. With table-level encryption granularity, the query performance of the system will be improved. Because the system performance will not be affected for the queries of unencrypted tables, and for the queries of encrypted tables, only the corresponding encrypted tables need to be decrypted instead of decrypting the whole database. Before the data reaches the target endpoint, the data has to be transmitted through a large number of link nodes. The serialization module completes the interconversion between byte stream and ABE data types, and it writes the structure variables output by other modules to byte arrays. When other modules need a specific type of variable, it reads the data from the byte stream to generate the variable of the target type. This encryption definition tool mainly defines the way to encrypt the data of each table in the database. After creating a certain database table, this tool defines the encryption method for the characteristics of that table.

Finally, all the possible attribute values are sorted according to some linear rule, and the corresponding statistics are then used to obtain a statistical histogram based on the attribute values. The entire record in the data table is encrypted and stored in the database file as a cipher text. The data block undergoes several data transformation operations, and each transformation operation produces an intermediate result, which is called state. Since the encryption and decryption keys are usually stored in the database, it also raises the risk of key management, which leads to the protection of the keys only through the access control of DBMS. The ciphertext data is decrypted, then encrypted using a different key, and then in transmission. In general,

TABLE 1: Comparison of the extra burden of different underlying encryption schemes when updating the underlying key.

Encryption mode	Communication burden	Client computing burden	Client computing burden
Stream cipher	17.26	$O(n)$	$O(1)$
DES	23.76	$O(n)$	$O(1)$
AES	31.74	$O(n)$	$O(1)$

the security of a cryptosystem can be divided into two forms: selective plaintext attack and selective ciphertext attack, depending on the attacker’s target. In a cryptosystem, we can prove that the cryptosystem is secure as long as the probability of a successful attack by an attacker is calculated to be negligible. The linear hybrid layer ensures a high degree of diffusion over multiple rounds, and the nonlinear layer consists of 20 S boxes juxtaposed to act as an obfuscation and the key encryption layer heterogeneous subkeys to intermediate states. Since the polynomial running time is feasible according to the Turing machine model, the probability is considered to be nonnegligible when it is the inverse of the polynomial.

Complexity of the cryptographic algorithm is as follows: the complexity of the cryptographic algorithm is a basic condition to ensure password security. If the cryptographic algorithm used in a cryptographic system is not complex, or it seems to be complex, but there are institutional weaknesses, it is easy to be exploited by attackers. In addition to the complexity of the cryptographic algorithm, key length is also the basic factor to ensure the security of cryptographic system. The simplest way to crack the key is to try all kinds of possible keys to see which one is actually used. In this attack, the number of passwords to be attempted is closely related to the entire key space to be retrieved.

3.2. Analysis of Big Data Encryption Algorithm. ECC, known as “ellipse curve cryptography”, is a public key encryption algorithm based on elliptic curve mathematics. The use of elliptic curves in cryptography is proposed independently. Different from the traditional encryption algorithm based on the problem of large prime number decomposition, this encryption method is based on the mathematical problem of “discrete logarithm.” The ECC algorithm is used to achieve encryption of big data, which has the advantages of low computational overhead and high encryption security performance in the encryption process, and is more suitable for research areas where computational power, space constraints, and bandwidth are limited. In describing the security proof, the simulator needs to simulate all realistic environments, including the use of realistic hash functions. It makes it possible to deceive the attacker to perform attack operations on the simulated environment. If the data is decrypted and then encrypted in the cloud, there is a risk of privacy leakage. The experimental analysis of the encryption algorithm and decryption algorithm on a laptop, encryption algorithm and decryption algorithm with num-

ber of attributes and number of keys, respectively, are shown in Figures 6 and 7 below.

First, when the receiver decrypts the encrypted data file, it needs to calculate the private key and the public key for decryption. The sender first encrypts the message using a symmetric cipher, then encrypts the key using the receiver’s public key, and then sends both to the receiver. When the user’s security level flag is higher than or equal to the security level flag of the data to be accessed, the data in the data table is decrypted using the three-level key, and the plaintext data is returned to the user. We can also use the same scheme to prove that if the attacker can recover the encrypted random number, then it can recover the information. The stream encryption algorithm theoretically encrypts increments of bits, but in reality, it usually encrypts every bit of the plaintext itself at a time. When decrypting the ciphertext, the user can only decrypt the message correctly if the set of attributes used for encryption matches the access control structure in the user’s private key, thus allowing the user to access the message content. To access the data in HDFS, the user needs to request a block access token from the Name Node, and only with the correct token can the user access the corresponding data block in HDFS, in order to test the relationship between mixed encryption time, attribute encryption time, and file size. In the experiment, the length of symmetric encryption secret key is fixed to 128 bit, and the encryption time varies with the file size as shown in Table 2 below.

Secondly, in the process of detection calculation, this paper will take the data file that needs to be encrypted individually as the granularity and detect the duplicate data in the data file. When an attribute revocation event occurs in the system, the system must ensure that the original private key of the user who owns the attribute is invalid; that is, the private key cannot decrypt the original ciphertext again. In CP-ABE, each user is divided into a group of describable attributes, and the unified attribute authorization center distributes the attribute private key to each user according to the attribute set of each user. The receiver decrypts the symmetric key using its private key and then decrypts the message using the key. For the user to write data to the database, when the security level flag of the user is lower than or equal to the security level flag to be written, a random key is generated and used as the tertiary working key to encrypt the data to be written to the database, the secondary key is used to encrypt the tertiary key, and the encrypted key is written to the database. In the initialization process, the main function of the key is to churn the S-box. Therefore, the application goal of the system is to ensure that it can support specific types of queries on ciphertext data and at the same time can protect the privacy of user data. The runtime of the Decrypt algorithm for IAI-CP-ABE1 and IAI-CP-ABE2 on a resource-constrained physical device was simulated in the Intel Edison development board, Outsourced Decrypt algorithm, and a comparison of the algorithmic efficiency of the Decrypt algorithm, and the big data encryption algorithm is shown in Figure 8 below.

Finally, the data files that do not duplicate each other in the complete file detection method are rearchived, and the

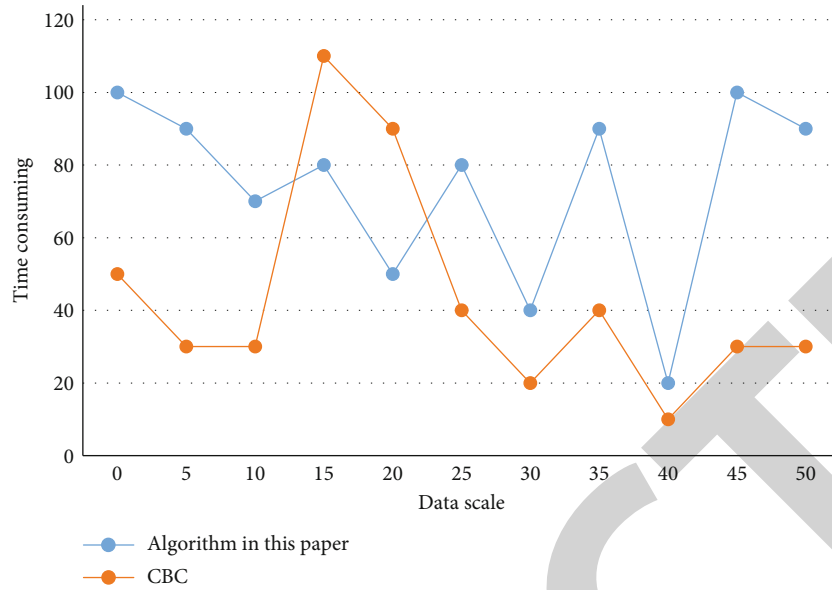


FIGURE 5: Comparison of time consumption of each algorithm when the key length is 256 bit.

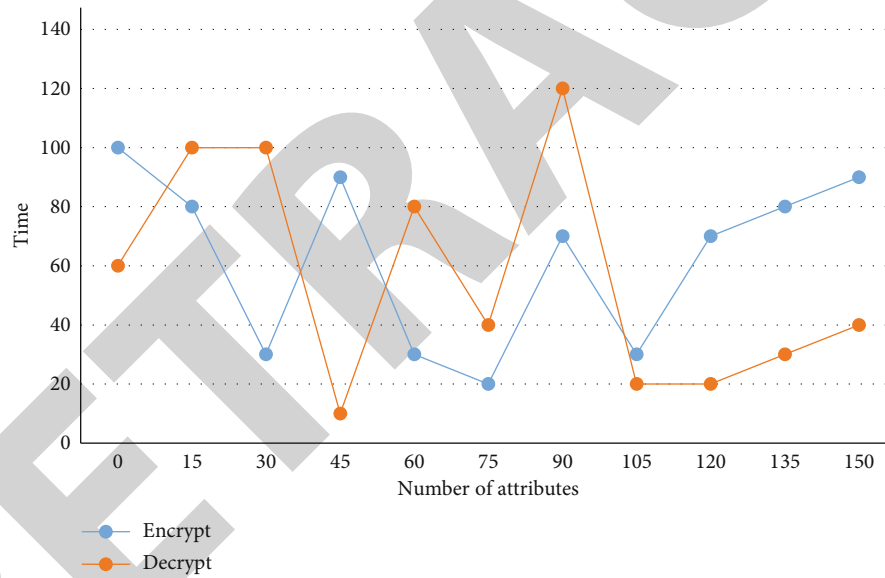


FIGURE 6: Relationship between Encrypt/Decrypt algorithm and the number of attributes.

TABLE 2: Experimental data of encryption time varying with file size.

File size (MB)	5	10	15	20	25	30
Encryption time (s)	17	28	37	42	58	63
Attribute encryption 128bit key time (s)	65					

process of this paper will use the CDC data block calculation method to archive them one by one from the file source. When the user accesses the data, the database system in the background must decrypt the data before returning it to the user. In order to respond to the user's request as soon as possible and reduce the user's waiting time, the speed of

data encryption and decryption must be fast. The data owner sets the access control structure and uses it to encrypt the information and then share it in the network, and other users can decrypt the information when and only when the set of attributes in their attribute private key matches the access control structure in the cipher text. In the process of cryptosystem security proof, we always end up with a hypothetical hard problem, such as BDH and CDH problem. If field-level, record-level, or data-item-level encryption granularity is used, the key is used to encrypt a three-level key. Due to the existence of partial weak keys, it makes the sub-key sequence to be completely duplicated in less than 10,000 bytes, and if it is partially duplicated, it may be able to be duplicated in less than 10,000 bytes. Therefore, the

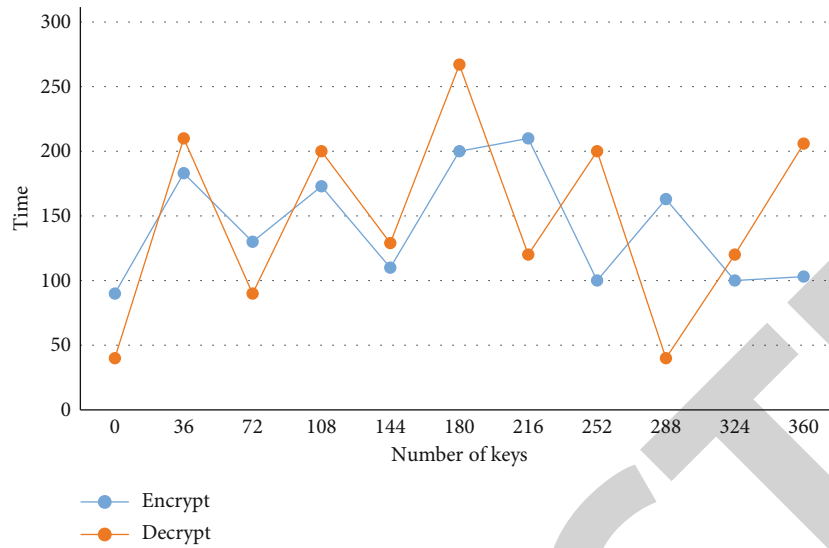


FIGURE 7: Relationship between Encrypt/Decrypt algorithm and the number of keys.

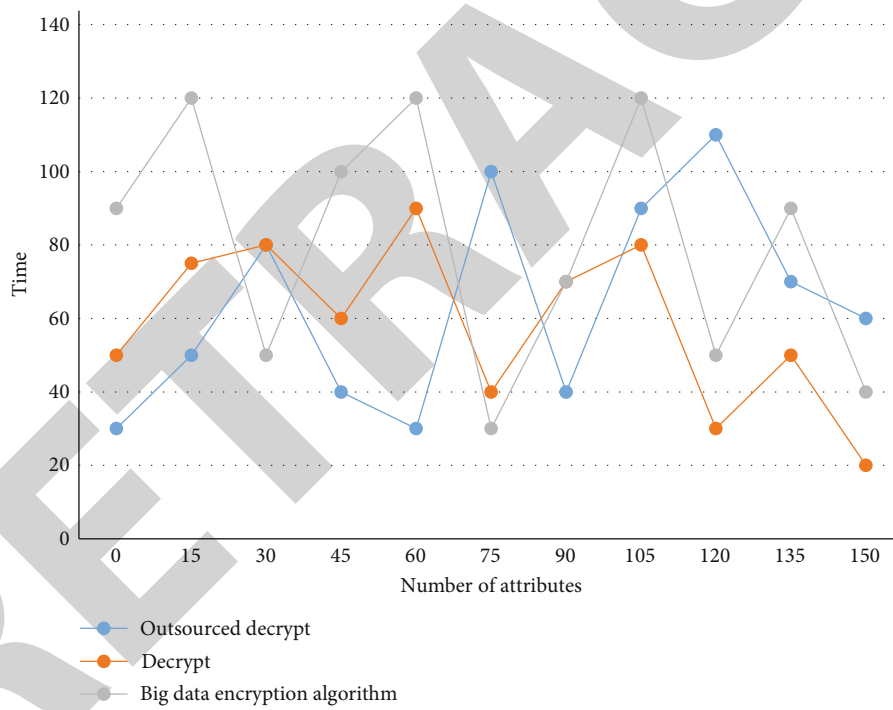


FIGURE 8: Comparison of algorithm efficiency between Outsourced Decrypt algorithm, Decrypt algorithm, and big data encryption algorithm.

encryption key must be tested to determine whether it is a weak key.

4. Conclusions

Network database is a new manifestation of database in the network era, which is a strong aggregate of network and database, and its application is very wide and has penetrated into various fields. Due to the huge amount of data in the big data environment, the traditional encryption scheme is inef-

ficient in the big data application, while the traditional access control mode is also difficult to apply to the big data environment. The degree of information security in the big data environment directly affects the privacy protection of users, and the superior features provided by encryption algorithms can compensate for the drawbacks caused by relying solely on software security prevention strategies, thus overcoming the difficulties and challenges faced by information security in a more efficient and stable manner. In this paper, we propose a mathematical modeling idea of static data attribute

Retraction

Retracted: Construction of a Knowledge Map Based on Text CNN Algorithm for Maritime English Subjects

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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 and A. Wei, "Construction of a Knowledge Map Based on Text CNN Algorithm for Maritime English Subjects," *Journal of Function Spaces*, vol. 2022, Article ID 6578682, 9 pages, 2022.

Research Article

Construction of a Knowledge Map Based on Text CNN Algorithm for Maritime English Subjects

Hui Wang  and Aimin Wei 

Marine School, Hainan Vocational University of Science and Technology, Haikou, Hainan 571126, China

Correspondence should be addressed to Aimin Wei; captalf@hvust.edu.cn

Received 16 May 2022; Revised 13 July 2022; Accepted 15 July 2022; Published 31 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Hui Wang and Aimin Wei. 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.

Knowledge map is a new method of knowledge management with the information revolution. This paper is aimed at forming a systematic and standardized huge redundant knowledge structure, which can be used to mine the knowledge structure and the relationship between knowledge and visualize it in a graphical way, in order to obtain more representative information and improve the classification accuracy of text classification model. In this paper, a knowledge map construction method based on the Text CNN algorithm is proposed for the subject of Nautical English. It is of practical significance and academic value to make use of knowledge map to study Chinese Maritime English, which is helpful to the development of Chinese Maritime English and provides guidance. In order to maintain the diversity of particle swarm optimization, the Text CNN algorithm is combined with the construction of Maritime English subject knowledge map, and the network parameters and structure are optimized. Using knowledge map to study China Maritime English has important practical significance and academic value and has certain guiding significance for the development of China Maritime English.

1. Introduction

Many foreign ship owners prefer to spend more money to hire Indian or Filipino crew with better English ability but less business ability [1]. As an important part of China's transportation industry, the maritime industry has a unique spatial and temporal scope of activities, is at the forefront of reform and opening up, and is characterized by participation in international logistics, international conventions and practices, internationalism with a profound world background, and extensive international commonality [2]. In the face of the increasingly fierce competition in the international crew labor market, the level of English proficiency of crew members has become an important factor to improve the competitiveness of Chinese crew members, which is directly related to the internationalization of China's maritime industry in the 21st century [3]. Accordingly, the teaching and training of listening and speaking English for maritime professionals require students to be able to use the target language to complete practical work tasks in a realistic working environment with the help of the knowl-

edge map of the maritime English subject and to improve the practical language skills [4].

The development of the discipline of Nautical English has now entered a mature stage; especially with the integration of postmodernist ideas, environmental-ecological perspectives, and ethical ideas, the content is constantly enriched, the number of literature increases all the time, and what is changing at the same time is the structure of the discipline's knowledge; the distribution of the author's academic community is gradually complicated and diversified with time and geography [5]. In addition, most students do not have a good memorization or learning method and just learn and memorize English knowledge points in a stereotypical and mechanical way, neglecting the intrinsic connection of knowledge points [6]. And the knowledge mapping which is aimed at discovering the historical evolution process of scientific theories and methods can be found by mining and analyzing the knowledge mapping relationships to discover: discipline structure characteristics, research hotspots, development sources, professional relevance and breakthrough achievements, and future development directions [7]. At present, informatization of teaching

knowledge can help more people learn more knowledge, and students can not only learn book knowledge in the classroom but also learn some other knowledge through the Internet after class time, which can well help students check the gaps [8].

Nowadays, an emerging cross-learning knowledge mapping, which is based on knowledge units and enables effective access to knowledge and quick grasp of the frontier areas of disciplinary knowledge, is emerging [9]. Knowledge mapping is a kind of graph showing the process of knowledge development and its interstructural relationships based on content analysis, citation analysis, and visualization [10]. It can describe human-owned knowledge resources and their carriers over time using visualization techniques, map, mine, analyze, and display scientific knowledge and their interconnections, and create a knowledge-sharing environment within organizations to promote collaboration and deeper scientific and technological research [11]. Text CNN is a machine learning model under deep supervised learning, and compared with traditional machine learning for manual feature extraction, Text CNN can automatically extract local features with shared weights, which is better than traditional machine learning algorithms for text classification. From the initial study of discrete models of TSP to the present, several generations of combinatorial mathematicians and operational researchers have studied this problem extensively and intensively and have achieved great achievements. Using the same global search strategy-based algorithm to train the neural network parameters can avoid the above-mentioned drawbacks of the genetic algorithm, because its velocity-displacement model is simple to operate and only decides the search direction according to its own velocity.

This paper combines quantitative analysis with qualitative analysis and introduces the theory and method of scientific knowledge mapping based on scientometrics. Scientometrics studies Maritime English and draws a series of knowledge maps of Maritime English, which is a new attempt to study the knowledge system in this field. The innovative contribution lies in the attempt to use the core and edge structure analysis function options in word2vec software to visualize the social network map of Maritime English researchers in the process of using social network analysis method to build knowledge map. This is a pioneering exploration on the technical details of knowledge map construction. At the same time, in order to maintain the diversity of particle swarm optimization, Text CNN algorithm is combined with the construction of Maritime English subject knowledge map, and the network parameters and structure are optimized at the same time. It is of practical significance and academic value to make use of knowledge map to study Chinese Maritime English, which is helpful to the development of Chinese Maritime English and provides guidance.

2. Thoughts on Constructing Knowledge Map of Maritime English Based on Text CNN Algorithm

2.1. Construction Framework of Subject Knowledge Map Based on Text CNN Algorithm. The disciplinary knowledge

mapping constructed in this paper is guided by the disciplinary ontology and takes the information of disciplinary entities and the relationship between entities as the core to construct the knowledge mapping within the scope of the disciplinary domain. At present, knowledge mapping is still limited to specific disciplines or fields, and studies are conducted for the knowledge contents within the disciplines or fields to meet the needs of people in specific professional fields and to conduct accurate and refined search for professional academic issues. It can display the tedious knowledge in the field through information processing, data mining, graphic drawing, and knowledge measurement and provide valuable and practical reference for discipline research to reveal the knowledge areas with dynamic development laws. The framework for the construction of disciplinary knowledge mapping is shown in Figure 1.

First, the extraction of subject entities and the relationships between entities are divided into data preprocessing, acquisition of labeled data, word vector representation, feature extraction, and relationship classification for implementation. The forward neural network function class N is dense in the space of vector-valued continuous functions, in a consistent paradigm sense, where the forward network function class is defined as a set of the following types.

$$N = \left\{ Y \in \mathbb{R}^m, Y = W^{(K+1)} N_K \{ N_{K-1} [\dots N_1(X)] \} \right\}, K = 1, 2, 3 \dots \quad (1)$$

CiteSpace is used as a construction tool, combined with scientometrics theories and methods such as word frequency analysis method and cooccurrence analysis method, to study the knowledge map of Chinese maritime English subject research subjects. According to the characteristics of the large sample data set, the Bayesian method is used to extract the prior information, which is coupled into the training neural network. Its input is W_x , output is x ; then the neural network is:

$$\frac{dx}{dt} = -WF'(x)^T F(x). \quad (2)$$

Suitable data objects in the field of Nautical English are selected and used as data sources for knowledge graph construction; then normalized preprocessing is carried out, including data text format conversion, Chinese word separation, deactivation word list filtering, indexing, data feature item extraction, and data annotation. We study and use theories, construction methods, and mainstream tools related to disciplinary knowledge mapping: CiteSpace, VOSviewer, etc., and functional tools Bibexcel (mainly for foreign language data text merging, data format conversion, field extraction and weighting, cooccurrence frequency calculation, and cooccurrence matrix generation), SATI (mainly for Chinese title field extraction, frequency statistics, and matrix generation), and Bicomb (mainly for bibliographic field extraction, frequency statistics, matrix generation). The first N words with greater weight are selected for output to get the concept set of the disciplinary knowledge map and to

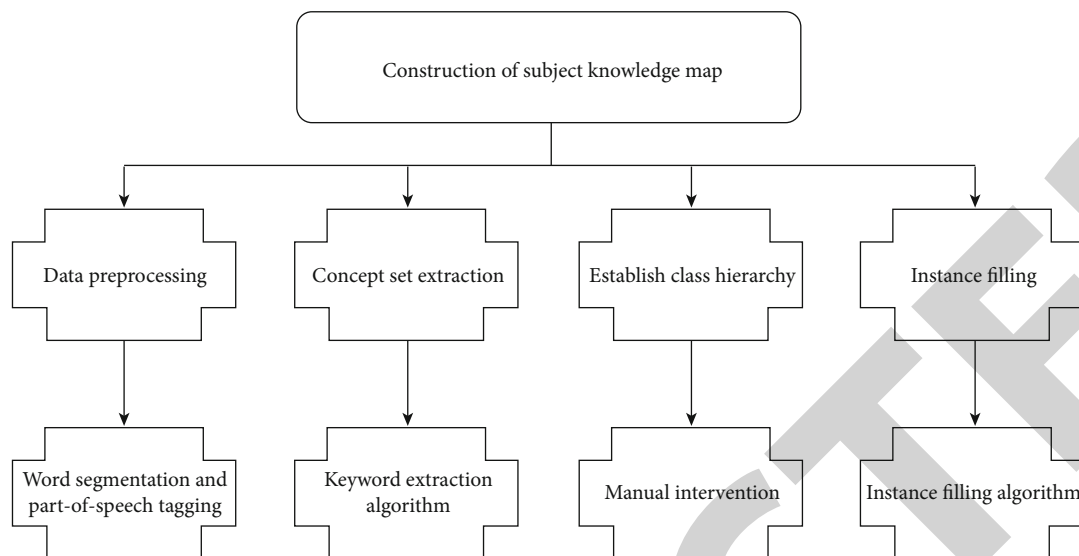


FIGURE 1: Framework of discipline knowledge map construction.

do the corresponding groundwork for constructing the disciplinary ontology later. The keyword extraction process is shown in Figure 2.

Second, data preprocessing is performed for the extraction of entities and their relationships in the subject area, that is, determining the scope of disciplinary knowledge graph construction-computer network disciplines, as well as performing manual culling, clause splitting, and other operations for computer network disciplinary texts. Using the interlayer local spatial correlation to connect the neuron node of each adjacent layer only to the upper layer neuron node that is close to it, i.e., local connection, a multilayer forward network is built with the function to calculate the output M of the hidden layer:

$$M_j = f\left(\sum_{i=1}^n W_{ij}X_i - b_j\right), j = 1, 2, \dots, m. \quad (3)$$

The data layer is an ontology-related data enrichment process that emphasizes the relationships between instances of English disciplines and their attribute-value pairs. Latent semantic algorithm analysis is performed on the data after preprocessing to obtain the results of the processed data. The two functional features extracted are upper and lower bounds and first-order derivatives, which are transformed into equivalent mathematical expression embeddings for training the neural network. Then the number of nodes is increased sequentially until the recorded error value is minimized. The formula for determining the number of nodes in the hidden layer using the trial-and-error method is as follows:

$$m = \sqrt{k+l} + a, \quad (4)$$

where m is the number of hidden layer nodes, k is the input node, and l is the output node.

In conjunction with the theoretical knowledge of discipline construction, the initial steps of constructing a disciplinary knowledge mapping tool reveal a complete picture of the practical and cognitive aspects of the discipline. The core is to identify and correlate the mainstream tools with the conceived disciplinary knowledge map construction scheme and iterate on them to improve the scheme step by step. The theoretical support and model of the disciplinary knowledge mapping scheme are also proposed.

Finally, the subject entities are obtained using hanlp, the labeled data with the relationship information between the entities are obtained using remote supervision, and the obtained corpus with the labeled information is used as the input of the neural network model. The relationship classification model is trained so that the model can automatically discriminate the relationship of entities in a sentence. The results of the data analysis obtained are used for visual technical mapping to obtain the disciplinary knowledge map. The feasibility and practicality of the constructed scheme are verified and optimized through the discipline of Nautical English, especially the construction method fusion. Relying on Text CNN to automatically extract features from textual information represented at the simple character level, the feature vectors are fed into the LSTM model via the highway network framework based on sequences thus achieving better classification prediction. Among them, the convolutional layer can well characterize the local features of the input data, while the pooling layer can further extract the most important part of the local features on this basis.

2.2. Entity Relationship Extraction Method. The most basic as well as the key to constructing a knowledge map within the subject area is to obtain the information of entities within the subject area and to realize the extraction of relationships between entities. Maritime English reading is a professional course with strong theoretical and practical aspects, which requires students to have a good professional basic knowledge before learning. The strong foreign-related

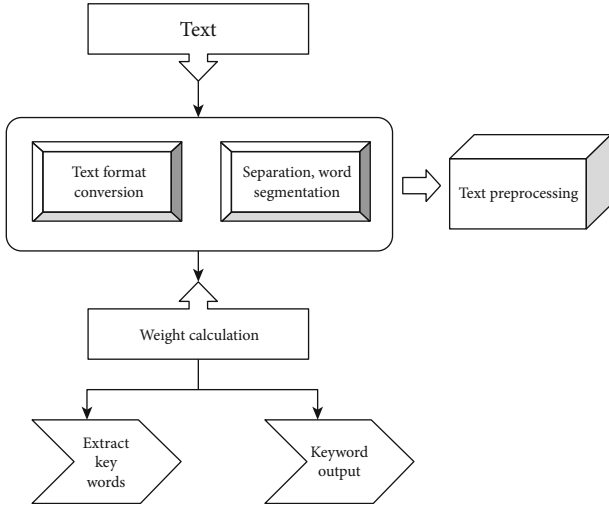


FIGURE 2: Keyword extraction process.

characteristics of the nautical profession and the strong professionalism of the nautical professional listening and speaking courses determine that the nautical professional listening and speaking teaching is different from the teaching methods of other courses. The convolutional neural network structure has been proved to be very effective in such tasks as image recognition and speech recognition. The process of subject entity and interentity relationship extraction is shown in Figure 3.

First of all, in the field of education, there are different subject entities for each subject, taking computer networks as an example, such as protocol names and device names, called named entities. However, in the current teaching of maritime English, due to the assessment method, the same paper-based examination as other courses is used. Therefore, many courses are set up to focus on how to improve students' reading ability, ignoring the training of speaking, listening, and writing ability, which leads to students' inability to understand and speak after they join the workforce. Therefore, the Chinese data literature catalogues are screened and saved in endnote format that SATI software can recognize, output as data files with txt format, and imported into SATI, and through data conversion processing, the similarity matrix and related data tables that can be used for analysis are generated for subsequent analysis. In the later stage of the search, the discovery probability is reduced in order to increase the convergence speed of the algorithm. Therefore, the discovery probability is improved as follows:

$$pa_t = \exp(t/t_{\max}) \times \cos(t/t_{\max}) \times pa_{\text{begin}}, \quad (5)$$

where pa_t is the discovery probability of t iteration and $\exp(t/t_{\max}) \times \cos(t/t_{\max})$ is the function dynamic decreasing factor.

The subsequent max-pooling layer allows for the secondary extraction of the most expressive features in each feature map. In order to implement a pure character-level

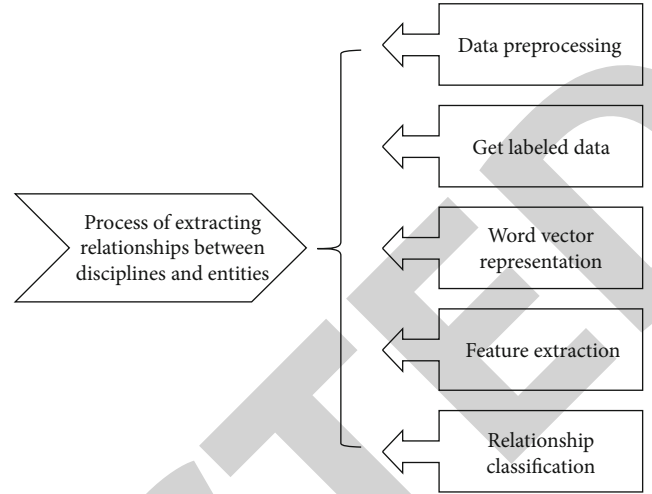


FIGURE 3: Process of extracting the relationship between subject entities and entities.

embedded convolutional neural network, the first thing to do is to construct the alphabet. The alphabet used in this paper is as follows, with 79 characters, for which one-hot encoding is used, plus an all-zero vector (for characters not in the character list, including unknown and null characters). The N -gram feature is a sequence model of every adjacent n bytes in a given malicious sample, which can be understood as a first-order Markov chain, that is, the order of occurrence of byte sequence is a random process of discrete events. By calculating the correlation between different granularity N -gram features, the different granularity features are correlated with each other, the corresponding weights are assigned to the different granularity features, and their representations are obtained by weighted summation:

$$\alpha_i^t = \frac{\exp(U_i^{tT} u_w)}{\sum_i \exp(U_i^{tT} u_w)}, \quad (6)$$

where c_i^t is the i word representation of t feature.

Second, the extracted keywords are used to align the existing remote knowledge base and extract the corresponding interentity relationship information to obtain the subject-wide entities, relationships, and entity triads. English teachers in maritime colleges and universities often do not have enough professional knowledge; thus, they are not able to teach maritime English. Therefore, the similarity matrix of high-frequency words is obtained by using the formula of low-frequency word boundaries of high-frequency words, that is, the matrix with a diagonal of 1. The data in the table indicates the frequency of any two keywords appearing together in the same literature, and the larger the value indicates the greater the degree of correlation or similarity between these two keywords. If the total input of the j th neuron N_j in the network is defined as u_j and the output state is v_j , then the state transfer equation of the network can be written as:

$$v_j = g(u_j) = g\left(\sum_t w_{ij}x_j - q_j\right), \quad (7)$$

where g is an S-type function, commonly used as:

$$g_1(x) = \frac{1}{(1 + e^{-\lambda x})}, \quad (8)$$

$$g_2(x) = th(\lambda x). \quad (9)$$

Next, Chinese pinyin sequences are introduced to semantically expand the original text, the Chinese text is converted by hanyu pinyin, and the pinyin-represented text is then embedded by character. After multilayer convolution and pooling operations, the obtained feature maps are expanded sequentially by rows, connected into vectors, and input into a fully connected network, which usually uses Softmax logistic regression as the feature classifier. As the large-scale search gradually stabilizes and the update speed slows down, the coefficients of the speed v_i^t at the t moment are adjusted by using the property that the inverse cosine function is monotonically decreasing on $[0, 1]$ as follows:

$$w(t) = \text{sqr}t[4/7 \times \arccos(t/t - \max)] \times [(t - \max - t)/(t - \max - 1)], \quad (10)$$

where $w(t)$ is called the inverse cosine function speed adjustment factor. The speed update formula is transformed into:

$$v_i^{t+1} = w(t) \times v_i^t + (x_i^t - x^*)F_i. \quad (11)$$

Finally, the entity pairs extracted from the textual information are aligned to the remote knowledge base, and the remote knowledge base is used to automate the annotation of the relationships between entities, or “transmission units,” to achieve the extraction of triples in the subject area, in order to automate the extraction of subject knowledge using neural networks. Therefore, it is necessary to convert the input statements into a form that can be recognized by the machine, i.e., a vectorized representation. Due to the inevitability of differences in experiential background (e.g., in terms of background knowledge of sailing acquired by students from coastal and inland regions), learners’ perceptions and understanding of problems are often very different. So, in order to ensure that the whole classification model does not have training difficulties due to excessive depth, the feature extraction stage does not use the stacked convolutional layers to extract high-dimensional sufficiently abstract features, but a novel framework of a single-layer Text CNN plus highway network is used instead. It refers to the fact that the nodes of the convolutional layer are only connected to some of the nodes of the previous layer and used to learn local features. This connection significantly reduces the number of parameters, speeds up the learning efficiency, and reduces the possibility of overfitting to some extent.

3. Application and Analysis of Text CNN Algorithm in the Construction of Knowledge Map of Maritime English

3.1. Training Analysis of Text CNN Algorithm. Up to now, the method used for training convolutional neural networks is still the traditional gradient descent method. Among them, if the batch gradient descent method is used, although the best convergence effect can be obtained, the convergence speed of the training process is severely limited because all training samples are required to participate in the operation in each iteration process. The advantage of Bi LSTM is that it can learn the dependence between observation sequences (input words) through two-way settings. In the training process, LSTM can automatically extract the features of observation sequences according to targets (such as recognition entities). Therefore, the Text CNN algorithm uses the Bi LSTM to process the input vector, and then at each time step, the output of the Bi LSTM is stitched with the corresponding word vector as the “semantic vector” of the current time step, which can well represent the contextual features of the text. The accuracy and loss curves of the dataset during the training process are shown in Figures 4 and 5.

First, each iteration process requires only a small number of samples to participate, and shipping samples can speed up the convergence while ensuring that the optimal solution is found as much as possible. In the training phase of the algorithm model, the input is labeled data, and the label of the document is the output. Planes of the same depth are called depth slices, and the same slice shares the same set of weights and biases. The repetition unit is able to identify both the feature, without considering its position in the viewable domain, helping the neural network to remain spatially invariant to the input. Individual genes are evaluated according to some criterion, resulting in a ranking table of importance for each gene, and then the gene with relatively high evaluation score is selected as the feature gene based on the ranking result. The attributes are conditionally independent of each other for a given target value. In order to avoid the information carried by other attributes being used by attribute values that never appear in the training set, the probability values are usually modified when estimating. Laplace correction is often used. The assumption of attribute conditional independence is relaxed to some extent. Directed acyclic graph is used to describe the dependency relationship between attributes, and conditional probability table is used to describe the joint probability distribution of attributes. The advantage is that few parameters are required to be estimated and less sensitive to missing data. In order to improve the accuracy of the network parameters, the number of iterations has to be increased, and then the computational effort is also greatly increased. The size of the response of the Text CNN network to the input depends on the distance between the input vector and the center of the network, and the smaller the distance between the input vector and the center, the larger the response of the neuron. So the center correction process of Text CNN is essentially the process of clustering the input samples based on the distance between them, the input vectors with small distances from

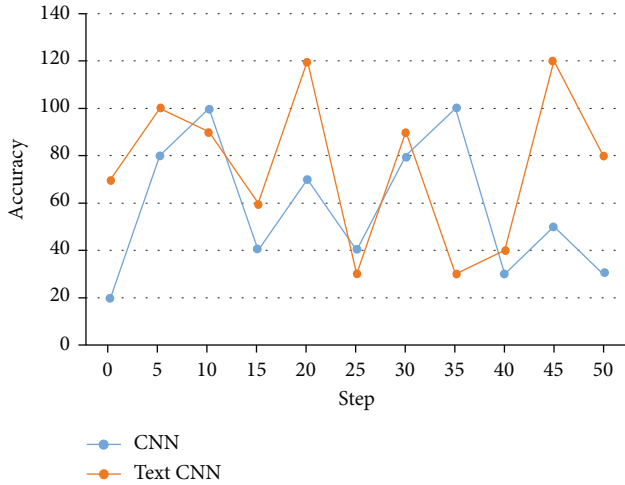


FIGURE 4: Curve of accuracy rate changing with time during training.

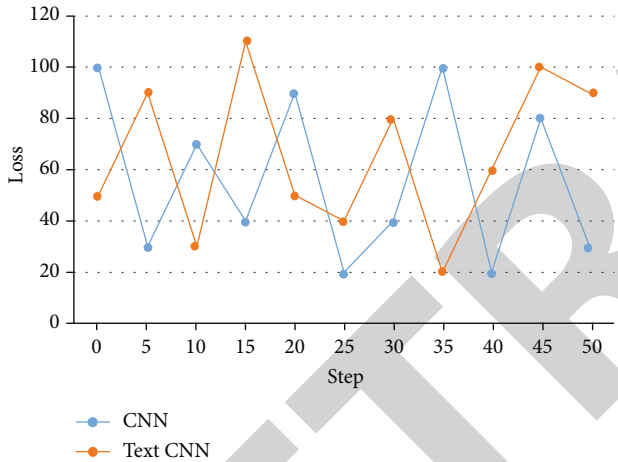


FIGURE 5: Curve of loss with time during training.

each other are grouped into one class, and the center of clustering is the network center. Three datasets of different sizes are selected for comparison experiments, and the details are shown in Table 1.

Second, to prevent overfitting in the training of the model, L2 regularization is used to constrain the parameters of the convolutional neural network. Some kind of metric is usually used to evaluate the importance or relevance of each gene to the classification, then the genes are ranked according to their importance, and finally, a certain number of top-ranked genes are selected as the feature genes; among such methods, the signal-to-noise ratio evaluation metric is the most commonly used. Corresponding initial search instructions should be made to form a unique search format by example queries, so that users can find the queried data more easily and quickly. Logistic regression is a generalized linear regression analysis model, which is used to deal with regression problems where the dependent variable is a categorical variable and is commonly used for dichotomous or

TABLE 1: Details of data sets of different sizes.

Data	Number of target classes	Average length of text	The maximum length of the text
MR	3	26	28
SST	5	45	65
Irony	2	37	84

binomial distribution problems, where the graph of the relationship between the dichotomous probability and the independent variable is mostly an S-shaped curve and can also deal with multicategorical problems. Based on the tensor flow deep learning platform, using one GPU and trained with a convolutional kernel width of 3, the results of Google Billion Word benchmark test of Text CNN trained on a single GPU are shown in Figure 6.

Finally, a dropout strategy is introduced for training the final fully connected layer parameters, i.e., a portion of the trained parameters is randomly selected for discarding at each update. In the hierarchical clustering algorithm, after the region division is completed in the lower layer, the path-finding optimization of each region in the higher layer actually uses only the center of gravity coordinates of the neighboring regions, and this information is obtained in the process of region division in the lower layer. In addition, the similarity matching process should be done, the queried features should be compared with the feature values in the database through certain algorithms, and the data can be returned to the user only after satisfying certain similarities to ensure that the user's retrieval needs can be met. Representing instances as points in space, the mapping makes instances of separate categories separated by distinct intervals as wide as possible, maps the new instances to the same space, and predicts the category to which they belong based on which side of the interval they fall. After that, make the number of hidden layer nodes plus one and still go through a certain number of gradient descent iterations as in the beginning, if after the iterations the objective function decreases by a value greater than the threshold value, this means that an additional hidden node has been added, which is more useful for the network, and the number of hidden layer nodes at this time does not make the objective function of the system extremely small, so make the number of hidden layer nodes plus one and continue the iterations as above.

3.2. *Feature Extraction Analysis of Text CNN Algorithm in Subject Knowledge Map.* The ultimate goal of teaching conversational maritime English is to equip students with the ability to communicate effectively in English in real work situations. Knowledge graphs are usually expressed using the semantic technology standard language RDF or ontology language, and the construction of domain knowledge graphs is often based on the concept of domain knowledge ontology. Therefore Text CNN algorithm is mainly to extract the content needed by users by processing the original media data, and the process is more focused on the accuracy of information retrieval. Then it needs to be

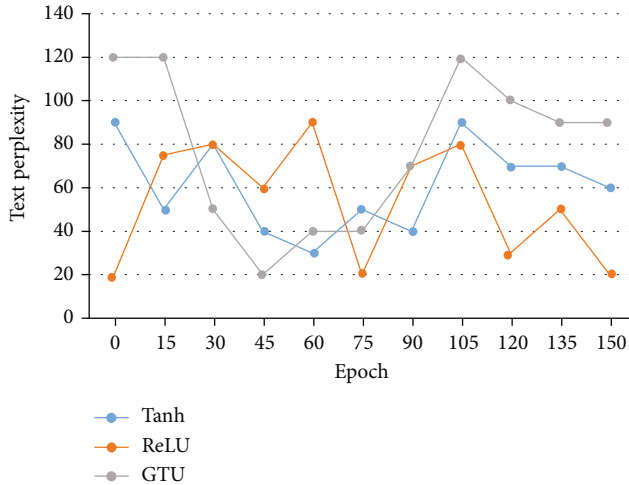


FIGURE 6: PPL change curve during training.

trained with data with category labels, and each sample data contains not only a number of features, i.e., the sample expression values in all genes, but also the category to which it belongs. The shipping makes the data itself already have better distribution characteristics. The classification performance on the English dataset will be compared with several common deep learning models in terms of accuracy. A comparison of the accuracy change curve between the Text CNN algorithm and Char Deep CNN during the training process is shown in Figure 7.

First, in a certain document, a word or phrase appears repeatedly, i.e., the word frequency TF is high, while distinguishing from other documents, the word rarely appears in other documents, i.e., the word frequency TF is low, which means that the word or phrase is more representative of the features of the document and can be used as feature extraction for classification. Training word2vec in advance, using it as input features for CNN, and updating it continuously during iterative training, shipping is equivalent to introducing certain prior knowledge that can guide the model to converge to the optimal solution in a better direction during the training process. It usually employs a classifier to directly evaluate the classification performance of a subset of feature genes and then adopts some strategy to adjust the subset according to the evaluation results to achieve the purpose of continuously exploring the optimal subset. Coword clustering analysis is to gather the ones with high similarity of subject words together to form a category with high intragroup similarity, low intergroup similarity, and relatively independent concepts. The existing methods for evaluating text classification are mainly judged by the model's prediction of correct text labels. Table 2 shows the mixture matrix built according to the Text CNN algorithm, which is used to introduce the calculation of evaluation metrics.

Second, text representation using vector space models requires a lot of work to annotate the text if lexical features need to be preserved. In contrast, word2vec simplifies the manual work by converting words in the text into vector representations based on contextual features. Feature extraction and model training are not completely separate, and the

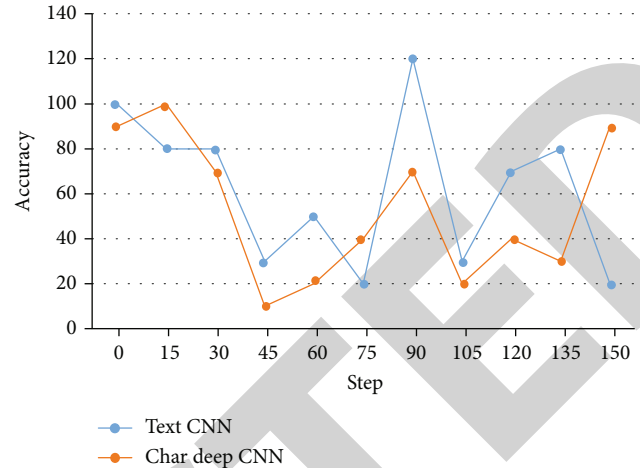


FIGURE 7: Curve of accuracy during training.

TABLE 2: Correlation of evaluation indicators.

	Predict	Not retrieved
Real	16	5
Irrelevance	31	7
Grand total	47	12

input distributed features are updated as parameters during the iterative training process. The performance of the classification is tested on the whole dataset, and then the change in performance after subtracting each gene is calculated. The gene with the smallest absolute value of association weight in the classification function is selected and removed from the training set, the process is repeated until the training set data is empty, and the subset of feature genes removed in the last step is the optimal classification subset. A trifold cross-validation method is used to select training and test samples, which allows the classification model to be fully learned. The relationship between Text CNN and SVM in terms of feature dimension and classification error rate is shown in Figure 8.

In the actual entity relationship extraction process, its individual sentence lengths are often inconsistent. It also means that this paper needs to extract the inconsistent local features from the individual sentences with inconsistent lengths for predicting the relationship types of the target entities. Using this method to train the network not only shows very intuitively and clearly the magnitude of the influence of each hidden layer central node on the deviation decline rate of the network. Moreover, the number of iterations is also very small, and generally only a few iterations are needed to finalize how many hidden layer nodes are needed.

Finally, for the analysis of large amounts of text data, we need a tool to represent the text as data understood by the computer; in other words, word2vec is a tool to convert text into a numerical representation. The skip-gram model for feature extraction and the convolutional neural network model for classification should be considered whole, and

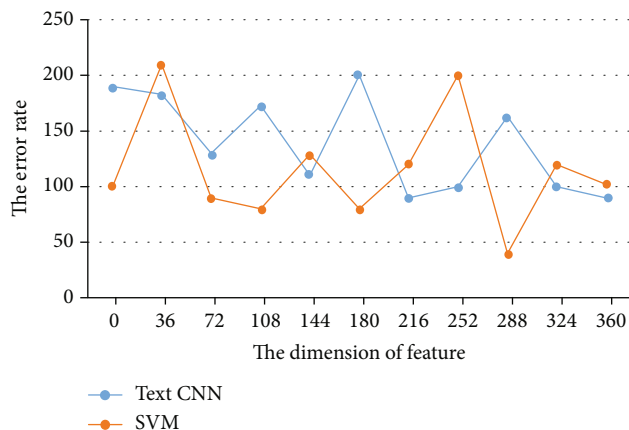


FIGURE 8: Relationship between feature dimension and classification error rate.

together, they perform the whole process of the short-text sentiment classification task. The recursive hierarchical feature gene selection method is repeatedly run according to different training sample distribution structures, guided by the classification accuracy, and then, the final selected subset of feature genes is obtained by synthetically integrating numerous feature selectors. The maximum pooling operation is used to perform secondary extraction of features, that is, for each filter set in the convolutional layer, the value of this paper selects the maximum of them as the final retained value, and the rest is discarded. The network accuracy is further improved with each additional network center, and the next center selected each time is the one that contributes most to the error reduction in the remaining part.

4. Conclusions

This paper presents a method of constructing Maritime English subject knowledge map based on Text CNN algorithm. The application of Text CNN algorithm in the construction of subject knowledge map is analyzed. It makes knowledge more relevant and hierarchical. At the same time, the knowledge retrieval technology of knowledge mapping realizes the accurate answer to relevant English questions and can effectively help students check and fill in the gaps of English knowledge points. In the process of using social network analysis method to build knowledge map, try to use the core and edge structure analysis function options in the word2vec software to visualize the social network map of Maritime English researchers. In order to maintain the diversity of particle swarm optimization, Text CNN algorithm is combined with the construction of Maritime English subject knowledge map, and the network parameters and structure are optimized. The knowledge map construction method based on text CNN algorithm proposed in this paper can reveal the English subject knowledge system and standardize the knowledge structure. It is of practical significance and academic value to make use of knowledge map to study Chinese Maritime English, which is helpful

to the development of Chinese Maritime English and provides guidance.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was financially supported by the Education and Teaching Reform Research Foundation of Higher Education Institution of Hainan Province of China: Evaluation of Nautical English Ability Based on Network Learning Platform and Teaching Improvement Research (Grant No. Hnjg2021-117) and the Hainan Province Philosophy and Social Science Planning Project: A Study on the International Ship Registration System under the Background of Free Trade Port (Grant No. HNSK(ZX)21-90).

References

- [1] Q. Kuo and X. Shijie, "The application of subject knowledge map in constructing middle school English curriculum map," *Modern Communication*, vol. 18, p. 2, 2017.
- [2] C. Hui, D. Hui, and W. Daoting, "Research on automatic construction of knowledge graph for instructional design discipline based on natural language processing," *China Education Informatization*, vol. 7, p. 5, 2020.
- [3] X. Zhu and Z. Xie, "Automatic knowledge graph construction algorithm based on massive text data," *Journal of Jilin University: Engineering Edition*, vol. 51, no. 4, p. 6, 2021.
- [4] X. Wu, "The current situation and countermeasures of China's maritime English teaching research," *Maritime Education Research*, vol. 35, no. 1, p. 7, 2018.
- [5] J. Wang, "On the design and application of digital teaching resources in navigational English listening and conversation," *Journal of Jiamusi Vocational College*, vol. 194, no. 1, pp. 212-213, 2019.
- [6] Z. Xiaolan, "Discussion on design elements of navigation English digital teaching resources," *Journal of Shandong Agricultural Management Cadre College*, vol. 35, no. 4, pp. 120-121, 2018.
- [7] A. Nugroho and I. Mutiaraningrum, "EFL teachers' beliefs and practices about digital learning of English," *EduLite: Journal of English Education, Literature and Culture*, vol. 5, no. 2, pp. 304-321, 2020.
- [8] D. Hongxin and L. Jiexiang, "Analysis of the application of structural metaphor cognition in vocabulary teaching and learning of Nautical English," *Caizhi*, vol. 1, p. 1, 2020.
- [9] Z. Xiaolan, "Constructing a digital learning platform for professional English—taking marine English as an example," *Journal of Shandong Agricultural Engineering Institute*, vol. 34, no. 10, p. 3, 2017.

Retraction

Retracted: Application of Wireless Sensor Network Model Based on Big Data Ecosystem in Intelligent Health Monitoring System

Journal of Function Spaces

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 2023

Copyright © 2023 Journal of Function Spaces. 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. Tu, "Application of Wireless Sensor Network Model Based on Big Data Ecosystem in Intelligent Health Monitoring System," *Journal of Function Spaces*, vol. 2022, Article ID 3179915, 10 pages, 2022.

Research Article

Application of Wireless Sensor Network Model Based on Big Data Ecosystem in Intelligent Health Monitoring System

Jiaqing Tu 

School of Emergency Technical, Zhejiang College of Security Technology, Zhejiang 325016, China

Correspondence should be addressed to Jiaqing Tu; tujiaqing@zjcst.edu.cn

Received 20 May 2022; Revised 13 July 2022; Accepted 15 July 2022; Published 29 July 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Jiaqing Tu. 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 enhancement of people's health awareness, the awareness of disease prevention began to improve, and the demand for intelligent health status monitoring is also increasing. At present, the health status monitoring mostly depends on hospital equipment, which has high cost and cannot realize remote analysis. Based on this, this paper studies and analyzes the application of wireless sensor network model based on big data ecosystem in intelligent health monitoring. ZigBee protocol is used to realize data transmission, wireless sensor is used to monitor human body temperature, heart rate, blood oxygen saturation, and other physiological indicators, human posture monitoring algorithm is used to monitor posture changes, which is combined with human physiological parameter information, and the long-term and short-term memory model is used to effectively use the time sequence information of data to realize the functions of physiological parameter measurement, wireless positioning, and information fusion. An experiment was designed to test the accuracy of the health monitoring system. The results show that the wireless sensor retention model can meet the design requirements in terms of the monitoring accuracy of body temperature, heart rate, and blood oxygen saturation index in intelligent health monitoring, with low cost and can be combined with human dynamic information.

1. Introduction

Modern medical models and medical concepts are gradually changing, from hospital centered to patient and family centered, from treatment centered to prevention and health care centered. The precise portable home monitoring system includes cholesterol, blood, blood pressure monitoring, and other portable health monitoring terminals. Users can take care of themselves at home and monitor their health status without completely relying on medical staff or being admitted to the hospital. However, the current level of medical informatization is still relatively low. Although families have electronic diagnostic apparatus, the doctors cannot obtain the first-hand data of blood pressure and blood glucose measured by themselves in time. Patients are still unable to get timely assistance. With the development of Internet of Things, cloud computing, sensor technology, and wireless communication technology, home health monitoring technology has been greatly developed in recent years. These health monitoring terminals allow patients to easily complete various health indicators at home.

The traditional health monitoring system mainly includes two kinds: one is the Holter system, which records ECG data for 4 hours. It can be carried by patients, but it cannot carry out real-time diagnosis. The other is the bedside monitoring system used in the ward, which can carry out real-time diagnosis, but can only be used in fixed places in the hospital. With the development of computer technology, sensor technology, and medical device technology, using micro sensors to monitor human health has become the focus of research in the field of medical instruments. With the improvement of people's living standards, people's attention to health is also gradually increasing. The World Health Organization survey shows that the proportion of healthy people in the world is only 5%, and most people are in subhealth state [1]. At present, people have changed the concept of going to the hospital only when they are ill, and began to pay attention to the maintenance of the body in their daily life. Therefore, health monitoring has become a daily demand [2]. In the current health monitoring system, some countries have begun to establish telemedicine system to realize health monitoring through long-distance

transmission and analysis of data [3]. Human life monitor also began to enter people's daily life. The indicators that can be monitored include important physiological indicators such as body temperature, blood pressure, and heart rate [4]. However, these technologies are basically based on wired transmission, and the data obtained is only physiological data in a specific state, which cannot realize the dynamic analysis of data. The instrument is inconvenient to carry, and the supervision is very inconvenient [5].

Based on this background, this paper studies the application of wireless sensor network model based on big data ecosystem in intelligent health monitoring system, which is mainly divided into four chapters. Section 1 briefly introduces the research of health monitoring system and summarizes the chapter arrangement of this study. Section 2 introduces the application status of health monitoring and wireless sensors at home and abroad and summarizes the shortcomings of the current research. In Section 3, the wireless sensor network model based on big data ecosystem is constructed and applied to the intelligent health monitoring system. The wireless communication technology and ZigBee communication principle are used to realize the long-distance transmission of data. The important physiological indexes in human medical monitoring, such as temperature, heart rate and blood oxygen saturation, were selected as the research objects. A wireless sensor network of physiological parameters is designed to monitor physiological indicators. In Section 4, the health monitoring wireless sensor network constructed in this paper is simulated and tested, and the application effect of wireless sensor network is determined by analyzing the measurement accuracy of temperature, heart rate, blood oxygen saturation, and other indicators. The experimental results show that the algorithm proposed in this paper can effectively process information with high accuracy. In the monitoring of temperature, heart rate, and blood oxygen saturation, the relative error is relatively small and can obtain accurate physiological data. The classification and analysis of physiological parameters of different behaviors can provide more scientific data for human health monitoring.

The innovation of this paper is to establish a wireless sensor network for health monitoring. These sensors can not only monitor the daily physiological indexes of the human body but also detect the changes of physiological parameters in different states of the human body and can locate them accurately in real time. The data obtained by the system can use ZigBee technology to send information to personal terminals to remind people to know the health status in time.

2. State of the Art

In human intelligent health monitoring, great research progress has been made from the previous hospital physical examination to the current real-time monitoring. With the development of big data mining technology, all kinds of cluster analysis technology can be applied to data monitoring. For example, Orimoloye et al. used remote sensing technology to analyze the impact of surface temperature and radiation on human body temperature [6]. In the research and analysis, Chu et al. developed an active pulse sensing system using flexible sandwich piezoelectric electret film, which is similar to the

pulse palpation of the trained traditional Chinese medicine doctors [7]. Guo et al. designed a simple and low-cost optical strategy of a stretchable strain sensor in the research and analysis. The sensor can measure 100% large strain, has low detection limit ($\pm 0.09\%$), fast response (12 ms), and high reproducibility (more than 6000 times), and can be assembled into clothes or installed on the skin surface [8]. In the research and analysis, Yoon et al. proposed a sensitive and selective sweat sensor based on flexible NICU (ooh)/polystyrene (PS) electrode for the detection of urea biomarkers [9]. Al-Temeemy proposed a descriptor for home health care monitoring system in research and analysis, which is generated by a color method to extract robust features from the image of the monitoring system and is robust to events that usually interfere with the monitoring system [10]. Luo et al. used paper substrate and spray deposited metal electrodes and traces to achieve low manufacturing cost. They used the capacitive sensing scheme of deformable triangular PDMS sensing film to study the trade-offs between different sensor array designs to achieve the best design. The fabricated sensor array can be comfortably attached to the individual's temporal and ankle areas [11]. In the research and analysis, Li et al. used porous graphdiyne (GDY) material to build a printed respiratory sensor suitable for life and health monitoring, which showed ultrasensitive response and rapid recovery performance under real-time respiratory conditions, and realized 40 cm long-distance noncontact detection [12]. Parrilla et al. introduced a wearable paper-based chemical resistor for monitoring sweating dynamics (sweating rate and sweating volume), which realizes in vivo detection of biological fluid based on single-layer carbon nanotubes and surfactant (sodium dodecylbenzene sulfonate) nanocomposites integrated in cellulose fibers of traditional filter paper [13]. Dai et al. prepared a self-healing hydrogel with a variety of shape memory characteristics through one-step polymerization in the study, which is used to manufacture resistance strain sensors and capacitive pressure sensors and can detect various types of human motion [14, 15].

To sum up, we can see that there are many researches on health status monitoring at home and abroad, most of which use sensor technology. At present, the research of this technology has gradually matured, but most of them use wired technology, which has great limitations in application. Although many scholars have analyzed wearable technology, the cost of such equipment is generally high, and the monitored indicators are also very limited. There are some research results in human posture detection algorithms, but these are only monitoring human posture and cannot get real-time physiological data changes, and it is difficult to meet people's needs for health monitoring. Therefore, it is of great practical significance to carry out the application research of wireless sensor networks based on big data ecosystem in intelligent health monitoring.

3. Methodology

3.1. Design of Wireless Sensor Network Model. At present, people pay more and more attention to health. There are many kinds of human health sensor monitoring equipment on the market, but the existing equipment can only monitor one index, and most of them adopt a wired transmission mode, which cannot realize real-time data acquisition and analysis.

In this paper, the wireless sensor network junction is used to design and realize intelligent health monitoring. See Figure 1 for details. In the design of the system, the wireless sensor installed on the terminal node is used to monitor the human physiological indexes in real time, covering the important physiological parameters of temperature, heart rate, and blood oxygen saturation. The transmission is realized by ZigBee wireless sensor network, and the microprocessor processes the data and displays the results in real time [16].

Considering that the system needs to monitor different human physiological indicators, each node needs to be miniaturized and does not affect people's daily life, the cost is as low as possible and can meet the economic requirements of most people, and the sensor can continuously stimulate human information, so it needs a low-power consumption, simple operation, and reliable results. In this paper, ZigBee technology is selected in wireless sensor networks. The system is divided into three parts: front-end acquisition part, information acquisition part, and information transmission part [17]. The front-end acquisition is realized by wireless sensors, and the information is received by the upper computer at the computer end. The transmission between the lower computer and the upper computer is realized by ZigBee communication protocol, and these collected data are stored at the same time.

The wireless sensor network used in this paper adopts the simplest star network topology, which mainly covers two parts: terminal node and network coordinator [18]. The terminal node mainly collects human body temperature data, including data acquisition module and wireless receiving module. The terminal node sends the collected data to the coordinator. The coordinator is the core of the whole structure. It can initialize data, open physical channels, allocate addresses, and add or delete individual terminal nodes.

In the setting of wireless communication network, the transmission situation will be affected by path loss, environmental loss, and external consumption. In the design of sensor nodes, we need to focus on the problem of RF antenna. If the design is unreasonable, it will increase the loss and even fail to communicate. In the design, the wireless node antenna selects a single-stage antenna in combination with the actual situation. In wireless communication networks, antennas play an important role. The selection and setting of antennas directly affect the working quality of the whole wireless communication network. The circuit board lead antenna is a wire printed on the circuit board, which can sense air waves and receive signals. Based on the research results of wireless sensor networks, the author designed a low-power, low-cost, and practical wireless sensor network node. This antenna type has high transmission effect. Its corresponding resonant antenna with electromagnetic wavelength of $1/4$ is determined by the following formula:

$$L(\text{cm}) = \frac{7125}{f(\text{MHz})}. \quad (1)$$

If the working frequency is 2.45 G, the length is about 2.9 cm. The design of the antenna can meet the requirement that the RF input/output impedance is 50Ω .

3.2. Collection of Health Data. At present, there are two kinds of wireless temperature sensors. The first kind of analog temperature sensor designs the corresponding conversion circuit to convert the signal into digital signal, and the single-chip micro-computer processes the digital signal. The data output from the second digital temperature sensor is a digital signal, which can be read directly through the corresponding instructions. At present, the thermistor sensitivity of temperature sensors in the market is relatively high, such as thermistor and DS18B20; among them, the consistency of thermistor is insufficient to realize the exchange of data [19]. DS18B20 resistor realizes data conversion and control through a network cable, but its accuracy is not very high. These traditional contact temperature measurement methods have their own limitations, and infrared thermometer has been applied in wireless sensor. Generally speaking, the contact thermometer is relatively simple, is reliable, and has high measurement accuracy. However, due to the delay of temperature measurement and the limitation of high-temperature resistant materials, it cannot be applied to a high-temperature measurement. This sensor can quickly calculate temperature data through thermal balance. The measurement range is wide, and the accuracy can reach $\pm 0.2^\circ\text{C}$. It is very suitable for human body temperature measurement [20]. The wireless sensor is small, is easy to carry, and can be networked to obtain real-time data. In application, infrared radiation is used. The wavelength of this light is between red light and microwave. Objects above absolute zero are considered to be emitting infrared radiation energy, and this radiation energy density is related to temperature, which is expressed as follows:

$$E = \varepsilon \delta T^4 = 5.67 \times 10^{-8} \varepsilon T^4, \quad (2)$$

where ε is the radiation coefficient, and the value range is 0~1. According to the above formula, it can be seen that the radiation power can be calculated after the temperature and emissivity of the object are obtained. Therefore, the temperature of the object can also be calculated until the radiation power is obtained. The output signal of the infrared sensor can be adjusted according to the target temperature and its own temperature. The formula is as follows:

$$V_{ir}(T_a, T_0) = A(T_0, T_a), \quad (3)$$

where A is the instrument constant.

In the monitoring indicators of human health, heart rate cannot be ignored. At present, there are great differences in the action mechanism of wireless sensors that care about heart rate. The PPG monitoring method used in this study uses infrared light to pass through the tissue and emit infrared light, and the heart beat produces contraction, which affects the refractive index of light. The amount of blood in the tissue also affects the amount of light detected by the detector [21]. The infrared diode transmits the infrared light to the human body, and part of the light is reflected. Therefore, each heart beat will cause the change of the amount of infrared light. At the same time, a high gain amplifier is added to increase this change, and the heart rate can be obtained through the change of the amount of red light.

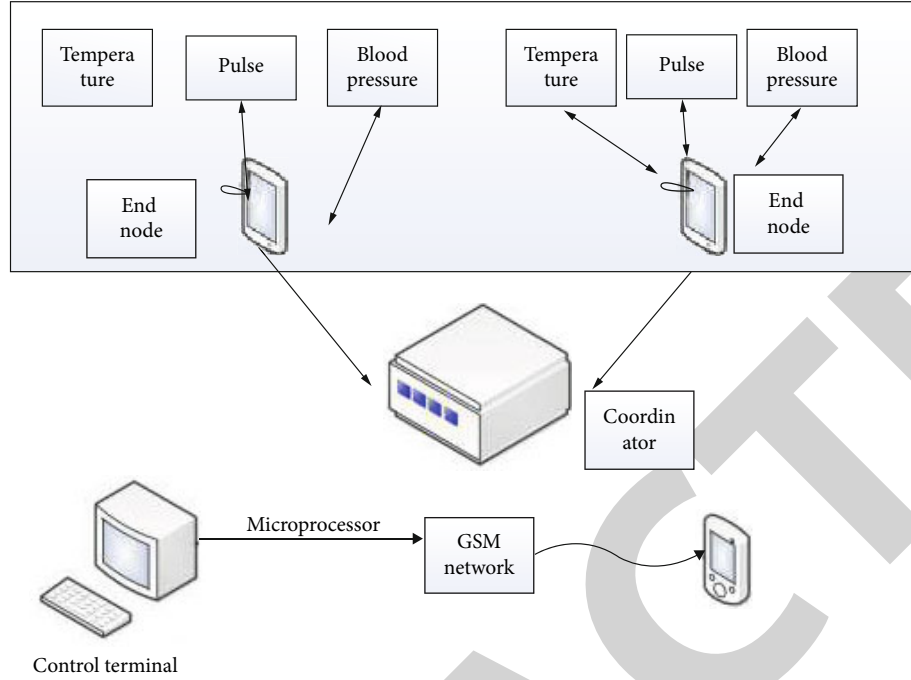


FIGURE 1: Intelligent health monitoring system.

Blood oxygen saturation reflects the ability of blood to carry oxygen and is an important nonparameter to evaluate human respiratory function. Therefore, by evaluating this index, human lung function can be monitored. The blood oxygen saturation wireless sensor used in this paper is based on Lambert Beer's law. It irradiates the human body with red light and near-infrared light and calculates the blood oxygen saturation through the periodic changes of the two lights [22]. The light will be absorbed after irradiating the human body, and the absorption of light by the tissues and blood vessels is basically unchanged. It is expressed by DC, and the pulsation component will change due to the pulsation of the blood vessels. It is expressed by AC. The light intensity is calculated according to the periodic change, and the formula is as follows:

$$I_{DC} = I_0 e^{-\epsilon_0 c_0 L} e^{-\epsilon_{HbO_2} c_{HbO_2} L} e^{\epsilon_{Hb} c_{Hb} L}, \quad (4)$$

where ϵ_0 represents the total absorption coefficient of nonpulse pulse, c represents the light absorption concentration, and ϵ_{Hb} represents the arterial absorption coefficient. Assuming that when the heart rate beats, the optical path length of the arterial blood vessels increases by ΔL , and the increase value of brightness can be expressed as follows:

$$I_{AC} = I_{DC} - I_{DC} e^{-(\epsilon_{HbO_2} c_{HbO_2} + \epsilon_{Hb} c_{Hb}) \Delta L}, \quad (5)$$

where I_{AC} represents the change value of light intensity. Calculate the natural logarithm with the formula:

$$\ln \left[\frac{I_{DC} - I_{AC}}{I_{DC}} \right] \approx -(\epsilon_{HbO_2} c_{HbO_2} + \epsilon_{Hb} c_{Hb}) \Delta L. \quad (6)$$

Since the ratio of AC to DC is far less than 1, the formula can be transformed into the following:

$$\ln \left[\frac{I_{DC} - I_{AC}}{I_{DC}} \right] \approx \frac{I_{AC}}{I_{DC}}, \quad (7)$$

$$\frac{I_{AC}}{I_{DC}} = -(\epsilon_{HbO_2} c_{HbO_2} + \epsilon_{Hb} c_{Hb}) \Delta L.$$

Considering that the optical path length is not clear, different wavelengths of incident light are selected for analysis in the analysis. Assuming that the wavelengths are λ_1 and λ_2 , respectively, it can be obtained by bringing them into the formula:

$$\frac{R_{\lambda_1}}{R_{\lambda_2}} = \frac{\epsilon_{HbO_2}^{\lambda_1} c_{HbO_2} + \epsilon_{Hb}^{\lambda_1} c_{Hb}}{\epsilon_{HbO_2}^{\lambda_2} c_{HbO_2} + \epsilon_{Hb}^{\lambda_2} c_{Hb}}. \quad (8)$$

Blood oxygen saturation is very different from the absorption spectrum of Hb. In order to further improve the measurement accuracy, the absorption coefficient of blood oxygen saturation should be increased. In this paper, red light with 660 nm central wavelength is selected. According to the definition of blood oxygen saturation, the calculation formula can be obtained:

$$SpO_2 = \frac{\epsilon_{Hb}^{\lambda_2} (R_{\lambda_1}/R_{\lambda_2}) - \epsilon_{Hb}^{\lambda_1}}{(\epsilon_{HbO_2}^{\lambda_1} - \epsilon_{Hb}^{\lambda_1}) - (\epsilon_{HbO_2}^{\lambda_2} - \epsilon_{Hb}^{\lambda_2}) (R_{\lambda_1}/R_{\lambda_2})}. \quad (9)$$

Physiological parameters will vary under different motion states, so it is necessary to monitor human motion. In this research and design, the wireless sensor for human state

detection is placed in the front end. Once the acceleration of human motion changes, the specific change will be found through the algorithm. At present, the commonly used recognition methods are mainly through threshold and pattern recognition. In pattern recognition, the data are mapped to high-dimensional space for analysis through machine learning algorithm, which requires too much calculation. Therefore, the monitoring algorithm based on threshold is selected in this paper. The natural coordinate system and dynamic coordinate system are established for the human body, which are expressed as $Ox_0y_0z_0$ and $Oxyz$, respectively. When the human body stands still, the vertical coronal front is x , the left is y axis, the horizontal down is z axis, and the neutral acceleration is consistent with the z axis direction. Whether the human body is at rest or in motion, the coordinate axis is established according to this. The acceleration of three axes under dynamic marking is the corresponding output. When stationary, the received acceleration is g and the direction is along the z axis. Assuming that the included angle between the z axis and the horizontal plane is θ , the acceleration can be expressed as follows:

$$a_z = g \sin \theta. \quad (10)$$

Therefore, the included angle can be expressed as follows:

$$\theta = \arcsin \left(\frac{a_z}{g} \right). \quad (11)$$

The differential acceleration and vector amplitude are needed to distinguish the motion state of human body. The larger the velocity vector amplitude is, the more intense the motion is. The calculation formula is follows:

$$SVM = \sqrt{a_x^2 + a_y^2 + a_z^2}, \quad (12)$$

where SVM represents the amplitude of acceleration vector. The differential acceleration amplitude is calculated according to the average representation of SVM absolute value. The larger the value is, the more obvious the state movement is. The formula is as follows:

$$MADS = \frac{1}{T} \int_0^T |SVM'| dt, \quad (13)$$

where MADS is the amplitude of differential acceleration. Calculating this value in the dynamic coordinate system can monitor the human movement, and the calculation of SVM value can also reduce unnecessary movement monitoring. MADS index can be used as the judgment basis of human fall.

In the instantaneous state of human fall, the direction cannot be predicted, so it is not suitable to calculate the coordinate axis acceleration, and the instantaneous spatial acceleration needs to be calculated separately [23]. The acceleration frequency of human body will not exceed 15 Hz. Therefore, in the sampling, it is set to 10 Hz. In order to obtain a more accurate value, SMV needs to be corrected. The formula is as follows:

$$\text{AverSMV} = \frac{1}{M} \sum_n^{M=1} \sqrt{A_{x(n)}^2 + A_{y(n)}^2 + A_{z(n)}^2}, \quad (14)$$

where M represents the acceleration sampling value and SMV represents the spatial acceleration vector at the moment of falling.

Based on the analysis of the channel characteristics with human body as the communication channel, in order to complete the transmission of sensor signal in human body channel stably and reliably, we need to choose a reasonable signal coupling mode. Unlike wireless communication, which transmits data through air, human communication channel is a human tissue structure with complex electromagnetic characteristics. How to take an effective way to couple the signal into the human body channel and realize the efficient transmission of the signal in the human body channel is the most critical problem to realize human body communication. Considering the motion state of human body, it is necessary to adopt wireless positioning method in human body positioning, consider the factor of saving components, and use the signal strength RSS of ZigBee protocol to realize positioning analysis. The distance between the two stages measures the received signal energy. This technology needs to combine multiple reference nodes. At present, there are three positioning methods based on RSS, among which there are some differences in approximate calculation methods, which need to be used by multiple routers at the same time, with insufficient accuracy, sensitive to noise data, and insufficient use [24]. Fingerprint location method is the most widely used, with an accuracy of 2~3 M and strong anti-interference ability. However, considering the indoor dynamic multipath effect, the accuracy will be reduced. Therefore, this location method needs to be improved. Multiple multichannel location can reduce the instability caused by the measurement of a single frequency band.

In the fusion analysis of human physiological indexes, it is necessary to collect human image information first, then combine the simulation data of human motion state, and fuse the data with physiological indexes after obtaining the data, as shown in Figure 2. The motion state of human body is divided into dynamic state and static state, and the most representative state is selected for analysis. Considering that human posture is not easy to obtain directly and changes at any time, it is necessary to extract human image data. In this paper, convolutional neural network is used. The convolution neural network inputs a picture with three channels. With the calculation, the feature size is gradually reduced and the number of channels is increased. With the deepening of the network, the obtained eigenvalues are also increasing [25]. It is difficult to recognize actions only by human posture, which also needs to be combined with human physiological indexes and position changes. It is difficult to analyze directly by using data [26]. Therefore, the long-term and short-term memory model is used for analysis, and the changes of human physiological indexes are tracked for a long time after recognition. Obtain the changes of physiological indexes in different states to realize the long-term monitoring of health status.

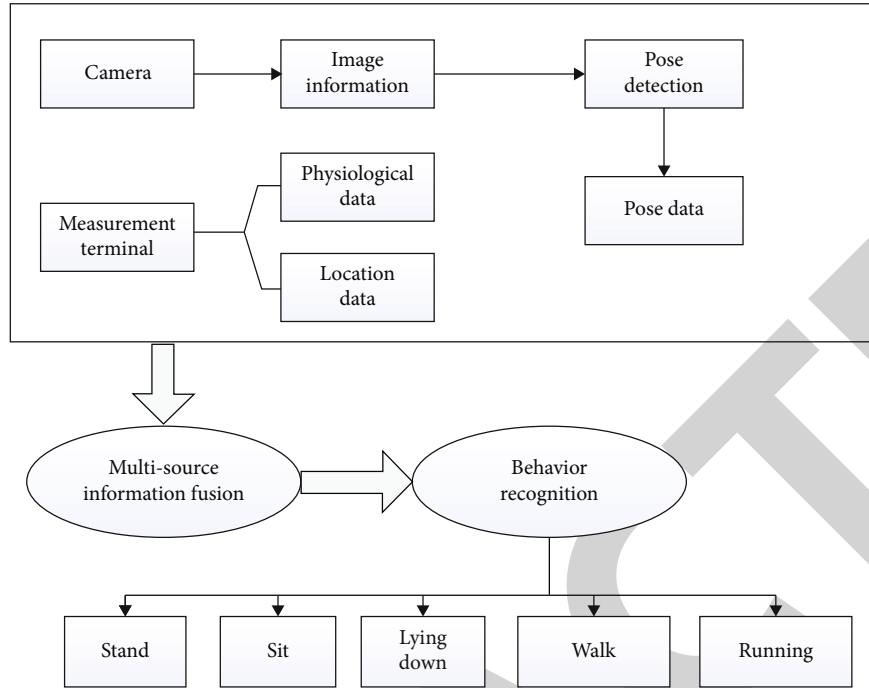


FIGURE 2: Multi-information fusion system.

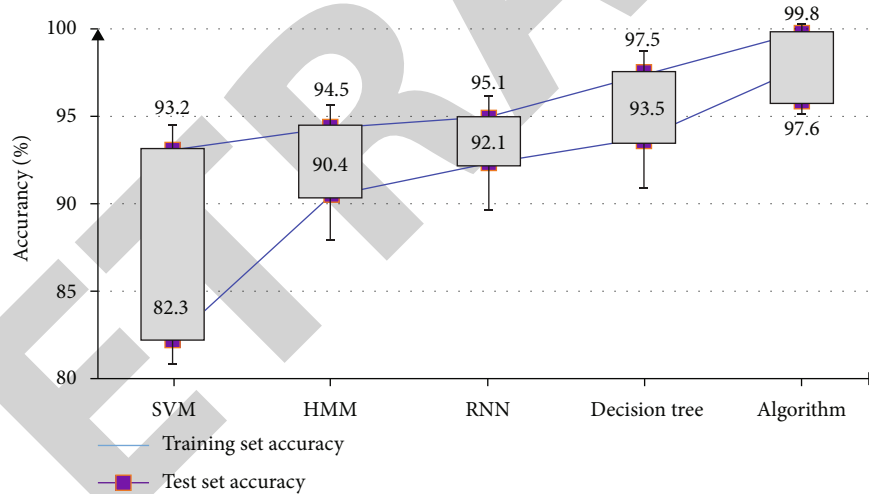


FIGURE 3: Comparative analysis of recognition rates of different algorithms.

4. Result Analysis and Discussion

4.1. *Simulation Experiment Analysis.* This paper collects the medical information of the ward through the distributed wearable sensor device through the health monitoring system platform. Data transmission through the public network can realize uninterrupted real-time health monitoring for anyone at anytime, anywhere. There are great differences in the unit and range of collected data, which cannot be analyzed directly. Therefore, it is necessary to standardize the data first. Using the zero mean standardized processing method, 80% of the obtained data are used as the training sample set, and the others are used as the test sample set

to test the application effect of different algorithms. In the analysis of physiological index and position change data, because classification is very important, the long-term and short-term memory algorithm (LSTM) used in this paper is compared with other classification algorithms. Support vector machine and multilayer sensing algorithm cannot directly use the data, so the data are spliced together. These data include body temperature, heart rate, blood oxygen saturation, and human posture data. In the test, different algorithms adopt the same training strategy, the learning rate is set to 0.001, 500 iterations, and the time sliding window length is 5. The accuracy test results of different algorithms are shown in Figure 3.

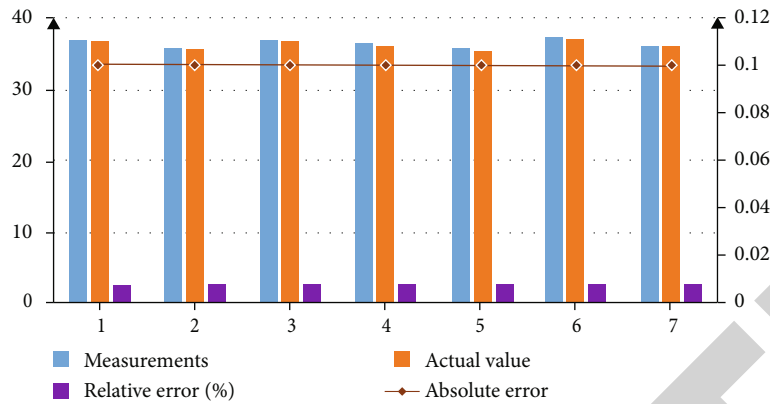


FIGURE 4: Body temperature measurement accuracy analysis.

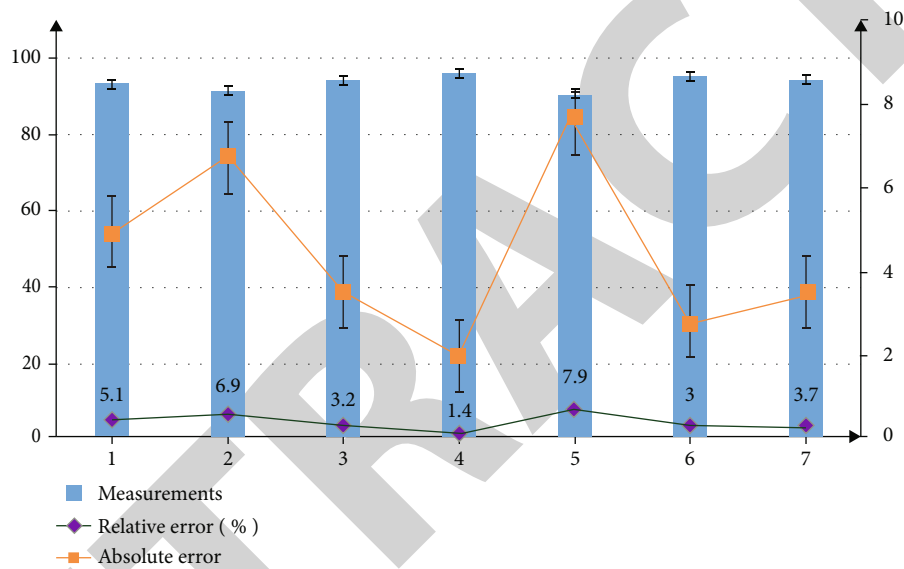


FIGURE 5: Accuracy analysis of blood oxygen saturation test.

It can be seen from the data in the figure that in data processing, machine learning algorithm and multilayer perceptron algorithm cannot effectively extract data, and the accuracy is lower than LSTM algorithm. Neural network algorithm and RNN algorithm will over fit, and the accuracy of multilayer sensing algorithm in the test set is too low. LSTM algorithm can effectively process information with high accuracy, which is because this algorithm can deal with the problem of gradient disappearance and improve the accuracy.

In the system design, the indicators of power consumption and cost cannot be ignored. The power consumption reflects the performance of the system design. The wireless sensor network used in this design system uses battery power supply, so it is necessary to reduce the power consumption as much as possible. In this group, low-power equipment is selected, and two dry batteries can provide voltage for the system. In terms of cost, the cost is also relatively low, meeting the requirements of low design cost.

4.2. Health Monitoring and Evaluation. The maximum transmission distance of a single wireless sensor node is 300 m in an open environment. If a larger distance is required, an attack

and release module needs to be added to cover a wider range. The wireless sensor is used to monitor the body temperature. Under normal conditions, it is tested for 5 times, and the measurement results are compared with the household electronic thermometer. The test results are shown in Figure 4. From the data in the figure, it can be seen that the wireless sensor network used in this paper has higher accuracy, more stability, and less absolute error.

After the ratio of red light to infrared light is obtained, the test results can be obtained according to the calculation formula of blood oxygen saturation. After the red light and infrared light output by the sensor pass through the method, the collected red light voltage value is 0.151 V and the infrared light voltage value is 0.157 V, and the measured current value and blood oxygen saturation value are calculated. The blood oxygen saturation test is carried out under normal conditions. Each person tests 10 groups of data to calculate the average value. The test results are shown in Figure 5. Compared with the normal value, the absolute error is small and meets the design requirements.

When measuring human heart rate, the signal output by the sensor is transformed into pulse signal after processing, and the heart rate value is obtained by counting. Under

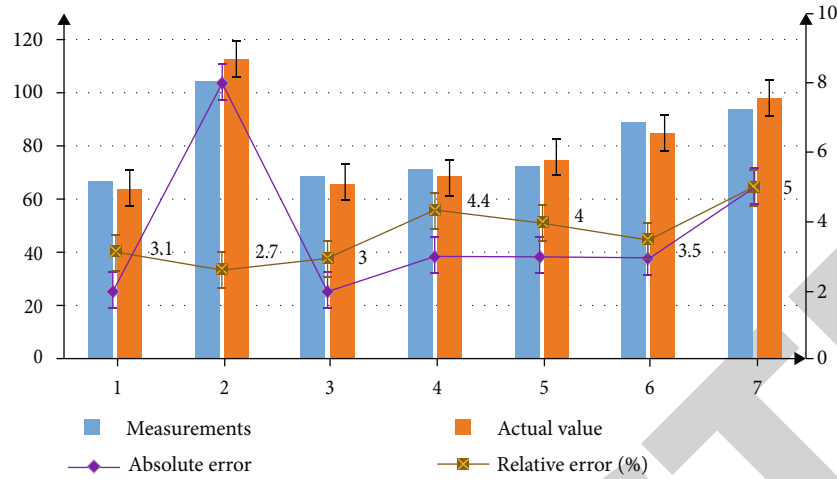


FIGURE 6: Analysis of the accuracy of heart rate measurement.

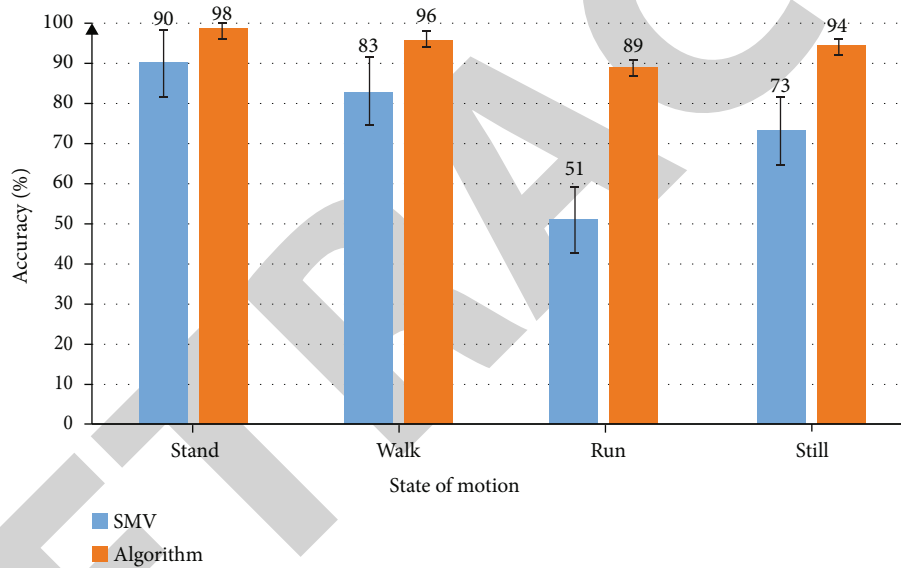


FIGURE 7: Analysis of posture monitoring accuracy.

normal circumstances, the test object is tested, and each person tests 10 groups of data to calculate the average value. The test results are shown in Figure 6. The absolute error does not exceed 6%, and the measurement accuracy is high.

In human posture monitoring, based on tiny OS platform, node sensors are used to collect simulated fall data, with a sampling rate of 10 times per second, simulate the condition of human body in walking state, and process the collected data. At the moment of falling, the acceleration of the three axes will change, so the falling behavior of the monitoring object can be judged. The algorithm in this paper is combined with other algorithms to collect 10 groups of falling data in different states, take the average value, and conduct simulation analysis. The results are shown in Figure 7. From the data change in the figure, it can be seen that several algorithms can monitor the dumping in a calm state, but in a moving state, only the algorithm used in this paper is ideal. This is because the algorithm

introduces the inclination eigenvalue as the reference object on the basis of SMV algorithm, so the accuracy can be improved.

4.3. Monitoring and Analysis of Human Health Data in Different States. After completing the monitoring of various indicators of health status, it is necessary to conduct a comprehensive analysis combined with behavior to analyze the changes of various physiological indicators under different behaviors. Select a test to test human physiological indexes and remove abnormal data values. Criteria for judging abnormal values of health evaluation of physical examination items are as follows: in the physical examination of residents, the abnormal values are judged according to the following criteria, and the corresponding health suggestions and health evaluation of abnormal values can be given in the general physical examination. If there are no symptoms, it may be physiological data changes, which do not need special treatment. Generally, it does not affect

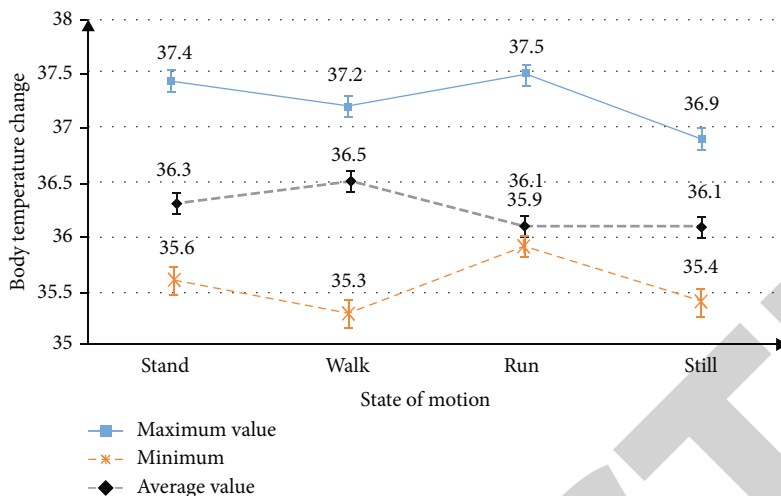


FIGURE 8: Temperature changes in different states.

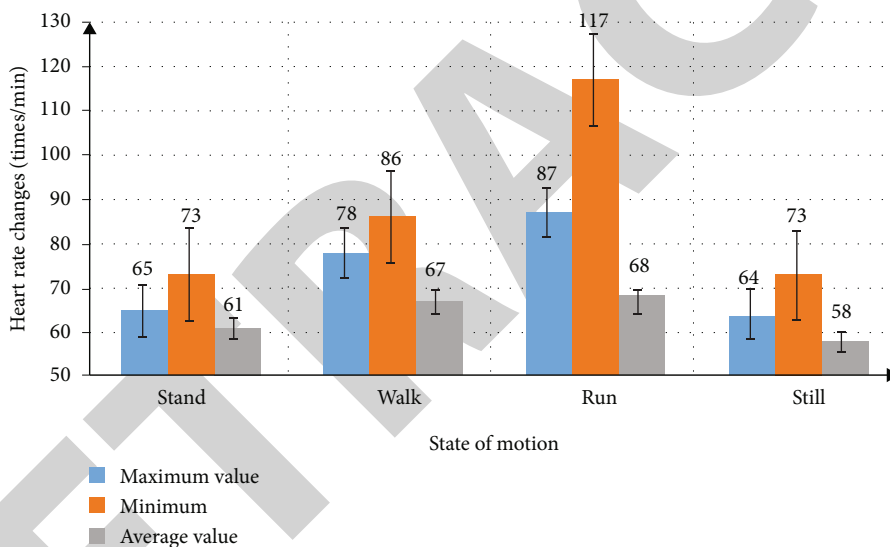


FIGURE 9: Heart rate changes in different states.

health. Take temperature change and heart rate change as examples for data analysis. The test results are shown in Figures 8 and 9.

From the data changes in the figure, we can see that under different behavior modes, physiological index parameters will change differently. For example, the heart rate in the static state is lower than that during running. The classification and analysis of physiological parameter indexes of different behaviors can provide more scientific data for human health monitoring.

5. Conclusion

This paper studies the application of wireless sensor network model based on big data ecosystem in intelligent health monitoring and designs ZigBee protocol architecture. The front-end nodes of wireless sensor networks mainly collect physio-

logical index data and send it to microprocessors. The control terminal collects and analyzes data and displays abnormal data changes. In human posture monitoring, through the comprehensive analysis of physiological indicators, we can obtain the physiological data changes under different exercise states. The functional test results show that the accuracy of the posture monitoring algorithm used in this paper is higher than the traditional algorithm, and the monitoring accuracy of each physiological index is relatively high. It should be pointed out that in the monitoring of physiological indicators, this paper selects more typical indicators. In addition to these indicators, physiological indicators such as pulse, blood pressure, blood glucose, and electrocardiogram also need routine monitoring, and more monitoring modules need to be added to the sensor node. At the same time, wireless cashier is used to monitor human body temperature, heart rate, and blood oxygen saturation, and common wireless sensors in the market are used to

Retraction

Retracted: Optimization Algorithm of Communication Resource Allocation in a Complex Network Based on an Improved Neural Network

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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 and Q. Liu, "Optimization Algorithm of Communication Resource Allocation in a Complex Network Based on an Improved Neural Network," *Journal of Function Spaces*, vol. 2022, Article ID 7309612, 8 pages, 2022.

Research Article

Optimization Algorithm of Communication Resource Allocation in a Complex Network Based on an Improved Neural Network

Haomiao Zhang¹ and Qing Liu²

¹School of Artificial Intelligence and Big Data, Chongqing College of Electronic Engineering, China

²School of Automation, Chongqing University of Posts and Telecommunications, China

Correspondence should be addressed to Haomiao Zhang; zhanghaomiao0902@126.com

Received 19 May 2022; Accepted 7 July 2022; Published 28 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Haomiao Zhang and Qing 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 traditional optimization algorithm of communication resource allocation in a complex network has the disadvantage of weak antijamming ability, and the communication quality decreases obviously when the number of users is large. In China's large urban network applications, mobile phones and other networks can have problems such as reduced network efficiency when there are more access users at some communication base stations, thus affecting user network usage. An optimization algorithm of communication resource allocation in the complex network based on an improved neural network is proposed. Increase inertia improves the traditional BP neural network algorithm, using the average path length, clustering coefficient, and connectivity distribution index analysis of the complex network; the improved Hopfield neural network is utilized to confirm each user volume size; it is concluded that their users are able to get the number of subchannels, through the instantaneous channel coarse pair gain dynamic channel allocation, calculating bit load matrix at the same time, minimize transmission power, and achieve bit loading and power allocation and communication resource allocation optimization. Experimental results show that the proposed method has better application performance by introducing the improved neural network and suppressing the external interference on the basis of enhancing the communication effect.

1. Introduction

A complex network is characterized by a complex structure, a large number of nodes, and multiple connection patterns, which refers to a network in which scale-free parts, small worlds, attractors, self-similarity, self-organization, or all properties exist. Also, in complex networks, the connection weights between nodes are completely different from the directionality. Therefore, the relevant network evolution is manifested as the disappearance and generation of connections or nodes; with the complexity of dynamics, the node sets are likely to be in a nonlinear dynamic system; in the diversity of nodes, the nodes can represent anything in a complex network. When communicating in a complex network, communication resources cannot be evenly allocated due to the excessive complexity of the

channel, and the information transmission is unstable, thereby causing poor user experience.

Therefore, according to the resource allocation problem generated when adding a cellular network based on D2D communication, literature [1] proposed a joint Hungary resource allocation method. By means of the transmission power and spectrum resource allocation problems for D2D users and cellular users, the algorithm of the proposed method can maximize the transmission rate in a heterogeneous network. Meanwhile, on the foundation of solving the incidence matrix, it is possible to transform into a nonconvex optimization problem through joint optimization and then obtain the solution through a continuous convex estimation strategy. Considering that the Space-Wire network is in the hotspot communication mode, literature [2] studies the cache resource allocation method, derives the

method for analyzing the average delay and full-load probability of network routing nodes, and calculates the key communication nodes inside the network. According to the resource allocation scheme for D2D communication in millimeter wave 5G networks mentioned in literature [3], the large communication capacity can be realized through combining device-to-device (D2D) technology with mmWave. Thus, the combination of the above two technologies can help complete resource allocation in an outdoor millimeter wave environment. With the goal of maximizing throughput, the admission set of each D2D user is selected through a linear correlation method. Based on this, the power of communication users is controlled, the communication resource allocation optimized model is constructed via multistage matching algorithm of bipartite graph, and multiple KM algorithms are introduced to solve the allocation model. Although the above three traditional methods are effective, poor communication quality or even communication interruption may appear once the interference becomes strong.

To this end, an optimization algorithm for communication resource allocation in a complex network based on an improved neural network is proposed. Since the neural network is a multilayer feedforward neural network trained by the error backpropagation method, the error drops in the direction of the gradient as being influenced by the connection strength between the input node and the hidden layer node as well as the connection strength between the hidden layer node and the output node. By means of repeated learning and training, the network parameters corresponding to the minimum error can be confirmed, and the accurate allocation requirements can be obtained according to channel differences. For this reason, an inertia term is added to the algorithm so that the network can input information into similar samples, handle the minimum nonerror output error independently, and even transform information linearly to avoid insufficient resource allocation. As can be observed from the experimental results, the proposed improved Hopfield neural network can fully consider the bit loading in the complex network and dynamically realize optimal communication resources in subchannels based on the instantaneous channel gain.

2. Improved Neural Network Optimization Algorithm

2.1. BP Neural Network Algorithm. As a layered feedforward neural network, the BP neural network can be divided into an input layer, hidden layer, and output layer, of which the hidden layer is composed of one or more layers of hidden layer nodes [3], as shown in Figure 1.

Among them, in addition to the input layer nodes, the network also has one or more layers of hidden layer nodes with no connection in the same layer. The input signal is transmitted from the input node to the hidden layer nodes in sequence and then to the output node. That means that the output of each layer node will only affect the output of the next layer node. After removing the equal input and output in the input layer, the unit structure of the remaining nodes [4] is shown in Figure 2.

2.2. Improvement of the BP Neural Network Algorithm. Since there is no connection between the nodes in the same layer, it is difficult to cope with the intricate internal structure of the complex network. Meanwhile, the output result of the BP neural network only has a vertical influence, so neither the optimal output result can be obtained according to the neighboring nodes, nor the subtle requirement of channel resource allocation in complex network can be met. Therefore, the weights and thresholds are adjusted and improved in this research so as to obtain a new Hopfield neural network. The specific formula is

$$\begin{aligned}\Delta v_{ji}(N+1) &= a_1 d_t^{(k)} b_j, \\ \Delta \gamma_t(N+1) &= a_1 d_t^{(k)}, \\ \Delta w_{ji}(N+1) &= a_2 e_j^{(k)} x_i^{(k)}, \\ \Delta \theta_j(N+1) &= a_2 e_j^{(k)}.\end{aligned}\tag{1}$$

In the above formula, N represents the number of nodes in the network, Δv_{ji} represents the neural network asynchronous working method that reaches the threshold, $\Delta \gamma_t$ represents the neural network asynchronous working method that does not reach the threshold, Δw_{ji} represents the neural network synchronous working method that reaches the threshold, $\Delta \theta_j$ represents the neural network synchronous working method where the threshold is not reached, $d_t^{(k)}$ represents the neural network asynchronous coefficient, $e_j^{(k)}$ represents the neural network synchronization coefficient, b_j represents the neural network asynchronous threshold, and $x_i^{(k)}$ represents the neural network synchronization threshold.

During the actual learning process, the learning rates a_1 and a_2 have a greater impact. The greater the a_1 and a_2 , the stronger the weight and threshold changes, which will lead to instability, namely, oscillation; the smaller the a_1 and a_2 , the more stable it is, but the speed of convergence will be slower. In practical application, generally, the values of a_1 and a_2 should be large under the premise of not causing oscillation, and in order to make the learning speed fast and not easy to oscillate, an ‘‘inertia term’’ shall be added. The specific formula is

$$\begin{aligned}\Delta v_{ji}(N+1) &= a_1 d_t^{(k)} b_j + \eta_1 \Delta v_{ji}(N), \\ \Delta \gamma_t(N+1) &= a_1 d_t^{(k)} + \eta_1 \Delta \gamma_t(N), \\ \Delta w_{ji}(N+1) &= a_2 e_j^{(k)} x_i^{(k)} + \eta_2 \Delta w_{ji}(N), \\ \Delta \theta_j(N+1) &= a_2 e_j^{(k)} + \eta_2 \Delta \theta_j(N).\end{aligned}\tag{2}$$

The choice of a_i and η_i has a greater impact on the speed of network convergence, so the formula is

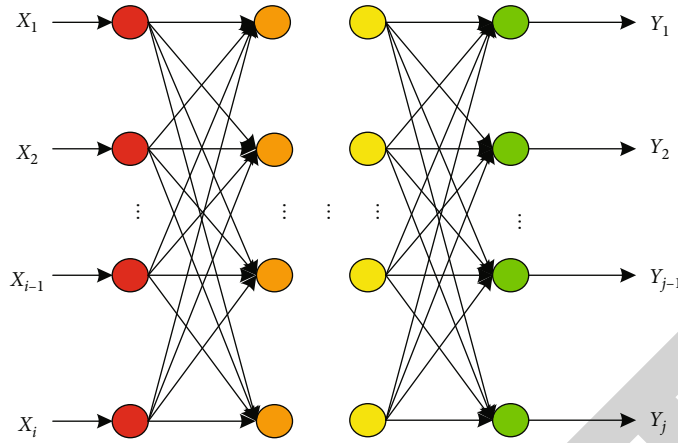


FIGURE 1: Schematic diagram of the structure of the BP neural network.

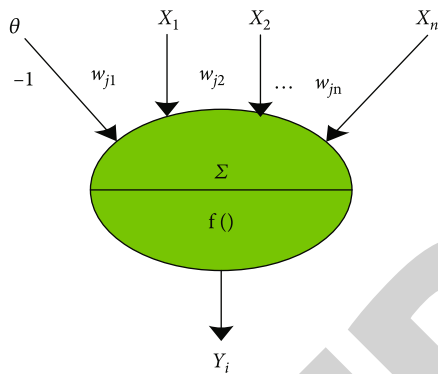


FIGURE 2: Schematic diagram of a single neuron.

$$\begin{aligned}
 a_i &= a_i \cdot \varphi, \eta_i = \eta_i, \Delta E(N) < 0; (i = 1, 2), \\
 a_i &= a_i \cdot \beta, \eta_i = 0, \Delta E(N) > 0; (i = 1, 2), \\
 \Delta E(N) &= E(N) - E(N - 1), \\
 \varphi &> 1 \beta < 1,
 \end{aligned} \tag{3}$$

If $\Delta E(N) > 0$, it means that the learning error will increase, and the current output value is deviating from the expected value. At this time, the weight adjustment amount shall be reduced and learning efficiency shall be lowered to get rid of the inertia term. If $\Delta E(N) < 0$, it means that the gradient modification direction is correct, so adding the inertia term will increase the learning rate and improve the learning efficiency. The value of a_i and η_i shall be between 0 and 1 [5].

3. Optimization Algorithm for Communication Resource Allocation in the Complex Network

3.1. *Communication Indicators in the Complex Network.* A complex network mainly contains the following communication indicators:

- (1) Average path length: the distance $d(x, y)$ between any two nodes in the network is the shortest path connecting the two nodes including the number of vector edges. The average path length is the average distance between all nodes in the network. The specific formula is

$$L = \frac{\sum d(x, y)}{N(N - 1)/2}. \tag{4}$$

The average path length can reflect the length of the communication link between network nodes.

- (2) Clustering coefficient/cluster coefficient: if there are at most $T(n) = k(n)[k(n) - 1]/2$ vector edges between the connection of a node n and the remaining $k(n)$ nodes in the network and if there are $E(n)$ vector edges between $k(n)$ nodes, then the clustering coefficient formula for node n is

$$C(n) = \frac{E(n)}{T(n)}. \tag{5}$$

The clustering coefficient is mainly used to measure the clustering of network nodes, and the clustering coefficient is to measure the nature of this network. The specific network clustering coefficient formula is

$$C = \frac{\sum C(n)}{N}. \tag{6}$$

The clustering coefficient can reflect the distribution characteristics of network nodes as a whole. Research has proven that regular networks have larger cluster coefficients and average distances, while random networks have smaller cluster coefficients and average distances [6].

- (3) Distribution of connectivity $p(k)$: connectivity of network node n is the number of edges connected to this node vector $k(n)$. By randomly selecting a

node in the network, the probability of connectivity k is p . The function $p(k)$ where p value changes with the change of k is the connectivity distribution

3.2. Improvement of Neural Network Algorithm Input and Output Calculation. The Hopfield neural network is a continuous Hopfield neural network used based on the allocation model of interference resources. Its input and output relationship formula is

$$\begin{aligned} S_j &= \sum W_{ij}v_j + I_j, \\ a_j u_j &= -b_j \frac{du_j}{dt} + S_j, \quad a_j, b_j > 0, \\ v_j &= f(u_j) \quad j = 1, 2, \dots, n, \end{aligned} \quad (7)$$

The input weight S_j of the above neuron and the input state u_j of the neuron can be expressed by a dynamic equation, and the function of neuron transfer is usually represented by $f(u) = 1/(1 + e^{-\lambda u})$. W_{ij} represents the feedback weight of the output neuron j to the input neuron, and u_j represents the neuron, the output state of the element [7].

3.3. Energy Function and Stability. The definition formula of the Hopfield energy function is

$$E = -\frac{1}{2} \sum_i^n \sum_j^n W_{ij} \cdot v_i v_j - \sum_j^n v_j I_j + \sum_j^n a_j \int_0^{v_j} f^{-1}(v) dv, \quad (8)$$

In the above formula, $f^{-1}(u)$ represents the inverse function of u , that is, $f^{-1}(u_j) = u_j$.

If the network satisfies $W_{ij} = W_{ji}$, $W_{ij} = 0$, i, j represents the number of neurons, and f^{-1} represents a monotonically increasing function, indicating that the above-mentioned network is stable, and the function of network energy corresponds to the objective function, so that the optimal solution to the problem is obtained when the function converges to the minimum value.

3.4. Communication Resource Allocation. First, the bit allocation plan and the subchannel allocation plan are confirmed. After the traffic size of each user and the instantaneous channel gain are known, the subchannels are allocated reasonably and dynamically to minimize the system transmission power P_T and then to complete bit loading as well as power allocation on each subchannel [8].

3.4.1. Dynamic Subchannel Allocation Analysis. Instead of considering the actual channel characteristics of all users, the traditional subchannel allocation solution only allocates the number of subcarriers based on users' size of traffic and allocates fixed subcarriers to each user, so it belongs to a static subchannel allocation method. In view that the actual channel characteristics of each user are not taken into account, the solution of dynamic subchannel allocation is realized in this research through the Hopfield neural network [9–11].

TABLE 1: Experimental data.

Experimental parameters	Experimental value
Radius of the cell (m)	600
Total bandwidth (MHz)	8
Path loss (dB)	128.1 + 37.6lgd
Noise power (dBm Hz ⁻¹)	-183
Base station to user SINR threshold (dB)	-10
User and user SINR threshold (dB)	-8.4
Base station to user distance (m)	100
Number of experiments	1000

The number of subchannels available to each user is identified by using the traffic size of each user, and then, the subchannels are dynamically assigned based on the instantaneous channel gain of each user (the required bit error rate is set to the same value for all users). The specific optimization function formula is

$$\min_{\rho_{k,n}} P_T = \min_{\rho_{k,n}} \sum_{k=1}^K \sum_{n=1}^N \frac{P}{a_{k,n}^2} \rho_{k,n}, \quad (9)$$

where P represents the transmission power of the next channel symbol in the fixed modulation method, which is processed by normalization [12, 13]. The specific objective function P_T formula is

$$P_T = \sum_{k=1}^K \sum_{n=1}^N \frac{P}{a_{k,n}^2} \rho_{k,n}. \quad (10)$$

And the formula for constraint condition is

$$1 = \sum_{k=1}^K \rho_{k,n}, \quad (n \in \{1, 2, \dots, N\}), \quad (11)$$

$$N = \sum_{n=1}^N \sum_{k=1}^K \rho_{k,n}. \quad (12)$$

Formula (11) satisfies that all subcarriers can only be used by one user, while formula (12) satisfies that all subcarriers can be used by users [14].

By combining the above algorithms, the subchannel distribution neural network energy function E is set; the specific formula is [15]

$$E = A \sum_{k=1}^K \sum_{n=1}^N \frac{P}{a_{k,n}^2} \rho_{k,n} + \frac{B}{2} \sum_{n=1}^N \left(\sum_{k=1}^K \rho_{k,n} - 1 \right)^2 + \frac{C}{2} \left(\sum_{n=1}^N \sum_{k=1}^K \rho_{k,n} - N \right)^2. \quad (13)$$

The neural network motion formula for solving the allocation of subchannels is given as

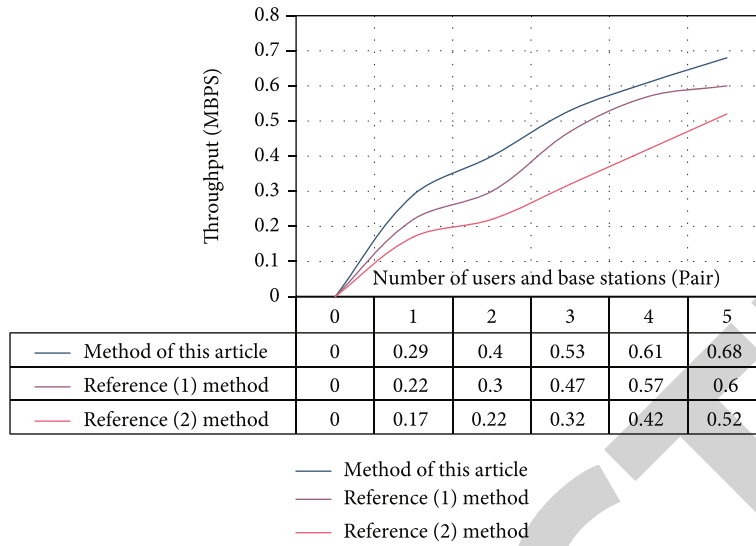


FIGURE 3: Comparison of throughput of different algorithms.

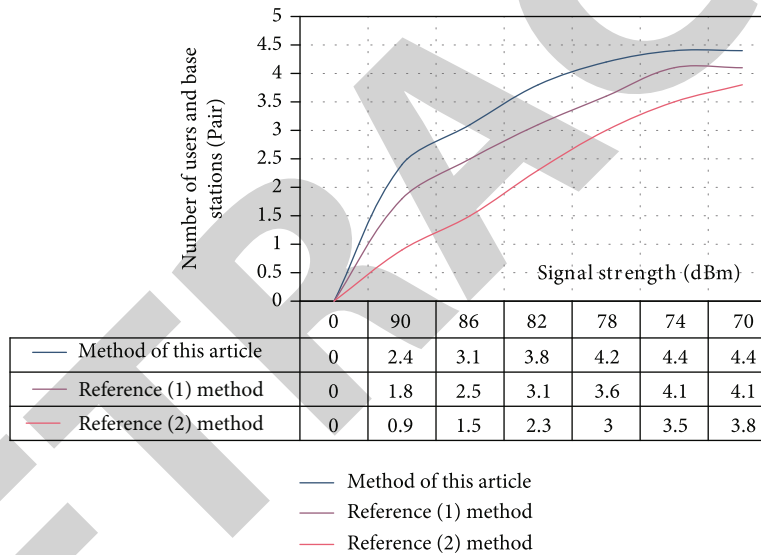


FIGURE 4: Comparison of user communication under different signal strengths.

$$\begin{cases} \frac{dU_{k,n}}{dt} = -A \sum_k \frac{P}{a_{k,n}^2} \rho_{k,n} - B \left(\sum_{k=1}^K \rho_{k,n} - 1 \right) - C \left(\sum_n \sum_k \rho_{k,n} - N \right), \\ V_{k,n} = g(U_{k,n}) = \frac{1}{2} \left[1 + \tanh \left(\frac{U_{k,n}}{U_0} \right) \right], \end{cases} \quad (14)$$

where A , B , and C represent the empirical output value [16], and the excitation function is $g(\cdot)$, which is a tangent function that approximates the S-shaped hyperbolic.

$$g(\cdot) \frac{\partial E}{\partial U_n} = - \frac{\partial U_n}{\partial t}. \quad (15)$$

The problem of dynamic subchannel allocation is con-

firmed via formulas (13) and (14). Starting from the network initial state U_0 , the network data [11] is obtained through continuous iteration of the motion equation, which is the subchannel allocation matrix ρ .

3.4.2. Power Allocation and Bit Loading of Subchannels. The above problems are solved by the Hopfield neural network, and ρ , v , and l are used to jointly represent the bit loading matrix c [17]. $\rho_{k,n}$ represents the n th subchannel allocated by the k th user; $v_{n,j}$ represents the allocation of j bits in the n th subcarrier, and $l_{n,j}$ represents the number of bits to be allocated in the n th subcarrier. The calculation method of the bit loading matrix is described as follows:

- (1) A new matrix $n \times n$ can be obtained by multiplying the v and l matrices with the corresponding elements

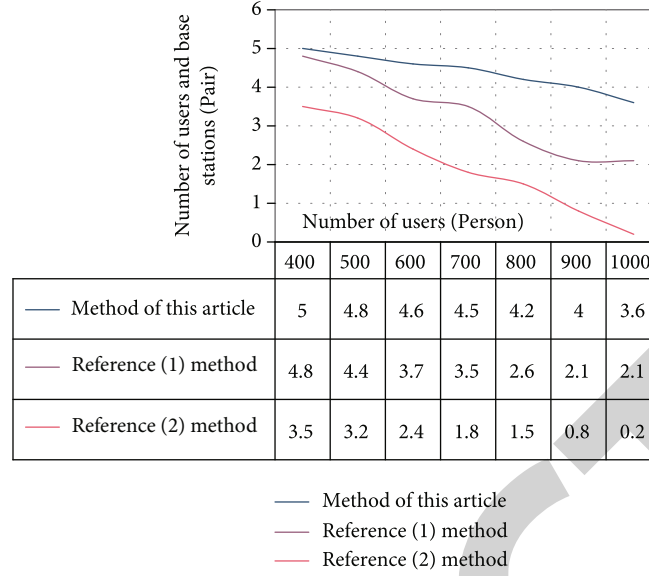


FIGURE 5: Comparison of communication signals under different distances of users.

- (2) A new matrix mpn [14, 18] can be obtained by multiplying the matrix of 1 row and j columns where all elements of 1 are with matrix npn
- (3) The bit loading matrix can be obtained by expanding the matrix mpn into k rows and then multiplying the elements corresponding to ρ [19]

Through the bit loading matrix calculation method, the specific formula can be obtained as

$$\min P_T = \sum_{k=1}^K \sum_{n=1}^N \frac{T \rho_{k,n} (2^{\rho_{k,n} v_{n,j} l_{n,j}} - 1)}{a_{k,n}^2}, \quad (16)$$

where $T = N_0/3[Q^{-1}(p_e/4)]^2$.

And the formula for constraint condition is

$$1 = \sum_{k=1}^K v_{n,j}, \quad (17)$$

$$N = \sum_{n=1}^N \sum_{j=1}^J v_{n,j}, \quad (18)$$

$$R_k = \sum_{n=1}^N \rho_{k,n} v_{n,j} l_{n,j}. \quad (19)$$

Formula (17) satisfies the fixed number of bits allocated by all subcarriers; formula (18) meets that all subcarriers can be allocated bits; formula (19) can meet the transmission rate of all users [20].

Based on this problem, the neural network energy function E for power allocation and bit loading is set. The specific formula is

$$E = A \sum_{k=1}^K \sum_{n=1}^N \frac{T \rho_{k,n} (2^{\rho_{k,n} v_{n,j} l_{n,j}} - 1)}{a_{k,n}^2} + \frac{B}{2} \sum_{n=1}^N \left(\sum_{j=1}^J v_{k,n} - 1 \right)^2 + \frac{C}{2} \left(\sum_{n=1}^N \sum_{j=1}^J v_{n,j} - N \right)^2 + \frac{D}{2} \sum_{k=1}^K \left(\sum_{n=1}^N \sum_{j=1}^J \rho_{k,n} v_{n,j} l_{n,j} - R_k \right)^2. \quad (20)$$

The movement of the neural network connected to the power allocation and bit loading is calculated. The specific formula is

$$\begin{cases} \frac{dU_{n,j}}{dt} = -A \sum_k \frac{T \rho_{k,n} (2^{\rho_{k,n} v_{n,j} l_{n,j}})}{a_{k,n}^2} \rho_{k,n} l_{n,j} - B \left(\sum_j v_{n,j} - 1 \right), \\ -C \left(\sum_n \sum_j v_{n,j} - N \right) - D \sum_k \left(\sum_n \sum_j \rho_{k,n} v_{n,j} l_{n,j} - R_k \right), \\ V_{n,j} = g(U_{n,j}) = \frac{1}{2} \left[1 + \tanh \left(\frac{U_{n,j}}{U_0} \right) \right]. \end{cases} \quad (21)$$

where D represents the value given by experience. The Hopfield neural network about the dynamic bit loading problem can be confirmed through formulas (20) and (21). Starting from the initial state U_0 [21], the network output v is obtained based on the continuous iterations of the network motion formula, and then, the transmission power can be obtained through formula (19), which is minimized to complete the optimal allocation of communication resources [22].

4. Verification of Experiment Simulation

4.1. Experiment Parameter. In order to verify the effectiveness of the proposed method, a base station is set in simulated community where users are evenly distributed and each user uses

only one communication resource node. The simulation is completed via MATLAB, which requires 200 Kb/s of data rate, $B = 1$ MHz of bandwidth, $N = 64$ of the number of subcarriers, and $BER = 10^{-3}$ of the bit error rate. The data is randomly generated each time, and the average value is taken. The specific parameters are shown in Table 1: the unit of distance between the transmitter and the receiver is km.

4.2. Analysis of Experiment Results. The proposed method is compared with the approaches in literature [1] and literature [2], as shown in Figure 3.

According to Figure 3, in the process of constantly increasing the number of network communication user nodes, the proposed method can output the minimum linear result by improving the neural network algorithm and ensure the fine scheduling requirements of complex network without waste of resources, so its throughput is always better than that of the other two methods.

Under different communication signal strengths, the result of comparing the number of connections between users is constructed; the RSRP is valued between -90 and -70 dBm. The details are shown in Figure 4.

As can be learned from Figure 4, the results of the proposed method are superior to those of the other two methods because it calculates complex network communication indicators, clarifies the subfactors that affect communication quality, focuses on analysis, and effectively suppresses the influence of the outside world on the distribution method. When the RSRP value is -90, the user communication connection is not interrupted, so that the data-related information is successfully transmitted and the timeliness is achieved.

In order to further prove the effectiveness of the proposed method, the number of users in the complex network at the same time is increased, as shown in Figure 5.

Figure 5 shows that as the number of users increases, the number of communication connections gradually decreases, mainly because the demand for communication signals must be guaranteed by increasing the transmission power. In view of this, the original interference to the users will increase, thereby affecting the constructed user communication connection. However, by means of introducing the improved neural network algorithm, the proposed method can suppress external interference based on enhanced allocation, and its decline rate is much lower than that of the approaches proposed in literature. Therefore, it is proven that the proposed method has strong applicability and can cope with the continuous access as well as output of big data under complex networks.

5. Conclusions

In the network applications in China's large cities, take the cellular network as an example. Despite many numbers of cellular network base stations in large cities, the load of cellular network base stations varies greatly in different time periods and locations. For example, during office hours, there are more numbers of mobile phone network users in the city subway, office buildings, and other areas. During city off-hours, the number of cell phone network users increases in areas such as city residential buildings and hotels. This will lead to differences in the

efficiency of network usage in the above-mentioned areas at different time periods, in order to improve the efficiency of complex network communication resource classification.

The proposed optimization algorithm for communication resource allocation in the complex network based on the improved neural network can properly allocate network communication resources with large network throughput and better communication effect. However, due to the influence of external interference factors, the communication connection in the proposed method will still be reduced, so further research is needed to achieve more reasonable allocation of communication resources.

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 state that there is no conflict of interest.

Acknowledgments

The manuscript is supported by the Science and Technology Project of Chongqing Municipal Education Commission: Research on Key Technologies of Big Data Service Platform Based on Gridded Smart Community (KJQN202003109).

References

- [1] Z. Xiaoyan, C. Yajun, and Y. Peiyan, "Joint algorithm of resource allocation and power optimization in D2D communication," *Microelectronics and Computers*, vol. 35, no. 5, pp. 1–4, 2018.
- [2] G. Quanming, Y. Mengting, J. Xiujie, and A. Junshe, "Cache resource allocation algorithm for SpaceWire network hotspot communication mode," *Journal of National University of Defense Technology*, vol. 40, no. 2, pp. 41–47, 2018.
- [3] Y. Bing, C. Yongli, and Z. Ruixue, "Resource allocation scheme for D2D communication in millimeter-wave 5G network," *Telecommunications Science*, vol. 35, no. 1, pp. 144–152, 2019.
- [4] L. Li, "Research on optimization of mobile communication network based on deep neural network," *Modern Electronic Technology*, vol. 43, no. 10, pp. 83–85, 2020.
- [5] X. Wenwang, S. Yunlian, Y. Shimin, W. Huayou, and X. Bingham, "A power line communication data processing algorithm based on improved VPGA optimized Elman neural network," *Power System Protection and Control*, vol. 47, no. 6, pp. 64–71, 2019.
- [6] Z. Long, G. Wei, W. Jianyong, and H. Jietao, "Network security evaluation model based on neural network algorithm," *Journal of Shenyang University of Technology*, vol. 40, no. 4, pp. 426–430, 2018.
- [7] X. Wang Ruyan and N., W. Dapeng, "Resource allocation mechanism of virtualized cloud wireless access network based on energy consumption and delay awareness," *Journal of Electronics and Information*, vol. 41, no. 1, pp. 83–90, 2019.
- [8] T. Shuzhu, Y. Peng, Y. Dawei, and Y. Shaowei, "Optimal design of near space communication network based on

Retraction

Retracted: Exercise Recommendation Model Based on Cognitive Level and Educational Big Data Mining

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Pu and H. Chen, "Exercise Recommendation Model Based on Cognitive Level and Educational Big Data Mining," *Journal of Function Spaces*, vol. 2022, Article ID 3845419, 11 pages, 2022.

Research Article

Exercise Recommendation Model Based on Cognitive Level and Educational Big Data Mining

Yongming Pu ¹ and Hongming Chen²

¹Teachers' College, Chengdu University, Chengdu 610106, China

²Chengdu Jing Furong Yi Du School, Chengdu 610061, China

Correspondence should be addressed to Yongming Pu; puyongming@cdu.edu.cn

Received 20 May 2022; Revised 7 July 2022; Accepted 12 July 2022; Published 27 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Yongming Pu and Hongming 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.

There are differences in the learning ability and cognitive ability of different learners. The unified exercises of traditional teaching ignore the differences of learners and cannot meet the personalized needs of learners. Previous recommendation systems focus on the optimization of recommendation performance, rarely clearly reflect the learning state of learners' knowledge points, and there are large errors in the recommendation results. This paper combines the comprehensive cognitive analysis module and the classified knowledge point cognitive analysis module to analyze the cognitive degree of learners' knowledge points. Based on the analysis results, appropriate exercises are selected from the educational resource data to form a list to be recommended. The experimental results show that the exercise recommendation algorithm based on cognitive level and data mining has better recommendation effect and accuracy than the other two recommendation models. The error between the actual difficulty of recommended exercises and the index value is very small. It can recommend an appropriate exercise list according to the actual situation of learners. The teaching comparison results show that the exercise recommendation algorithm can meet the personalized needs of students, recommend targeted exercises, and effectively and greatly improve the learning effect and test scores in a short time. When the motion recommendation algorithm based on cognitive level and data mining has the best recommendation effect, the cognitive module of classifying knowledge points accounts for a large proportion in parameter adjustment. Compared with other recommendation systems, this model has higher accuracy and recommendation effect.

1. Introduction

One of the ultimate goals of school education is to prepare students for employment. For students in school, they have relatively little understanding of the skills required for future work. Therefore, the choice of courses is often lack of pertinence, which cannot play a good support for future employment. In order to avoid the blindness of students' curriculum selection, schools need to combine the specific situation and similar situation of students in school. The learning process of graduates recommends appropriate learning plans for them and makes dynamic adjustments according to the actual situation of students in the learning process. This tailored course teaching method enables every student to get timely guidance in the learning process. It is conducive to the sustainable development of students' learn-

ing interests, helps students finally complete the study of this subject, masters the knowledge of relevant fields, and plays a good auxiliary role in the future employment process.

The purpose of knowledge learning lies in application. Continuous application of knowledge is not only an important means to deepen learners' understanding of knowledge points and enhance their mastery of knowledge points but also an indispensable part of education. Limited by educational resources, in the past, educators not only needed to teach dozens of students at the same time but also provided relatively limited exercise resources for students to practice. Educators could not pay enough attention to each student, analyze each student's current mastery of knowledge points, and solve each student's existing problems. Popular education methods and popular education and exercise resources make many learners spend more time finding their own

problems and solving them in the sea of invalid questions [1]. The development of information technology has broadened the way for learners to obtain learning resources. The exercise resources available to learners have increased exponentially. The diversification of exercise types helps learners optimize the whole knowledge system [2]. However, in the face of a large number of exercises, learners also need to pay the cost of mental strength and time to choose exercises, which will lead to the wrong evaluation of the cognitive level of current knowledge points due to the improper selection of exercises, increase the learning burden, and reduce the learning efficiency [3]. Therefore, based on educational data mining, providing personalized educational resources for educators and learners has become a research hotspot in the field of education. The researchers draw lessons from film and television resource recommendation and shopping recommendation and introduce various recommendation algorithms into educational resource recommendation. Among them, the most common recommendation algorithms are collaborative filtering algorithm based on user and item history interaction information, content-based recommendation algorithm, and knowledge information-based recommendation algorithm [4]. These algorithms can provide corresponding recommendations according to different emphases, but with the explosive increase of information and data, the recommendation function of basic recommendation algorithms and models has not met the expectations of users, and further improvement and perfection are an inevitable trend.

The research innovation lies in introducing a cognitive level model to analyze the cognitive status of learners' knowledge points based on educational big data. Through the established exercise recommendation model to achieve effective exercise table recommendation, the test and comparative experimental results of the exercise recommendation model are analyzed. The experimental results show that the motion recommendation algorithm based on cognitive level and data mining has better recommendation effect and accuracy than the other two recommendation models. The error between the actual difficulty of recommended practice and the index value is very small. It can recommend appropriate exercise tables according to the actual situation of learners. The teaching comparison results show that the exercise recommendation algorithm can meet the personalized needs of students, recommend exercises pertinently, and effectively and significantly improve the learning effect and test scores in a short time.

2. Research, Development, and Current Situation of Educational Data Mining and Recommendation System

The explosive growth of information data promotes the research of related retrieval technology. Recommendation system is one of its branches. At present, it has a relatively perfect and mature theoretical organization [5]. The key content of the recommendation system is the recommendation algorithm, which is related to the recommendation

method, performance and results, and the effect of meeting the objectives and requirements. According to the method and algorithm rules, the content-based recommendation algorithm, through collaborative filtering algorithm and the combination of the two are common systems [6]. In order to catch up with the demand of information and data processing, some scholars continue to improve it. Some scholars integrate big data technology into the system and carry out information data mining and recommendation with the help of association rules [7]. Ball proposed that the item recommendation algorithm should consider the practicability of items to users and sort and recommend on this basis [8]. Jiang and Yang establish fuzzy sets in the learning resource recommendation system and use the knowledge reasoning model to extract and recommend the required information [9]. With the diversified development of user needs, Huo et al. have introduced context aware model into mobile social networks to achieve the purpose of user recommendation [10]. In addition, Yunita et al. pay attention to the personalized needs of users, meet the purpose of personalized resources and information data recommendation through algorithm improvement and optimization, and apply it in the fields of film, education, and so on [11].

Educational data mining originated in the late 1980s. In the primary stage of its development, the level of information data acquisition and processing technology is low, which cannot meet its needs to achieve a large number of data acquisition and research. The research methods are limited, and the results obtained are relatively few [12]. The development of computer information technology provides a driving force for educational data mining. At the same time, the rise and scale expansion of online education provide more channels and space for obtaining educational data information. The scale of relevant data expands rapidly, which provides effective and large amounts of data for educational data mining research, and the results are gradually enriched [13]. Wang and Fu take students as the starting point and use cluster model analysis to conclude that there is a large gap in the level of mastering knowledge points among student groups [14]. Other scholars found through the analysis results that the teaching effect will be largely affected by the teaching content and students' own preferences, and the important influencing factor of students' own learning effect is their learning habits [15]. For the research of teaching resources, Bhat et al. pointed out that teaching resources have a hierarchical structure, and there is a correlation between this structure and curriculum characteristics [16]. Zhang et al. pointed out that the main reason for the differences in learning effects lies in the differences of learners. For different learners, educational resources with different emphasis should be provided to realize accurate and personalized retrieval and recommendation [17]. Based on this, Fan et al. have introduced collaborative filtering recommendation algorithm into personalized education recommendation to complete education resource recommendation according to similar interactive information between learners [18]. Through the research on students' learning methods and state, D'Agostino et al. pointed out that exercise is not only an important means to improve learning

effect but also a necessary process. Improving the retrieval of exercise information is to improve learners' learning efficiency [19]. Aziz combines analytic hierarchy process and semantic correlation analysis model to obtain learners' personalized information and complete exercise recommendation on this basis [20]. Chen et al. believe that the problem of exercise recommendation is the transformation of information recommendation. The target area is the students who need exercise recommendation, and the auxiliary scope is the users who have completed the search of relevant historical information. With the help of the auxiliary area information, we can obtain more accurate exercise classification results for the target area, which greatly avoids the recommendation error [21]. Cheng and Bu put forward new recommendation rules from the perspective of learners' probability of doing the right exercise, that is, if learners have more than 50% probability of doing the right exercise, they will make recommendation [22]. The probability setting of this algorithm is enlightening and can achieve the optimal selection in a certain range, but it is not the global optimal. Before recommending exercises, Rianto and Fachrie analyze the current knowledge learning status of students through expert evaluation, form the recent development area with relevant data, and use the gambling machine algorithm to select the exercises with the best effect in this range [23]. The development of deep learning theory has opened up new algorithmic ideas for researchers. MC et al. have introduced deep reinforcement learning theory into the review system, and the constructed learning system can provide interval repetitive learning [24]. In addition, some scholars optimized the network model based on the review model to enhance the performance. To sum up, at present, the optimization and improvement of many exercise personalized recommendation systems focus more on the performance of algorithms, and there are few models that can truly and accurately show learners' current knowledge mastery and learning state, which makes exercise recommendation lack effectiveness and has weak advantages in improving learners' sense of learning experience. In the future development of exercise recommendation algorithm, the position of learners will be improved. The system will not only mention improving performance but also continuously increase personalized service, emphasizing the central role of learners.

3. Construction of Exercise Recommendation Algorithm Based on Cognitive Level and Data Mining

3.1. Idea of Establishing the Overall Framework of Exercise Recommendation Algorithm. The current motion recommendation algorithms of data mining focus on mining association rules in data. By effectively discovering, understanding, and applying association rules, we can make complex and useful knowledge hidden in a large number of sources make greater contributions to the construction of modern education system. The data required by these systems come from students' examination scores accumulated in the teaching process over the years. Through the in-

depth mining of these data, it is not difficult to find that the level of students' grades not only depends on the curriculum itself but also is affected by many aspects, such as the curriculum of the discipline, the formulation of teaching plans, and the order of each course.

Although the era of big data has reduced the time cost of obtaining information, it has increased the time cost of effective information screening and the difficulty of information processing. The purpose of exercise recommendation algorithm is to reduce the time for learners to choose exercises, help learners choose appropriate exercises from massive resources, and gradually enhance their cognitive level. Exercise is the practical link of learners in the process of education. It is an important link to deepen learners' understanding and application of knowledge. It is an unavoidable part for both learners and educators. Many educators believe that learning is a process in which practice makes perfect. Exercises are a tool to help learners improve their proficiency. They should complete as many exercises as possible in a limited time, which will help students deepen the memory of knowledge points. According to the relevant research results, the increase in the number of exercises does not necessarily achieve the effect of improving academic performance. On the contrary, it is very likely that too many exercises will lead to the reduction of learners' thinking time on each question and only know one of the learning knowledge points. In addition, educators need to face a certain number of learners. They are not allowed to give targeted guidance to learners in terms of time and energy. They can only take the situation of most learners as the main basis for teaching feedback and arrange exercises accordingly. For advanced and backward learners, the difficulty of exercises does not meet their current learning state, which greatly affects the learning effect. The same exercises cause different difficulties for different learners. The unified explanation of educators can only solve most of the problems, and some individual problems still become obstacles for learners.

To sum up, the exercise recommendation model not only needs to select suitable exercises for learners in a short time but also needs to clearly distinguish between completed and unfinished, correct, and wrong exercises. This requires the establishment of the exercise recommendation model to truly and accurately reflect the cognitive state of learners and realize effective exercise recommendation for current users in combination with other learners' data information in the relevant range. Figure 1 shows the flow chart of exercise recommendation algorithm based on learners' cognitive level.

The exercise recommendation algorithm based on learners' cognitive level establishes the completed exercise module, marked exercise module, and knowledge point information module, respectively, which provide the input original information. According to the relevant cognitive theories and models, calculate the comprehensive cognitive degree of each learner's knowledge and the cognitive degree of knowledge in different modules, and obtain relevant data information, combined with collaborative filtering algorithm, it can predict the probability that each exercise in the recommended exercise set may be correct. On the basis

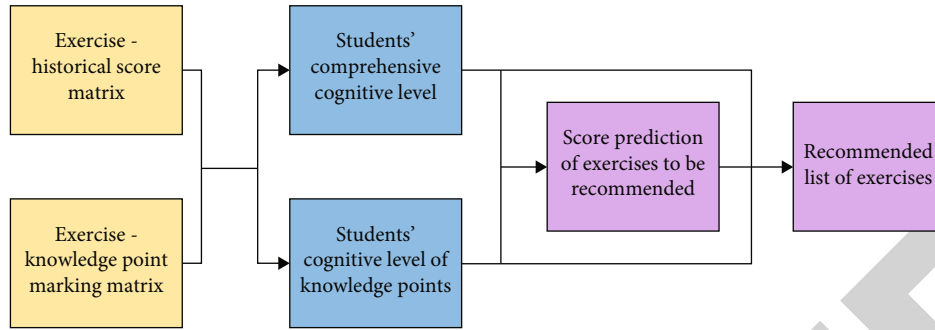


FIGURE 1: Flowchart of exercise recommendation algorithm based on learners' cognitive level.

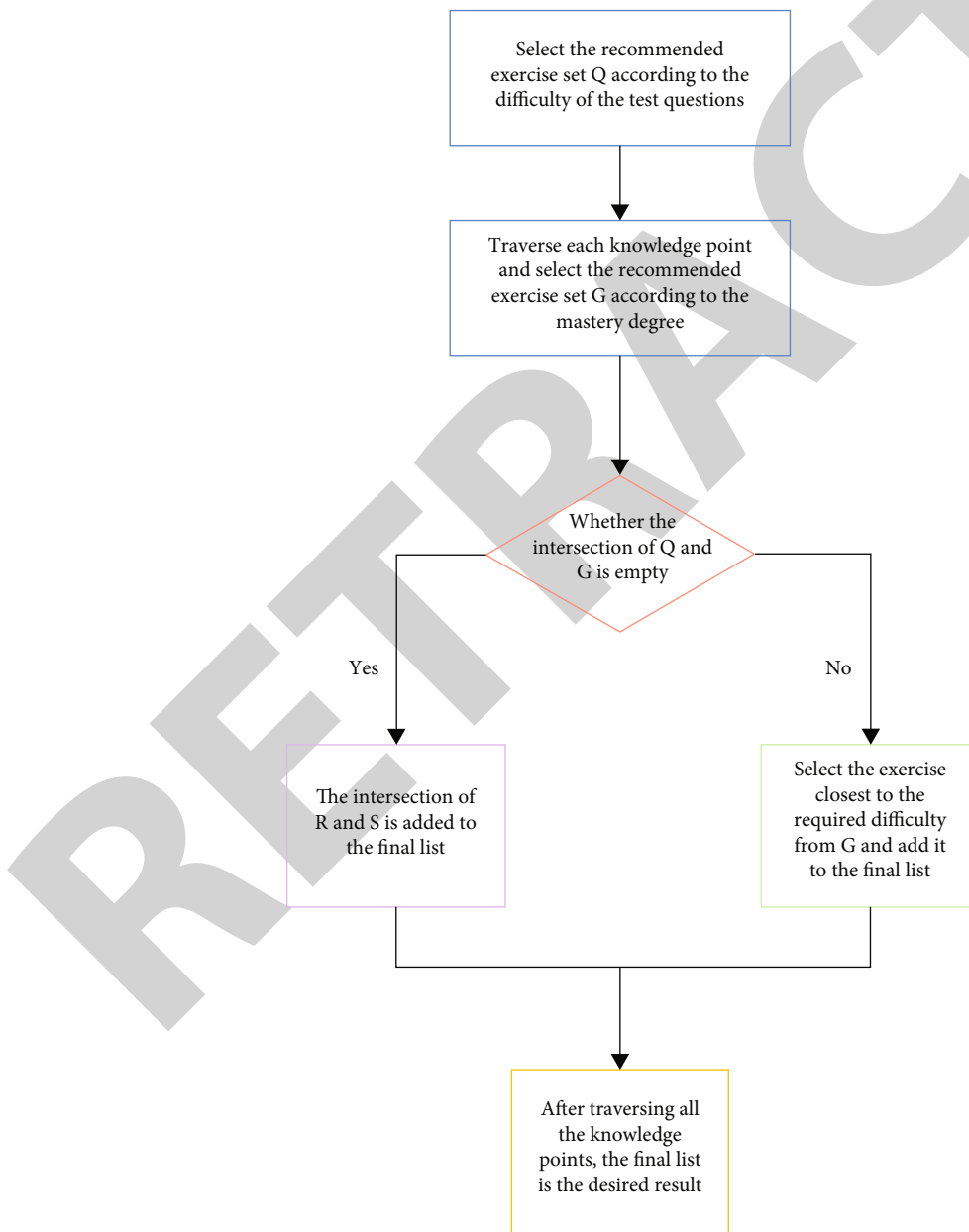


FIGURE 2: Exercise recommendation algorithm based on cognitive level and data mining.

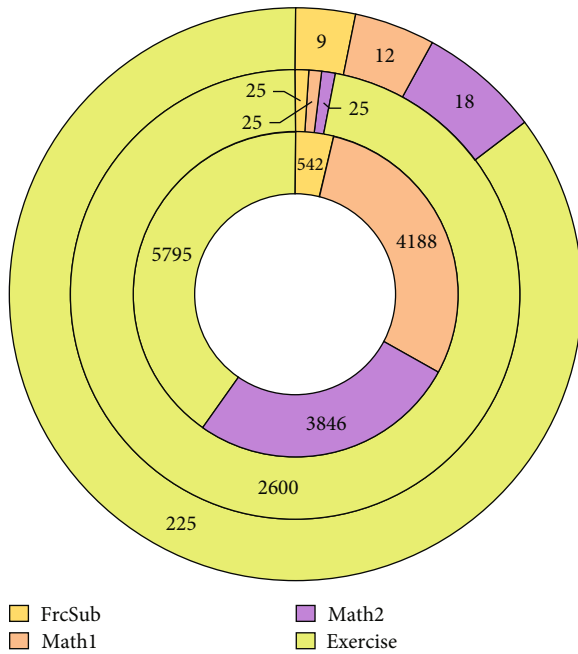


FIGURE 3: Details of four data sets.

of three aspects of data information, the system can obtain the recommended set of exercises and combine the output.

3.2. Establishment of Exercise Recommendation Algorithm Based on Learners' Cognitive Level Model. The comprehensive cognitive level of learners is their ability to apply the content. Learners' self-evaluation and mutual evaluation can not only promote students' learning of knowledge but also improve students' evaluation ability. Such evaluations should therefore be actively encouraged. Ask students to browse each other's works and put forward revision suggestions according to the rubric. According to relevant cognitive theories and models, the calculation formula of the correct rate of learners a completing exercise b is set as shown in the following formula:

$$P_{ab} = \lambda_b + \frac{1 - \lambda_b}{1 + e^{-1.7\beta_b(D_a - \varepsilon_b)}}, \quad (1)$$

where λ_b represents the probability of learners guessing the exercises correctly without any relevant knowledge; β_b represents the difference degree of exercises; ε_b indicates the difficulty of the exercise; D_a reflects learners' comprehensive cognitive level.

By solving the parameters in formula (1), the learners' cognition of the comprehensiveness of knowledge can be obtained. Let $R_{ab} = 1$ represent the learners' correct answer to the exercise, and $R_{ab} = 0$ represent no correct answer. The new calculation formula of the correct probability is shown in the following formula:

$$P_{ab}^{R_{ab}} (1 - P_{ab})^{1 - R_{ab}}. \quad (2)$$

Set the number of exercises as B , and its maximum like-

lihood function is established as shown in the following formula:

$$L(\beta_b, \varepsilon_b, D) = \prod_{a=1}^A \prod_{b=1}^B P_{ab}^{R_{ab}} (1 - P_{ab})^{1 - R_{ab}}. \quad (3)$$

Take the corresponding log likelihood function for the next derivative, as expressed in the following formula:

$$\ln L = \sum_{a=1}^A \sum_{b=1}^B (R_{ab} \ln P_{ab} + (1 - R_{ab}) \ln (1 - P_{ab})). \quad (4)$$

The derivative of the obtained result with respect to the unknown parameter is obtained and its derivative is zero, as shown in the following equations:

$$\frac{\partial \ln L}{\partial \beta_b} = 0, \quad (5)$$

$$\frac{\partial \ln L}{\partial \varepsilon_b} = 0, \quad (6)$$

$$\frac{\partial \ln L}{\partial D_a} = 0, \quad (7)$$

$$1 \leq a \leq A, 1 \leq b \leq B. \quad (8)$$

The cognitive level of various knowledge modules shows the level of learners' mastery of each type of knowledge points. The results of learners' exercises based on this prediction are more targeted. Let the number of knowledge points contained in exercise b be expressed as V and predict the correct state according to the in-depth learning of learners' knowledge points. The calculation is shown in (11):

$$\eta_{ab} = \prod_{v=1}^V C_{av}^{J_{bv}}. \quad (9)$$

Among them, the examination status of knowledge takes you is J_{bv} , and its status is divided into examination and nonexamination, that is, when $J_{bv} = 1$ or $J_{bv} = 0$. The state of learners' mastery of knowledge points is described by C_{av} . when the value is 1, learners have mastered knowledge points, and when the value is 0, learners have not mastered knowledge points. In fact, when learners have mastered the knowledge points, they still guess the correct answer due to careless mistakes, or when learners have not mastered the knowledge points. Considering this situation, introduce the careless parameter w into the model and guess the parameter l correctly, as shown in the following formula:

$$P_{ab} = P(R_{ab} = 1 | C_a) = w_b^{1 - \eta_{ab}} (1 - l_b)^{\eta_{ab}}. \quad (10)$$

Among them, the probability of learners not doing right due to carelessness is w_b , and the probability of correctly guessing the answer is l_b . It is also necessary to solve the position parameters in the above formula to obtain the

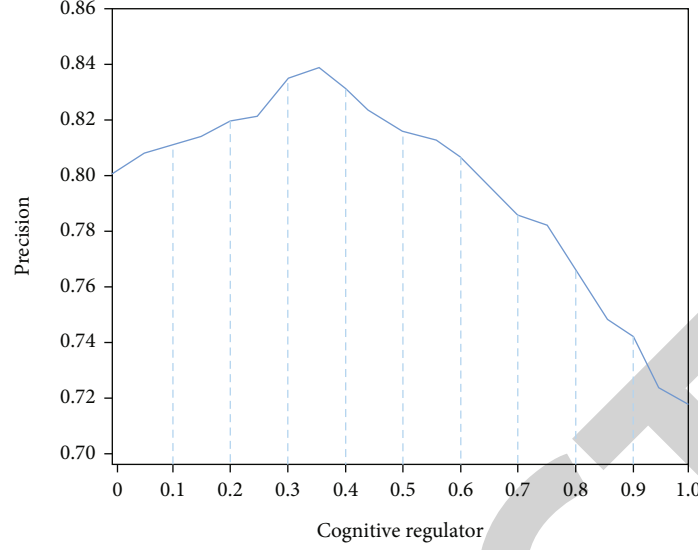


FIGURE 4: The results of frcsub dataset exercise recommendation model under different parameter values.

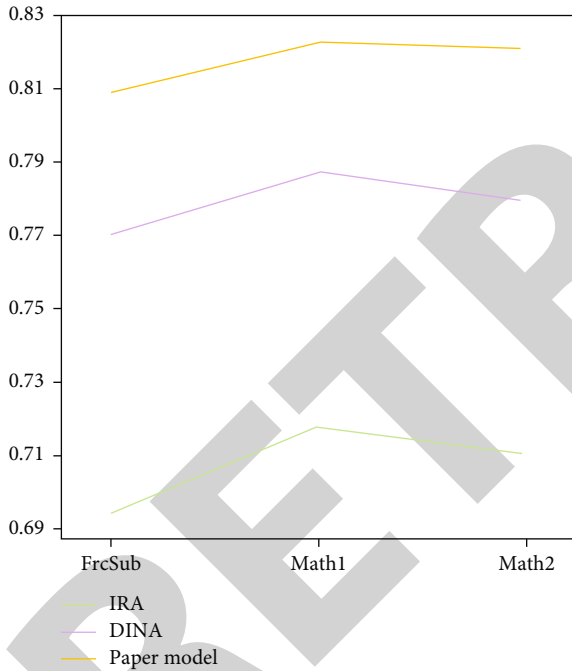


FIGURE 5: Prediction and test results of exercise accuracy of three exercise recommendation models in three open data sets.

learners' cognitive level in the classification knowledge points.

There are only two cases for the value of cognitive degree parameters of classified knowledge points obtained through the above. In the actual situation, learners' cognitive state has always changed, which not only has tortuous progress but also is vulnerable to forgetting. Therefore, learners' mastery state of knowledge points cannot be judged by two points, but needs to reflect its continuity. The processing method is to expand the amount of data contained in the learner set with similar knowledge mastery level, and the

parameter C_{av} is processed continuously by taking the set mean. The calculation formula of the level similarity of learners' mastery of knowledge points is as follows:

$$\text{sim}(i, j) = \frac{|M| + |N|}{|M \cup N|} \cdot \frac{1}{|D_i - D_j|}. \quad (11)$$

Among them, the set of exercises with correct answers by both learners is M , and the set of exercises with wrong answers by both learners is N . The comprehensive cognitive level of learners i is described by D_i , sorted according to the similarity calculation results between learners i and other learners, and the top h is selected to form its set of adjacent learners, which is set as H . Replace the original value of the understanding level of continuous classification knowledge points obtained through the following formula:

$$C'_{av} = \frac{\sum_{m \in H} C_{mv}}{h}. \quad (12)$$

Among them, the numerator represents the number of current knowledge points mastered by learners in the adjacent learner set.

After calculating the comprehensive value of learners' cognitive level and the cognitive level value of classified knowledge points, the correctness of the answers to the exercises that have not been done by learners shall be predicted, and the prediction probability of learners' correct answers to the exercises is set as shown in (15):

$$R_{ab} = (1 - \delta) \sum_{bv} \sqrt{\prod_{v=1}^V C'_{av} + \delta} \frac{(D_a + 3)}{6}. \quad (13)$$

The cognitive regulation parameter is expressed as δ .

The list of recommended exercises of the model is based on the prediction accuracy and learners' understanding of

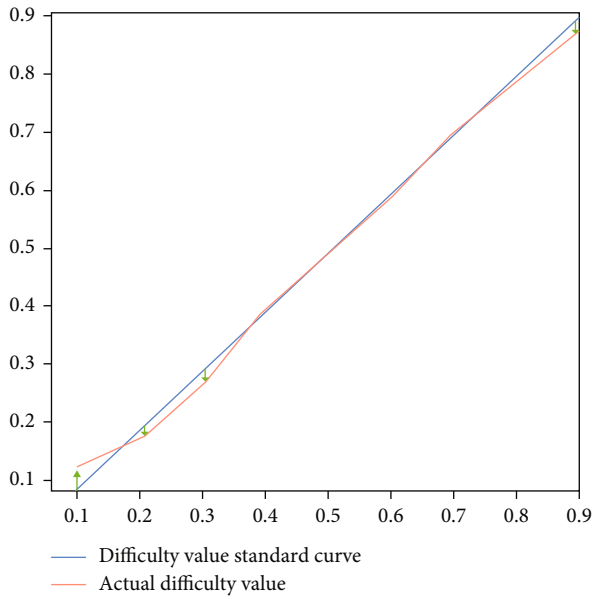


FIGURE 6: Comparison results of actual recommended exercise difficulty and theoretical difficulty index values of exercise recommendation model.

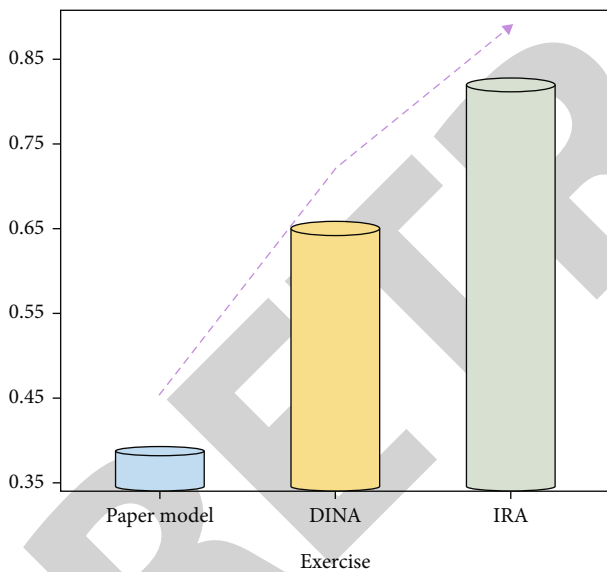


FIGURE 7: Comparison results of exercise list recommendation effects of three exercise recommendation models.

classified knowledge points in the above. The combination of the two can ensure learners' good practice experience, recommend targeted and systematic exercises, take care of all knowledge points, and promote learners' enthusiasm. According to the above, the result range is $[0,1]$. Set the upper and lower line of the accuracy value of the recommendation list within its range, that is, θ_1, θ_2 , and the accuracy range of all the exercises to be recommended in the recommendation list G is $[\theta_1, \theta_2]$.

For the list Q of exercises to be recommended generated from classified knowledge points, the average value of

learners' understanding of knowledge points shall be obtained before obtaining the results. This is the threshold of recommended exercises. The calculation method is as follows:

$$\mu = \frac{\sum_{v=1}^h C_{av}}{h}. \quad (14)$$

The mastery level of all exercises in list Q is lower than the threshold. Compare these exercises with those in list G to be recommended, and obtain the exercises jointly owned by them and put them into the final exercise recommendation list, as shown in Figure 2. If there are no exercises jointly owned by the two, the exercises with the closest prediction accuracy to $[\theta_1, \theta_2]$ in all relevant exercises of the knowledge point will be included in the final list.

The advantages and disadvantages of the recommendation model are judged by the following indicators. The first is the prediction index of learners' score accuracy. The calculation method is shown in (17):

$$\text{Precision} = \frac{FM}{RM}. \quad (15)$$

Among them, the number of exercises whose predicted results are consistent with the actual results is FM , and the number of all recommended exercises is RM .

The difficulty degree of recommendation list of exercise recommendation model is calculated, as shown in the following formula:

$$SR = \frac{TM}{RM}. \quad (16)$$

Among them, the number of exercises that learners actually answered correctly is described by TM .

The calculation formula of recommended efficiency of exercises with full coverage of model knowledge points is shown in (17):

$$MP = \frac{RM}{\text{Total}}. \quad (17)$$

4. Experimental Results of Exercise Recommendation Algorithm Based on Cognitive Level and Data Mining

Previous recommendation systems focus on the optimization of recommendation performance, rarely can clearly reflect the learning state of learners' knowledge points, and there are large errors in the recommendation results. This paper combines the comprehensive cognitive analysis module with the classified knowledge point cognitive analysis module to analyze learners' cognitive degree of knowledge points. Based on the above excellent theory and practice, inherit and carry forward the previous research results. It is intended to introduce data mining methods into problem-solving solutions. So that it can scientifically guide

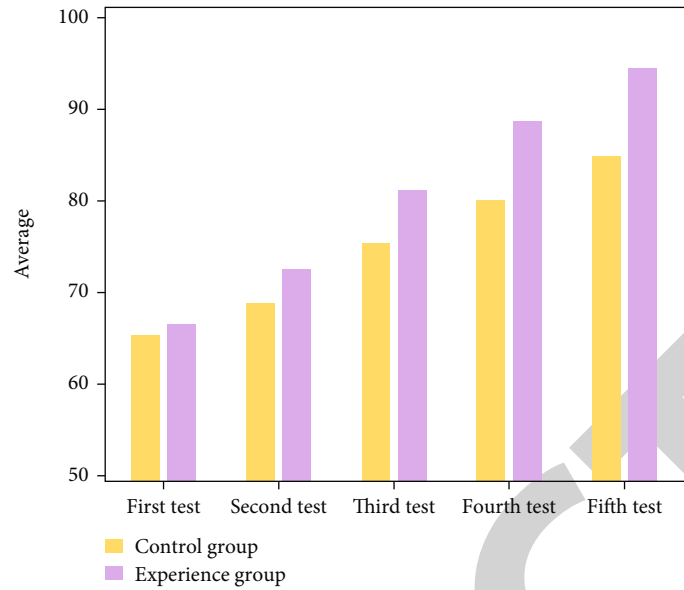


FIGURE 8: Comparison of average scores of five tests between experimental class and control class.

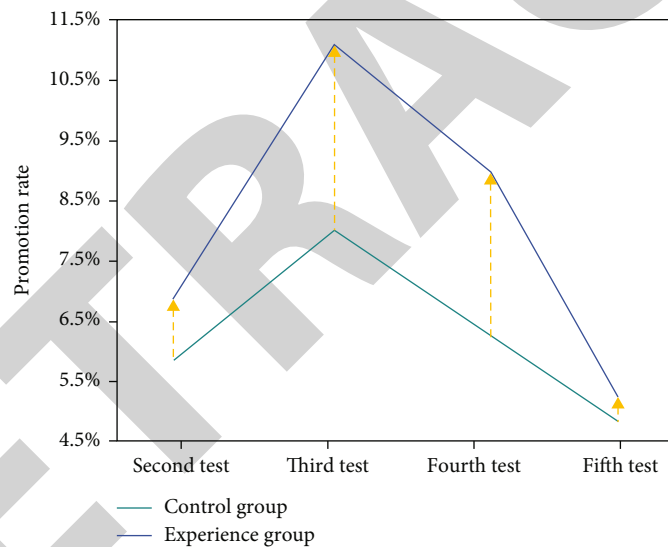


FIGURE 9: Comparison of the average score improvement rate of four tests between two classes.

students to arrange courses according to students' employment satisfaction and other factors.

In this paper, four data sets are selected for the recommended comparative experiment of mathematical exercises. Three data sets are open, but the scale of relevant data is small, and the amount of mathematical exercises is not enough to ensure the experimental effect. Therefore, the fourth data set is used to make up for it, as shown in Figure 3.

After the preliminary preparation for model training, call the fit interface to start the training process. You need to specify at least three key parameters: training data set, training rounds, and single training data batch size. The classification accuracy will still shake or fluctuate when the overall trend declines. If you stop when accuracy begins to

decline, you will definitely miss a better choice. So a good solution is to terminate when the classification accuracy is no longer improved within a certain period of time. Of course, it is OK to use loss in this area, and loss is also a criterion. The cognitive adjustment parameters in the exercise recommendation model have a regulatory effect on two cognitive degree modules, that is, in $\delta = 0$, the learners' cognitive degree of classified knowledge points determines the prediction results of accuracy, on the contrary, in $\delta = 1$, the learners' comprehensive cognitive degree plays a major decisive role. Figure 4 shows the result changes of exercise recommendation model of frctest dataset under different parameter values. The overall data in the figure shows a state similar to the positive Pacific distribution, that is, the values on both sides are lower than the middle value. The higher

the value, the better the recommendation effect. Therefore, the performance of exercise recommendation only based on any learner's cognitive level is not good. The exercise recommendation results obtained by the combination of the two are higher than those in a single case, which shows that the combination of comprehensive cognition and the cognitive degree of classified knowledge points can better show the learners' actual cognitive state from multiple angles. In addition, in the case of single cognition, the cognitive degree of classified knowledge points is higher than that of comprehensive cognition. When the data performance is the best, the cognitive degree of classified knowledge points accounts for more proportion in the value of adjustment parameters, which shows that learners improve the cognitive degree of classified knowledge points, even if the granularity of knowledge points is smaller, so as to promote the accuracy of exercises.

The adjustment parameter of the model is $\delta = 0.35$. The recommended model, project reflection theoretical model, and classical discrete model are tested for exercise accuracy prediction in three open data sets. The comparison of the results is shown in Figure 5. Among the three models, the recommended model in this paper shows the best results, the gap between the results of the classical discrete model and the model in this paper is small, and the accuracy of the project reflection theoretical model is the least ideal. This shows that the proposed model is more accurate in reflecting learners' cognitive level. In addition, the size of the data set has a certain impact on the test results. The larger data set provides more effective data for the model, which is conducive to the model to obtain finer prediction results.

In practice, learners with different cognitive states need different exercise difficulties. The exercise recommendation model can select and recommend exercises within a certain difficulty range according to the middle value of different difficulty needs. Exercise recommendation model is used to describe the internal and external learning characteristics of learners, and it is the premise and foundation of learning analysis. Practice recommendation model is of great value to teachers, learners, and learning system managers. Whether the various characteristics of learners contained in the practice recommendation model are complete and accurate is related to whether teachers can classify learners with similar learning characteristics according to the model and then provide students with personalized learning content, strategies, and learning resources. In addition, the practice recommendation model is conducive to learners' in-depth understanding of their learning status and shortcomings and then correct their learning behavior in advance.

Figure 6 shows the comparison results of the actual recommended exercise difficulty and theoretical difficulty index values of the exercise recommendation model. The data in the figure shows that the recommended difficulty of practical exercises is basically consistent with the theoretical difficulty on the whole, and there is only a small error at both ends, that is, when the difficulty of exercises is in the range of high or low, the actual difficulty of exercises will have a certain deviation, which is mainly because the coverage of model compromise processing knowledge points has an impact

on the actual presentation of exercise difficulty. On the whole, this model can recommend a list of exercises with appropriate difficulty according to learners' actual needs and cognitive state.

As shown in Figure 7, the comparison results of exercise list recommendation effects of three exercise recommendation models are shown. The recommendation efficiency of exercises with full knowledge coverage is taken as the test index. The larger the value of this index, the worse its efficiency is. At this time, the difficulty coefficient is uniformly set to 0.55. In order to ensure the accuracy and effectiveness of the experimental results, this part of the experiment is carried out in the fourth data set with large data scale. The results show that the index value of this model is the smallest, that is, the recommendation efficiency is the highest, and the gap between the index values is relatively large compared with the other two models, indicating that the efficiency of this model has been greatly improved. Therefore, the model design in this paper has a good effect on the coverage of knowledge points and abnormal data processing and can ensure a good effect of exercise recommendation.

In this paper, two classes with the same number and academic achievements in a middle school are selected for the comparative experiment of the application of exercise recommendation model. One class is the control class, which adopts the traditional educational exercise practice method, and the other is the experimental class, which adopts the exercise recommendation model practice in this paper. There are five performance tests during the comparative experiment. The first is the preexperiment test. The results are shown in Figure 8. The test results before the experiment show that the average score gap between the two classes is very small, and the average score keeps increasing during the experiment. In the subsequent tests, the average scores of the experimental class are higher than those of the control class. This shows that the exercises in traditional teaching are helpful to improve learners' performance. The exercises recommended by the exercise recommendation model are more targeted and perform better in improving learning efficiency and test performance.

As shown in Figure 9, the results of the improvement rate of the average score of the four tests of the two classes are compared. The improvement rate of the average score of the four tests of the experimental class is higher than that of the control class. The improvement rates of the second and last tests of the two classes are relatively low, mainly because of the low mastery and application ability in the primary stage of learning knowledge points. In the fifth test, the learning knowledge points have entered the optimization period. In this stage, the score improvement difficulty is high, and the promotion rate is reduced. In the third and fourth tests, the promotion rate of the experimental class is significantly higher than that of the control class, which shows that in the knowledge point foundation consolidation and application stage, the exercise recommendation algorithm based on cognitive level and data mining meets the needs of learners in different stages and strengthens the practice of learners' knowledge points in a targeted and personalized manner in a short time to quickly improve learning efficiency.

5. Conclusion

This paper establishes an analysis model of learners' cognitive level of knowledge points, selects appropriate exercises according to the analysis results, and forms a recommendation list. According to the teaching comparison experiment, in the practical application of the recommended exercises, the exercise recommendation system meets the personalized needs of learners, maintains good accuracy, effectively helps learners lay a solid foundation, improves learning efficiency and effect, and greatly improves examination performance. In order to comprehensively analyze learners' cognitive state from multiple perspectives, the cognitive model includes a comprehensive cognitive analysis module and a cognitive degree analysis module of classifying knowledge points. The problem recommendation algorithm based on cognitive level and data mining proposed in this paper needs further systematic improvement and research on the impact of time factors on learners and the connection factors between knowledge points. The experimental results show that when the motion recommendation algorithm based on cognitive level and data mining has the best recommendation effect, the cognitive module of classifying knowledge points accounts for a large proportion in adjusting parameters. Compared with other recommendation systems, this model has higher accuracy and recommendation effect.

However, it is difficult to obtain fine-grained recommendation data information from the interaction between users and exercise information in this study, so the accuracy of exercise recommendation results needs to be improved. This needs to be 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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] L. I. Hao-Jun, Z. Zhang, and P. W. Zhang, "Personalized learning resource recommendation method based on three-dimensional feature cooperative domination," *Computer Science*, vol. 46, no. 1, pp. 471–477, 2019.
- [2] E. Fernandes, M. Holanda, M. Victorino, V. Borges, R. Carvalho, and G. Van Erven, "Educational data mining: predictive analysis of academic performance of public school students in the capital of Brazil," *Journal of Business Research*, vol. 94, pp. 335–343, 2019.
- [3] A. Noorian, A. Harounabadi, and R. Ravanmehr, "A novel sequence-aware personalized recommendation system based on multidimensional information," *Expert Systems with Applications*, vol. 202, p. 117079, 2022.
- [4] D. Mo, X. G. Chen, S. Duan et al., "Personalized resource recommendation based on collaborative filtering algorithm," *Journal of Physics Conference Series*, vol. 1302, no. 2, article 022025, 2019.
- [5] Y. S. Su and C. F. Lai, "Applying educational data mining to explore viewing behaviors and performance with flipped classrooms on the social media platform Facebook," *Frontiers in Psychology*, vol. 12, 2021.
- [6] B. M. M. Alom and M. Courtney, "Educational data mining: a case study perspectives from primary to University Education in Australia," *International Journal of Information Technology and Computer Science*, vol. 10, no. 2, pp. 1–9, 2018.
- [7] E. Kartikadarma, S. Jumini, N. Ismail, B. Fachri, D. Sudrajat, and R. Rahim, "Educational data mining to improve decision support on the ratio of students and study groups in elementary schools in Indonesia using K-means method," *Ilkogretim Online*, vol. 20, no. 1, pp. 691–698, 2021.
- [8] N. Ball, "Using educational data mining to identify and analyze student learning strategies in an online flipped classroom," *Education Sciences*, vol. 11, no. 11, p. 668, 2021.
- [9] Y. Jiang and J. Yang, *Research and Design of Personalized Learning Resource Recommendation System Based on Deep Neural Network*, China Computer & Communication, 2019.
- [10] Y. Huo, D. F. Wong, L. M. Ni, L. S. Chao, and J. Zhang, "Knowledge modeling via contextualized representations for LSTM-based personalized exercise recommendation," *Information Sciences*, vol. 523, pp. 266–278, 2020.
- [11] A. Yunita, H. B. Santoso, and Z. A. Hasibuan, "Research review on big data usage for learning analytics and educational data mining: a way forward to develop an intelligent automation system," *Journal of Physics: Conference Series*, vol. 1898, no. 1, article 012044, 2021.
- [12] L. I. Haojun, Z. Zhang, H. Guo, and D. Wang, "Personalized learning resource recommendation from the perspective of deep learning," *Modern Distance Education Research*, vol. 4, pp. 92–103, 2019.
- [13] P. Beibei, G. Juanqiong, and M. Wenxin, "Extracting topics and their relationship from college student mentoring," *Data Analysis and Knowledge Discovery*, vol. 2, no. 6, pp. 92–101, 2018.
- [14] H. Wang and W. Fu, "Editorial: physical layer security and wireless access control (QSHINE 2017)," *Mobile Networks and Applications*, vol. 25, no. 1, pp. 1–3, 2020.
- [15] H. E. Aouifi, M. E. Hajji, Y. Es-Saady, and H. Douzi, "Predicting learner's performance through video sequences viewing behavior analysis using educational data-mining," *Education and Information Technologies*, vol. 26, no. 5, pp. 5799–5814, 2021.
- [16] M. Bhat, M. Zaman, and M. Butt, "An intelligent prediction system for educational data mining based on ensemble and filtering approaches," *Procedia Computer Science*, vol. 2, no. 167, pp. 1471–1483, 2020.
- [17] X. Zhang, J. Shen, P. Wu, and D. Sun, "Research on the application of big data mining in the construction of smart campus," *Open Access Library Journal*, vol. 8, no. 11, pp. 1–10, 2021.
- [18] J. Fan, M. Zhang, A. Sharma, and A. Kukkar, "Data mining applications in university information management system development," *Journal of Intelligent Systems*, vol. 31, no. 1, pp. 207–220, 2022.
- [19] J. V. D'Agostino, E. Rodgers, and S. Konstantopoulos, "The effects of HEROES on the achievement levels of beginning readers with individualized education programs," *The Journal of Educational Research*, vol. 114, no. 5, pp. 433–444, 2021.

Retraction

Retracted: E-Commerce across Boarder Logistics Risk Evaluation Model Based on Improved Neural Network

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Qiao, "E-Commerce across Boarder Logistics Risk Evaluation Model Based on Improved Neural Network," *Journal of Function Spaces*, vol. 2022, Article ID 2355298, 10 pages, 2022.

Research Article

E-Commerce across Boarder Logistics Risk Evaluation Model Based on Improved Neural Network

Wei Qiao ^{1,2}

¹Graduate School, Management & Science University, Shah Alam 40100, Malaysia

²Department of Management, Wuxi Institute of Technology, Wuxi 214121, China

Correspondence should be addressed to Wei Qiao; qiaow@wxit.edu.cn

Received 19 May 2022; Revised 4 July 2022; Accepted 8 July 2022; Published 25 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Wei Qiao. 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.

BP neural network is a typical algorithm in artificial intelligence network. It has strong nonlinear mapping ability and is the most prominent part to solve some nonlinear problems. In the traditional BP algorithm, the coincidence initialization of weights and thresholds is random, which reduces the efficiency of the algorithm on the one hand and affects the accuracy of the algorithm results on the other hand. In order to solve these problems, this paper studies an e-commerce cross-border logistics risk assessment model based on improved neural network. This model can help merchants engaged in cross-border e-commerce to select appropriate third-party settlement platforms, so as to reduce the cost of merchants in the process of capital settlement. The key information in BP neural network algorithm is stored in weights and thresholds, which is enough to prove the importance of weights and thresholds for the effective operation of the whole network. The e-commerce cross-border logistics risk assessment model based on improved neural network aims to solve the problem of low level of risk assessment and the bottleneck of logistics risk assessment. The improved e-commerce cross-border logistics risk assessment model based on neural network can be used for risk rating before business development, so as to adopt different risk management methods for different risk levels.

1. Introduction

With the continuous development of economic globalization and the rise of mobile Internet, big data, cloud computing, and other information technologies, e-commerce across boarders has come into being. Under the background of the national macrodevelopment strategy of “One Belt, One Road” and “Pilot Free Trade Zone,” China’s e-commerce across boarders has become a “dark horse” in foreign trade [1]. E-commerce across boarders has also become a new hot spot, while the development of e-commerce across boarders is hot, and the development of e-commerce across boarder logistics can no longer meet the needs of e-commerce across boarder business development [2]. China’s cross-border trade is developing at a fast pace, but the constraints of real factors such as capital, scale, and management level directly lead to the dilemma that Chinese cross-border e-merchants generally face high costs in the settlement link, poor timeliness, and difficulty in repayment [3]. While being highly valued, the instability

and high failure rate of alliances make enterprises have to weigh the benefits and risks when joining them [4]. In addition, the lack of sufficient funds to build a modern information management system has led to “indigestion” and order loss in most logistics enterprises from time to time. Especially after the “double 11” promotional activities, it often highlights the weak links in China’s e-commerce logistics distribution [5]. The commodity itself and its price are no longer the most important factor affecting consumer shopping, replaced by the merchant’s customer service, logistics, and distribution of these services [6].

The government has continuously introduced relevant policies to support the development of e-commerce and has introduced laws and regulations and formulated regulatory systems at the level of standards and supporting systems, with a view to promoting the sustainable and stable development of e-commerce across boarders [7]. As more and more foreign logistics companies with advanced logistics technologies enter the Chinese market, domestic logistics companies are facing

increasingly fierce competition and have to develop innovative services and find new profit points. The provision of logistics risk evaluation services can not only strengthen the cooperation with upstream and downstream enterprises in the supply chain but also gain new profits [8]. Logistics risk evaluation is a complex multifactor comprehensive analysis process with incomplete information, involving many evaluation objects [9]. Cross-border logistics in the context of international trade belongs to the concept in a broad sense. Cross-border logistics in a broad sense refers to the logistics service activities between two or more countries, and cross-border logistics is a manifestation of logistics service development to an advanced stage [10]. With the support of the government, China's e-commerce across borders will certainly make great development and then promote the development of China's foreign trade. So in order to avoid the difficulties caused by incomplete survey information, reduce the influence of subjective factors in the evaluation, and avoid the problems caused by correlation among indicators, this paper adopts an improved neural network evaluation method with high nonlinear approximation ability, strong fault tolerance, self-learning, and easy to use [11].

Artificial neural network is a hot research field in recent years, involving many disciplines, and its application areas include modeling, time series analysis, pattern recognition, and control and are constantly expanding. The neural network will train the input according to its own network structure and learning rules and then perform specific tasks according to specific applications. Analyzing the risk indicators from the perspective of logistics enterprises is in line with the development of the times, which helps logistics institutions to rate the logistics of loan enterprises and reduce the loan risk. It avoids setting index weights artificially, effectively avoids some subjective and random factors, and makes the evaluation of logistics risk as objective as possible, so as to better avoid the risks in the implementation of logistics projects.

The innovations of this paper are

- (1) This paper analyzes the obstacles affecting the growth of China's e-commerce across borders—cross-border logistics—by using the latest developments in the industry and reviewing the current situation according to the environment and development of China's e-commerce across borders
- (2) The initial weights and thresholds of the improved neural network are used to conduct a comprehensive analysis of the parameters involved in the neural network, and the training model is used to determine the optimal parameters and select the most suitable network structure for empirical analysis
- (3) By analyzing the basic principles of neural networks and addressing the problems of long search time for extremes and falling into local optima of BP neural networks, the improved neural network algorithm is analyzed in this paper. Initialize multiple neural networks with different parameter values, and take the smallest as the result. Using random gradient descent is not the same as using standard gradient descent to

accurately calculate gradients. Random gradient descent method adds random factors to the calculation of gradient. So even if it falls into a local minimum, the calculated gradient may not be. In this way, it is possible to jump out of the local minimum and continue the search

2. Thoughts on the Construction of E-Commerce across Boarder Logistics Risk Assessment Model Based on Improved Neural Network

2.1. Establishment Method of Index System. The selection of logistics risk evaluation indicators faces more complex customer relationships.

First of all, for the establishment of logistics risk evaluation system, it is a systematic project, and each evaluation index subsystem consists of a set of indicators, which are independent of each other and linked to each other and can reflect the whole risk early warning system. Individual risk and overall risk are to determine the evaluation benchmark, which can be called individual evaluation benchmark and overall evaluation benchmark, respectively. Both risk suppression and risk compensation are measures to reduce losses after the occurrence of risks by adopting early warning programs and total programs. Through this series of methods, it constitutes the whole process of risk management, as shown in Figure 1.

Affecting logistics and transportation is also the degree of impact of natural disasters and changes in customs clearance policies of countries for e-commerce across boarder parcels; in addition, in terms of logistics capital chain settlement, the degree of impact of macroeconomic fluctuations of countries will also affect logistics costs to a certain extent. In the neural network, the neural favor nodes between layers are no longer in fully connected form. Using the local spatial correlation between layers, the neuron nodes of each adjacent layer are connected only to the upper layer neuron nodes that are close to it, i.e., locally connected. A multilayer forward network is built with the function to calculate the output M of the hidden layer.

$$M_j = f \left(\sum_{i=1}^n W_{ij} X_i - b_j \right), j = 1, 2, \dots, m. \quad (1)$$

In the e-commerce across boarder logistics system, as the cross-border logistics chain covers several supplier enterprises and provides logistics services to customers all over the world, the complex supplier and customer groups make high requirements on logistics and transportation capacity and risk control ability. In summary, the selected primary evaluation index system is shown in Figure 2.

Once the learning samples are input to the neural network, they start to pass forward in the direction of the input, hidden, and output layers, based on the weights and bias vectors of each layer, and finally, the actual output is obtained. An input vector (provided by the training samples) is transformed through a series of hidden layers to obtain an output vector,

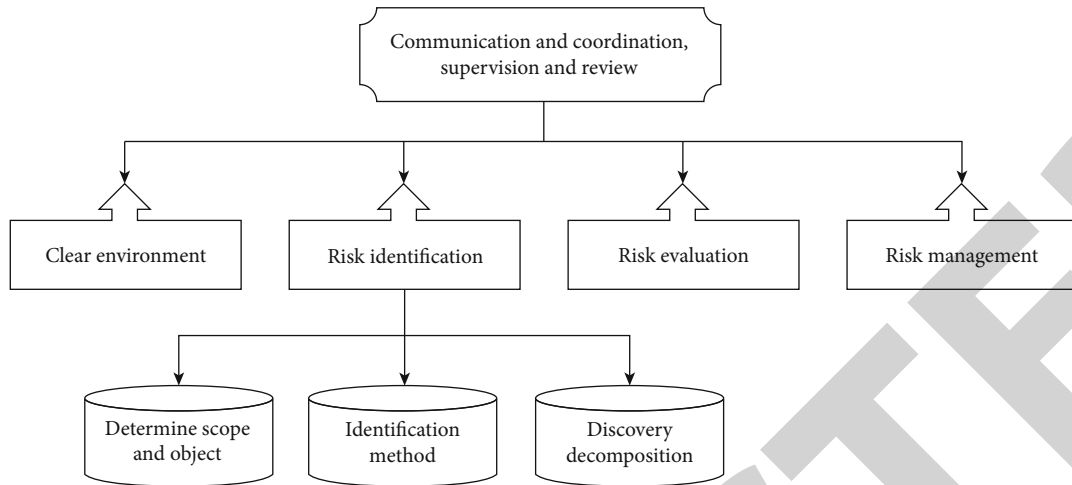


FIGURE 1: Risk management process.

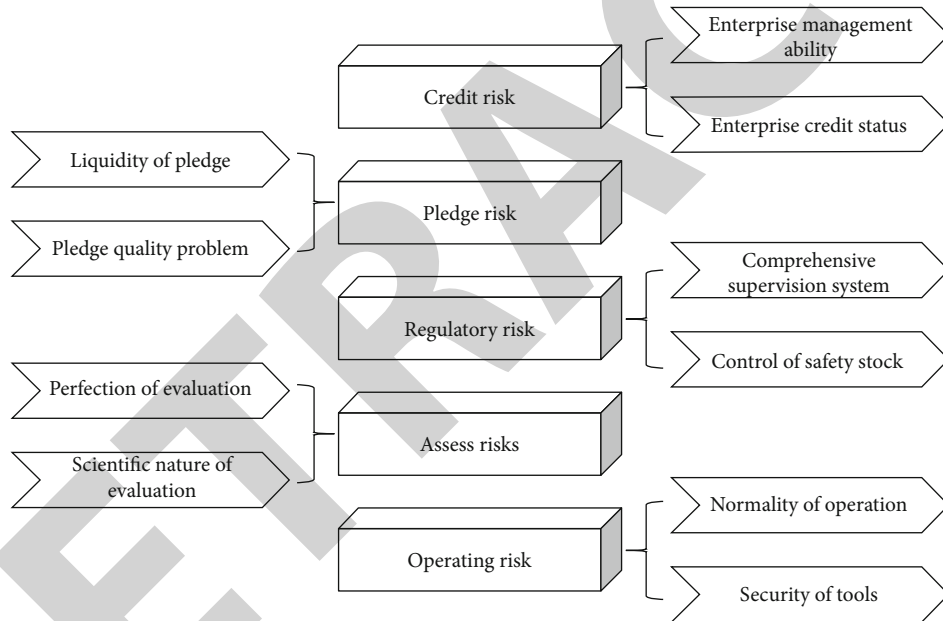


FIGURE 2: Initial evaluation index system of logistics risk.

thus achieving a mapping relationship between the input data and the output data. The Swish function is chosen as the activation function of the network model to improve the classification accuracy of the images. Its mathematical expression is given by the equations:

$$f(x) = x \cdot \sigma(\beta x), \tag{2}$$

$$\sigma(x) = \frac{1}{1 + \exp(-x)}. \tag{3}$$

$\sigma(x)$ is Sigmoid activation function number.

Second, the scientific nature of index selection is reflected in the combination of theory and practice, which should have both theoretical basis and also reflect certain objective reality.

In the management of overseas warehouse, the increase of product return rate, product life cycle fluctuation, and inventory risk will bring certain risks to the overseas warehouse storage link. To determine the overall risk level of logistics, we need to analyze the factors affecting the development of logistics, identify the logistics risks, and clarify the relationship between various risks, their interaction, and the extent and consequences of their impact on logistics implementation. The presence of a target in the region is determined by combining a set of weak classifiers through a simple linear boosting classifier with a decision function of the following equation:

$$\hat{y}_i = \sum_{m=1}^M \omega_m h_m(T_i). \tag{4}$$

h_m is a weak classifier. ω_m is the weight of the classifier should be weak. \hat{y}_i is the probability of existence of target. M is the number of weak classifiers.

In the process of transportation and distribution, cross-border logistics involves multiple transit and multiple batches, which requires longer in-transit transportation time. From the management aspect, higher transit rate of goods will bring greater risk of lost and damaged parts. Therefore, the neural network is used to calculate the difference in error among the actual output and the ideal output, and if the magnitude of the error does not satisfy the preset requirements, the signal of the error is spread along the direction of the output layer, the hidden layer, and the input layer to gradually update the model parameters. The positive spread of the information on the input and the negative spread of the error on the output constitute the information loop of the network, and the information loop function is

$$x' = \frac{2[x - 0.5(\max + \min)]}{\max - \min}. \quad (5)$$

x' is the normalize the calculated value. x is the value of the current sample point. \min is the minimum value of sample data. \max is the maximum value of the sample.

It is similar to the traditional neural network model in that the input is added to the network, and the output is calculated for each input vector and then the value of each correlation weight W_{ji} is corrected according to the following equation:

$$W_{ji}(t+1) = W_{ji}(t) + \Delta W_{ji}(t). \quad (6)$$

Finally, the indicators selected for logistics risk indicators should be available for quantitative analysis, and each indicator should have a strong realistic operability and comparability. The utilization rate of facility resources indirectly affects the size of transportation risks, and high transportation costs can also increase the risk of enterprises. Long-distance transportation vehicle traffic safety, transportation equipment turnover bumps, and other factors are potential risks. Since the transformation between the hidden layer output to the output layer is a linear transformation, people more often use the RLS algorithm which has been relatively mature. In the context of e-commerce across boarders, cross-border logistics has significant e-commerce characteristics, and commodity transportation is no longer manifested as cross-border spatial displacement of bulk commodities, but small batch and multifrequency cross-border spatial displacement of commodities through cross-border logistics mode, so this belongs to the category of cross-border logistics in a narrow sense. In solving the actual problem, the network composed of single neuron can hardly meet the requirements because its structure is too simple. In addition, we can also make simple improvements to the above modified equations like BP networks. Therefore, neural networks usually need to be composed of multiple neurons, and when the problem is complex, it is also necessary to extend the single-layer neural network into a multilayer neural network. That is, the predictability of risk, the probability of occurrence, and the prediction of consequences are carried out in

three aspects, and the influencing factors of logistics implementation risk are analyzed in a comprehensive and multilevel way.

2.2. Establishment of Risk Assessment Model. The learning process of BP network model consists of forward propagation and error backward propagation process, which embodies the best part of artificial neural network. Because of its better self-learning and self-association function among various neural network models, it has become the most widely used artificial neural network at present. E-commerce across boarder logistics is reasonably divided according to different customer states and provides reasonable and effective logistics services such as warehousing, distribution, and transportation. The node structure of e-commerce across boarders logistics supply chain is shown in Figure 3.

First, the fitting accuracy of the network is positively correlated with the number of layers and the number of nodes per layer, and increasing the number of layers can improve the fitting accuracy, but it complicates the network and increases the training time. Forward propagation is used to compute the forward network, i.e., to compute the output of a certain input information by the network. The original $n \times n$ size convolutional kernel now needs approximately only $2a + n + 1$ numbers to preserve the sparse structure, where a is the number of connection weights being preserved, n is the number of convolutional kernel rows or columns, and the remaining one digit preserves the position of the weight matrix in the whole filter set. Thus, the total compression ratio for network pruning is

$$CR_p = \frac{2a + n + 1}{n \times n}. \quad (7)$$

In order to reduce the transportation distance and transportation time of overseas e-commerce shipments, the overseas e-commerce goods can be transported and stored in bonded warehouses in the free trade zone in advance, and when the orders are generated, the goods will be cleared directly from the bonded warehouses into the Chinese territory. In subtractive clustering, each of its data points will be considered as a potential cluster center, and then, the probability of the point being considered as a cluster center is calculated based on the data density around each data point its. For an arbitrary signal $f(t)$ or function that satisfies $f(t) \in L_2(R)$ and $\psi(t)$ satisfies the wavelet tolerance condition, the continuous wavelet transform of $f(t)$ is defined as

$$WT_f(a, b) = |a|^{-1/2} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt, a \neq 0. \quad (8)$$

In this process, the update parameters (weights and bias values) of the model are gradually applied starting from the output layer to the input layer, and the error function is used to find the gradient of the weights and update the weight vector from the output layer to the concealed layer. There is a weight matrix w between the input layer and the hidden layer. The value of the hidden layer is obtained by multiplying the input X by the weight matrix. There is also a weight matrix from the hidden layer to the output layer. Each value of the

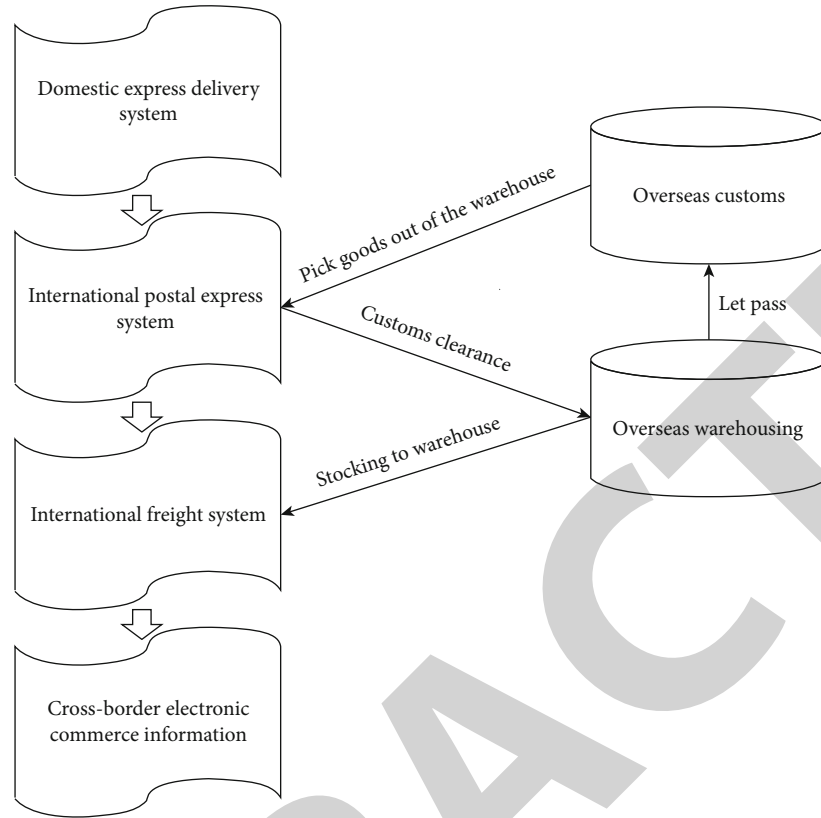


FIGURE 3: E-commerce across borders logistics supply chain node structure diagram.

output layer vector y is actually the vector point of the hidden layer multiplied by each column of the weight vector W . The class of forward neural network functions is dense in the space of vector-valued continuous functions, and in a consistent parametric sense, the class of forward network functions is defined here as a set of the following type:

$$N = \left\{ Y \in R^m, Y = W^{(K+1)} N_k \{ N_{k-1} [\dots N_1(X)] \} \right\}, K = 1, 2, 3 \dots \quad (9)$$

Second, the input data is processed so that the data changes between 0 and 1. The middle part of the tansig function changes more obviously and can better distinguish the input data, so the tansig function is used as the transfer function of the implicit layer. The selection probability proportional to the individual fitness value is usually used to randomly select individuals, so the selection probability is defined as

$$P_i = \frac{f_i}{\sum_{i=1}^n f_i}, i = 1, 2, \dots, n. \quad (10)$$

P_i is the selected probability. f_i is the individual fitness value. n is the population size.

Back propagation is used to pass the error layer by layer, modifying the connection weights and thresholds between neurons so that the output obtained by the network for the input information after calculation can meet the desired error

requirements. In addition to overseas e-commerce customers, the group's bonded warehouse can also provide short-term warehousing services to other international trading companies or domestic e-commerce companies or even other third-party logistics companies. All neurons between two adjacent layers are interconnected, while neurons that are on the same layer cannot be linked. The number of convolutional kernels in the lower layers is small because the structures in the lower layers are generally less diverse and smaller in size, while the higher convolutional layers are used to extract structural information in the higher layers, which are more diverse and larger in size, so the higher convolutional kernels are more numerous and larger in size. Therefore, it is also necessary to train a layer of convolutional network to get the region correction parameters. Suppose the a priori region parameters are defined as

$$P\{P_x, P_y, P_w, P_h\}. \quad (11)$$

P_x, P_y are coordinates of the central point of the prior area. P_w, P_h are the width and height of prior area.

Finally, a momentum term is added to the training function of the conventional algorithm as a damping term to reduce the tendency of oscillation in the learning process. Moreover, the learning rate of this function is adaptive, thus varying the learning rate according to the complexity of the problem to control the training time of the network. The input signal is transmitted from the input layer through the hidden unit to the output layer, and the output signal is generated at

the output end, which is the positive propagation of the working signal customs declaration and inspection is an important part of the customs clearance process. Inspection business before the customs clearance business, the object of inspection and quarantine departments, only the goods required by law need to be declared. In artificial neural networks, these units are called nodes, and each unit has a threshold value that is reached to receive and transmit information. By learning from valid data and training themselves, the processing units can then mine the information in the data and save it in the form of connection weights and thresholds. Usually, the customs clearance of the goods can be issued only after the inspection and quarantine department's inspection business is completed, and the customs will accept the application for customs clearance and examine and tax or detain the goods according to their category, value, quantity, and other factors. As import and export proceeds, the prices of the two commodities change accordingly, thus affecting consumption. Equilibrium is reached when the relative prices of the same commodity in both countries are the same, and the result of equilibrium is that both sides of the trade raise the level of domestic consumption and receive the benefits from international trade.

3. Application Analysis of Improved Neural Network in Logistics Risk Evaluation Model

3.1. Analysis of Learning Process of Neural Network. First of all, the learning algorithm of the neural network is actually the method of finding the minimum value of the error function, which uses the most rapid descent method, so that it is repeatedly trained to learn multiple samples and modify the connection weight coefficients by back propagation of the error. The objective function and network parameters are modified in the same way as the gradient descent method, except that a threshold value must be given. The optimal number of hidden nodes is selected by training the number of hidden nodes between [5, 10] and [10,15] one at a time. The variation of the mean squared deviation of the network structure between [5, 10] and [10,15] for different functions is shown in Figures 4 and 5.

Since, the cross-border logistics process is more complex, including not only the traditional logistics links and international freight but also the customs and commodity inspection of the output country, customs and commodity inspection of the input country, logistics and distribution of the input country, reverse logistics, and other links. Therefore, the pruning rate has a great impact on the overall accuracy of the network, and the higher the proportion of retained connections, the smaller the loss of network accuracy, especially the pruning of the initial network layer conv1 causes a great loss of accuracy. Neural networks can have multiple layers, each layer can have multiple neurons, the number of neurons in each layer can be unequal, and even each neuron in the same layer can have a different transfer function. Through the variance variation graph of each network structure above, the mean square error and the number of training steps for different number of hidden layer nodes are compared as shown in Table 1.

Next, the neural network is changed along the negative gradient direction of the output error function, and the error function is made to converge to the minimum point of the function at the end. The number of nodes in the hidden layer is taken as one, and then, the value of the objective function at the last time is recorded after a certain number of iterations using the gradient descent method. The target classification and detection layers have great fluctuations in the network pruning rate, mainly because this part is mainly divided into two parts of the network, where the target classification network is not sensitive to the network pruning degree. The original input of the multilayer neural network is the input layer of the network, the last layer is the output layer of the network, and the middle layer is the hidden layer. The output of each neuron is connected to other neurons, thus forming a dynamic feedback relationship, and the network structure has the ability of self-searching for superiority regarding the energy function. Besides, e-commerce across borders and other species such as payment and customs will also produce synergy, cross-border logistics and other species such as payment and customs will also produce synergy, and both e-commerce across borders and cross-border logistics will face the influence of internal environment and external environment. So in neural networks, there are two ways to reduce the error and improve the accuracy, one is to increase the number of network layers, and the other is to increase the number of neurons in the hidden layer; the former tends to make the network too complex, and the network training time is greatly increased; compared with the former, the training effect of the latter is easier to observe and adjust than the former. Searching for the sample individuals with the optimal fitness, obtaining the optimal decoding value, and applying it to the established neural network structure are the initial weight threshold between the connections of each layer of this network, and the comparison of the actual output of the training samples with the fit results of the desired output values is shown in Figure 6.

Finally, with the logistics risk level as the output, the BP neural network uses MATLAB to randomly generate the weights and bias values in the initial stage, and after training the input and output data, the BP neural network finally obtains the weights and bias values. According to the results of the comprehensive model evaluation, it is found that the evaluation result of e-commerce across boarder ecosystem synergy is "average," which indicates that the overall e-commerce across boarder ecosystem synergy is weak at present, which is a significant lack of synergy. Therefore, the number of nodes in the hidden layer is increased by one, and it is still iterated with a certain number of gradient decreases as in the beginning. If, after the iteration, the objective function decreases by a value greater than the threshold value, this indicates that an additional hidden node has been added, which is more useful for the network. At this point, the number of hidden nodes does not make the objective function of the system extremely small, so the number of hidden nodes is added by one, and the iterations are continued as above. And its corresponding target region transformation parameter prediction network needs to get more accurate data values, so it is extremely sensitive to network pruning, and a smaller connection retention rate will cause a rapid decrease in the overall network detection

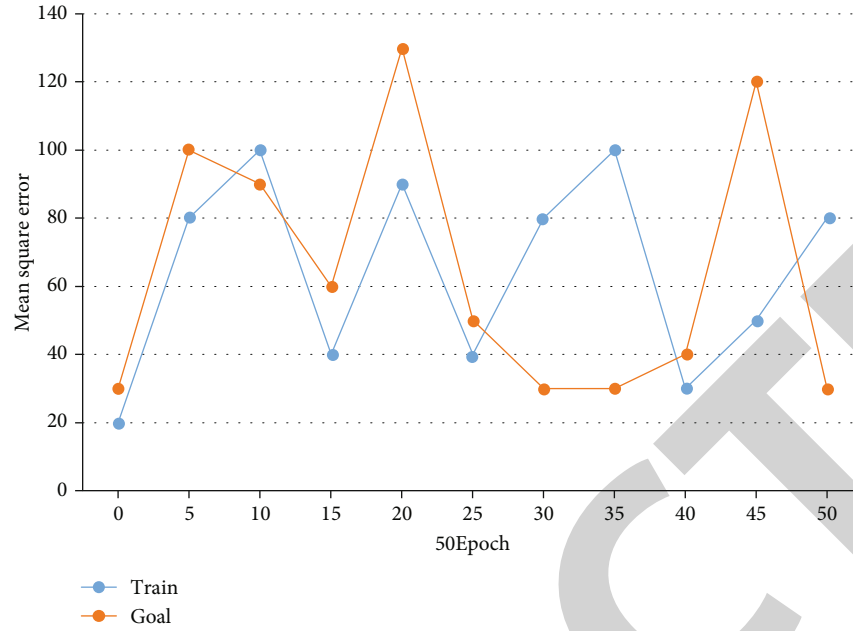


FIGURE 4: Mean square deviation of network structure on [5, 10].

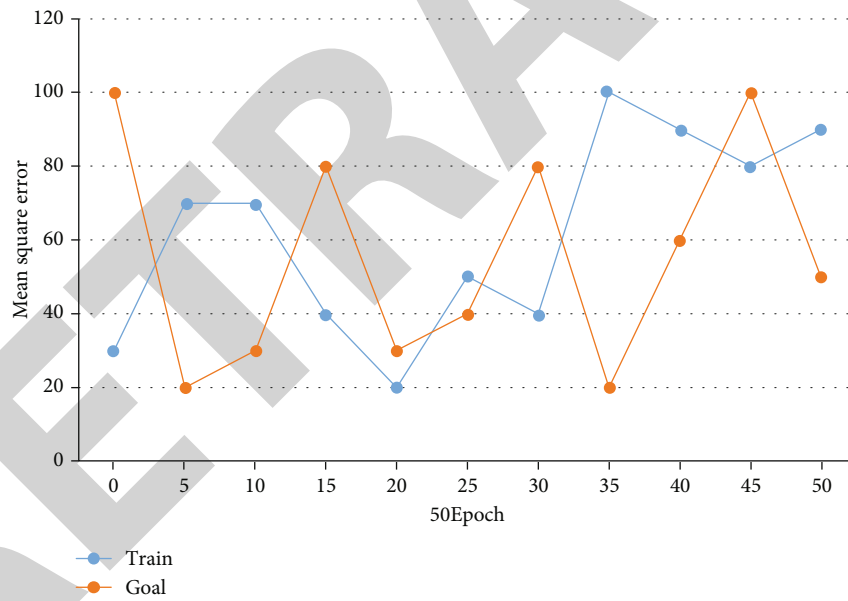


FIGURE 5: Mean square deviation of network structure on [10, 15].

TABLE 1: Comparison of training results of different numbers of hidden nodes.

Hidden node number	6	8	10
Mean-squared error	0.00007765	0.00003561	0.00001462
Training steps	187	129	84

accuracy. Since the input layer does not contain any neuron, but only an input vector, the input layer is not counted in the number of layers of the neural network in this paper. In this network structure, there are interconnections between the same layers, and there are mutual constraints between neurons, but it is still a feed-forward network structure in terms of the relationship between layers, and many self-organizing neural networks mostly have this structure. It indicates that the path is invalid, i.e., e-commerce across boarder synergy with other species has no positive effect on cross-border logistics chain synergy.

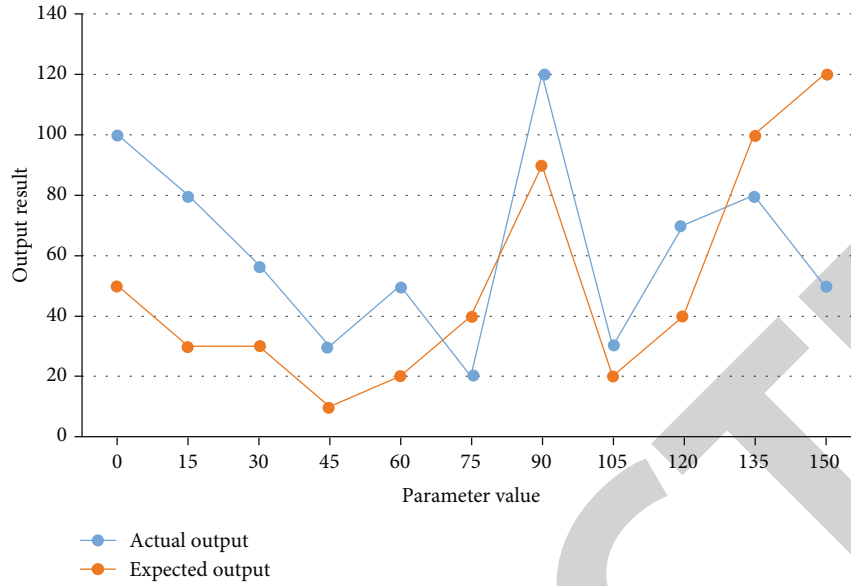


FIGURE 6: Comparison of the fitting results between the actual output of the training sample and the expected output value.

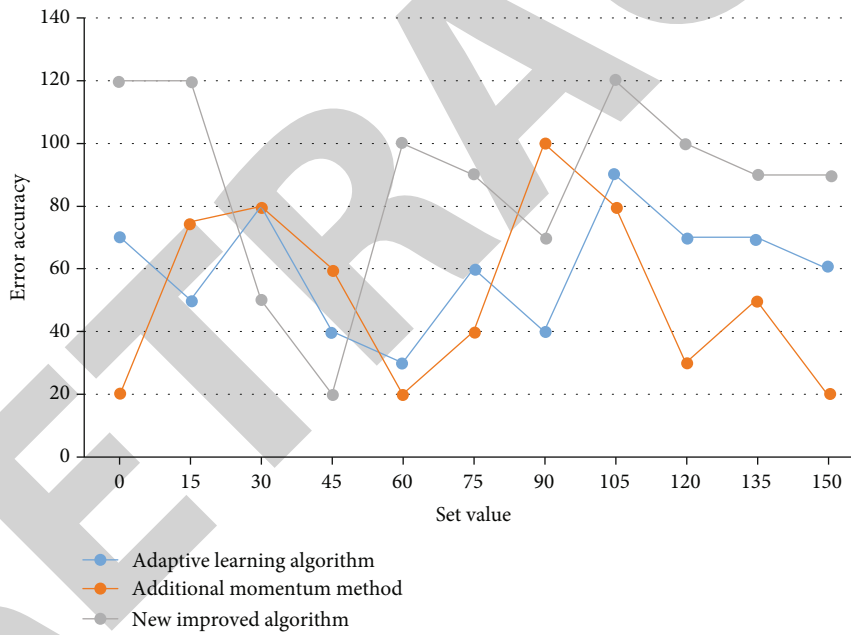


FIGURE 7: Error curves of new improved algorithm, additional momentum method, and adaptive learning rate algorithm.

3.2. Analysis of Improved Neural Network Algorithm. According to the shortcomings of neural network algorithms, there are now a variety of improved algorithms, and these improved algorithms are broadly classified into two main categories: one is heuristic learning methods, and the other is numerical optimization methods. The optimal individual is found mainly by calculating the fitness, and the weights and thresholds carried by the optimal individual are given to the network for training, in order to have faster convergence and higher prediction accuracy of the model. The improved neural network has differentiability and saturated nonlinear properties, which can enhance the nonlinear mapping ability of the network and

TABLE 2: Robustness detection.

Infected input	Expected output	Actual output	Mean square deviation
0.0265	0.0187	1	3.71
0.0261	0.0173	0	1.45
0.0342	0.0165	1	4.62

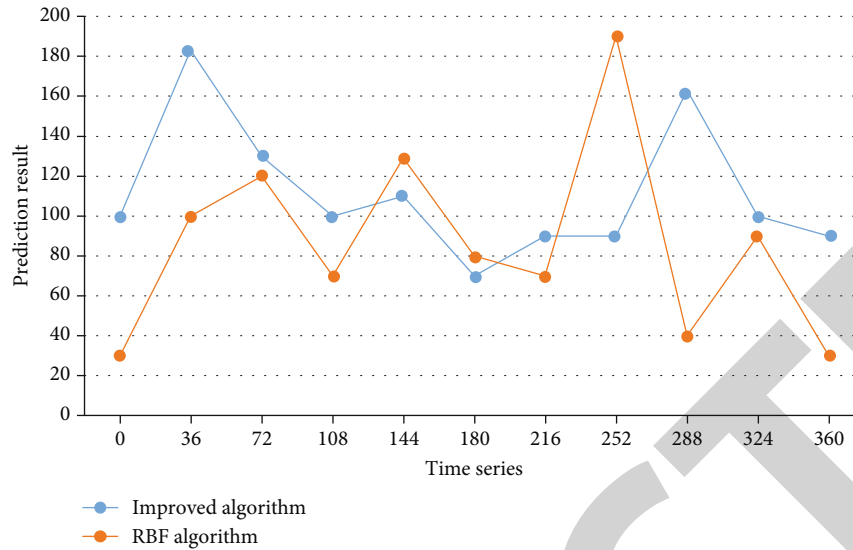


FIGURE 8: Comparison chart of prediction effect.

meet the requirement of continuous differentiability of the activation function in the hidden layer. In the algorithm, the amount of variation of the connection weight coefficients is determined by both the learning rate and the gradient of the error function, however, the learning rate is a homogeneous value in the standard algorithm. After several runs of the algorithm, the error profile plots of the new improved algorithm, the additional momentum method, and the adaptive learning rate algorithm were derived. This is shown in Figure 7.

First, check and judge whether the corrected weights have achieved the effect of reducing the value of the error function. If the value of the error function is indeed reduced; then, it means that the learning rate value chosen by the algorithm is small, and the learning rate can be increased on the basis of the original one with appropriate treatment. The size of the response of the improved neural network to the input depends on the distance between the input vector and the center of the network, and the smaller the distance between the input vector and the center, the larger the response of the neuron will be. So the center correction process of the improved neural network is essentially a process of clustering the input samples based on the distance between them, and the input vectors with small distances from each other are grouped into one class, and the center of the clusters is the network center. In order to retain the target detection accuracy of the overall network as much as possible, it is necessary to reduce the pruning for network layers with higher pruning sensitivity and increase the pruning for network layers with lower sensitivity. Network pruning after pruning, it is the most critical to ensure that the network accuracy remains unchanged. It can be clearly said that after deleting some network connections directly, the network accuracy will certainly decline. Therefore, in order to keep the accuracy of the network unchanged, it is necessary to retrain the pruned network. After repeated retraining, the accuracy of the network will be improved to reach the accuracy of the original network.

In order to test the robustness of the new algorithm, we assume that the input samples are contaminated by random

noise. The robustness test results of the improved neural network algorithm are shown in Table 2.

E-commerce across boarder platforms exist as core species in the e-commerce across boarder ecosystem, and the lack of synergy within themselves will affect the synergy of the e-commerce across boarder ecosystem. In addition, the e-commerce across boarder platform and key species such as suppliers and consumers also have the problem of missing synergy. Therefore, it is sufficient to take the most compact structure as possible under the premise of satisfying the accuracy requirement, i.e., adding up to one neuron to speed up the decrease of error as long as it can solve the problem.

Second, the adjustment of the connection weight coefficients will all contain a portion of the previous adjustment of the connection weight coefficients. The clustering of all the input vectors is performed according to the k -means clustering method to obtain the cluster centers, which are the centers of the improved network; this is followed by a supervised weight determination process, which uses the LMS method to determine the weights of the network based on the actual output values of the system and the network center values obtained in the previous step. On the other hand, as the network deepens, the number of parameters is more enormous for deeper network layers because the number of channels in the feature map increases. To pass the sensitivity vectors from backward to forward, we need to obtain a recursive equation between the sensitivity vectors of different layers. The fitness value of each individual in the population is then used to perform the search, and the fitness value is used to determine the degree of excellence of the individual. If the risk is caused by controllable factors, under the condition that the risk control cost allows, such as the price volatility of pledges is high, the risk warning should be strengthened, and the risk should be controlled by reducing the pledge rate and setting the inventory risk warning in many aspects. We use the test sample to test the trained network and compare the effect of the improved neural network algorithm with the traditional RBF network, as shown in Figure 8.

Retraction

Retracted: Certificateless Batch Authentication Scheme and Intrusion Detection Model Based on the Mobile Edge Computing Technology NDN-IoT Environment

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Sun, "Certificateless Batch Authentication Scheme and Intrusion Detection Model Based on the Mobile Edge Computing Technology NDN-IoT Environment," *Journal of Function Spaces*, vol. 2022, Article ID 5926792, 9 pages, 2022.

Research Article

Certificateless Batch Authentication Scheme and Intrusion Detection Model Based on the Mobile Edge Computing Technology NDN-IoT Environment

Jianzhao Sun 

School of Computer Engineering, Henan Institute of Economics and Trade, Zhengzhou 450018, China

Correspondence should be addressed to Jianzhao Sun; jzsun@henetc.edu.cn

Received 27 May 2022; Revised 5 July 2022; Accepted 8 July 2022; Published 22 July 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Jianzhao 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 rapid development of mobile communication technology, the data transmission rate of mobile communication has been significantly improved and emerging Internet of things applications need a lot of computing resources to meet their own computing needs. However, the existing intelligent terminals have limited computing power and cannot meet the low-latency computing requirements with limited energy consumption. Mobile edge computing technology is considered to be one of the technical solutions that can efficiently solve this problem. Through mobile edge computing technology, intelligent terminals can actively divert some computing tasks to the computing servers deployed in edge network nodes and return the computing results to intelligent terminals after the edge computing servers have finished computing, thus greatly reducing the computing delay of users. In this paper, aiming at multiple edge server networks, combined with dual-connection technology, based on the research of a single edge computing server network, it is further extended to two edge computing servers and a joint optimization problem of computing task allocation and security performance requirements of intelligent terminals and edge computing servers is proposed, so as to minimize the global time delay for completing the computing tasks of intelligent terminals. By layering the problem, this paper proposes a corresponding algorithm, which can efficiently obtain the optimal solution of the initial problem. Compared with the linear search algorithm, the algorithm proposed in this paper can significantly reduce the computation time.

1. Introduction

With the wide application of the computer network and information confidentiality and authentication, network security and protection have attracted more and more attention from all walks of life. A password is a key technology in information security and plays a very important role in information security [1]. The certificateless signcryption scheme combines the advantages of certificateless cryptosystem and signcryption, which can not only complete the functions of signing and encrypting messages in the same logical step at a lower communication cost but also avoid the storage problem of the public key certificate in the public key infrastructure and the key escrow problem in the identity-based

cryptosystem. Huge digital data are stored in computer databases and then transmitted between cumbersome communication networks. Without the protection of security technology, these data will be easily intercepted in the transmission process, which will lead to the leakage of secrets and the risk of data being extracted or copied in storage [2]. Cryptography is the key technology to ensure information security and network security and plays an important role in the field of information security.

Firstly, this paper deeply studies the mobile edge computing service deployment and information interaction mode in the multiterminal multi-intelligent base station environment and designs the system architecture and resource management algorithm based on multibase station cooperation.

The characteristics of all kinds of mobile terminal tasks are abstracted as ordered vectors, and the characteristics and signal transmission rate of each component of the system are quantified. The optimization problem of overall task processing delay is decomposed into two parts: calculating delay and transmission delay, which are modeled separately. The optimization algorithm based on the genetic algorithm is combined with the resource management algorithm based on multibase station cooperation. Further considering the scenario of multiple intelligent terminals and a single edge server, how does the edge server choose intelligent terminals to provide computing services with limited energy resources [3]? Aiming at the abovementioned problems, this paper proposes an accurate and effective algorithm to find the optimal intelligent terminal selection scheme and a large number of simulation results prove the accuracy of the proposed algorithm [4]. By layering the problem, this paper proposes a corresponding algorithm, which can efficiently obtain the optimal solution of the initial problem. Compared with the linear search algorithm, the algorithm proposed in this paper can significantly reduce the computation time.

Although the mobile edge computing technology has all the abovementioned advantages, due to the data transmission between the intelligent terminal and the edge server, it is necessary to jointly optimize the wireless resource allocation and computing resource allocation in order to obtain an effective computing diversion scheme. In view of the scarcity of wireless network resources and the problems of application efficiency, the importance of network resource allocation utility can be improved. Further solve the scenario limitations of existing wireless network resource allocation methods. On this basis, the general efficient wireless network resource allocation architecture adopts the utility-based resource allocation strategy algorithm, which has strong scenario adaptability and high utility. This paper further considers the security problems in edge computing while considering the uplink and downlink transmission and indirectly reflects the security degree of computing diversion data by quantifying the security overflow probability defined in the physical layer security. Through three scenarios, the amount of uploaded data, and energy distribution of intelligent terminals, the overall delay is divided into two parts: edge server processing delay and local computing delay, in which the server processing computing time is equal to the sum of uplink and downlink transmission delay and computing delay. By jointly optimizing the flow scheduling and energy distribution of intelligent terminals, the total delay consumed by intelligent terminals in completing computing tasks is minimized [5]. Although the joint optimization problem in this scenario is a nonconvex optimization problem, by exploring and utilizing the convex optimization characteristics of the energy allocation subproblem, this paper decomposes the joint optimization problem into the bottom-level problem and the top-level problem to be optimized and solved, respectively. Compared with the heuristic algorithm, the accuracy of the algorithm proposed in this paper is verified. Through a large number of data tests and simulations, the effectiveness of the algorithm proposed in this paper is verified.

2. Related Work

At present, the research on mobile edge computing is relatively scarce and mainly focuses on computing migration. However, it has been widely regarded as the development direction of mobile cloud services and has initially laid a foundation for the implementation of subsequent projects. The edge computing technology can not only reduce the computing delay of intelligent terminals and improve the utilization rate of resources but also meet the requirements of large-scale device access in the 5G era, and the business sinks to the edge to relieve the pressure of the core network [6].

Li et al. studied binary computing diversion of a single user in a random wireless channel. A decentralized binary computing distributary game method is proposed [7]. Zhao studied the joint optimization of multiuser binary computing diversion decision and resource optimization [8]. Zhu et al. studied the problem of maximizing the computing rate of binary computing with wireless power supply. This paper analyzes the energy delay dynamic shunting balance of mobile edge computing technology with energy collection devices and proposes an incentive compatible auction mechanism for resource transactions between mobile devices and cloud base stations to stimulate binary computing shunting [9]. Wen-He et al. proposed a strategic computing diversion algorithm based on the deep Q network, which is used to learn the optimal binary computing diversion strategy in a superdense network [10]. Chu and Leng used the dynamic voltage scale to study partial calculation shunt. An energy-efficient resource allocation scheme is proposed based on the edge calculation part to calculate the diversion. Considering the application with limited resources, the energy-saving part of wireless resources and computing resources is jointly optimized to calculate the shunting method, through wireless energy transmission [11]. Wang et al. proposed a joint optimization scheme of the edge computing system based on full-duplex relay wireless power supply and proposed a joint optimization method of energy consumption and delay based on fog computing [12]. Yang et al. put forward a strategy called PRIMAL, which selectively migrates computing tasks to their best location, considers the benefits and costs of migrating users' tasks to the appropriate mobile edge cloud according to their location, and optimizes the tradeoff between migration benefits and migration costs. However, this document designs a real-time model to calculate each computing task separately, which will bring a lot of overall task processing delay [13]. Li et al. introduced mobile edge computing into blockchain, taking edge computing as the network startup factor of mobile blockchain. Literature has exhausted the resource management and pricing of mobile edge computing to support the application of mobile blockchain. Among them, the mining process of miners can be migrated to edge computing service providers [14]. Alferaidi et al. have optimized the algorithm overhead. In order to reduce the transmission time of the program status on the system network, the mobile edge cloud is designed as a tree hierarchy of geographic distribution servers. Firstly, the scenario that each

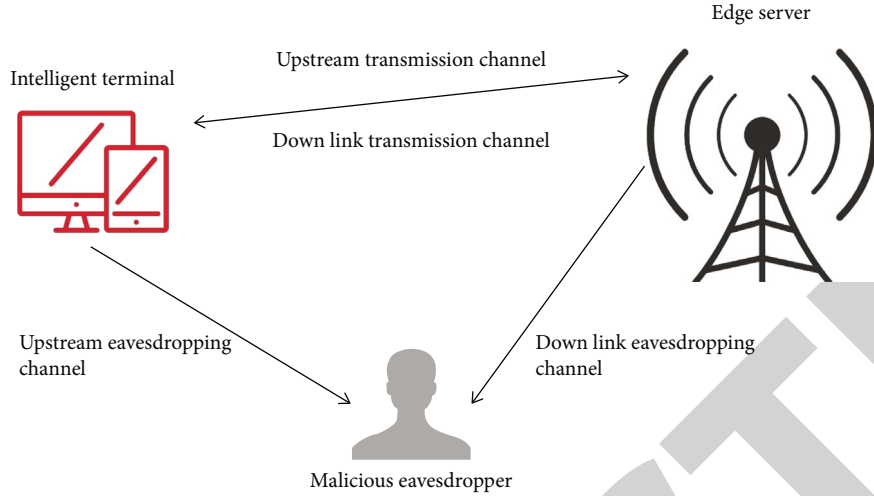


FIGURE 1: Downlink transmission system model.

layer of the mobile edge cloud has only one server is considered, and then, the workload placement decisions of different mobile edge cloud servers are summarized together. At the same time, a workload allocation algorithm is proposed, which determines which mobile edge cloud service the mobile program is placed on and how much computing power is provided for it. However, despite the optimization of the algorithm overhead, the running process of these algorithms still needs to take up a certain amount of computing time for mobile applications. If large-scale user migration or large-scale computing tasks occur, the overhead of real-time algorithms will be one of the main obstacles to the development of delay-sensitive mobile applications [15]. Koo and Lim proposed an implementation scheme of block flow application, which can remotely retrieve the application from the server as required and run it locally on the Internet of things device. In particular, after investigation, it is found that the design of computing diversion based on edge computing is closely related to mobile data diversion based on the wireless network [16]. Song et al. put forward a method to find the best diversion strategy based on learning by using the energy collection system to supply power to data. The new paradigm based on Hong Jizhan double-join technology is applied to efficient data streaming [17]. Wang and Xu provide a scheme for data distribution in the Internet of vehicles to use edge computing servers for effective storage [18].

However, none of the abovementioned studies have considered the problems existing in wireless transmission. Based on this problem, this paper decomposes the joint optimization problem into the bottom-level problem and the top-level problem and optimizes them. Compared with the heuristic algorithm, the accuracy of the algorithm proposed in this paper is verified.

3. Multireceiver Signcryption without a Certificate

3.1. Bilinear Pair. A bilinear pair is also called bilinear mapping, and its predecessor is the Weil pair and Tate pair on

the algebraic curve. They are mainly used in cryptography to attack the elliptic curve or hyperelliptic curve cryptosystems. It is mainly used to provide a higher access rate for users. Cooperative networking of base stations can not only increase the coverage of base stations but also facilitate the concentration of base station resources, thus serving more users. Double-connection technology mainly involves two different base stations covered by two different cells, macro cell and smaller cell [19]. The downlink transmission system model is shown in Figure 1.

The intelligent terminal uploads some computing tasks to the edge computing server for computing, and during the transmission, it is eavesdropped by a potential malicious eavesdropper. Dual connection (DC) means that the mobile phone can simultaneously use the wireless resources of at least two different base stations (divided into the master station and slave station) in the connected state. The dual connection introduces the concept of “shunting bearer,” that is, the data is shunted to two base stations at the PDCP layer. The PDCP layer of the master station user interface is responsible for PDU numbering, data shunting, and aggregation between master and slave stations. Through the dual-connection technology, the intelligent terminal can transmit data with Hong Jizhan and a small base station at the same time, so it is particularly important to divide the data between Hong Jizhan and a small base station reasonably according to its own location, channel condition, and security performance requirements [20].

According to the principle of physical layer security, the security throughput that the terminal can upload to the edge server is shown in formula (1).

$$C_{iB}^{\text{sec}} = \left(W_b \log_2 \left(2 \frac{p_{ib} g_{ib}}{n_b} \right) - W_B \log_2 \left(2 \frac{p_{ib} g_{iE}}{n_E} \right) \right). \quad (1)$$

W_B is the bandwidth of the edge server, and p_{ib} represents the channel gain from the intelligent terminal to the edge server and eavesdropper. Because of the randomness

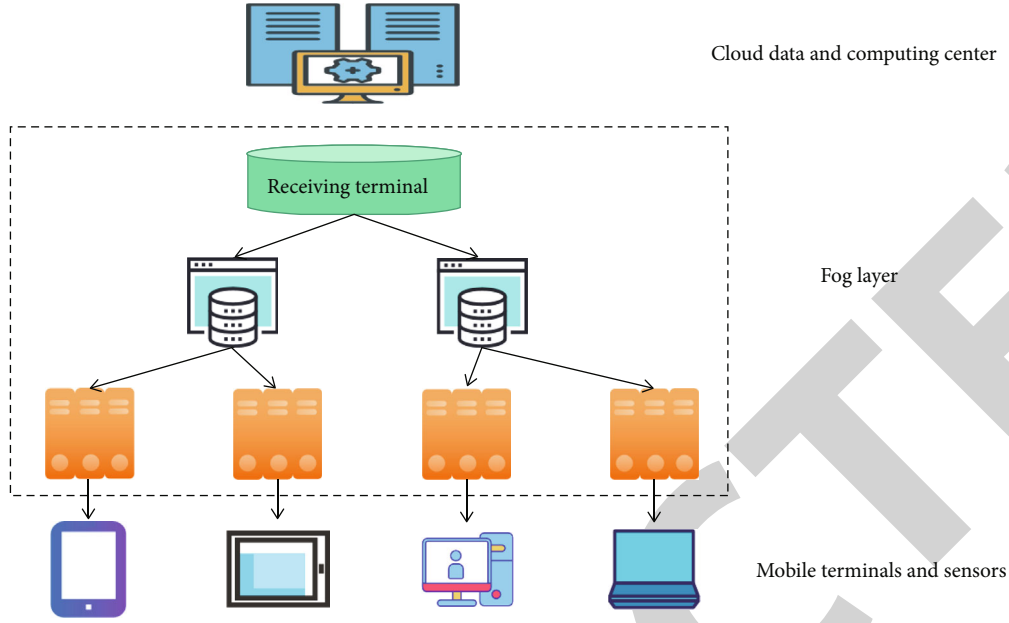


FIGURE 2: Schematic diagram of the fog computing structure.

and uncertainty of the eavesdropper's location, we cannot get the exact value of its channel gain. Assuming that the channel gain from the intelligent terminal to the eavesdropper obeys the exponential distribution with a constant mean value and given a fixed shunting rate from the intelligent terminal to the eavesdropper, the throughput of uplink transmission can be obtained as shown in formula (2).

$$x_{iB} = W_B \log_2 \left(1 - \frac{P_{ib}}{g_{ib}} (1 + e^{g_{ib}/g_{ie}}) \right),$$

$$P_{\text{out}} = pr \left\{ (x_{iB} | W_B \log_2 \left(1 - \frac{P_{ib}}{g_{ib}} \right) + W_B \log_2 \left(1 + \frac{P_{ib}}{g_{iE}} \right)) \right\}. \quad (2)$$

There are two key parameters in the abovementioned formula, one is the security overflow probability of intelligent terminal, and the other is the strength of the eavesdropping channel. With the increase of throughput, the overflow probability decreases, which means that it is a relatively relaxed security protection environment. The security overflow probability of intelligent terminals increases with the decrease of overflow probability, which means that the ability to withstand eavesdropping is stronger at this time [21]. The time taken to upload the computing task to the edge server in this paper represents the uploading speed of the intelligent terminal, as shown in formula (3).

$$P_{iB} = \frac{x_{iB}(1 + c_{ib})t_i}{g_{ib} - 2W_B(1 - c_{ib})}. \quad (3)$$

After calculating the task distributed by the intelligent terminal, the edge server returns the calculation result to the intelligent terminal. In this model, it is assumed that

the intelligent terminal will compress the data of the calculation task. In this model, the downstream security throughput can be expressed as shown in formula (4).

$$C_{Bi}^{\text{sec}} = \left(W_B \log_2 \left(1 + \frac{P_{ib}}{g_{ib}} \right) \right). \quad (4)$$

We set the safe overflow probability as shown in formula (5).

$$x_{Bi} = W_B \log_2 \left(1 - (1 - e^{g_{ib}/\alpha_{\text{sec}}})(1 + \alpha_{\text{sec}}) \right). \quad (5)$$

Next, this paper models the global delay of the intelligent terminal. The overall time includes two parts: calculation delay and transmission delay. The downlink transmission power of the edge server is shown in formula (6).

$$P_{Bi} = \frac{n_E 2^{s_i n^E / W_B}}{G_{B_i} - n_E 2^{s_i n^E / W_B}}. \quad (6)$$

3.2. Fog Computing and Cloud Computing. Mobile edge computing allows the data needed by mobile terminal computing tasks to be processed at the edge of the mobile network without further migration to mobile cloud. The components of the mobile edge network mainly include the following: mobile terminal equipment, mobile edge equipment, mobile edge computing server, and mobile edge computing intelligent base station. These components need to have corresponding computing power and storage capacity to support mobile edge computing. In a word, mobile edge computing does not need the active assistance service of the cloud and it focuses more energy on the edge of mobile cloud. Mobile cloud integrates all the advantages of mobile computing, cloud computing, and mobile Internet, while mobile cloud can provide

corresponding resources for mobile terminals according to their needs. In the MCC infrastructure, the centralized cloud server cluster is far away from the mobile terminal equipment, so the computing efficiency of MCC is low in the environment with high computing pressure. Fog computing can support most common connection devices on the market [22]. The structure diagram of fog computing is shown in Figure 2.

Fog computing is characterized by its distributed architecture, in which data are collected and processed by distributed fog computing devices from different sensors. Compared with cloud computing far away from mobile terminal users, fog computing greatly reduces system latency. However, FOG computing has certain limitations because of its dependence on wireless connection. In order to ensure the execution of complex operations, wireless connection has high requirements for real time. However, in the mobile edge computing environment, the intelligent computing, communication capability, and data processing capability are node based, rather than the hierarchical network of fog computing, so the development trend of mobile edge computing in 4G and future 5G networks is on the rise. The radio resources in MEC resource management mainly include power and subchannels. The power refers to the uplink transmission power of the terminal equipment when the user unloads the task data. The uplink transmission energy consumption is calculated as the product of uplink transmission power and uplink transmission delay, and the uplink transmission delay is also affected by uplink transmission power, so the uplink transmission power affects the uplink transmission energy consumption. Subchannels are obtained by evenly dividing the system bandwidth, which can be understood as different frequency bands. The bandwidth of the subchannel affects the uplink transmission rate and then affects the uplink transmission delay and the uplink transmission energy consumption. In addition, when considering the MEC system with multiple base stations, intercell interference should be considered, that is, users occupying the same subchannel in different cells will interfere with each other. Interference will reduce the uplink transmission rate and increase the transmission delay and energy consumption. Therefore, it is necessary to allocate channels reasonably for users and reduce intercell interference.

3.3. Security Model. The mobile edge computing intelligent base station proposed in this paper has both computing function and storage function, so the intelligent base station needs to schedule computing tasks and data caching tasks according to the current situation of each mobile terminal and its own service. If the control unit receives a computing task request, the function of the control unit is to determine the execution location of the computing task. If the control unit receives the data request required by the task, the function of the control unit is to determine whether the data needs to be cached and where it is cached. In the certificateless cryptosystem, adversaries are divided into type 1 and type 2. According to the types of adversaries, confidentiality and unforgeability games are divided into two types. Therefore, the function of the computing unit is to calculate the computing tasks requested by some mobile terminals cov-

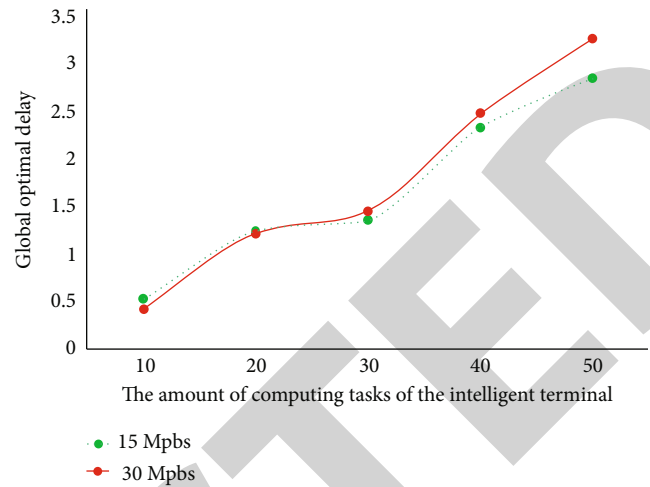


FIGURE 3: Ratio of global optimal delay.

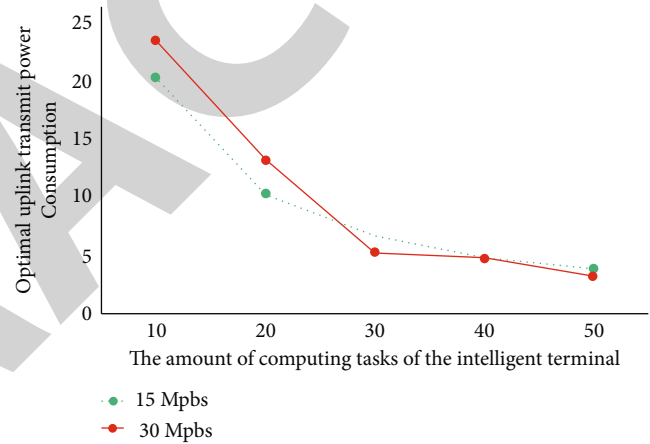


FIGURE 4: Optimal uplink transmission power consumption.

ered by the mobile edge computing intelligent base station according to each collaborative resource management algorithm, so as to reduce the pressure of computing load in the mobile cloud. In this paper, the global enumeration method is used to solve the original problem type 1. The computing speed of the intelligent terminal is 120 Mbps, the upper limit of energy consumption is 4J, the compression rate is 0.25, and the computing power consumption and the computing power consumption of the server are both 0.1.

3.3.1. Simultaneous Comparison. When the local computing speed of the server is different (15 Mbps and 30 Mbps), the minimum transmission delay solved by the global enumeration algorithm is shown in Figures 3 and 4.

The comparison is about the optimal uplink transmission power and the change under different calculation tasks. All the abovementioned results show the consistency between our proposed algorithm and enumeration method, which also shows the accuracy of this algorithm from the side. In order to model the problem, this paper considers

TABLE 1: Comparison of this scheme with other schemes.

Program	The time of the signcryption process (ms)	The time of the decryption process (ms)
A bilinear pairing operation	$29.63n + 99.425$	232.428
A dot multiplication operation on intrusion detection	$15.62n + 99.384$	$83.86n + 180.362$
Addition operation on an intrusion detection	$23.66n + 88.241$	135.328
An exponentiation on intrusion detection	$26.25n + 76.634$	134.248
Scheme of this paper	$4.26n + 61.265$	118.823

introducing a binary variable to indicate whether the server selects the intelligent terminal for access. In this paper, the problem of maximizing edge server revenue represents the total transmission delay of all intelligent terminals, including uplink transmission delay and downlink transmission delay. The purpose of this problem is to solve the problem of maximizing revenue by reasonably selecting users' access under the limitation of time slot and energy consumption. The efficiency of this scheme is compared with other proposed certificateless multireceiver signcryption schemes, and the comparison results are shown in Table 1.

In the actual network transmission, the communication between devices is very complicated and different intelligent terminals have different security requirements and different energy consumption restrictions, which lead to different data splitting strategies for computing tasks. Because of the coexistence of edge computing technology and dual-connection technology, the resource allocation between the intelligent terminal and a single edge server becomes the resource allocation among multiple computing members, which further increases the difficulty coefficient of the already complicated problems, so it is even more necessary for this paper to allocate computing resources reasonably.

4. Simulation and Result Analysis

4.1. Unloading Decision Subproblem Analysis. Before the closed expression of the power optimization strategy and computing resource allocation is obtained, this paper adopts the particle swarm optimization algorithm to optimize by iterative search and defines fitness function to evaluate the solution. The algorithm starts iterative search from an initialized population, which contains several particles, and each particle has three attributes, namely position, speed, and fitness value. The position coordinate of each particle is a candidate solution of the problem to be solved, while the velocity of the particle controls the update of the particle position, and the fitness value is used to evaluate the solution. After many iterations, the optimal solution corresponding to the optimal position of all particles is found. The setting of inertia weight plays a key role in the performance of the algorithm. Each particle converges according to its updated speed. When the inertia weight is set larger, the algorithm has better global search ability but the convergence speed will slow down. If the inertia weight is set smaller, the algorithm will have stronger local search ability but this will also lead to the search easily falling into local

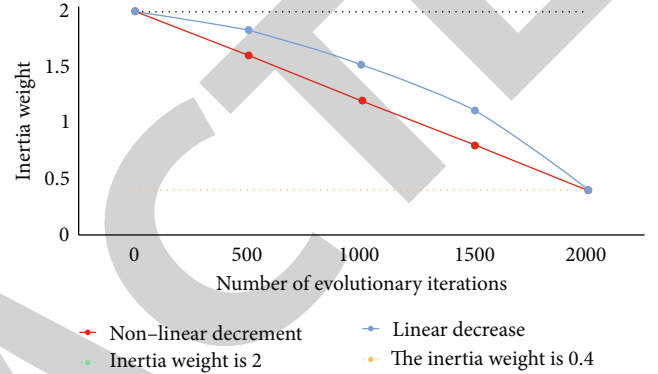


FIGURE 5: Scheme changes of different inertia weights.

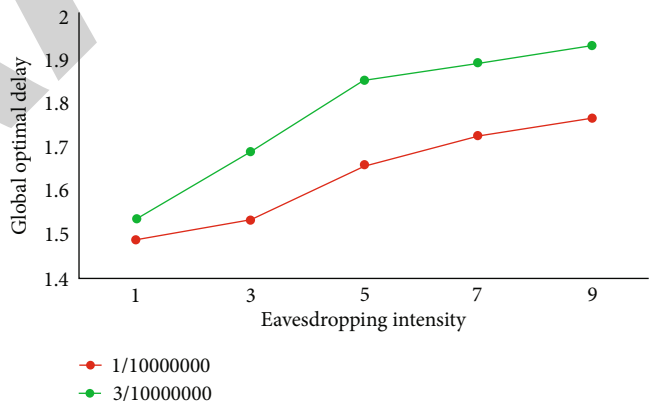


FIGURE 6: Influence of eavesdropping intensity on global optimal transmission delay.

optimum. The inertia weight of the particle swarm optimization algorithm is closely related to the size of population, spatial dimension, and the decline rate of inertia weight. Through the experimental study of several representative functions, the results show that properly changing the inertia weight can quickly converge, improve the search efficiency, and avoid falling into local optimization. In order to coordinate the ability of global search and local search, this paper uses the nonlinear decreasing adaptive inertia weight scheme and its formula is shown in (7).

$$w(t) = \frac{w_{\text{start}}}{t} - (w_{\text{end}} - w_{\text{start}}). \quad (7)$$

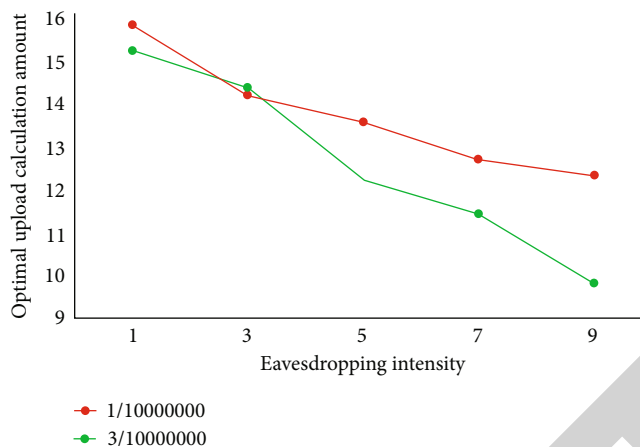


FIGURE 7: Influence of the change of eavesdropping intensity on the optimal upload calculation.

t represents the current evolution algebra, and w_{start} and w_{end} represent the maximum and minimum inertia weight, respectively.

Inertia weight decreases with the increase of evolution points and decreases slowly in the early stage of iteration and rapidly in the late stage of iteration. This makes the algorithm have a large inertia weight in the early stage of iteration to enhance the global search ability, while in the late stage of iteration, the inertia weight is small and the local search ability is enhanced, which improves the search accuracy and accelerates the convergence speed. As shown in Figure 5.

It can be seen in Figure 5 that the unloading decision variables are binary variables from 0 to 1, so this paper uses the binary particle swarm optimization algorithm to solve the problem, when the global optimal delay decreases and the calculation amount of optimal upload increases. The abovementioned results are reasonable. When the intelligent terminal's own security performance is relatively loose, the intelligent terminal can upload more workload to the server by taking advantage of computing split, and then, the overall optimal delay will decrease.

4.2. Optimal Diversion Scheme for Multiuser Scenarios. Edge computing is a computing facility that deploys data resources outside the data center and close to users. Through this facility, a series of network devices link edge computing devices to users or processes, such as the Internet of things. Therefore, the deployment of edge computing devices does not have the physical security of the data center and cannot adopt the access, network, and data security measures applied by the software or hardware residing therein. The security challenge of edge computing is to provide the additional security required to make the security of edge computing facilities meet the security and compliance of data center standards. In many cases, this means that it is necessary to securely access edge computing devices. Whether it is physical access or through the user interface, some key security measures must be taken. For each intelligent terminal, the binary particle swarm optimization algorithm can be used to solve its optimal security computing diversion strat-

egy. In this section, from the point of view of the edge computing server, this paper mainly studies how the edge computing server selects the access of an intelligent terminal under limited computing resources and a certain time limit and makes use of the edge server to maximize its own revenue under the condition of limited channel time slots and limited energy consumption for transmission and computation.

In order to maximize the revenue, this section will study the influence of eavesdropping intensity on the optimal diversion scheme, setting the eavesdropping intensity at 0.2, and changing it within the interval $[10^{-7}, 9 \times 10^{-7}]$ to verify the minimum transmission delay and the influence of the optimal diversion scheme. By using the method of comparative analysis, two datasets are set as 10^{-7} and 3×10^{-7} . The comparison results are shown in Figure 6.

It can be observed in Figures 6 and 7 that when the eavesdropper's eavesdropping intensity on the user increases, the global optimal transmission delay increases, while the optimal upload workload decreases. This result is reasonable, because as the channel strength of the eavesdropper to the user increases, it means that the eavesdropper's eavesdropping ability is increasing. At this time, the environment in which the intelligent terminal is located is unsafe, so the calculation workload of uploading should be reduced and the risk of being eavesdropped should be lowered. Therefore, the optimal uploading workload is decreasing, which leads to the increase of the delay of the intelligent terminal.

Through qualitative analysis, this paper proposes a joint intelligent terminal computing task splitting, wireless resource allocation (including time delay and energy consumption), and security overflow probability to minimize the overall time delay of the system. The main goal of this paper is to minimize the overall time delay (including transmission time delay and calculation time delay) under the limited energy consumption budget while completing the computing task of the intelligent terminal and meeting the security requirements of the intelligent terminal. Based on the consideration of eavesdropping, this paper also puts forward the concept of overflow probability, that is, the transmission rate of the intelligent terminal when actually

distributing data is less than the probability that the intelligent terminal allocates the data transmission rate to the edge server. Finally, by comparing various situations of each layer, we can find the global optimal solution and finally solve the initial optimization problem. Finally, the accuracy of this algorithm is verified by comparing with other heuristic algorithms.

5. Conclusions

Traditional intrusion detection systems have some problems in the detection algorithm and decision-making control mechanism. In the field of game theory, there are already a set of mature theories and methods to solve the problem of competition and mutual influence among people with different goals in the same system. These methods have been successfully applied in many fields such as political economy, decision theory, and control theory.

In order to improve the accuracy of the intrusion detection system, it is necessary to analyze the attacks on network layer routing, data link layer, transport layer, and application layer and the layers need to cooperate with each other. Based on the mobile edge computing technology, this paper puts forward the security computing diversion strategy in two scenarios and obtains a series of research results. A multireceiver cryptosystem can send the same ciphertext to multiple receivers, which is more efficient than encrypting the same message and then sending the ciphertext to the corresponding receivers. Firstly, this paper combines the advantages of a multireceiver cryptosystem and signcryption cryptosystem and constructs a certificateless multireceiver signcryption scheme, which realizes that each independent receiver in the designated multireceiver can recover the message.

This paper still has some limitations. Because of the complexity of network information, it is difficult to determine the parameters in the revenue function. How to combine the fuzzy game theory with fuzzy control and multiobjective evaluation methods to improve the accuracy of the intrusion detection decision control model is also the next task to improve the intrusion detection system based on the game theory.

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 they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Science and Technology Project of Henan Province: innovative application of big data in the field of industrial digital upgrading (no. 222102210338).

References

- [1] G. Thumbur, N. B. Gayathri, P. V. Reddy, M. Z. U. Rahman, and A. Lay-Ekuakille, "Efficient pairing-free identity-based ADS-B authentication scheme with batch verification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 5, pp. 2473–2486, 2019.
- [2] S. Shamsad, M. Rana, K. Mahmood, M. K. Khan, and M. S. Obaidat, "On the security of a secure anonymous identity-based scheme in new authentication architecture for mobile edge computing," *Wireless Personal Communications*, vol. 124, no. 1, pp. 283–292, 2022.
- [3] W. Liu, J. Song, H. Wu et al., "Non-crypto authentication for smart grid based on edge computing," *Journal of Physics: Conference Series*, vol. 1646, no. 1, article 012060, 2020.
- [4] F. J. Mora-Gimeno, H. Mora-Mora, D. Marcos-Jorquera, and B. Volckaert, "A secure multi-tier mobile edge computing model for data processing offloading based on degree of trust," *Sensors (Basel, Switzerland)*, vol. 1, no. 4, pp. 58–66, 2018.
- [5] C. C. Lee, "ES-HAS: ECC-based secure handover authentication scheme for roaming mobile user in global mobility networks," *Cryptography*, vol. 5, no. 4, p. 35, 2021.
- [6] H. Huang, Y. Wu, F. Xiao, and R. Malekian, "An efficient signature scheme based on mobile edge computing in the NDN-IoT environment," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 1, pp. 1108–1120, 2021.
- [7] D. Li, Z. Cai, L. Deng, X. Yao, and H. H. Wang, "Information security model of block chain based on intrusion sensing in the IoT environment," *Cluster Computing*, vol. 22, no. 1, pp. 451–468, 2019.
- [8] J. Zhao, "Research on intrusion detection system based on joint mobile agent technology in cloud computing environment," *Automation & Instrumentation*, vol. 8, no. 2, pp. 46–55, 2017.
- [9] X. Zhu, H. Xu, Z. Zhao et al., "An environmental intrusion detection technology based on WiFi," *Wireless Personal Communications*, vol. 119, no. 2, pp. 1425–1436, 2021.
- [10] M. Zhong, Y. Zhou, and G. Chen, "Sequential model based intrusion detection system for IoT servers using deep learning methods," *Sensors*, vol. 21, no. 4, article 1113, 2021.
- [11] X. Chu and Z. Leng, "Multiuser computing offload algorithm based on mobile edge computing in the Internet of things environment," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 6107893, 9 pages, 2022.
- [12] X. Wang, X. Wang, and Y. Li, "NDN-based IoT with edge computing," *Future Generation Computer Systems*, vol. 115, no. 2, pp. 397–405, 2021.
- [13] Y. C. Yang, F. Ali, and S. Nazir, "Selection of devices based on multicriteria for mobile data in Internet of things environment," *Mobile Information Systems*, vol. 2021, Article ID 2117915, 7 pages, 2021.
- [14] J. Li, X. Li, G. Li, and R. Zhang, "Non-cooperative game forwarding leveraging user trustworthiness in mobile edge networks," *Sustainability*, vol. 14, no. 8, p. 4473, 2022.
- [15] A. Alferaidi, K. Yadav, Y. Alharbi et al., "Distributed deep CNN-LSTM model for intrusion detection method in IoT-based vehicles," *Mathematical Problems in Engineering*, vol. 2022, 8 pages, 2022.
- [16] S. Koo and Y. Lim, "A multi-objective computation offloading algorithm for dependent tasks based on a mobile edge

Retraction

Retracted: Intelligent City 3D Modeling Model Based on Multisource Data Point Cloud Algorithm

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Wu and Z. Zhou, "Intelligent City 3D Modeling Model Based on Multisource Data Point Cloud Algorithm," *Journal of Function Spaces*, vol. 2022, Article ID 6135829, 10 pages, 2022.

Research Article

Intelligent City 3D Modeling Model Based on Multisource Data Point Cloud Algorithm

Youping Wu  and Zhihui Zhou 

Department of Applied Engineering, Gandong College, Fuzhou, Jiangxi 344000, China

Correspondence should be addressed to Youping Wu; 199460011@ecut.edu.cn

Received 25 May 2022; Revised 29 June 2022; Accepted 4 July 2022; Published 21 July 2022

Academic Editor: Miaocho Chen

Copyright © 2022 Youping Wu and Zhihui Zhou. 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 smart cities, intelligent navigation, and autonomous driving, how to quickly obtain 3D spatial information of urban buildings and build a high-precision 3D fine model has become a key problem to be solved. As the two-dimensional mapping results have constrained various needs in people's social life, coupled with the concept of digital city and advocacy, making three-dimensional, virtualization and actualization become the common pursuit of people's goals. However, the original point cloud obtained is always incomplete due to reasons such as occlusion during acquisition and data density decreasing with distance, resulting in extracted boundaries that are often incomplete as well. In this paper, based on the study of current mainstream 3D model data organization methods, geographic grids and map service specifications, and other related technologies, an intelligent urban 3D modeling model based on multisource data point cloud algorithm is designed for the two problems of unified organization and expression of urban multisource 3D model data. A point cloud preprocessing process is also designed: point cloud noise reduction and downsampling to ensure the original point cloud geometry structure remain unchanged, while improving the point cloud quality and reducing the number of point clouds. By outputting to a common 3D format, the 3D model constructed in this paper can be applied to many fields such as urban planning and design, architectural landscape design, urban management, emergency disaster relief, environmental protection, and virtual tourism.

1. Introduction

Cities are the most information-intensive, most frequently changing, and most important area on the earth's surface, and urban 3D information plays an increasingly important role in urban planning, urban change monitoring, public safety, and other fields [1]. The application of two-dimensional data as the main body can no longer meet the needs of various professional applications in cities, and a more intuitive, WYSIWYG 3D spatial data is gradually becoming a new and enthusiastic data expression for customers, which will become the core data of digital cities and even digital earth [2]. The existing fine-grained 3D city modeling techniques mainly include two categories of manual interactive modeling based on multiple data sources and semiautomated modeling based on images or laser-scanned dense point clouds [3]. Thus, urban 3D modeling is highly valued as an important means to obtain or produce 3D spatial data [4]. It is an essential and important part

of the construction of the digital city infrastructure framework [5].

As an important part of smart cities, the fine 3D models of buildings play an important role in urban planning and emergency command [6]. The common method of 3D information acquisition in the past was to use measuring equipment to collect data of points and surfaces of target objects, but the disadvantages of this method are that the collection of 3D data is time-consuming and particularly inefficient, the operation is too cumbersome, and the collected data are easily affected by the complexity of the surface [7]. Traditional surveying and mapping techniques mainly acquire building surface information by means of single-point measurements through measuring instruments such as latitude and longitude instruments and total stations, which have long data collection cycles and high costs, and the acquired 3D coordinate points are not dense enough, making it difficult to meet the needs of rapid construction of smart cities [8]. Intelligent city 3D modeling

model refers to the establishment of a city information model, which collects, organizes, and summarizes information from every corner of the earth and establishes a complete information model according to the geographical coordinates of the earth and connects them with a network, so that everyone on the earth can quickly, completely, and graphically understand the macro- and microconditions of the earth and give full play to the role of these data.

Various types of 3D model data (tilt photography models, point clouds, fine models, etc.) with different accuracy represent different levels of detail in the expression of the real world and are applied to urban 3D scenes at different spatial scales [9]. In order to meet the current growing public demand for geospatial location services, multisource data point cloud algorithms are gradually developed from dimensional to dimensional and even the current popular dimensional direction with the support of computer science and technology and virtual reality technology [10]. For users, the representation of the objective world in 3D city models can convey a more realistic and direct feeling, enabling them to make more accurate analysis and decisions. The description of building complexes is becoming more and more detailed, modeling building types from simple rectangles to universally meaningful spires, herringbones, and flat-topped polygons to complex and unique building models, thus allowing for a more detailed depiction of building structures and gradually increasing the measurable accuracy of the model. Laser point cloud data is collected by laser scanning equipment, which uses the principle of laser ranging to obtain the relative coordinate information of a point on the surface of an object. Since the laser scanner has only a limited scanning angle and the mutual occlusion of different parts of the scanned object, in practical application, we must scan the object from different angles and then transform the scanned multiple point clouds to the same coordinate system and merge them into a complete point cloud model. Along with the continuous development of geospatial science and technology, the promotion of intelligent urban 3D model based on multisource data point cloud algorithm will lead the mapping field to a more modern direction, so that it can better serve the development of national economic construction, which will surely bring rich economic and social benefits in the near future.

The innovation points of this paper are as follows.

- (1) Compared with the traditional modeling means, based on the multisource data-based point cloud algorithm proposed in this paper, it can significantly reduce the labor cost and improve the modeling efficiency, thus providing technical support for the corresponding engineering construction
- (2) Point cloud data, due to its fast acquisition speed and high accuracy with the advantage of real relative position, can be applied to intelligent city 3D modeling to present the spatio-temporal transformation of the earth and its related social activities as well as the environment virtually in the form of data through computer technology, so that it can serve human social activities
- (3) The research results will promote the construction of digital city 3D spatial data infrastructure, so that it can be more widely used in many fields such as urban planning and design, architectural landscape design, urban management, emergency disaster relief, environmental protection, and virtual tourism

2. Three-Dimensional Modeling of Intelligent City Based on Multisource Data Point Cloud Algorithm

2.1. Regular Shape True Three-Dimensional Model Method. Intelligent city 3D modeling based on multisource data point cloud algorithm is to use 3D data to model the real 3D objects or scenes in the computer and finally realize the simulation of real 3D objects or scenes on the computer. The digital surface model of the city generated by the point cloud is superimposed on the orthophoto image, and combined with the field survey information, the height of the building is measured from different angles several times to obtain the arithmetic average worth to the real height information of the building and its top obvious shape as marked by the blue rectangle box in the figure, to carry out accurate modeling. The alignment process of multisource data point cloud data is shown in Figure 1 below.

True 3D modeling with regular shape mainly refers to texture acquisition for the purpose of this modeling, and the texture information is used as first-hand information for texture recovery of the model or becomes texture mapping.

Firstly, for true 3D modeling of features with regular shapes, the feature point cloud data is usually imported into the relevant CAD software. Defining the sketch plane is to determine the reference plane for establishing a 2D sketch, and the sketch plane is usually one of the faces of the object. In order to establish the mapping relationship between the point cloud data and the image data, the conversion parameters between the scanner coordinate system and the camera coordinate system are determined first. In addition to having fixed x - y coordinates, elements on a 2D grid can actually have infinite expansion in the vertical direction, so a 2D grid can express 3D information. A grid cell represents a 3D voxel with infinite expansion in the vertical direction, while a grid point represents a vertical line segment passing through the point, and the length of the vertical line segment is the value of the z coordinate of the point. In order to compare the difference between MFPPH and each point FPPH, the feature points are selected by calculating a distance metric. The most commonly used distance metrics are Manhattan and Euclidean, etc., which are formulated as follows:

$$\begin{aligned} \text{Manhattan} : dm &= \sum_{i=1}^f |p_i - u_i|, \\ \text{Euclidean} : de &= \sqrt{\sum_{i=1}^f |p_i - u_i|^2}. \end{aligned} \quad (1)$$

dm, de : And Manhattan Euclidean distance measure

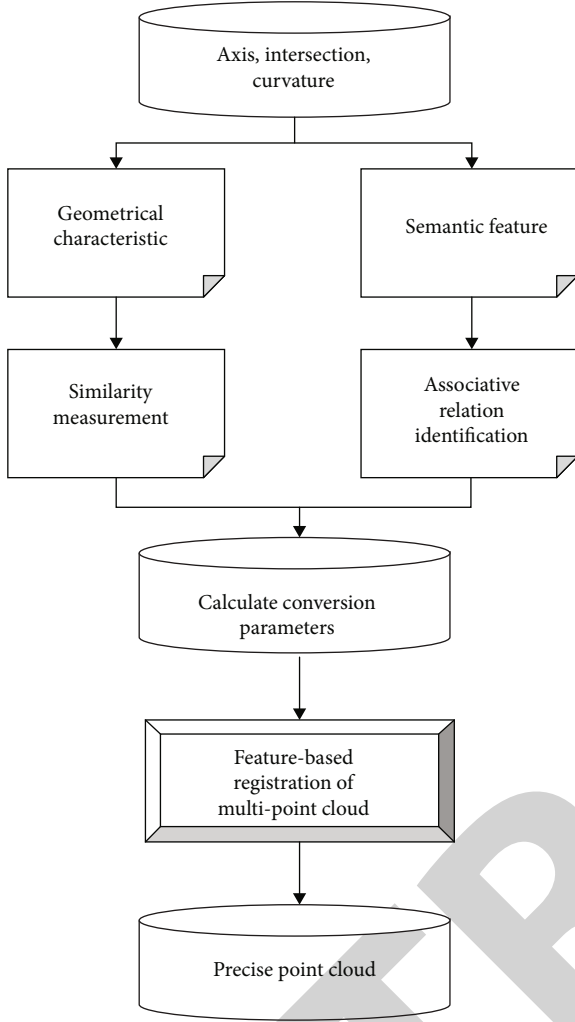


FIGURE 1: Multisource data point cloud data registration process.

p_i, u_i :FPFH and MFPFH values for the i th subinterval
 f :Number of neutron intervals in histogram.

By defining and storing some basic geometric forms in advance in the computer running the modeling system, the set of basic voxels or deformation operations produce more complex object models according to the actual needs. When the Euclidean distance is greater than the threshold, the point is considered to be a feature point; otherwise, it is a nonfeature point, and the screening formula is as follows:

$$p_i = \begin{cases} \text{True}, & d_m > \sigma_h, \\ \text{False}, & d_m < \sigma_h. \end{cases} \quad (2)$$

σ_h :Threshold value

P_i :A point in the set of points P .

When the surface Hermitian data samples of two grid points connected by a grid edge belong to unused layers, the vertical wall connecting these two layers also divides their projected images in the x - y plane. Before selecting the initial four points, first 1 points are randomly selected from the target point set, which is controlled by setting the minimum distance

between any two points, and the minimum distance between any two points satisfies the following equation:

$$\sqrt{(x_{p_i} - x_{p_j})^2 + (y_{p_i} - y_{p_j})^2 + (z_{p_i} - z_{p_j})^2} > d. \quad (3)$$

d :Minimum distance threshold.

Most of the urban monolithic buildings have floors, and the CSG and B-rep expressions of the monolithic building 3D modeling, which is the hierarchical combination expression of the monolithic building model, are shown in Figure 2.

Secondly, in the point cloud processing module in CAD software, the 3D of the feature in any direction and angle can be displayed, so the contour lines of the feature are extracted using manual capture. The premise of multipoint cloud fusion is to clarify the superior information and complementary needs of different point clouds. That is, a sketch of object contours is drawn based on the contours projected by the defined sketch reference plane. Since we mainly aim to obtain the basic CSG elements of a single building, we are able to obtain the building CSG elements by decomposing them according to the principle of combining the shapes of building components and the overall shape of the building. Based on the base geographic grid model, the mapping relationship between multisource 3D model data and the base geographic grid model based on location and spatial scale is established. And the coding of the base geographic grid unit is used to uniquely identify the multisource 3D model data, and the geographic grid is used as the basic unit to express the spatial scope and spatial scale where the 3D model is located. For this purpose, the candidate point sets in point set P are constructed separately, and the candidate point sets in Q are constructed as follows:

$$C = \{c_i | c_i = \langle p_i, q_{i1}, q_{i2}, q_{i3}, \dots, q_{ik}, 1 \leq i \leq 4 \rangle\}. \quad (4)$$

c_i : i th point

p_i :Candidate point set

k :Number of candidate points.

Finally, the creation of a true 3D model of the feature can be done in the sophisticated CAD modeling module. The rigid body transformation is obtained by minimizing the following error function:

$$E = \sum_{i=1}^m \|Rp_i + t - q_{ci}\|^2, p_i \in P, q_{ci} \in Q. \quad (5)$$

The shape and size of the object outline sketches drawn above are edited and rectified to obtain accurate outline shapes. The geometric shape model of the single building is abstracted, and then, the single building is decomposed according to certain rules, the purpose of which is to obtain simple and basic building CSG elements by decomposing the single building. The geometric features of point cloud data mainly refer to the points, lines, and surfaces that can reflect the geometric shape and texture characteristics of the target object, and these geometric features are the basis of 3D modeling, running through the whole process of 3D modeling and

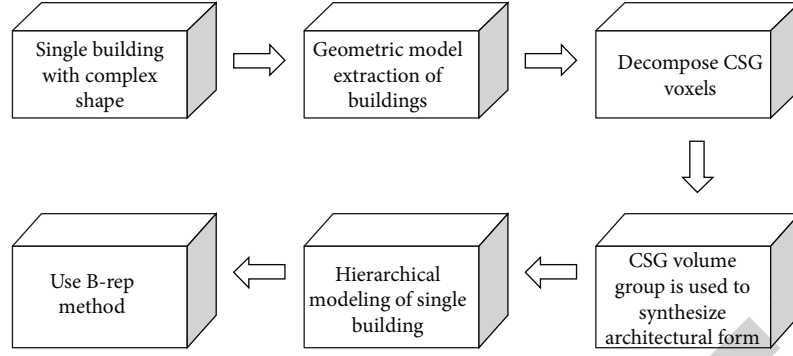


FIGURE 2: General idea of CsG/B-rep hybrid modeling for single building.

affecting the accuracy and credibility of each step. Constructing solid geometry is to make use of the basic geometry we are familiar with through Boolean operations to construct interesting volumes and spaces. On the one hand, it has a unique architectural language on the facade, and on the other hand, it constructs a nontraditional curved interior space. The concept of constructing solid space can increase the sense of hierarchy of the whole building and emphasize the clarity and systematicness of the design idea, by eliminating the inconsistency of spatial datum, scale, and semantic expression between different point clouds, such as the conflict of spatial location, structural semantics, and topological relationship, in order to achieve accurate data and minimum redundancy. The mapping relationship between the 3D model and the benchmark geographic grid cells is constructed according to the accuracy of the 3D model, to describe the relationship between the multisource 3D model data in terms of the adjacency and association relationships between the grid cells.

2.2. True Three-Dimensional Modeling Method of Irregular Shape. Three dimensional laser scanning has the characteristics of high efficiency, high precision, high security, and automatic operation. The high efficiency is reflected in the fact that the 3D laser scanner can shoot millions of points per second and quickly form spatial information images. Compared with the traditional manual measurement method, it greatly saves man hours. High precision which is reflected in the past manual measurement is often in the form of point to area. The three position laser scanner is equipped with a precision sensor, which can adjust the resolution to meet the accuracy requirements of the project. It can also make space imaging, form the coordinates of space points, and form a more delicate display of the target. The application of laser 3D scanning data depends on the quality of the point cloud data modeling, at present for the point cloud data true 3D modeling which has been the focus of research in the application of laser scanning data. In 3D modeling, the main problem is to model irregular shapes in true 3D and make the drawn model have a three-dimensional and realistic feeling, to achieve the ideal visual effect; at the same time, we also need to organize the data, reduce the storage space, facilitate the transmission of data, and speed up the display speed. The display of 3D graphics

includes geometric transformation, projection transformation, shear transformation, and view area transformation. The process is shown in Figure 3.

First, the point cloud units are dispersed into the 3D ellipsoid by using the basis function to interpolate the long places in the multidimensional space, with the interpolation center point as the center of the sphere and the support domain as the radius to build the ellipsoid. After the rigid body transformation of P , the center of mass of P should be the same as the center of mass of S . Therefore, we first calculate the centers of mass of P and S as follows:

$$\begin{aligned}\bar{P} &= \frac{1}{m} \sum_{i=1}^m p_i, \\ \bar{s} &= \frac{1}{m} \sum_{i=1}^m s_i.\end{aligned}\tag{6}$$

Using stereo image data and digital photogrammetry, we obtain the coordinates of feature points based on the interrelationship between images to build a digital surface model and then build a building model by texture mapping. We have previously converted all point cloud data projections into 2D regular grid, given a quadtree cell c in 2D regular grid, the set of surface Hermitian points of all grid points in c is denoted as S , and the set of boundary Hermitian points of initial grid edges on c is denoted as B . The focus is to find the interpolation center and local support domain to build the ellipsoid, and the selection of interpolation center points. The principle is that the ellipsoid formed by the support domain can contain all point clouds. Moreover, the number of ellipsoids obtained is most appropriate when the overlap of the support domains is greater than a threshold value. The unified description rules of multisource 3D model information are designed to structure the expression of geometric information, appearance information, and attribute information, so that the multiscale reference geographic grid model is constructed to cover the spherical geospatial space where the survey area is located. Then, the residuals of each point are squared and summed up as the total residuals of the whole fitted plane. The residuals from a point to a plane $aX + bY + cZ = 1$ are calculated as shown in the following equation.

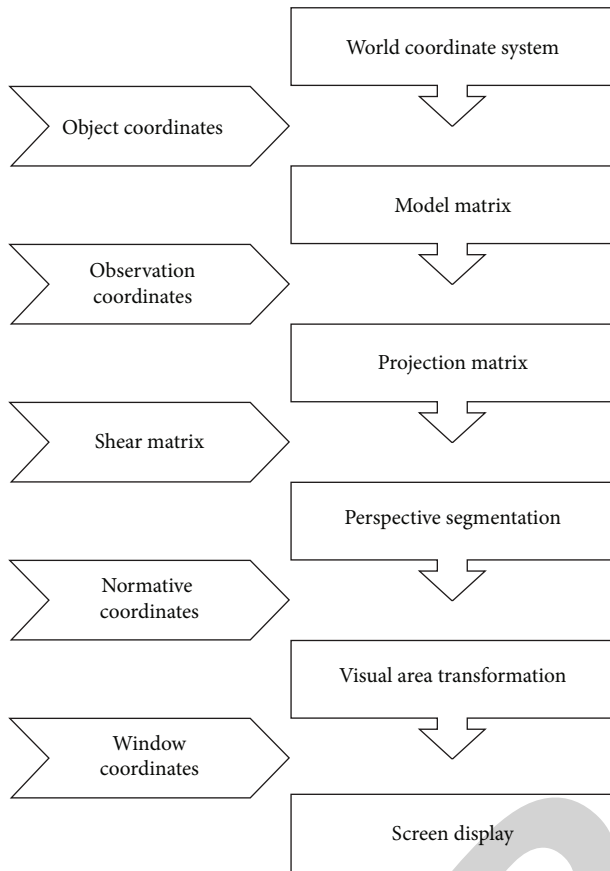


FIGURE 3: Display flow of three-dimensional graphics.

$$\delta_i = \frac{|ax_i + by_i + cz_i + 1|}{2\sqrt{a^2 + b^2 + c^2}}, \quad i = (1, 2, \dots, n). \quad (7)$$

The fitting residual of the plane is:

$$\sigma = \sum_i^n \delta_i^2, \quad i = (1, 2, \dots, n). \quad (8)$$

Secondly, the point cloud data within the ellipsoid is approximated by polynomial to the surface, and the local surfaces between adjacent perched spheres are weighted by radial basis functions to fit, to obtain a 3D triangular mesh model. The spatio-temporal reference and accuracy consistency processing aims to establish a uniform point cloud model for the whole scene, and the scale consistency processing aims to cut down the scale differences between point clouds of different densities and accuracies for the same target representation. The semantic consistency processing aims to synthesize the representation of different detail features of the same target by different point clouds. If the inner surface of a single building is modeled in 3D, the direct consequence is the surge of data volume. The mapping relationship between the 3D model data and the reference geographic grid cells is established according to the spatial extent and data accuracy. Thus, it can describe the spatial location and expression scale of 3D model data by using the reference geographic grid unit and

complete the transformation from the scattered expression of multisource 3D model data to the unified expression based on the reference geographic grid unit. At the same time, it can also use the “body” expression to introduce the octree representation mechanism, so that it can propose a 3D data model similar to the raster vector integration of the 2D plane. By judging whether the angle of the local normal vectors between two adjacent points is within the specified threshold, it is possible to determine whether these two points belong to the same plane. The calculation method of the angle of the normal vector between discrete points refers to:

$$\cos \theta = \frac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} + \sqrt{x_2^2 + y_2^2 + z_2^2}}. \quad (9)$$

Finally, the surface is approximated locally by polynomials under the constraint of the support domain. Then, normalized RBF weighted fitting is used to form an adaptive PU implicit surface, and another part of the function is normalized RBF-weighted fitting to improve the adaptive PU implicit surface to obtain a more satisfactory 3D model. A unified description rule is designed to realize the structured expression of 3D model information and complete the unified expression of multisource 3D model data based on location. The top geometric features of the building are measured in the orthophoto image. The feature points of the contour lines are extracted by the contour line extraction algorithm, but the geometric structure of these feature points is not a regular rectangular structure, so we can further regularize the initial contour lines. When the viewpoint is close to the object, the model details can be observed in abundance; when the viewpoint is far away from the object, the observed model details are gradually blurred, and the program selects the corresponding details for display in real time according to certain conditions of judgment, to avoid wasting time by drawing those details that are relatively less meaningful. For street-level buildings or landmark buildings, important public buildings with local detail features above 0.3 m are represented by the model, and those smaller than that size are expressed with the help of textures, but the steps need to be represented in full.

3. Application Analysis of Multisource Point Cloud Algorithm in Intelligent City 3D Model

3.1. Preprocessing Analysis of Point Cloud Data. The postprocessing process of point cloud data plays the most critical role in order to accurately express the condition of the scanned area in order to obtain a large amount of point cloud data of a scene in a city using a ground-based LiDAR scanning system. The original point cloud data is collected by the 3D laser scanning system, and its point information generally includes 3D coordinate information, point cloud texture information, and reflection intensity information. Point cloud alignment is used to stitch the point cloud dataset collected several times into the same scene, and the basic principle is to find the same name point pair and solve the transformation parameters according to the same name point pair. The point cloud

dataset is selected as a point set with a sampling angle of 40 degrees for the synthetic point cloud and 20 degrees for the other point clouds and rotated uniformly at 10-degree intervals from 0 degrees and used as the alignment point set, and the alignment results of the target point set along the parallel and vertical directions are shown in Figures 4 and 5.

First, in the three-dimensional modeling of features, noise point data has a great impact on the three-dimensional modeling of features, so as in the point cloud data modeling before the need to eliminate noise points. In the scanning measurement, external objects on the measured object caused by partial occlusion, such as in the city of a building for scanning measurement, vegetation, passers-by, etc. on the measured object of the occlusion, resulting in the scan to obtain the measured object point cloud data in the false points and uneven points point cloud data acquisition by the three-dimensional laser scanner in different positions under the respective acquisition, its point cloud data collected at each station is the current attitude of the scanner. Therefore, the point cloud data collected at each station need to be aligned to the same scene by eliminating noise points. By finding the closest point in the target point set for each point in the source point set to form a homonymous point correspondence, and solving the rigid body transformation matrix based on the principle of least squares, the source point set is made to be close to the target point set under continuous iterative operation. In order to reduce the complexity of the search algorithm, an effective method is to select a smaller number of target feature points as candidate points to reduce the search range.

The next step is to simplify the point cloud data, where the scanner works by first performing a vertical deflection angle of the laser beam in the vertical direction, followed by a preset horizontal rotation with horizontal angular resolution, and then a vertical deflection angle of the laser beam in the horizontal direction. The most common strategy is to use only those feature points that can effectively represent the point cloud or to obtain a subset of feature points that keep the target shape as constant as possible. In the process of urban modeling, it is important to make full use of existing topographic maps of the city at scale and design drawings of planned buildings as a data base for modeling, for example, the inflection or fold points of boundary lines, intersection points between surface boundaries, and common points of multiple adjacent surfaces. Through these points, the topological relationship between each local surface of the learning point cloud can be analyzed. The key to this type of algorithm is to define suitable features for each scanned point such that the feature points in the overlapping regions of the two point clouds are identical, so that the features of the defined points must remain constant under rigid body transformations. The relative running time changes as the rotation angle increases, thus indicating that the rotation angle has an effect on the efficiency of the point cloud algorithm. A comparison of the efficiency based on the rotation angle is shown in Figure 6.

Finally, the point cloud is aligned, and two adjacent point cloud data are aligned. Point cloud alignment is the process of calculating the precise mapping of spatial geometry between different point cloud collections, finding the coordinate trans-

formation parameters, and transforming the dataset to be transformed into a rigid body. The curvature and the change of curvature in the region of the same name point in the collocated point cloud data are constant, only the normal vector changes, which is related to the vector angle and curvature change value, so they can be used as the basis for finding the same name point. We extract the four corner points of the contour line and connect the two opposite corner points to obtain two diagonal lines. After that, we use the intersection point of the diagonal lines as the center of the circle, draw a circle with the farthest distance from the center of the circle as the radius of the four corner points, extend the diagonal lines of the contour line, and intersect the circle to obtain the new four corner points. And the discrete point cloud within each grid cell is statistically analyzed to obtain a PDF for each grid cell. PDF can be used to represent the distribution of discrete points within each grid cell and likewise to represent the process of generating each point within the grid. The key point of this algorithm is to convert the 3D boundary point cloud into 2D road segment shape file form, to realize the association between the sequence of satellite positioning track record points and 3D road boundary through map matching, and to get the road dynamic information based on the satellite positioning track data on the road.

3.2. Analysis of 3D Point Cloud Data Registration. Combine the building outline in GIS with the building height (calculated by the number of floors or other methods). Simple geometry is used to express the shape features of buildings. This method is the most convenient and has the least amount of three-dimensional data, but it is also the most different from the actual situation. Because the aerial image truly reflects all the top information of urban buildings, it also reflects part of the side information of buildings and most of the ancillary information of buildings. Therefore, the shape features of buildings can be obtained by means of digitalization and human interaction. This method can obtain the required information more realistically, but the workload is quite large due to the need for manual intervention.

If there is noise in the point cloud data at the time of acquisition, the topology of the point cloud data obtained from two acquisitions in the same area is not strictly consistent. Therefore, there are errors between the eigenvalues of the corresponding points of the 3D point cloud data, and there are many errors in the final matched point pairs. A comparison of the parameters of the SAC-IA algorithm and the multisource data point cloud algorithm is shown in Figures 7 and 8.

First, matching point pairs are filtered based on Hausdorff distance. The point cloud rotation matrix constructed by seven parameters or quaternions can be obtained by using local features such as point-line identification and matching to achieve multipoint cloud fusion. This requires the construction of a baseline geographic grid model based on the spatial extent of the survey area to virtually divide the geographic space where the survey area is located and provide a unified positioning datum; unlike descriptors, the use of RGB values as point features does not need to consider the neighborhood relationship of points, so its computational complexity in the

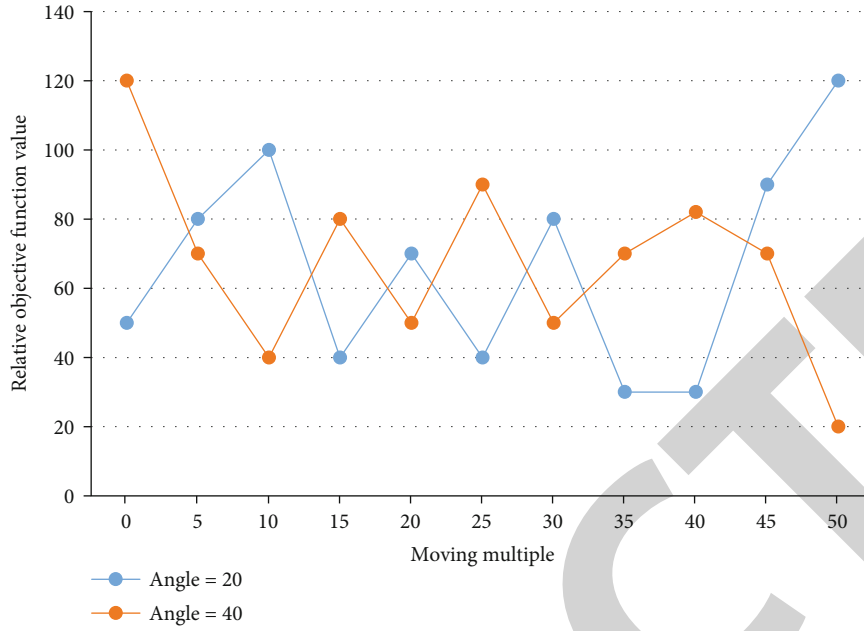


FIGURE 4: Registration results in parallel direction.

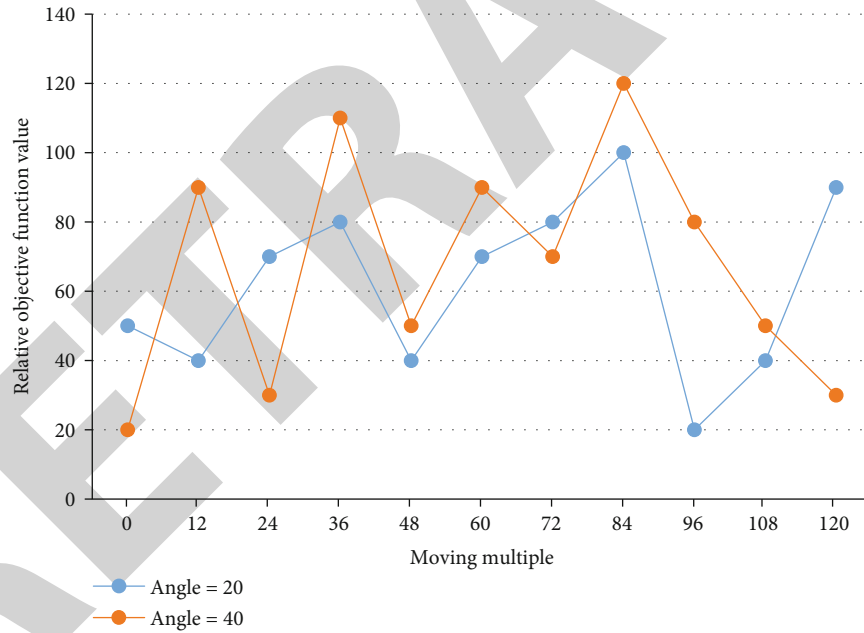


FIGURE 5: Registration results in the vertical direction.

process of constructing the candidate point set is $O(n)$, much smaller than that of descriptors. Once the point-to-point correspondence is determined, the rigid body transformation matrix for global alignment can be solved using the correspondence, and then, the final point cloud alignment can be completed by local alignment operation. When the overlap region is small and the features in the overlap region are obvious, we need to rely on the feature region to recover the global transformation. The sampling mode using normal vector space collects more data at the features than the random sampling mode, which can better reflect the feature information of

the point cloud set, and this feature is the key of the collocation algorithm, so it will inevitably affect the speed and collocation accuracy of the collocation algorithm. Moreover, as the number of point clouds increases, the advantage of Hausdorff distance in alignment efficiency becomes more obvious. The SAC-IA algorithm and the multisource data point cloud algorithm are compared and evaluated, and the two are studied quantitatively in terms of two performances, namely, alignment error and alignment time consumption, respectively, and the relevant performance comparison parameter information is detailed in Tables 1 and 2.

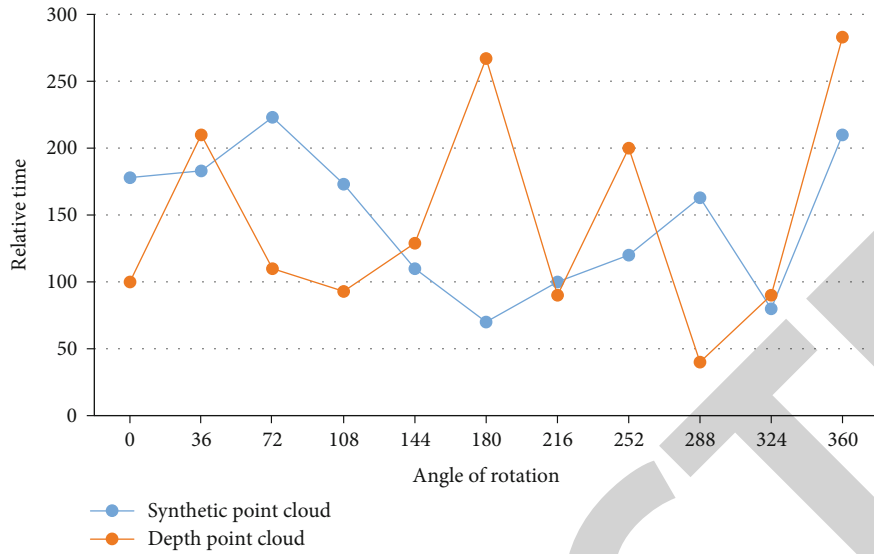


FIGURE 6: Efficiency comparison based on rotation angle.

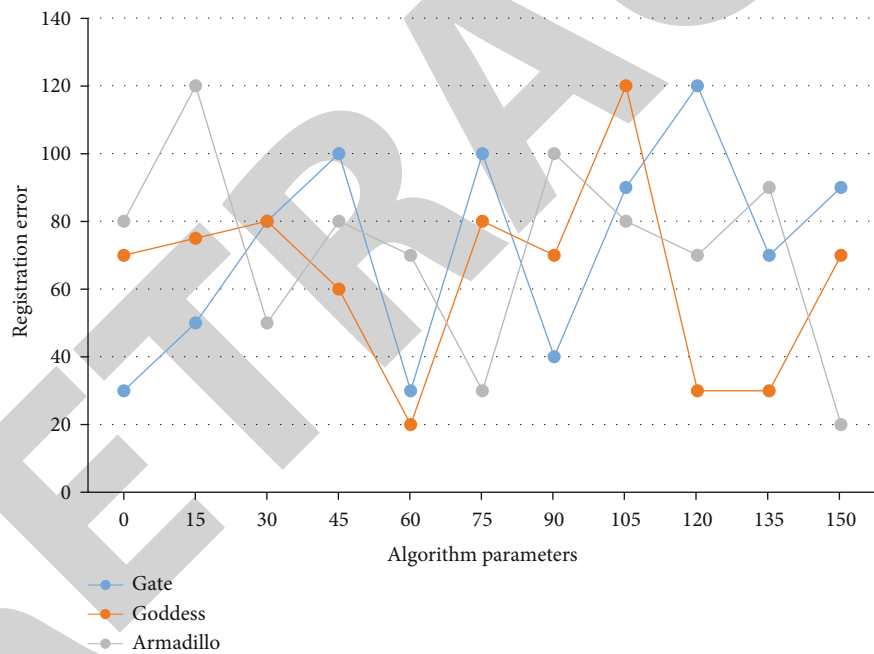


FIGURE 7: Bar chart of SAC-IA algorithm parameter comparison.

Secondly, the random sampling consistency algorithm is used to reject the incorrectly matched point pairs by combining the rigid distance constraint to obtain more accurate matched point pairs. The parameters of the mathematical model are estimated using the minimum set of sampled points that meet the requirements of the algorithm; then, based on the estimated parameters of the mathematical model, more points are selected in the 3D point set to expand the set of sampled points, and the proportion of internal points that fit the mathematical model is calculated. This uses a variety of 3D point cloud descriptors such as fast

point feature histograms to extract features, or to identify geometric feature points and feature lines of building edges. This is because the geometric features of building edges usually satisfy stability and specificity, and executing the algorithm after global alignment can obtain high point cloud alignment accuracy while eliminating the above disadvantages. For the case of incomplete photographs of a one-story building model, the textures of each face should be mapped by taking textures similar to those of the surrounding buildings. This is because the descriptors are expressed in the form of high-dimensional histograms, which are more

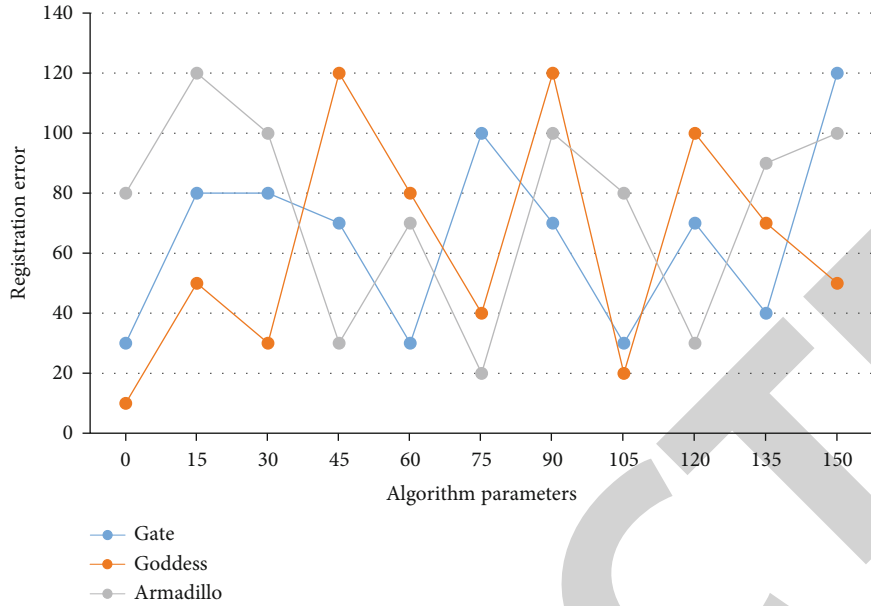


FIGURE 8: Bar chart of parameter comparison of multisource data point cloud algorithm.

TABLE 1: Comparison of registration errors between SAC-IA algorithm and multisource data point cloud algorithm.

	Gate	Goddess	Armadillo
SAC-IA	2.9×10^{-3}	3.1×10^{-3}	4.3×10^{-3}
Multisource data point cloud algorithm	1.7×10^{-5}	3.6×10^{-5}	4.5×10^{-5}

TABLE 2: Comparison of registration time between SAC-IA algorithm and multisource data point cloud algorithm.

	Gate	Goddess	Armadillo
SAC-IA	76	82	49
Multisource data point cloud algorithm	38	52	17

sensitive to the distinction between points, and thus, the generated candidate point set is more accurate than the RGB candidate point set.

Finally, the improved nearest point iteration algorithm is used to accurately align the 3D point cloud data. Focus on the geometric relations of the same correspondences, then sample a large number of these same correspondences, and rank the sampled obtained correspondences according to a certain ranking method, and finally, select the optimal correspondence. Whether it can fully reflect the comprehensive information of the subject depends on the good or bad splicing effect, and the good or bad splicing effect is directly affected by the matching accuracy of the same name point during the splicing. Therefore, we can use the matching scale of the same name points in the iterative process as the judgment factor of the matching accuracy of the same name points. After the candidate point set is determined, the FIPP algorithm can be used to search for homonymous point pairs between two point cloud datasets. It is mainly because the feature values are only limited to a smaller area, and there will be many points with similar features. Moreover, due to the existence of noise in

the point cloud acquisition process, the topology of the point cloud data acquired from the same region twice is not strictly consistent, so it will cause errors in the feature values of corresponding points between the point cloud data. Each pair of regions is aligned to obtain a rigid body transformation and an LCP metric to measure the effect of alignment, i.e., the proportion of overlap between two regions after alignment. Because of the high-dimensional characteristic of FPFH, the probability that the candidate points contain points with the same name is high, so the time to search for new pairs of points is short.

3.3. Data Acquisition. The basic geographic information center is responsible for the collection of basic terrain data, orthophoto data, and 3D coding data of buildings in the whole experimental area. The data source is color aerial image, and the data acquisition means is jx4a Digital Photogrammetry Workstation DPW, which is also one of the most advanced digital photogrammetry operation methods in the world. The resolution of DEM and orthophoto of digital elevation model are 2 meters and 0.5 meters, respectively, covering the

Retraction

Retracted: Prediction and Risk Assessment of Extreme Weather Events Based on Gumbel Copula Function

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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.-H. Yang, Y. Yu, F. Gu, M.-J. Qu, and J.-M. Zhu, "Prediction and Risk Assessment of Extreme Weather Events Based on Gumbel Copula Function," *Journal of Function Spaces*, vol. 2022, Article ID 1438373, 13 pages, 2022.

Research Article

Prediction and Risk Assessment of Extreme Weather Events Based on Gumbel Copula Function

Peng-Hui Yang ¹, Yao Yu ¹, Feng Gu ², Meng-Jie Qu ³, and Jia-Ming Zhu ¹

¹School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China

²School of Finance, Anhui University of Finance and Economics, Bengbu 233030, China

³School of Accounting, Anhui University of Finance and Economics, Bengbu 233030, China

Correspondence should be addressed to Jia-Ming Zhu; zhujm1973@163.com

Received 1 May 2022; Accepted 28 June 2022; Published 21 July 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Peng-Hui Yang 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.

Damage caused by climate catastrophes is severe, especially for the 1-in-100-year events. This study is aimed at assessing the frequency and spatiotemporal regularity of extreme weather events. Based on the selected Gumbel copula function, a joint trivariate distribution of weather events is established. In this study, different univariate return periods and return periods of the joint trivariate distribution are calculated separately. Second, the Moran index is used to determine whether there is a spatial correlation between weather events. In this paper, the spatial and temporal patterns of weather events are determined based on a geographically weighted regression model. The suggestion of adding Bayesian information to the model measurements to improve the model accuracy is presented. Finally, a wavelet neural network model is constructed to predict the probability of extreme weather events throughout the Americas.

1. Introduction

Disasters have shown significant impacts on all types of business in both developed and developing countries. Both direct and indirect impacts of natural disasters are devastating to business activities and their continuity. These catastrophic events have created a significant negative impact on most of the business entities during recent years [1]. In addition, according to the latest findings of the nonprofit German Observatory, nearly half a million people have died from diseases related to climate disasters in the past two decades the past 20 years.

In late March 2021, people living on the east coast of Australia experienced a rare meteorological event. Record-breaking rainfall in some areas and very heavy and sustained rainfall in others led to severe flooding. However, in different places, the disaster was described as a once-in-thirty-year, once-in-five-year, or once-in-100-year event. For meteorologists, every 100 years means that one or more events occur every 100 years on average. The exact probability still varies

from place to place. In parts of the United States, events that occur more than once in 100 years are more frequent than events that occur once in a century.

In this paper, we determine the frequency of weather events based on the average return period. The average return period is the reciprocal of the probability of occurrence per year. For an event with an annual probability of 0.01, the average return period is 100 years, i.e., the once-in-100-year event mentioned in the data, i.e., a once-in-100-year event is not the same as an event that occurs once or at least once every 100 years.

Weather events are mainly expressed by the three aspects of damage degree, damage extent, and duration of weather events. Damage extent refers to the damage to people and property brought about by disasters like typhoons, rainstorms, and earthquakes in a nonman-made force majeure [2]. The extent of damage refers to the extent of damage impact brought about by natural disasters in the spatial dimension, for people and property. Weather event duration refers to the duration of the weather event in days [3].

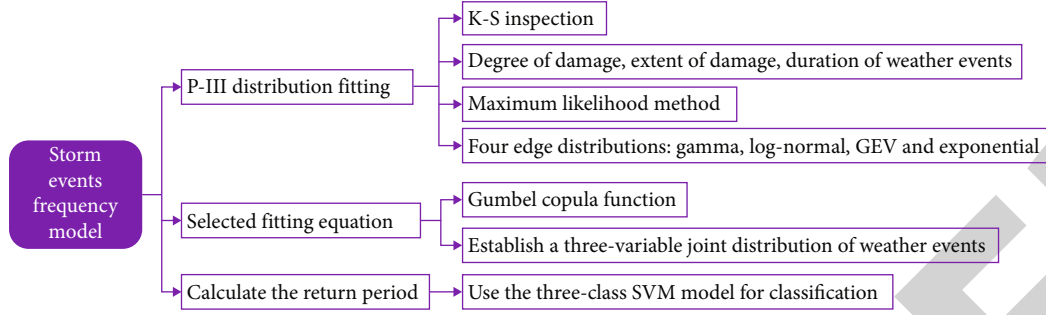


FIGURE 1: The idea of the storm event frequency model.

In the study of weather events, early scholars focused on the impact of disasters [4, 5] and later extended to regional and economic studies [6–8]. Ordinary least squares (OLS) can only estimate the parameters in a global or average sense and cannot reflect the spatial local variability; so, it cannot reveal the spatial dependence; geographically weighted regression (GWR) can estimate each parameter spatially and can better reflect the spatial dependence among the factors affecting the occurrence of extreme weather events.

In this paper, we established a weather event frequency analysis model, analyzed the return period of weather events, and considering the spatial heterogeneity in the country, such as the latitude and longitude of each state, natural environment, and social environment, and a geographically weighted regression model is established to deduce the temporal and spatial pattern of weather time development in the United States and predict the relative frequency of weather events in different regions. The research method in this paper is well generalized and innovative in the frequency prediction of extreme events.

2. Basic Assumptions

To simplify the problem, we make the following basic assumptions. (i) We define that all research objects are randomly distributed in space. In Moran index analysis, we can calculate the index value by placing the observation index in the same space state through this assumption. (ii) All observation indicators are not independent in space, and the spatial relationship is nonstationary. For the spatial geographic weighted regression model, because of the correlation between the indicators, different spaces are heterogeneous and nonstationary and thus have different effects on the observation indicators. (iii) The observation index of each unit can be seen as a point in space. (iv) For spectral clustering analysis, because points in different areas of space constitute a point set, therefore, we can use distance to measure the degree of spatial correlation.

3. Storm Event Frequency Model Based on Copula

3.1. Concept Introduction. There is a certain relationship between the intensity of a natural event and its probability of occurrence, as discussed elsewhere [9]. The greater the

intensity, the less likely or the lower the probability of occurrence, such as the rare Australian meteorological event in March mentioned in the data, which caused a very serious flood. Such events are called low-probability events and are usually obtained by extrapolation of extreme distribution functions, but the error range becomes larger as the degree of extrapolation expands. In addition, the results obtained by using different distribution functions are not the same. We assume that the probability of occurrence of such an event is P , and the time interval between its reoccurrence and the initial time is T ; we call it the average return period, which can be expressed as follows:

$$\begin{cases} T = \sum_{n=1}^{\infty} np(n), \\ p(n) = p(1-p)^{n-1}. \end{cases} \quad (1)$$

Among them, n is the return period, $p(n)$ is the probability that the return period is n , and the average return period can be obtained from the above formula. That is, the average return period $T = 1/p$ is the inverse of the annual probability of occurrence, as discussed elsewhere.

For example, for an event with an annual probability of 0.01, the average return period is 100 years, that is, the once-in-100-year event mentioned in the data; that is, the once-in-a-hundred-year event is not equal to an event that will occur once or at least once in 100 years [10]. However, a once-in-a-hundred-year event does not mean an event that will occur once or at least once in 100 years. The probability of such a small event needs to be extrapolated from the extreme distribution function; that is, the lower the probability of occurrence, the greater the intensity of the event. Based on this idea, we build a frequency analysis model which is shown in Figure 1.

3.2. Weather Event Frequency Model Based on Copula Function. To construct a multivariate copula function, it is necessary to determine the marginal distribution function of the variable. Based on the domestic and foreign research on the marginal distribution of flood events in various weather events [11–14], this paper first adopts the P-III distribution for fitting calculation. To further improve the fitting accuracy and select the best fitting function, this paper selects four marginal distributions of gamma, log-normal, GEV, and exponential to fit the damage degree (S), damage

range (D), and weather event duration (T) sequence [15]. First, the maximum likelihood method is used to estimate the marginal distribution parameters, and the K - S test method is used to verify the fit.

The Kendall rank correlation coefficient is used to describe the correlation degree of different variables [16, 17]. The calculation formula is as follows:

In the formula, n is the length of the weather event sequence, sign is the sign function, and $\{(x_1, y_1) \cdots (x_n, y_n)\}$ is the random sample in the data event.

3.3. Construction of Joint Distribution of Weather Event Variables. Copula function is a multivariate function subject to uniform distribution, and its domain is $[0,1]$. By supposing F_1, F_2, \dots, F_n is a continuous random variable, then there is a unique Copula function C which is generated that satisfies the formula for any $X \in R$:

$$F(x_1, x_2, \dots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} = C[F_1(x_1), F_2(x_2), \dots, F_n(x_n)], \quad (2)$$

where x_1, x_2, \dots, x_n is a different sample, and $F(x)$ is a marginal distribution function.

Based on the data characteristics of weather events from 2001 to 2021, this paper summarizes and extracts three one-dimensional index data, namely, the degree of damage, the extent of damage, and the duration of the weather event. These three levels of indicators can more fully reflect the characteristics of any weather event in the data, and the joint distribution function is selected based on this characteristic data. In this article, two elliptical copula functions (t and Gaussian) and three symmetrical Archimedean functions are used to construct the two-dimensional joint distribution of damage degree-damage range, damage degree-weather event duration, and damage range-weather event duration and adopt two kinds of symmetry, and three asymmetric distributions are used to build the three-dimensional joint distribution of damage degree-damage range-weather events; the maximum likelihood method is used to estimate parameters; root means square error R_{RMSE} criterion and A_{AIC} information criterion method are used to judge the goodness of fit. The smaller the value, the higher the degree of fit [18]. The calculation formula is follows:

$$\begin{cases} M_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n (p_{ei} - p_i)^2, \\ A_{\text{AIC}} = n \ln(M_{\text{MSE}}) + 2k, \\ R_{\text{RMSE}} = \sqrt{M_{\text{MSE}}}. \end{cases} \quad (3)$$

In the formula, n is the weather event sequence length, k is the number of Copula parameters, i is the sequence number in descending order, p_i is the theoretical value of the i -th joint distribution, p_i is the i -th empirical value, and M_{MSE} is the mean square error.

3.4. Return Period Calculation. The calculation formula of the univariate return period T is $T = 1/[1 - F(x)]$. In the for-

mula, $F(x)$ is the univariate marginal distribution function. The calculation formula for the joint return period $T_a(x, y)$ and the co-occurrence return period $T_b(x, y)$ of the two variables X and Y is as follows:

$$\begin{cases} T_a(x, y) = \frac{1}{1 - F(x, y)} = \frac{1}{1 - C(u, v)}, \\ T_b(x, y) = \frac{1}{1 - u - v + c(u, v)}. \end{cases} \quad (4)$$

Among them, U and V are marginal distribution functions; $C(u, v)$ is the two-variable joint distribution function.

The calculation formula for the joint return period $T_a(x, y, z)$ and the co-occurrence return period $T_b(x, y, z)$ of the three variables x, y and z is as follows:

$$\begin{cases} T_a(x, y, z) = \frac{1}{1 - C(u, v, w)}, \\ T_b(x, y, z) = [1 - u - v - w + C(u, v) + C(u, w) + C(v, w) - C(u, v, w)]^{-1}. \end{cases} \quad (5)$$

Among them, u, v, w is the marginal distribution function; $C(u, v), C(u, w), C(v, w)$, is the two-variable joint distribution function; $C(u, v, w)$ is the three-variable joint distribution function.

3.5. Model Construction. The fitting test and correlation coefficient of the marginal distribution function of each univariate damage degree (S), damage scope (D), and weather event duration (T) are shown in the following table. Table 1 shows that the characteristic variables of weather events are more correlated well, the degree of damage has the strongest correlation with the duration of the weather event, $\varphi = 0.6239$, followed by the degree of damage and the scope of the damage, $\varphi = 0.6007$, and the degree of damage has the weakest correlation with the duration of the weather event, $\varphi = 0.5981$.

At the confidence level $\alpha = 0.05$, the statistic value of each variable K is less than the critical value; so, the K - S test is accepted. The higher the ρ value and the lower the k statistic value, the distribution function is selected as the univariate marginal distribution function of the weather event in the data, that is, the degree of weather event damage. The sequence selects the gamma distribution, the weather event range selects the log-normal distribution, and the weather event duration selects the P -III type distribution. The parameter values are shown in Table 1.

3.6. The Establishment of the Optimal Copula Function. The two-variable and three-variable joint distribution fitting test and parameter values are shown in Table 2. Table shows the best fitting function of the two-variable and three-variable joint distribution. As can be seen from Table 2, the Gumbel Copula function is the best fit for the joint release of the three variables. The fitting effect of the joint distribution of variables is the best. Therefore, Gumbel copula is used to analyze the three-variable return period of damage degree (S), damage scope (D), and weather event duration (T).

TABLE 1: Test statistics and correlation coefficient of marginal distribution function.

Variable	Fitting function	p	Statistic value k	Critical value d	Correlation coefficient r
Extent of damage	Gamma	0.8701	0.1821	0.3226	S-D (0.6007)
	Log-normal	0.7557	0.1524	—	—
	GEV	0.1265	0.2849	—	—
	exp	0.2782	0.2523	—	—
	P3	0.7249	0.1049	—	—
Damage scope	Gamma	0.4721	0.3426	0.3226	S-T (0.6239)
	Log-normal	0.7932	0.2576	—	—
	GEV	0.3425	0.1027	—	—
	exp	0.4192	0.4026	—	—
	P3	0.2491	0.2849	—	—
Weather event duration	Gamma	0.8047	0.1239	0.3226	D-T (0.5981)
	Log-normal	0.8245	0.1597	—	—
	GEV	0.9042	0.9034	—	—
	exp	0.8201	0.1942	—	—
	P3	0.9625	0.0892	—	—

*Note: the boldface in the table is the optimal marginal distribution function.

TABLE 2: Three-variable joint distribution of weather events.

Function type	Parameter	R_{RMSE}	A_{AIC}
Clayton	2.74	0.062	-40.7
Frank	7.52	0.051	-52.4
Gumbel	2.42	0.046	-51.9
Joe	2.97	0.069	42.5
NC	4.420/4.252	0.057	-49.2
NJ	3.024/2.563	0.049	-42.5

TABLE 3: Weather event return period.

Weather event type	Return period	Weather event type	Return period
Thunderstorm wind	30	Ice storm	50
Flash flood	30	Wildfire	50
Winter weather	30	Extreme cold	50
High wind	30	Coastal flood	50
Marine hail	100	Seiche	100
Marine tropical depression	100	Sneaker wave	100

3.7. *Model Results.* Establish the three-variable joint distribution of weather events based on the selected Gumbel Copula function, calculate the joint and co-occurrence return periods, and calculate the three-variable joint distribution return periods under different univariate returns periods [19, 20]. The results are shown in Table 3.

We further classify the return periods of weather events in the data through the three-class SVM model, and the results are shown in Figure 2.

4. Spatial Correlation Analysis: Moran Index

To further study the features of extreme weather events, we analyze the spatial regularity of several weather events using the Moran index and geographically weighted regression models [21].

Moran index is divided into global Moran index and local Moran index, which are used to judge the degree of aggregation and dispersion of the index in the global and local, respectively.

Moran I (global Moran index) is defined as follows:

$$\text{MoranI} = \frac{\sum_{i=1}^n w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{s^2 \sum_{i=1}^n w_{ij}}. \quad (6)$$

In it,

$$s^2 = \frac{1}{n} (y_j - \bar{y}), \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (7)$$

\bar{y} is the selected index value of the place (observation value), n is the total number of units in the whole area, and w_{ij} is the binary adjacent space weight matrix; according to whether the two units are adjacent or not, the value of w_{ij} is

$$w = \begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{bmatrix}. \quad (8)$$

$$w_{ij} = \begin{cases} 1 (i \text{ is adjacent to } j) \\ 0 (i \text{ is not adjacent to } j) \end{cases},$$

$i = 1, 2, \dots, 48, j = 1, 2, \dots, 48 (i \neq j).$

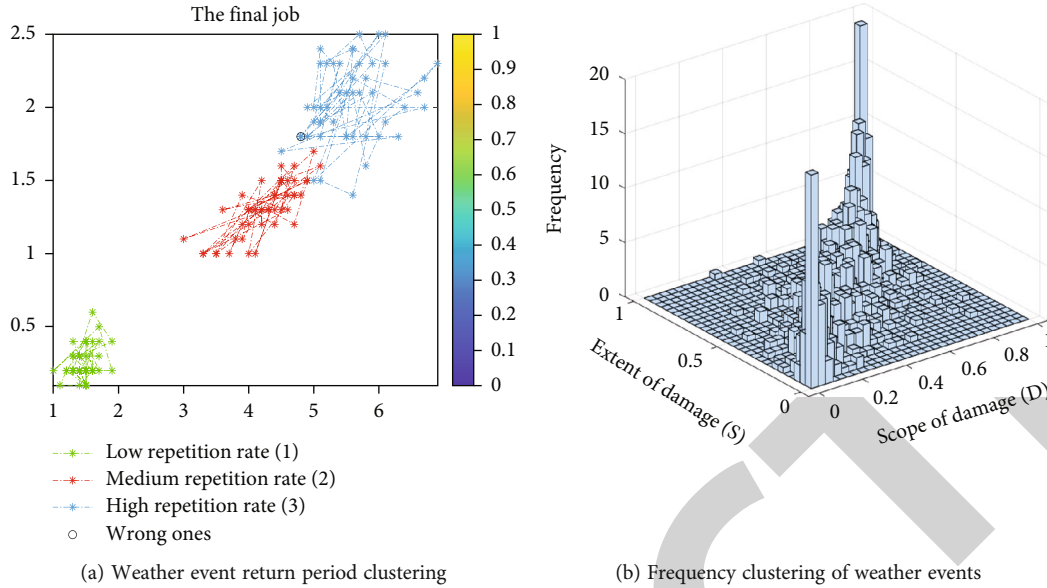


FIGURE 2: Weather event clustering.

Moran I is the product sum of observations in various regions, and its value range is in $[-1, 1]$ [22].

Moran

$$I \begin{cases} >0 (\text{The larger the value, the more significant the aggregation benefit}), \\ =0 (\text{No spatial correlation}), \\ <0 (\text{The smaller the value, the more significant the discrete benefit}). \end{cases} \quad (9)$$

It needs to be tested for the H_0 hypothesis, the hypothesis test that all the research objects are randomly distributed in the space. Next, the Z test is used for verification.

$$Z = \frac{I - E(I)}{\sqrt{\text{var}(I)}}. \quad (10)$$

In it, the calculation formula of $E(I)$ and $\text{Var}(I)$ is as follows:

$$E(I) = \frac{1}{n-1},$$

$$\text{Var}(I) = \frac{n^2(n-1)1/2\sum_{i \neq j} (w_{ij} + w_{ji})^2 - n(n-1)\sum_{i=1}^n (w_{kj} + w_{jk})^2 - 2(\sum_{i \neq j} w_{ij})^2}{(n+1)(n-1)^2\sum_{i \neq j} w_{ij}(\sum_{i \neq j} w_{ij})}. \quad (11)$$

Under the standard of significance of 0.05, as long as $|z| > 1.96$ or $(p < 0.05)$, the H_0 null hypotheses can be rejected, and all research objects are randomly distributed in space [22].

We calculated the global Moran index for the six weather events, and the results are shown in Figure 3. The abscissa represents the described variable, and the ordinate represents the spatial lag vector of the described variable. The four quadrants correspond to the four spatial aggregation effects

of high-high clustering, low-high clustering, low-low clustering, and high-low clustering. Figure 4 corresponds to the permutation test results of each variable Moran index and the corresponding statistics.

According to the p value of the global Moran index, among the six weather events, thunderstorm wind, flash flood, extreme cold/wind chill, and dense fog have significant spatial aggregation effects; blizzard and frost/freeze have lower spatial aggregation effects.

5. Spatial Law Exploration: GWR Model

The geographical weighting model (GWR model) is a model for spatial nonstationarity caused by changes in the relationship or structure of variables caused by changes in geographic location [23–25].

Here are the stats for weather events and latitude and longitude by state.

The general form of the model is as follows:

$$y_{\{i\}} = \beta_0(i) + \beta_1(i)x_{1,\{i\}} + \beta_2(i)x_{2,\{i\}} + \dots + \beta_k(i)x_{k,\{i\}} + \mu_{\{i\}}. \quad (12)$$

Among them, (i) represents the region, and $\{i\}$ represents the sample set included in the estimation; here, we select 48 states in the United States.

According to the principle of borrowing points, each local point obtains data from the surrounding area to form a different sample set for each region. Here are the stats for weather events and latitude and longitude by state. Establishing a new spatial weight matrix is as follows.

$$A_{\text{ICC}} = 2n \ln(\sigma) + n \ln(2\pi) + n \left(\frac{n + \text{tr}(s)}{n - 2 - \text{tr}(s)} \right). \quad (13)$$

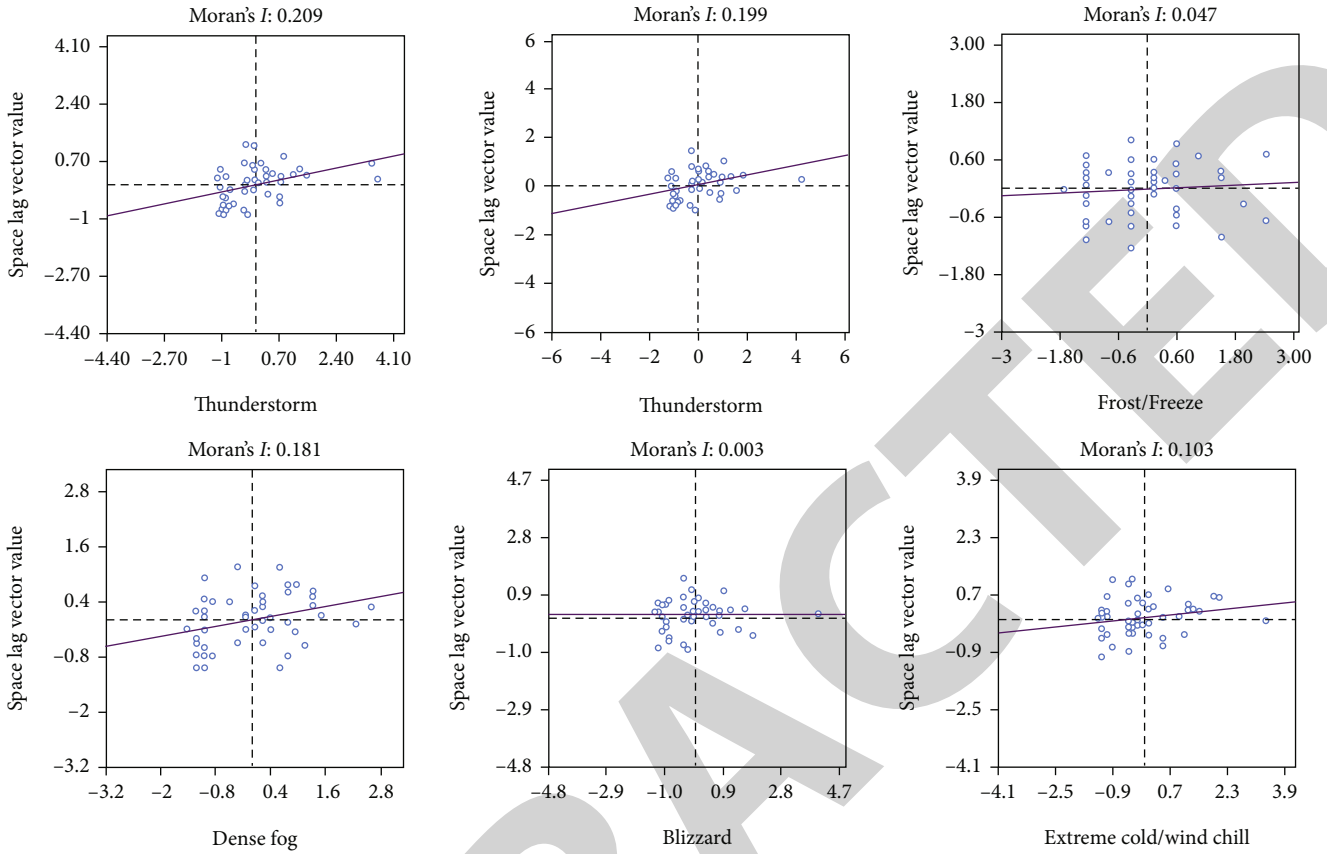


FIGURE 3: Global Moran index results for six types of weather.

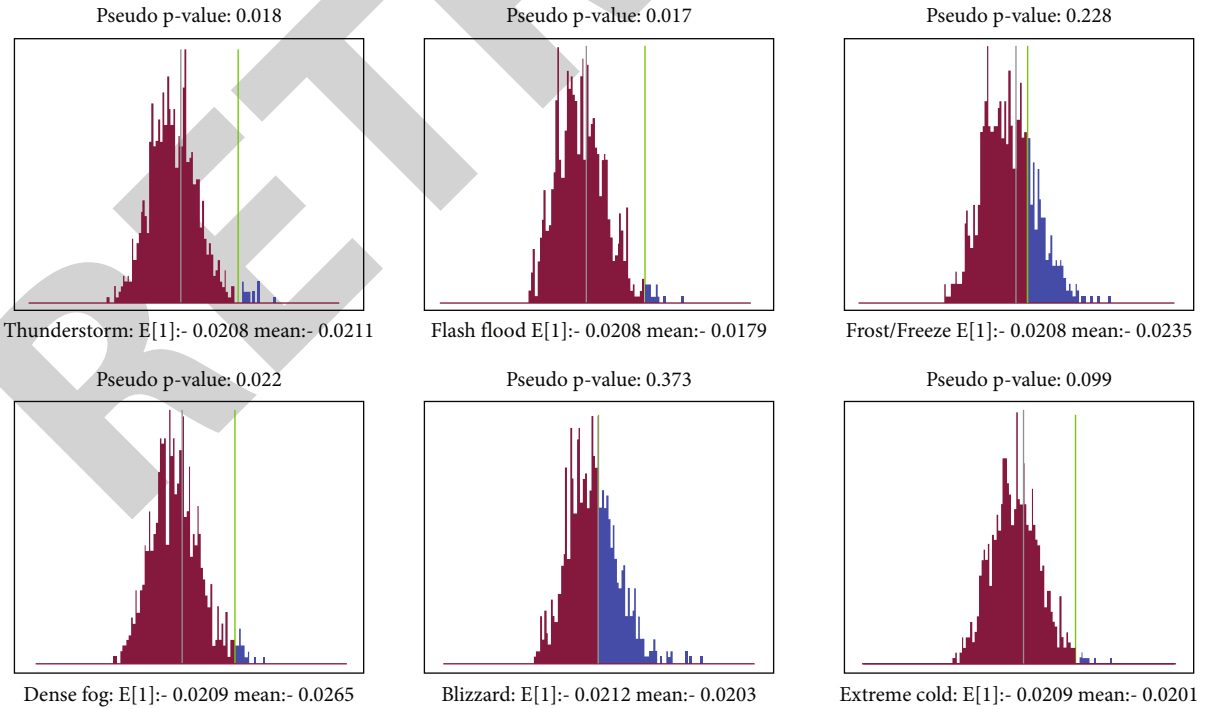


FIGURE 4: Test result of global Moran index after permutations.

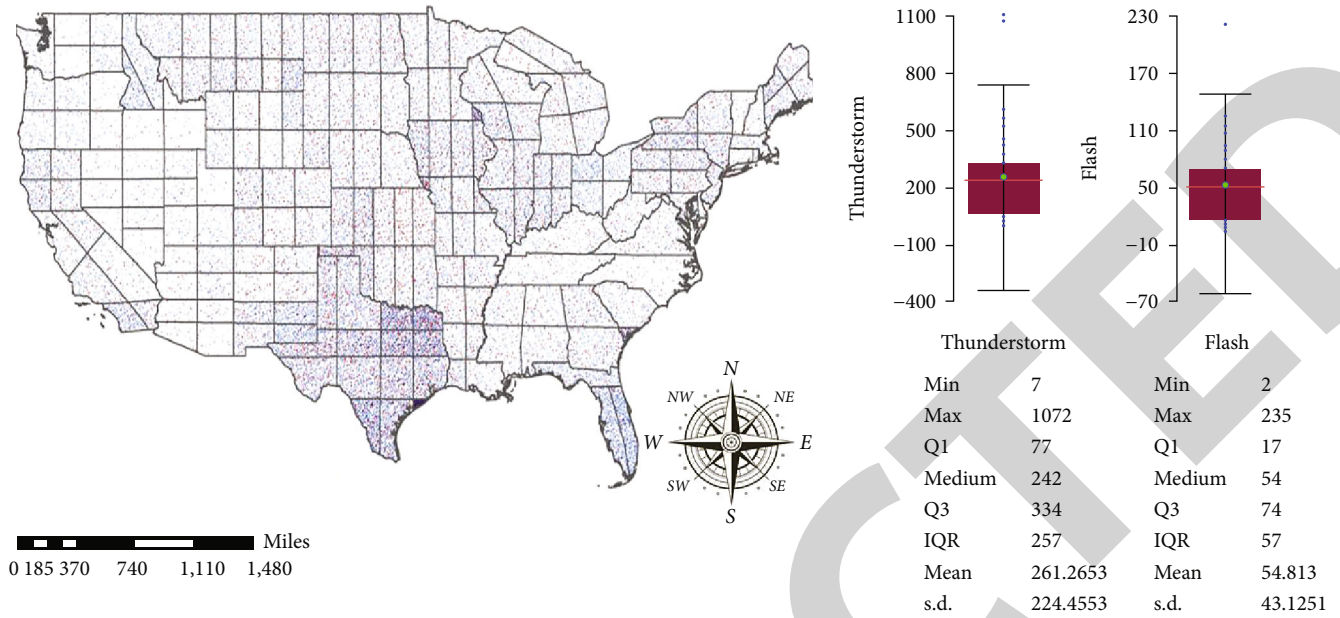


FIGURE 5: Thunderstorm wind and flash flood event correlation diagram.

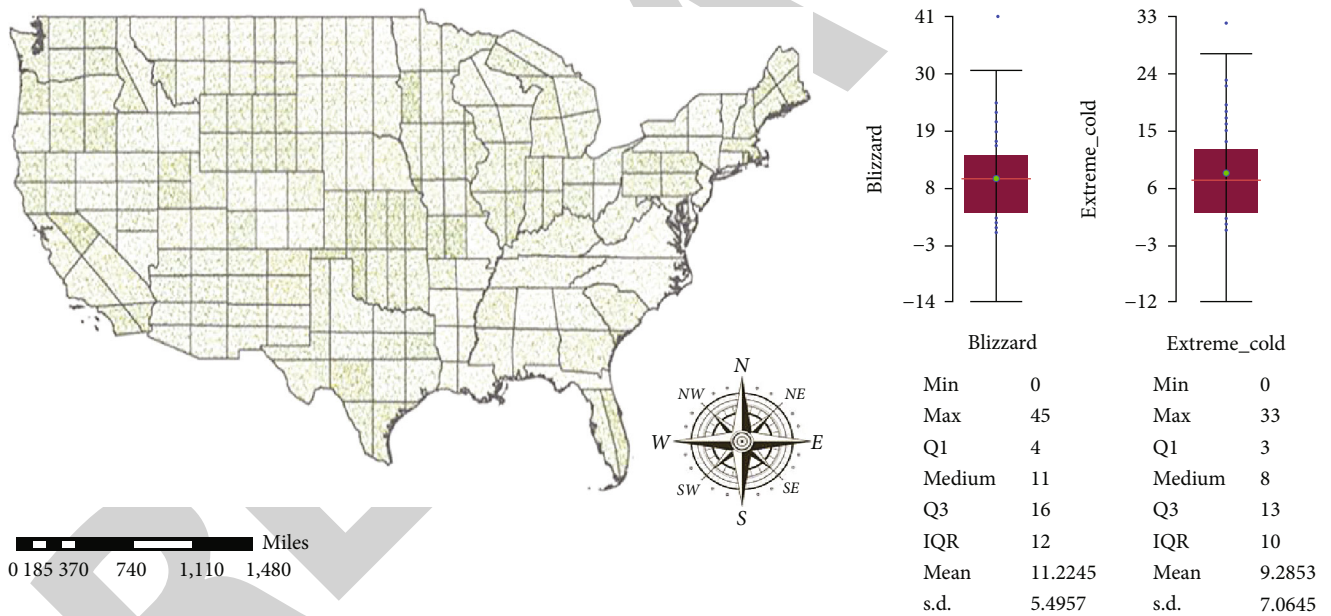


FIGURE 6: Blizzard and extreme cold/wind chill event correlation diagram.

The value of A_{ICC} is determined by the number of independent parameters and the maximum likelihood function of the model.

We count the relevant data of 2005, 2010, 2015, and 2020 to establish a geographically weighted regression model. The model results are shown in Figures 5–7.

Table 4 shows the model test results reflect that the GWR model exhibits larger goodness of fit R^2 , a smaller AICC value, residual sum of squares, and residual estimated standard deviation sigma value. Based on the results of the GWR model, the following conclusions were found.

Thunderstorm wind and flash flood have the highest frequency in the United States, showing a spatial trend of increasing from west to east and from north to south. Extreme cold/wind chill and blizzard occur in the middle frequency, and the frequency of extreme cold/wind chill on the west coast is much lower than that of the central and east coasts. The frequency of blizzard in the western states of California, Nevada, Arizona, and Iowa is much lower than the frequency in the central and eastern states. Dense fog and frost/freeze have a relatively stable distribution throughout the United States. Except for the four states on the east coast of New Jersey, Connecticut, Vermont,

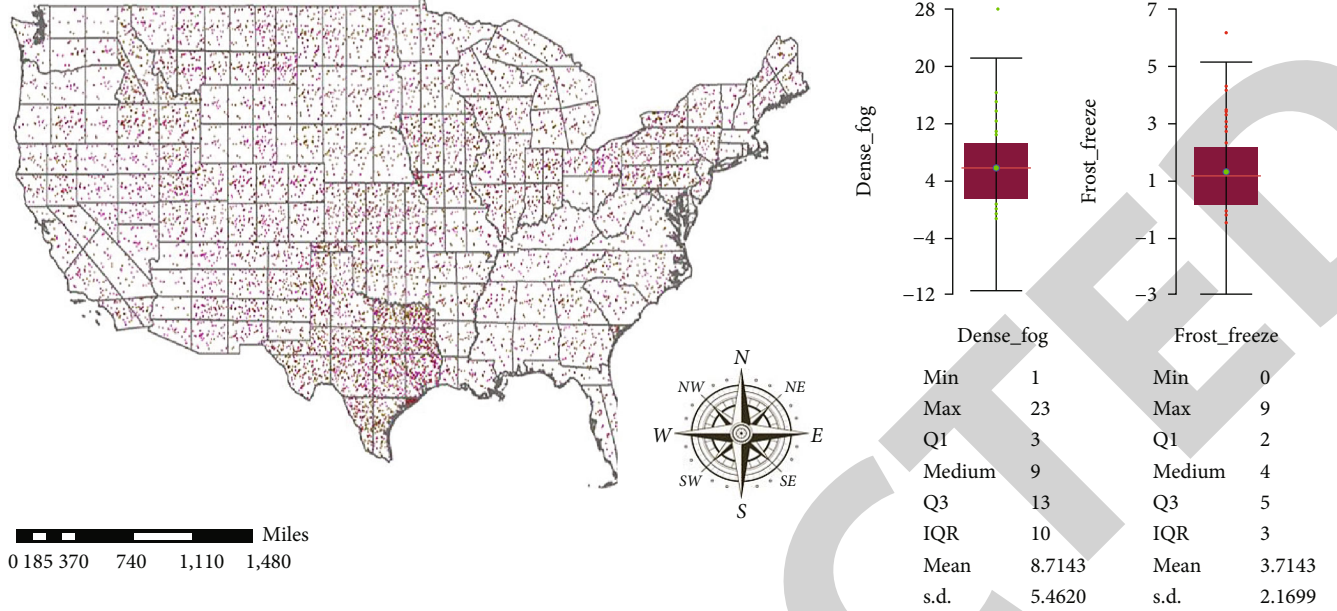


FIGURE 7: Dense fog and frost/freeze event correlation diagram.

TABLE 4: Geographically weighted regression model test results.

R^2	-2log-likelihood	A_{ICC}	Residual squares
0.8221	0.8491	0.9021	0.8743
-192.43	-187.65	-74.63	-117.65
-138.12	-137.82	-30.41	-42.84
762.84	692.29	721.62	965.62

and Maryland, where the frequency is higher, and the frequency of Texas is lower, the other states as a whole presents a relatively stable probability of occurrence.

Here, we divide the United States into five regions and perform descriptive statistics on the data to help illustrate the existence of certain spatial patterns in weather events. We select two typical weather events: flash floods and tornadoes and count their occurrences from 2010 to 2020, which is shown in Figure 8.

From the results of the descriptive statistics, it can be seen that the number of weather events changes with the time series, and the changing trends of regional weather events are different. In recent years, the frequency of flood disasters has increased in the east coast region, while the frequency in the central region has been almost unchanged. In contrast, tornado disasters have shown a relatively stable frequency of occurrence in recent years. Here, we should also note that different weather events are likely to show different spatial and temporal patterns in different regions and even globally, and not all-weather events are affected by climate change and become frequent.

6. Prediction of Frequency of Extreme Weather Events

6.1. Research Ideas. For this weather event, this article first uses an improved wavelet neural network model to predict

the total probability of the occurrence of various extreme weather events from 2010 to 2020. Perform this intelligent algorithm to replace the neurons in the traditional artificial neural network with wavelet elements based on wavelet analysis. Through mathematical transformation, the weights from the input layer to the hidden layer are transformed into new expansion parameters and the critical value of the hidden layer [26–29].

6.2. Research Method. The activation function of the hidden layer in the network diagram can be expressed as

$$g_j(x) = \prod_{i=1}^p \Psi \left(\frac{x_i - b_{ij}}{a_{ij}} \right). \quad (14)$$

Among them, Ψ represents the corresponding wavelet operation, X is the network input, i represents the different input wavelet elements in the network, and j is the network middle layer code and represents the new expansion and translation parameters after transformation. Therefore, the output function of the wavelet neural network can be expressed as

$$f_i(x) = \sum_{j=1}^h \omega_{ij} g_j(x). \quad (15)$$

In the formula, h is the number of levels of the wavelet network, and ω_{ij} is the output weight.

Although the functions and parameters of the traditional wavelet neural network are obtained after wavelet transformation, the transformation method is single and fixed, which cannot adapt to the complex and changeable conditions of weather events. At the same time, it is easy to cause

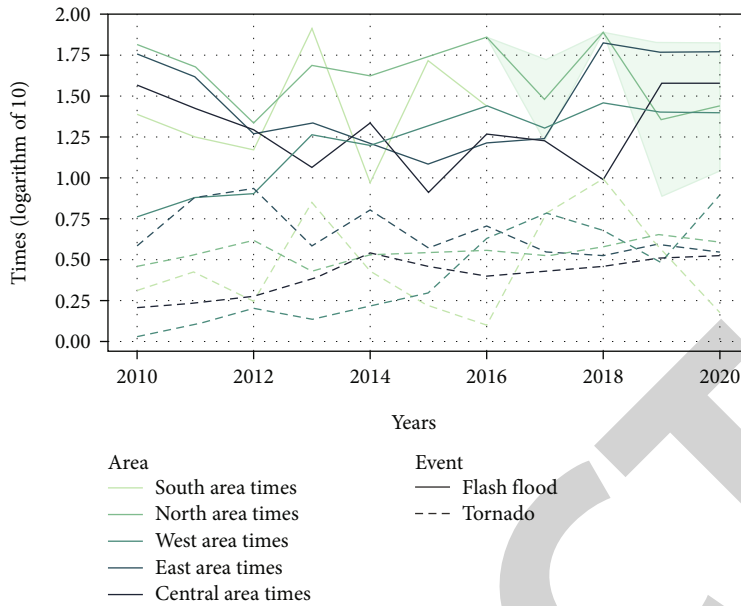


FIGURE 8: Frequency map of weather events in different regions over time.

the algorithm to reduce the approximation rate. Make random improvements:

To solve the problem of approximation rate, the improved excitation function and output function are, respectively,

$$\begin{cases} g'_j(x) = \prod_{i=1}^p \Psi, \left(\frac{\sum_{k=1}^n \omega_{jk} x_k(t) - b_j}{a_j} \right), \\ f'_i(x) = \sum_{j=1}^N \omega_{ij} \prod_{i=1}^p \Psi \left(\frac{\sum_{k=1}^n \omega_{jk} x_k(t) - b_j}{a_j} \right). \end{cases} \quad (16)$$

In the above formula, k represents the output layer code, N is the total number of wavelet elements, and p is the number of training sample spaces. According to the numerical characteristics, the expression of the error function can be obtained as

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{i=1}^N (d_i^p - y_i^p)^2. \quad (17)$$

In the above formula, d represents the mathematical expectation of the output layer, and y_i represents the actual network output value.

6.3. Result Analysis

Step 1. Determination of input and output. Through the descriptive analysis of the factors affecting the occurrence of extreme weather events, it is found that the frequency of extreme weather events was as follows.

Affected by precipitation factors, climatic factors, and personnel activity factors, it has a certain degree of randomness.

Therefore, the above three parameters are used as the input of the random wavelet network extreme weather event prediction model. In order to simplify the model, the year is the smallest unit, and the once-in-a-hundred-year event obtained in the first question is regarded as the most destructive weather event, that is, the extreme weather event. The ratio of the number of extreme weather events in the Americas to the number of all-weather events is taken as the probability of extreme weather events, and this probability is taken as the output of the random wavelet network extreme weather event probability prediction model.

Step 2. Hidden layer unit determination. In wavelet neural networks, the choice of the number of hidden layer units is also critical. The number of hidden units is too small, and the entire network cannot be well [30].

Information processing is as follows: too many hidden units will directly lead to structural redundancy and fall into local minimums. To balance the relationship between the two, the following formula is usually used to determine the number of hidden units of the wavelet neural network.

$$Z = \sqrt{mn + 1.68n + 0.93}. \quad (18)$$

Among them, Z is the number of hidden layer units, n is the number of network inputs, and m is the number of network outputs. Combining the number of inputs and outputs of the extreme weather event prediction model, substituting $n = 3$ and $m = 1$, $z = 2.99$ can be obtained. Therefore, the number of hidden layer units of the foundation pit settlement prediction model based on random wavelet network is set 3 which is appropriate.

This article uses MATLAB software to check the fitting and prediction results of the wavelet neural network model

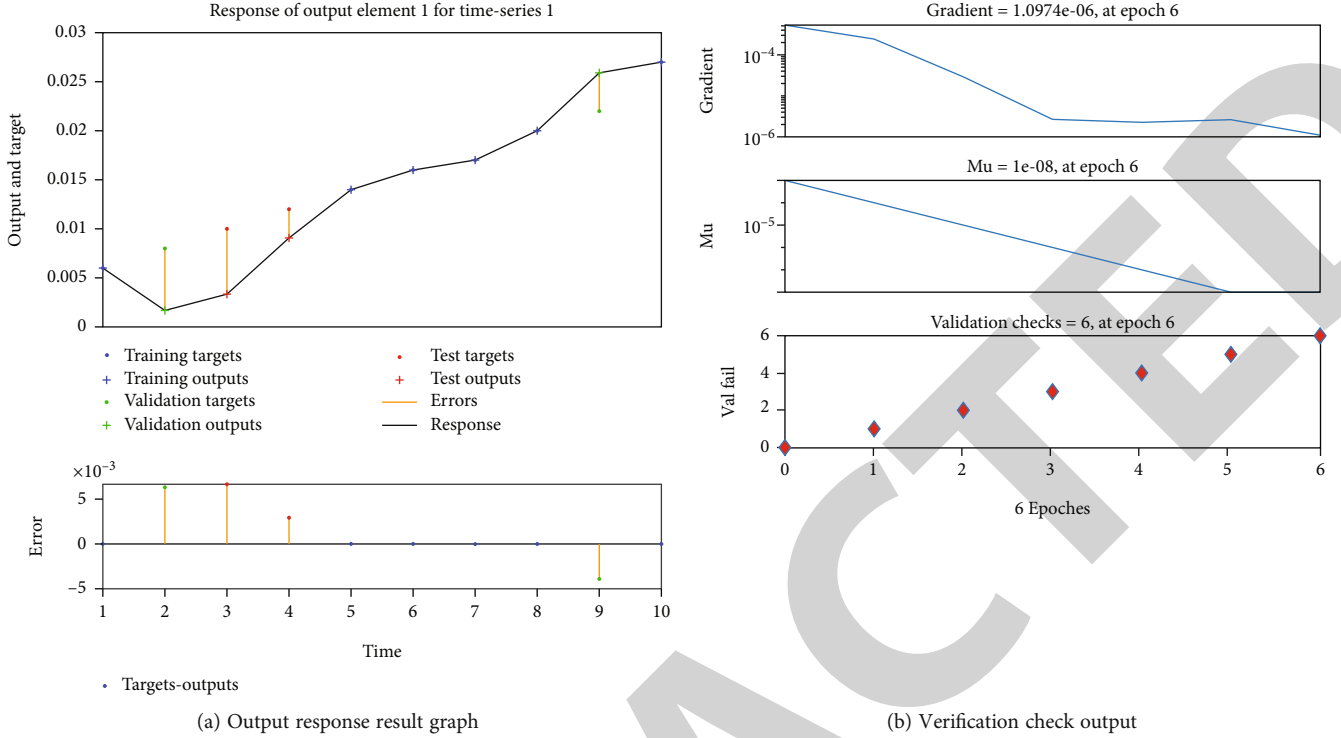


FIGURE 9: Model result graph.

[31–33]. The results can be seen in Figure 9. It can be seen from the results of validation checks that with the training of the network, the error of the confirmation sample has basically no longer been reduced, and it has been 4 times in a row. In the iteration, the error curve no longer drops, and the condition for the termination of training is generated at this time. From the fitting prediction results of the wavelet neural network model on the probability of extreme weather events, it can be concluded that the total probability of extreme weather events in the past ten years has shown a roughly rising trend, and from the predicted value, it is known that there is a high probability of extreme weather events in 2021. The probability of weather occurring in the entire Americas can reach around 0.016. The reason is that as forests and other vegetation have been destroyed on a large scale, the population has increased rapidly and is caused by global warming. Global warming means that the evaporation of water on the earth's surface increases, and a large amount of water vapor melts into the air, forming raindrops, rainstorms, and floods, and the temperature of the ground is getting higher and higher, which leads to droughts and sandstorms in some areas. The frequency and intensity of disasters such as droughts and floods will also increase. A wide range of extreme weather and climate events have severely affected life and production.

7. Evaluation and Spread of the Model

Although we analyzed the spatial laws of national weather events based on considering spatial heterogeneity, we still need to improve the model accuracy and outliers due to the shortcomings of the GWR model itself. For the accuracy

of the model, we should modify the fixed bandwidth in the model by using smaller bandwidths in regions with dense data points and larger bandwidths in regions with data point coefficients and by adding Bayesian information to the model measurements. Thus, the accuracy of the model will be further improved [34–36]. For outliers in the results, we should add the local Moran index, compare the global and local Moran index results, and eliminate outliers. The flow of the model improvement is shown in Figure 10.

In the process of using Bayesian inference methods to deal with sudden changes in extreme weather events, it is first necessary to calculate its posterior expectation with the following equation.

$$E(a|h) = \int_{\theta} a \cdot \theta \cdot P(\theta|h) d\theta, \quad (19)$$

where a is any distribution function assuming a model containing θ parameters, θ is the model parameters, and $E(a|h)$ is the posterior expectation of extreme events. The MCMC algorithm is used to carry out the solution problem of integration. The MCMC algorithm is also known as the Monte Carlo simulation algorithm of Markov chain. The formula of this algorithm is as follows.

$$E(a|h) \approx \frac{1}{N} \sum_{i=1}^N a(\theta^{(i)}). \quad (20)$$

In the formula, $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(N)}$, is the mean of a sample from a simulated Markov chain that obeys a priori

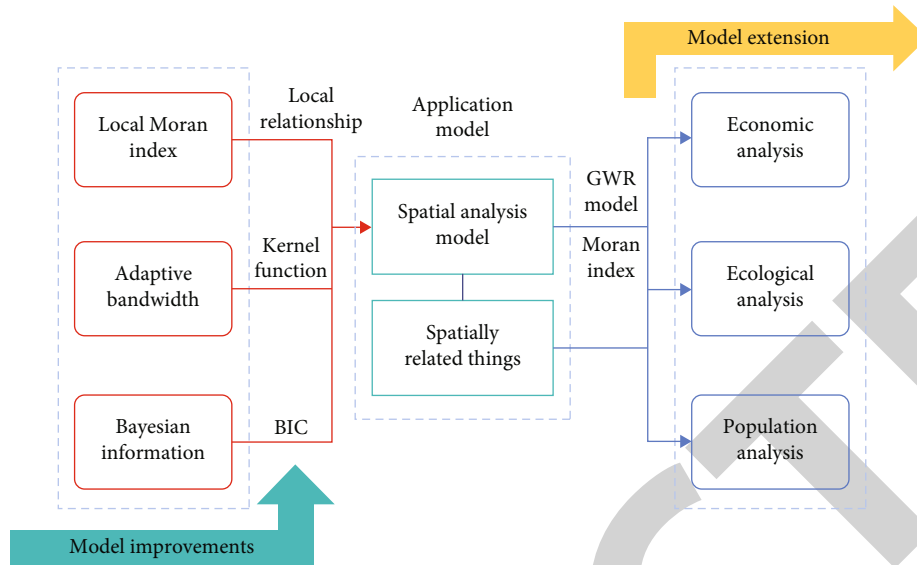


FIGURE 10: Model improvements and extensions.

probability distribution of $P(\theta|h)$ and whose value is an unbiased estimate.

The calculation of the complete reversible jump Markov Monte Carlo method BJM-CMC algorithm needs to be done by the following steps. First, the jump probability matrix between different model assumptions is determined $J(H_k \rightarrow H_{k'})$ and sampled from the simulated distribution of weather events $Q(u_k|H_k, H_{k'})$ to obtain u_k . Next, we set the transformation matrix $Q_{k'}, u_{k'} = g_{k,k'} \theta_k u_k$. Calculate the concession ratio m from H_k to $H_{k'}$. The probability that $H_{k'}$ is better than H_k is defined as h_{im1} , m . If h_{im1} , $m = 1$ then the jump to $H_{k'}$ is rejected. If h_{im1} , $m < 1$, then reject H_k and retain the original hypothesis $H_{k'}$, where the expression for m is as follows:

$$m = \frac{P(h|\theta_{k'}, H_{k'})P(\theta_{k'}|H_{k'})P(H_{k'})J(H_{k'} \rightarrow H_k)Q(u_k|\theta_{k'}, H_{k'}, H_k) \left| \frac{\partial g_{kk'}}{\partial(\theta_k, u_k)} \right|}{P(h|\theta_k, H_k)P(\theta_k|H_k)P(H_k)J(H_k \rightarrow H_{k'})Q(u_k|\theta_k, H_k, H_{k'})} \quad (21)$$

The spatial model of this paper, in addition to statistics and prediction of the spatial laws of weather events in the country, can also predict the property losses, casualties, and so on indirectly caused by disasters. If we need to make the statistical results more accurate, we can choose a smaller spatial unit, such as selecting counties as the basic spatial units, then we can analyze the spatial regularity of weather events in a certain region or state. Further, we can change the research object and choose all things that may have spatial laws, such as economy, ecology, and population, and analyzing the spatial laws of such things will be of great significance to the scientific development of society [33].

8. Conclusion

To build a weather event frequency analysis model. In this paper, four edge distributions of gamma, log-normal, GEV, and exponential are selected for fitting, and it is concluded that

the Gumbel Copula has the best fitting effect. Second, the maximum likelihood method was used to estimate the marginal distribution parameters, and the K -S test method was used to verify the fit. Then, based on the selected Gumbel Copula function, the three-variable joint distribution of weather events is established, and the joint co-occurrence regression period and the three-variable joint distribution regression period under different univariate regression periods are calculated, respectively. Finally, the SVM model is used to classify the obtained results to better present the data features.

The combination of the Moran index and the geographically weighted regression model used in this article can well predict the indicators and data with spatial distribution characteristics under the premise of considering spatial heterogeneity and nonstationarity. Studies have found that the United States has the highest frequency of thunderstorms and flash floods, showing a spatial trend of increasing from west to east and from north to south. Finally, through the establishment of an ARIMA model to predict the frequency of weather events, it is found that the total number of thunderstorms is on the rise. It can be seen that as the years go by, thunderstorms have become more frequent. Torrent data is close to stable. The frequency of blizzards, dense fog, and frost/icing has generally increased.

Then, a wavelet neural network model is established to predict the probability of extreme weather events across the Americas. It is concluded that the total probability of extreme weather events in the past ten years shows a roughly rising trend, and from the predicted value, it can be known that by 2021, the probability of extreme weather events in the entire Americas can reach around 0.016.

Data Availability

Data for this paper were obtained from the National Oceanic and Atmospheric Administration's SPC report and NOAA's National Weather Service input for the period January 1950 to October 2021.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Peng-Hui Yang contributed to the methodology, conceptualization, supervision, and leadership. Yao Yu contributed to the conceptualization, visualization, software, validation, and writing manuscript. Feng Gu contributed to the data collation, visualization, verification, and investigation. Meng-Jie Qu contributed to the software, method design, validation, and data analysis. Jia-Ming Zhu contributed to the verification, supervision, and writing review and editing. All authors read and approved the final manuscript.

Acknowledgments

This study was funded by the National Social Science Fund Project of China (21CTJ024), the Teaching and Research Fund Project of the Anhui University of Finance and Economics (acxkjs2021005 and acyljc2021002), Anhui Quality Engineering Project Teaching Demonstration Course “mathematical modeling” (2020SJJXSF0018), and Provincial Online and Offline First-Class Course “Advanced Algebra” (2020xsxxkc018).

References

- [1] G. Samantha, “The impact of natural disasters on micro, small and medium enterprises (MSMEs): a case study on 2016 flood event in Western Sri Lanka,” *Procedia Engineering*, vol. 212, pp. 744–751, 2018.
- [2] J. Fang and P. Shi, “A review of coastal flood risk research under global climate change,” *Progress in Geography*, vol. 38, no. 5, pp. 625–636, 2019.
- [3] N. Bhattacharya, J. Elizabeth, D. Proverbs, and F. Hammond, “Development of conceptual framework for understanding vulnerability of commercial property values towards flooding,” *International Journal of Disaster Resilience in the Built Environment*, vol. 4, no. 3, pp. 334–351, 2013.
- [4] Z. Chi, M. Jaboyedoff, M.-H. Derron, and C. J. Van, “An interactive web-GIS tool for risk analysis: a case study in the Fella River basin, Italy,” *Natural Hazards and Earth System Sciences Discussions*, vol. 16, no. 1, pp. 85–101, 2016.
- [5] M. Haggag, A. Siam, W. El-Dakhkhni, P. Counlibaly, and E. Hassini, “A deep learning model for predicting climate-induced disasters,” *Natural Hazards*, vol. 107, no. 1, pp. 1009–1034, 2021.
- [6] Z. Zhang, C. Gao, Q. Liu et al., “Risk assessment of rainstorm and flood disasters in the Huaihe River basin in different recurrence periods,” *Geographical Research*, vol. 33, no. 7, pp. 1361–1372, 2014.
- [7] G. F. Zhang, X. C. Zha, and G. P. Wang, “Risk assessment of flood disasters in different return periods along the Ankang reach of the Hanjiang River,” *Journal of Lanzhou University (Natural Sciences)*, vol. 55, no. 5, pp. 571–577, 2019.
- [8] X. Yang and Q. Liu, “Forecast of economic loss from storm surge disaster based on KPCA-RBF model,” *Marine Sciences*, vol. 45, no. 10, pp. 32–39, 2021.
- [9] Y. Li and R. Hu, “Research on Heilongjiang Province heavy rain event disaster assessment and pre-evaluation model based on grey correlation,” *Journal of Catastrophe*, vol. 31, no. 2, pp. 78–83, 2016.
- [10] Y. H. Huang, Y. H. Zhu, Y. Hu, and X. L. Zhou, “Time series prediction based on wavelet analysis,” *Smart City*, vol. 7, no. 18, pp. 128–130, 2021.
- [11] M. Gui and L. Liu, “Prediction of the total retail sales of social consumer goods in Hainan Province based on the ARIMA model,” *Mathematics in Practice and Understanding*, vol. 47, no. 3, pp. 25–30, 2017.
- [12] G. Cheng, Y. Zhang, J. L. Huang, J. B. Liu, J. W. Zhang, and S. Zhang, “Vulnerability analysis of aqueduct structure based on boundary method,” *Earth and Environmental Science*, vol. 567, no. 1, article 012036, 2020.
- [13] P. Riley, B. Dan, Y. Liu, P. Verronen, H. Singer, and M. Güdel, “Extreme space weather events: from cradle to grave,” *The Scientific Foundation of Space Weather*, vol. 214, no. 1, pp. 7–9, 2018.
- [14] J. Huang, J. Jiang, J. Wang, and L. Hou, “Crop diversification in coping with extreme weather events in China,” *Journal of Integrative Agriculture*, vol. 13, no. 4, pp. 677–686, 2014.
- [15] P. Stott, N. Christidis, F. Otto et al., “Attribution of extreme weather and climate-related events,” *Climate Change*, vol. 7, no. 1, pp. 23–41, 2016.
- [16] N. S. Diffenbaugh, D. Singh, J. S. Mankin et al., “Quantifying the influence of global warming on unprecedented extreme climate events,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 19, pp. 4881–4886, 2017.
- [17] M. D. Smith, “An ecological perspective on extreme climatic events: a synthetic definition and framework to guide future research,” *Journal of Ecology*, vol. 99, no. 3, pp. 656–663, 2011.
- [18] W. Cai, A. Santoso, G. Wang et al., “Increased frequency of extreme Indian Ocean Dipole events due to greenhouse warming,” *Nature*, vol. 510, no. 7504, pp. 254–258, 2014.
- [19] N. C. Duke, J. M. Kovacs, A. D. Griffiths et al., “Large-scale dieback of mangroves in Australia,” *Marine and Freshwater Research*, vol. 68, no. 10, pp. 1816–1829, 2017.
- [20] M. Trnka, R. P. Rötter, M. Ruiz-Ramos et al., “Adverse weather conditions for European wheat production will become more frequent with climate change,” *Nature Climate Change*, vol. 4, no. 7, pp. 637–643, 2014.
- [21] T. Oshan, Z. Li, W. Kang, L. J. Wolf, and A. Fotheringham, “MGWR: a Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale,” *International Journal of Geo-Information*, vol. 8, no. 6, pp. 269–300, 2019.
- [22] J. Tu, “Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression,” *Applied Geography*, vol. 31, no. 1, pp. 376–392, 2011.
- [23] S. Li, C. Zhou, S. Wang, S. Gao, and Z. Liu, “Spatial heterogeneity in the determinants of urban form: an analysis of Chinese cities with a GWR approach,” *Sustainability*, vol. 11, no. 2, pp. 479–495, 2019.
- [24] A. F. Jan and H. Chan, “A wavelet neural network conjunction model for groundwater level forecasting,” *Journal of Hydrology*, vol. 407, no. 1–4, pp. 28–40, 2011.
- [25] A. Perera, V. M. Nik, D. Chen, J. L. Scartezzini, and T. Hong, “Quantifying the impacts of climate change and extreme climate events on energy systems,” *Nature Energy*, vol. 5, no. 2, pp. 150–159, 2020.

Retraction

Retracted: Optimization Calculation Method and Mathematical Modeling of Big Data Chaotic Model Based on Improved Genetic Algorithm

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Zhang and Y. Zhang, "Optimization Calculation Method and Mathematical Modeling of Big Data Chaotic Model Based on Improved Genetic Algorithm," *Journal of Function Spaces*, vol. 2022, Article ID 6983242, 9 pages, 2022.

Research Article

Optimization Calculation Method and Mathematical Modeling of Big Data Chaotic Model Based on Improved Genetic Algorithm

Zhicheng Zhang ¹ and Yan Zhang ²

¹School of Science, Henan Institute of Technology, Xinxiang, Henan 453003, China

²College of Computer and Information Engineering, Henan Normal University, Xinxiang, Henan 453007, China

Correspondence should be addressed to Zhicheng Zhang; zhangzc@hait.edu.cn

Received 23 April 2022; Revised 14 June 2022; Accepted 6 July 2022; Published 20 July 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Zhicheng Zhang and Yan 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.

In order to find a chaotic trajectory sequence with strong global optimization ability to help the genetic selection of direction after the reversal of chemotaxis, an improved genetic algorithm based on chaos optimization is proposed by combining the characteristics of chaotic motion with the improved genetic algorithm. The optimal coverage problem in sensor networks can carry out fine optimization search on local areas. The results show that the overall trend of fitness and optimization efficiency is relatively stable. The optimization efficiency will be gradually improved with the continuous progress of time and genetics, and the error analysis will be reduced. This will greatly improve the impact of various adverse factors in the optimization process. In addition, the change rate of fitness is basically kept at a high change rate, which also reflects that the basic framework of the model is very excellent, and the whole algorithm structure and data processing are improved by 54%. The improved genetic algorithm proposed in this paper is used to adjust and optimize the controller parameters. When the uncertain parameters change greatly, the control system still has good control quality and strong robustness.

1. Introduction

For big data, efficient optimization calculation process is very important. The optimization process cannot be represented by any mathematical conditions [1]. With the development of computer technology, information technology, and system engineering technology, optimization technology has become an important branch of engineering and an important subject in applied mathematics [2]. Training data with sufficient data, proper distribution, and excellent performance is the premise for the above methods to successfully optimize fuzzy controllers, but the main training data are not always available [3]. Considering that the GA is suitable for multiparameter optimization, it does not need to know the local information of the object to be optimized, nor does it need good training data, but in practical control problems, uncertainty is ubiquitous [4]. The uncertainty may come from the modeling error of the described control object, or from the disturbance signal of the control system itself and the outside [5]. GA is a kind of global optimal

probability search method, which is based on the evolutionary law of biology and evolved from the genetic mechanism of survival of the fittest.

The basic idea of GA starts from a population that represents the possible potential solutions of the problem, and a population is composed of a certain number of individuals encoded by genes, and each individual is actually an entity with characteristics of chromosomes [6]. Then, the genetic evolution process of these groups is simulated, and the evolution direction of the algorithm is known by the fitness value of the individual [7]. Although this method has solved the above problems well, the interpretability of the system is still not guaranteed [8]. The reason is that the goal of these algorithms is to obtain the optimal system response performance, but there is no effective guidance for the number of fuzzy set partitions and the selection of membership function parameters [9]. No matter which field the optimization technology is applied to, the optimization problem to be solved can generally be described in mathematical language, that is, to establish the corresponding mathematical model

[10]. Linear programming is one of the important fields in optimization problems. The linear optimization algorithm is applied to the optimal power flow problem of the power system. Usually, the whole problem is decomposed into two suboptimization problems of the active part and the reactive part and then iteratively solves them alternately or solves them separately. One of the more popular methods is the fusion of fuzzy logic and GA, resulting in genetic fuzzy algorithm.

Conventional genetic fuzzy system is to add a learning process based on evolutionary computation to a fuzzy system, and this evolutionary computation can be GA, genetic programming, or other evolutionary algorithms, and the system formed by it is called genetic optimization system [11]. This adaptive optimization controller is based on the conventional optimization controller, uses the neural network model to approach the actual controlled object, and uses the GA to continuously optimize the control rules of the controller online, even if these rules can keep up with changes. It can improve the control ability and control efficiency of the optimized controller [12]. Traditional GAs usually use time domain index, frequency domain index, error integral index, and their combination as the objective function. The optimization model is based on the design of nominal controlled object and does not consider the reality of parameter uncertainty of controlled object. Therefore, it is often difficult for the system to meet the requirements when the parameters of controlled object change [13]. In this paper, an optimization calculation method of big data chaotic model based on improved GA is proposed. The experimental results are analyzed and processed. After genetic operation of the original parameters, the error and performance indices of the parameters are set as important parameters, and the minimum objective function is introduced. After screening, the range data is obtained, and then, the optimal index and fitness function are calculated.

The innovative contribution of this paper is that the improved genetic algorithm is proposed to adjust and optimize the controller parameters. When the uncertain parameters change greatly, the control system still has good control quality and strong robustness. In addition, the algorithm does not destroy the structure and essence of traditional genetic algorithm, so it inherits the advantages of genetic algorithm and takes into account the advantages of random algorithm. In the improved genetic algorithm, the individuals in the new generation population only calculate the individuals involved in crossover and mutation operations, which shortens the calculation time to a certain extent. A chaotic trajectory sequence with strong global optimization ability is found to help reverse chemotaxis after genetic selection.

The research is divided into four sections. The first section describes the learning process of traditional genetic fuzzy system calculation and the fusion background of fuzzy logic and genetic algorithm. Section 2 describes the materials and methods. The analysis and research contents of genetic algorithm are improved. The chaos model based on big data is analyzed. The general analysis and mathematical modeling of the design are optimized. In the mechanism part of

optimization design, the establishment of mathematical model and the application of simulation algorithm are analyzed. The results are analyzed and discussed in Section 3. Finally, the full text is summarized. The results show that in the improved genetic algorithm, the individuals in the new generation population only calculate the individuals involved in crossover and mutation operations, which shortens the calculation time to a certain extent.

2. Materials and Methods

2.1. Analysis and Research of Improved GA. GA is a computational model that simulates the natural selection and population genetic mechanism of biological evolution [14]. Genetic operation in GA is based on coding mechanism. Coding has a great impact on the performance of the algorithm, such as search ability and population diversity. As a new global optimization algorithm, it has almost no restrictions on the optimization problems to be solved and does not need to involve the complex and cumbersome mathematical solution process like the conventional optimization algorithm [15]. According to different examples, it is only necessary to properly adjust the operator parameters and make minor modifications to adapt to new problems, and the program can be universal. The mapping system is very sensitive to the initial values and parameters, and the subtle differences between the initial values and parameters may make the final chaotic random sequences very different. The distribution diagram of the trajectory sequence generated by studying the trajectory of the chaotic motion is visually represented, and the random sequence of chaos is added at the beginning of the inheritance, which improves the search diversity of the genetic iteration and the ability to optimize the global region. GA does not require continuous variables, which is also convenient for dealing with discrete variables. At the same time, it can introduce various constraints, generate new solutions by random search, eliminate solutions with poor performance, avoid local optimal solutions by mutation, and search the global solution space to achieve global optimization or approximate global optimization. Figure 1 is the basic operation flow chart of GA.

To some extent, the quality of fitness function determines the search range and global optimization ability of GA. The traditional GA takes a single objective function as the fitness function, so it is not comprehensive enough to consider the problem, and sometimes, it is impossible to find the optimal solution [16]. Generally, GA basically includes four main operations: chromosome encoding, population setting, calculation of fitness function value, and genetic operation. The quality of the coding method is very important to the GA. A good coding method may make the genetic operations such as crossover operation and mutation operation simply run [17]. A poor coding method may make genetic operation difficult to achieve. Therefore, this paper believes that the coding method largely determines how to carry out the genetic evolution operation of the population and its optimization efficiency [18]. Assuming that the value range of a parameter is $[U_{\min}, U_{\max}]$ and the parameter is

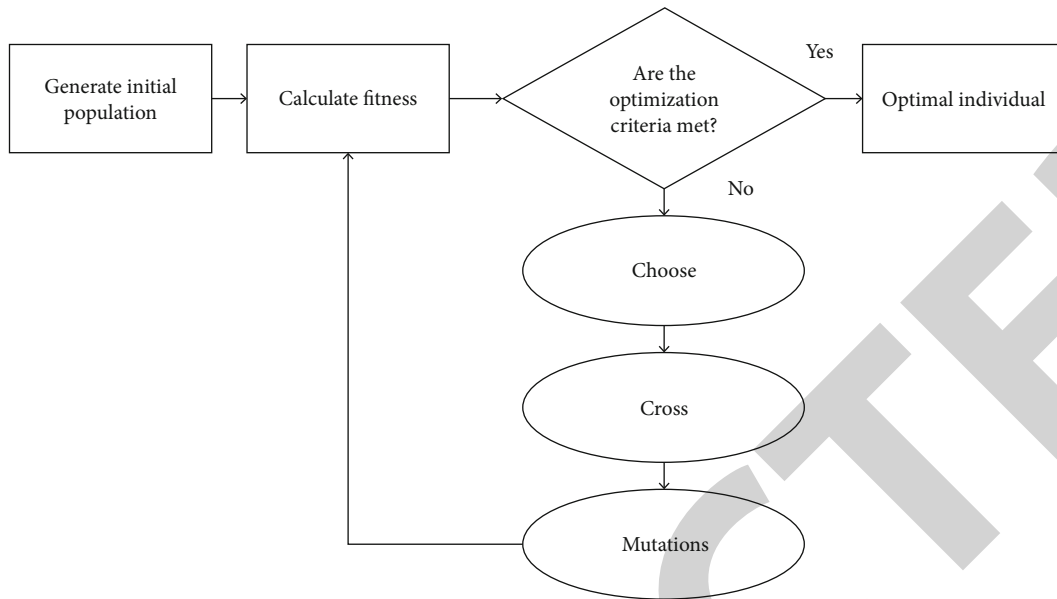


FIGURE 1: The basic operation flow chart of the GA.

represented by a string of binary coded symbols with length l , the coding accuracy of binary coding is as follows:

$$\delta = \frac{U_{\max} - U_{\min}}{2^l - 1}. \quad (1)$$

Assuming that the code of an individual is $X : b_l b_{l-1} \dots b_2 b_1$, the corresponding decoding formula is

$$x = U_{\min} + \left(\sum_{l=1}^l b_l \cdot 2^{l-1} \right) \cdot \frac{U_{\max} - U_{\min}}{2^l - 1}. \quad (2)$$

The main processing data are shown in Table 1.

At present, there are usually two methods to generate the initial population. One is generated according to the completely random method, which is applicable to the case where there is no prior knowledge of the solution of the problem [19]. Another method is to transform some prior knowledge into a set of requirements that must be met and then randomly select samples from the solutions that meet these requirements. By selecting the initial population, the GA can avoid premature convergence and find the optimal solution faster.

Since the GA basically does not participate in the analysis of external information, when the fitness function is used as the judgment basis, the required value is nonnegative [20]. Therefore, the objective function to be solved is directly converted into a fitness function:

$$\text{Fit}(f(x)) = f(x). \quad (3)$$

Maximum value when the objective function is negative:

$$\text{Fit}(f(x)) = -f(x). \quad (4)$$

Minimum value when the objective function is negative:

$$\text{Fit}(f(x)) = \begin{cases} f(x) - c_{\min}, & f(x) > c_{\min}, \\ 0. & \end{cases} \quad (5)$$

c_{\max}, c_{\min} in the above formula is the maximum and minimum estimates of $f(x)$, respectively.

2.2. Chaos Model Based on Big Data. The calculation of this parameter is an important step to extract its chaotic characteristics. Differential action reflects the rate of change of the system deviation signal and has predictability, which can predict the trend of deviation change, so it has a superprevious control effect. It can reflect the change rate of system deviation signal through differential action, which is predictive, can predict the trend of deviation change, and can carry out advance control and improve the dynamic performance of the system. It can reduce overshoot and adjustment time. The Lyapunov exponent can describe the chaotic intensity of big data, which has various description forms [21]. This paper adopts the following description form: set β_{n_1}, β_{n_2} as two points close to the limit in space, and then, its distance is expressed as $\beta_{n_1} - \beta_{n_2} = \delta \leq 1$. After Δn time, the trajectory of the two points can be expressed as $\delta_{\Delta n} = \beta_{n_1 + \Delta n} - \beta_{n_2 + \Delta n}$ [22]. Then, the maximum Lyapunov exponent can be described as

$$\delta_{\Delta n} = \delta_{0e}^{\lambda \Delta n}. \quad (6)$$

If λ is set to a positive number, then the parameters between the ringing tracks are separated, which means chaos. However, because the two tracks are generally close to each other, the above formula is only valid when the distance value is relatively small. If the distance is large, the separation of the tracks will be greatly reduced [23].

TABLE 1: Computational comparison of main models.

Model	Improved GA model	Big data chaotic model	Optimization model
Objective function 1	0.021	0.012	0.047
Objective function 2	0.024	0.032	0.365
Objective function 3	0.024	0.014	0.235
Objective function 4	0.025	0.045	0.333
Objective function 5	0.036	0.022	0.452
Objective function 6	0.022	0.036	0.324

Therefore, the GA is improved on this basis, and the flow chart of the improved GA as shown in Figure 2 is obtained.

2.3. General Analysis and Mathematical Modeling of Optimization Design

2.3.1. Mechanism and Analysis of Optimization Design. This paper designs the optimization calculation model under the framework of big data chaos model. For the traditional deterministic system, the inertial motion produced by it will show some observable characteristics and eventually return to static. When the external system is excited by a deterministic rule, the response of the system fed back to the outside world should also be deterministic [24]. However, for chaotic systems, this phenomenon is not true, and it may produce unpredictable, irregular, and never-repeated chaotic phenomena after being stimulated by deterministic rules [25]. In the process of using GA to optimize fuzzy control rules, many papers use the method of randomly generating initial populations, which will generate many populations that are not reasonable and do not conform to the process control experience, thus increasing the optimization of GA. The algebra reduces its convergence speed and is not conducive to us finding the optimal solution. Chaos models are generally divided into continuous and discrete models, as shown in Figure 3 for a schematic diagram of the discrete model.

The continuous model is mainly expressed as follows:

$$\begin{cases} \frac{d_x}{d_\tau} = a(y - x), \\ \frac{d_y}{d_\tau} = cx - xz - y, \\ \frac{d_z}{d_\tau} = xy - bz. \end{cases} \quad (7)$$

Discrete model uses difference equation to describe discrete chaotic system, which can generate discrete chaotic signal in time domain from an initial value and mapping formula and finally form digital chaotic sequence without the value of each sequence point, which is generally expressed in the form of nonlinear difference variance:

$$Z_{i+1} = \rho Z_i (1 - Z_i). \quad (8)$$

For $\rho \in (0, 4]$, $Z_i \in [0, 1]$, the above formula will not be directly used and then improved:

$$Z_{i+1} = 1 - \rho Z_i^2, \quad (9)$$

in which $\rho \in (0, 2)$ and $Z_i \in [-1, 1]$. The modified formula has the effect of simplifying calculation and has good practical significance in practical operation.

2.3.2. Construction of Mathematical Model and Application of Simulation Algorithm. Since the above parameters selected in this paper belong to the basic range, the overall design of the optimal algorithm is required in the specific mathematical model design, so that the subsequent experimental simulation can be registered to achieve the purpose [26]. Chaotic motion exists widely in nonlinear systems, it has excellent properties such as ergodicity and randomness, and it can go to each state in a certain domain irreproducibly; this feature makes chaotic search inherently random and orbital history. It can ensure that all possible states can be traversed without repetition in the global scope, which is conducive to overcome the limitations of the general random algorithm with distributed traversal as the search mechanism, which will be more conducive to the optimization of the chaotic model of byte big data, and solve the problem [27]. In order to overcome the convergence problem of the basic GA, this paper also adopts the elite selection strategy. That is, after crossing and mutation, the optimal individual in the new generation population is generated and compared with the optimal individual of the previous generation. If the former is less than the latter, the poor individual of this generation will be replaced with the optimal individual of the previous generation; otherwise, no operation will be carried out. It realizes the communication and cooperation between individual information and group information by simulating various excellent characteristics of different biological groups, so that the algorithm can constantly revise its own optimization conditions through the interactive information between individuals, in order to find the optimal solution or approximate optimal solution of the optimization algorithm, and at the same time, it also improves the search accuracy of the intelligent algorithm. First of all, on the initialization population, an evaluation needs to be performed. At this time, the average error, mean square error, and determination coefficient R^2 need to be used

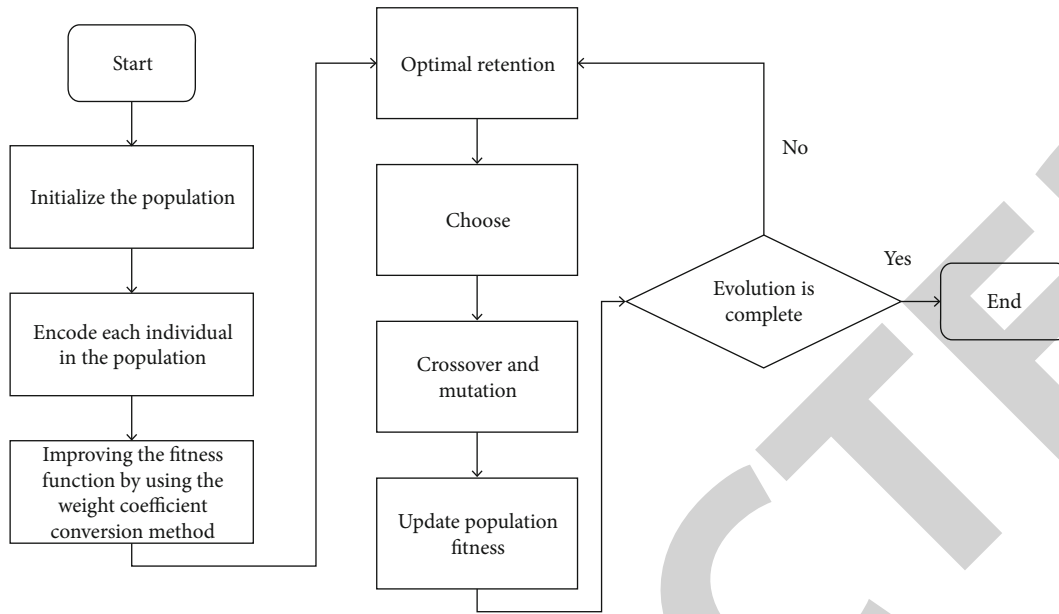


FIGURE 2: The basic flow chart of the improved GA.

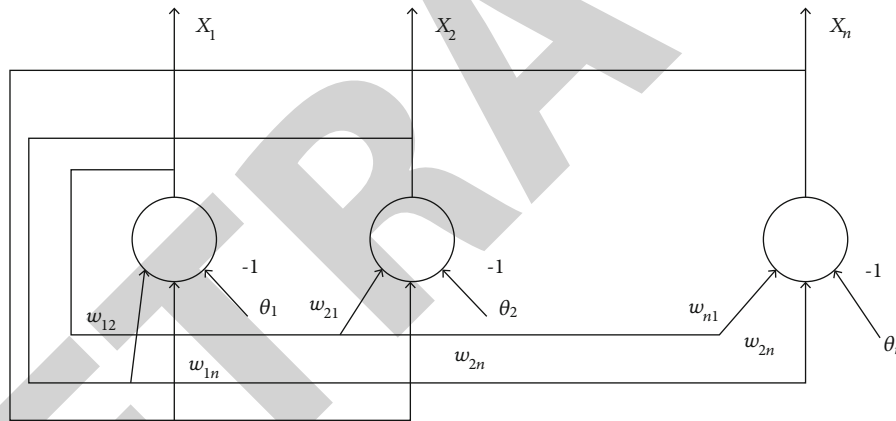


FIGURE 3: Schematic diagram of discrete model.

as the objective function. The specific calculation formula is as follows:

$$\begin{aligned}
 \text{MAE}(\hat{y}_i, y_i) &= \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|, \\
 \text{MSE}(\hat{y}_i, y_i) &= \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2, \\
 R^2(\hat{y}_i, y_i) &= 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2},
 \end{aligned} \tag{10}$$

where \hat{y}_i is the predicted value, y_i is the actual value, and \bar{y}_i is the center value of y_i . After comprehensively evaluating the whole initial population, the above formulas are cited as the multiobjective function, and the weight coefficient is added to the calculation to give weight to each

objective function. After getting the fitness of an individual, it is necessary to start screening the probability of being selected and determine whether the individual can inherit the next generation according to this item. Therefore, assuming that the fitness of individual i is the above objective function value f_i and the population size is PS, the probability of the individual i being selected is

$$P_i = \frac{f_i}{\sum_{i=1}^{PS} f_i}. \tag{11}$$

The last step is the crossover mutation process, in which the purpose of crossover operation is to retain genes with excellent traits, while the purpose of mutation operation is to increase the diversity of genes and improve the probability of finding the global optimal solution in the solution space. The range of variation probability used in this paper is between thousandths and percentiles. In

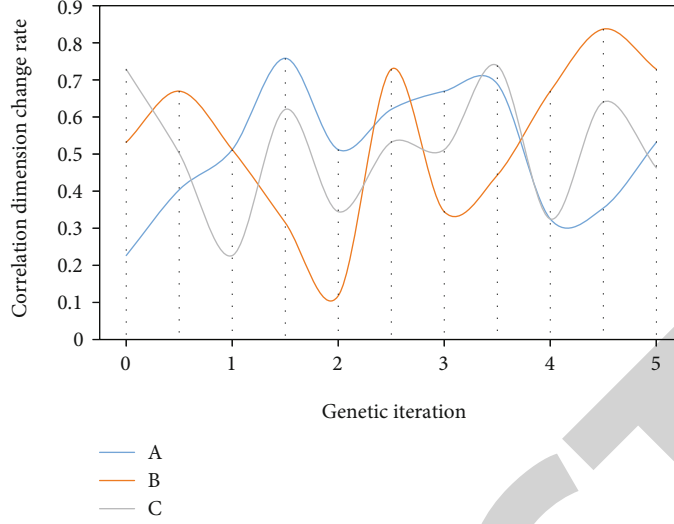


FIGURE 4: Genetic analysis of association dimension.

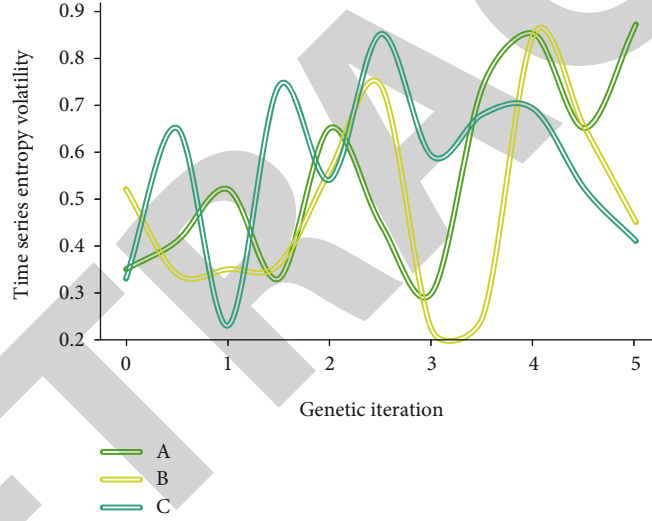


FIGURE 5: Genetic analysis of the volatility of time series entropy.

order to avoid deleting individuals with high fitness value, this paper stipulates that individuals with high fitness value in each generation do not need genetic operation and can directly enter the next generation, so as to improve the overall operation efficiency. Therefore, in the basic algorithm framework, the optimal selection is very effective.

Assuming a subset S of the initial rule set, the number of rules it contains is $N(S)$, and the classification accuracy is $E(S)$. Then, the simplified objectives can be considered as $N(S)$ and $E(S)$. This is a two-objective combinatorial optimization problem. Introducing the weight $0 < \omega < 1$, the fitness S of the rule set $f(S)$ is defined as

$$f(S) = \begin{cases} \omega \frac{E(S)}{E_0} + (1 - \omega) \frac{N(S)}{N_0}, & E_0 \neq 0, \\ \omega E(S) + (1 - \omega) \frac{N(S)}{N_0}, & E_0 = 0, \end{cases} \quad (12)$$

TABLE 2: Comparison results of the three parameters with traditional methods.

Parameter	Method	Mean and standard deviation
Associative dimension	Traditional method	7.12 ± 0.32
	The method of this paper	7.12 ± 0.112
Lyapunov index	Traditional method	210 ± 0
	The method of this paper	210 ± 0
Time series entropy	Traditional method	3 ± 0
	The method of this paper	1.36 ± 8.745

where E_0 is the classification accuracy using the initial rule base and N_0 is the number of rules contained in the initial rule base. The above fitness function is used to evaluate the

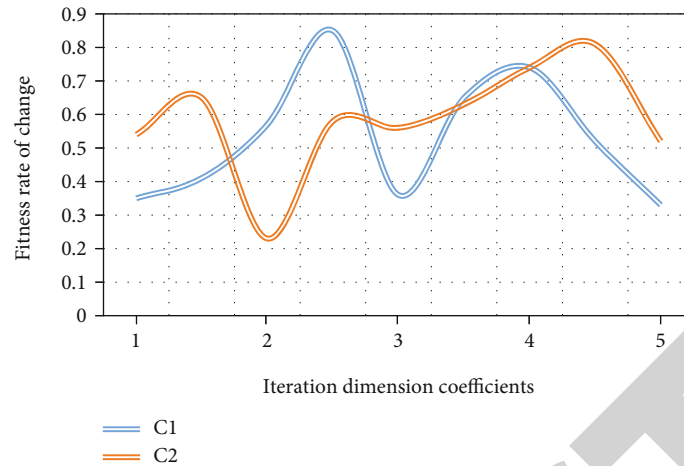


FIGURE 6: Iterative relationship diagram of fitness change rate.

individuals in the contemporary population. If the initial population is set as the initial matrix, in order to ensure the correlation and accuracy of each inheritance, in the process of optimal selection of circular inheritance, this paper makes a unified division in the objective function. Almost all machine learning algorithms come down to solving the optimization problem in order to achieve the goal we want the algorithm to achieve. In order to achieve a certain goal, it is necessary to construct an objective function, and then, let the function take the maximum or minimum value (that is, optimization), so as to obtain the model parameters of the machine learning algorithm. Constructing a reasonable objective function is the key to establish machine learning algorithm.

$$\min_{H \in \mathfrak{S}_\mu} \|G - H\|_F^2 \quad (13)$$

where H is the update matrix and G is the fixed matrix. After the above approximate problem of the initial matrix is calculated, a convex set projection can be obtained. At this time, it is willing to replace the off-diagonal elements of the matrix:

$$h_{ij} = \begin{cases} g_{ij}, & |g_{ij}| \leq \mu, \\ \mu \cdot \text{sign}(g_{ij}), & |g_{ij}| > \mu. \end{cases} \quad (14)$$

According to the above calculation, the elements can be arranged one by one in order, which can ensure that the matrix can be well iterated in the next update, and all the nondiagonal elements in the matrix that are greater than the threshold will be constrained. In this way, the optimal item can be obtained in the sequence. Of course, because interference items or coordinate items sometimes appear in genetics, it is necessary to pay attention to the different transformation of genetic matrix in actual calculation. Originally, the whole population evolved and decomposed into independent subpopulations. Each subpopulation evolves independently and cooperates to obtain the optimization target value. Because the dimension of each suboptimization

problem is reduced, the burden on the subpopulation is reduced, and the optimization efficiency is improved.

3. Result Analysis and Discussion

In order to verify that the model designed in this paper has the characteristics of scientific, feasible, and efficient, this paper designs some relevant experiments again for analysis. Based on the characteristics of the chaotic model under big data, the improved GA can obtain the accurate global optimal value on all parameters. At the same time, the optimization performance is extremely stable. The experiments designed for this purpose also verify the practical operability of the model from the aspects of correlation dimension, time series entropy, fitness change rate, and optimization efficiency. Figures 4 and 5 are the analysis diagrams of correlation dimension and time series entropy under three different heritages: A, B, and C.

Each subpopulation evolves independently and cooperates to obtain the optimization target value. As the dimension of each suboptimization problem is reduced, the burden on the subpopulation is reduced, and the optimization efficiency can be seen from the analysis of the above figure. It can be seen from the analysis of the above figure that the trend is unstable in three different heredity, whether in the correlation dimension or in the entropy of time series, which also directly shows that there will be great differences in the search for the optimal term in different heredity due to different genetic methods, genetic algebra, and other factors. Table 2 is the result data of correlation dimension, Lyapunov exponent, and time series entropy compared with traditional methods.

To improve, in terms of correlation dimension, if the correlation dimension is large, the overall trend will tend to be stable. This is because the amount and accuracy of correlation will be greatly enhanced when the algebra increases. The richness will also reduce errors. On the entropy of time series, because each matrix is different for time inheritance and there are too many uncertainties on the time axis, the objective function Jin Xu should be discussed and distinguished in concrete operation, and the algorithm designed

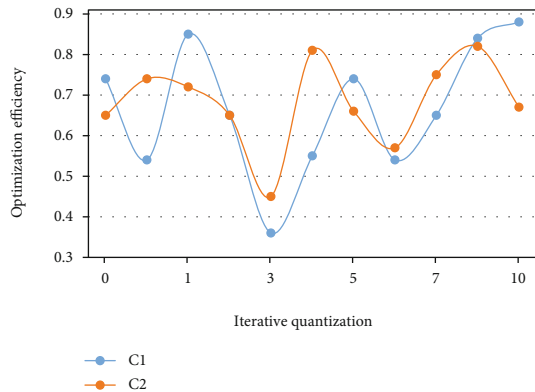


FIGURE 7: Optimization efficiency analysis diagram.

in this paper will optimize and deal with this problem. Figures 6 and 7 are the analysis figures of fitness change rate and optimization efficiency under two different dimensional coefficients C1 and C2.

It can be seen from experiments that the overall trend of fitness and optimization efficiency is relatively stable, and with the continuous progress of time and inheritance, the optimization efficiency will gradually increase, and the error analysis will be reduced by 75.4%, which will greatly improve the influence of various unfavorable factors in the optimization process, and because the rate of change of fitness basically maintains a high rate of change, it also reflects that the basic framework of the model is extremely excellent, for the entire algorithm structure and data processing soy sauce 54%. promote. The optimization process of GA is to force the penalty term to gradually approach 0, so that the penalty function reaches the minimum value, which also gradually approximates the originally out-of-bounds variable until it is pulled back to the constraint range of the variable. The whole algorithm spends most of its time on the calculation of the objective function, which reduces the calculation time. In the improved GA, only the individuals involved in crossover and mutation operations are calculated for the individuals in the new generation population, which shortens the calculation time to some extent.

4. Conclusions

Aiming at the disadvantage of premature GA, an improved GA is proposed. The improved GA introduces the population differentiation mechanism, adopts the dynamic adaptive crossover mutation operator, organically combines the improved heuristic crossover method and the chaotic mutation method degenerated with the number of iterations, and designs a mechanism to make the algorithm jump out of the local optimization. It can be seen from the experiment that the intelligent controller optimized by GA shows good dynamic and static characteristics when dealing with time-varying, time-delay, nonlinear, and model-uncertain systems. However, it should be pointed out that GA can find the global optimal solution when solving simple systems or some functions. The improved GA proposed in this paper is used to adjust and optimize the controller parameters.

When the uncertain parameters change greatly, the control system can still have good control quality and strong robustness. In addition, the algorithm does not destroy the structure and essence of traditional GA, so it inherits the advantages of GA and takes into account the advantages of random algorithm.

However, there are some limitations in this paper. In practical application, genetic algorithm is prone to premature convergence. We should not only keep the excellent individuals but also maintain the diversity of the group. Therefore, it is necessary to analyze the genetic algorithm in the 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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] X. Luo, J. Wang, and X. Luo, "A regression model optimization method based on GA [J]," *Software Engineering and Applications*, vol. 10, no. 1, p. 7, 2021.
- [2] H. Jin, Y. Fu, G. Yang, and X. Zhu, "An intelligent scheduling algorithm for resource management of cloud platform," *Multimedia Tools and Applications*, vol. 79, no. 7, pp. 5335–5353, 2020.
- [3] M. Yi, L. Mo, and Q. Shen, "Study on generation scheduling of cascade hydropower stations based on SAPSO," *Journal of Coastal Research*, vol. 104, no. sp1, pp. 371–378, 2020.
- [4] T. Qihua, Q. Dong, and D. Yixian, "Optimization of task distribution scheme for two-stage iterative model based on GA [J]," *Mechanical Design*, vol. 35, no. 3, p. 6, 2018.
- [5] B. Denkena, F. Schinkel, J. Pirnay, and S. Wilmsmeier, "Quantum algorithms for process parallel flexible job shop scheduling," *CIRP Journal of Manufacturing Science and Technology*, vol. 33, pp. 100–114, 2021.
- [6] V. Benifla and F. Adam, "Development of a Genetic Algorithm Code for the Design of Cylindrical Buoyancy Bodies for Floating Offshore Wind Turbine Substructures[J]," *Energies*, vol. 15, no. 3, p. 1181, 2022.
- [7] B. He, Z. Jiexin, and Z. Fuqiang, "An improved GA for solving job shop scheduling problem [J]," *Manufacturing Automation*, vol. 40, no. 8, p. 5, 2018.
- [8] L. Xiang, L. Dongsheng, and R. Hu, "Application of improved GA in cooperative interference resource allocation [J]," *Journal of Detection and Control*, vol. 40, no. 5, p. 7, 2018.
- [9] C. Gao, M. Shen, X. Liu, L. Wang, and M. Chen, "End-point Prediction of BOF Steelmaking Based on KNNWTSVR and LWOA," *Transactions of the Indian Institute of Metals*, vol. 72, no. 1, pp. 257–270, 2019.
- [10] S. Liangshan and L. Chenhao, "Extreme learning machine classification model based on improved pollen algorithm [J]," *Computer Engineering and Applications*, vol. 56, no. 1, p. 8, 2020.

Retraction

Retracted: Construction of Holographic Immersion Chamber Based on Multisource Information Fusion and Interactive Virtual Reality Technology

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Wang and Y. Liu, "Construction of Holographic Immersion Chamber Based on Multisource Information Fusion and Interactive Virtual Reality Technology," *Journal of Function Spaces*, vol. 2022, Article ID 6406211, 10 pages, 2022.

Research Article

Construction of Holographic Immersion Chamber Based on Multisource Information Fusion and Interactive Virtual Reality Technology

Chunying Wang ¹ and Yin Liu ²

¹College of Applied Technology, Changchun University of Technology, Changchun 130012, China

²Changchun Mingxuan Culture Technology Co., Hulun Beiwen, Inner Mongolia 130000, China

Correspondence should be addressed to Chunying Wang; wangchunying@ccut.edu.cn

Received 11 May 2022; Revised 27 June 2022; Accepted 1 July 2022; Published 18 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Chunying Wang and Yin 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.

Nowadays, many visual artists actively respond to the drastic changes taking place in today's society, build a communication bridge between real places and virtual places, and use holographic image technology to organically integrate them with each other. We are increasingly shuttling through the network of unbounded space. These spaces have no fixed geographical location. With the help of modern science and technology, they often fall into the field of vision of the viewer. With the development of the times, virtual reality technology is a computer simulation system that can create and experience the virtual world. It mainly uses computers to generate a simulated environment. It is a multisource information fusion interactive virtual reality technology to help users immerse themselves in the virtual environment. This paper studies the holographic immersion chamber based on multisource information fusion interactive virtual reality technology. The approximation index of the transformation matrix calculated by the registration algorithm based on SIFT+MSA is 1.7073, which is less than 4.0017 obtained by SIFT matching algorithm. It can be seen that the error of CCD image registration parameters calculated by this algorithm is smaller. In the process of the development of multisource information fusion interactive virtual reality technology, the technical feature of virtual reality is gradually integrated into the user and application object, becoming more closely connected and finally causing the user of the technology the feeling of "immersion" in the use experience.

1. Introduction

With the development of modern science and technology, the construction of places has become simpler and simpler. Nowadays, many visual artists actively respond to the drastic changes taking place in today's society, build a communication bridge between real places and virtual places, and use holographic image technology to organically integrate them with each other [1]. We are increasingly shuttling through the network of unbounded space. These spaces have no fixed geographical location. With the help of modern science and technology, they often fall into the field of vision of the viewer. One of the best is holographic image technology [2, 3]. We can apply holographic projection to all indoor spaces for display. Holographic projection system can turn

the objective environment into an "invisible" display interface and make anything existing in the virtual environment happen in real life visually and psychologically. The following professional categories can use holographic projection to complete a set of visual information transmission scheme [4]. The sense of immersion and the behavior of immersion are integrated. From the perspective of consciousness, immersion is inseparable from the creation of authentic and credible situations. Holographic image technology is used to restore all the relevant information of the recorded object in the real physical space. The three-dimensional image presented by it has a real visual effect combined with the depth of field information of the physical environment [5, 6]. In terms of the three-dimensional presentation effect of art, it subverts people's cognition of the traditional concept

of “place”—implanting three-dimensional virtual images and creating a virtual immersive environment place, which greatly expands the expressiveness of visual art and deeply affects people’s daily life.

Virtual reality emerged at the end of the 20th century. It is an information integration technology system that comprehensively uses computer multimedia, digital image processing, pattern recognition, artificial intelligence, sensors, and high-resolution display technology to achieve tactile virtual senses on the basis of vision and hearing and generate realistic three-dimensional virtual environment [7]. Virtual reality technology is a computer simulation system that can create and experience the virtual world. It mainly uses computers to generate a simulated environment. It is a multisource information fusion interactive virtual reality technology to help users immerse themselves in the virtual environment [8, 9]. Virtual reality technology truly realizes human-computer interaction, so that people can immerse themselves in a wonderful environment. At present, on the basis of multisource information fusion interactive virtual reality technology, many new concepts including virtual reality city and virtual reality tourism are also put forward. It can be said that virtual reality technology has become one of the most promising technologies [10]. The use of multisource information fusion interactive virtual reality can be divided into consumer and enterprise markets. The former focuses on immersive video and digital game experience, while the latter mainly focuses on military, medical, architecture, education, and other scenes [11, 12].

Holographic immersive image technology creates a virtual environment. This technical immersion environment can be divided into sensory immersion and participatory immersion. First, sensory immersion mainly appeals to the viewer’s sensory organ experience. Since the integration of the three-dimensional image presented by the holographic image and the real physical world is realized in the state of almost transparent interface, it has a strong sense of reality and presence, which can better promote the viewer to obtain an immersive live narrative experience [13]. Holographic immersive live experience can attract viewers’ attention, enhance their participation, let them actively rather than passively integrate into the narrative, and establish a deeper understanding and memory of the narrative content and significance. An education column is set up on the website to provide users with training services for the use skills of multisource information fusion interactive reality technology. Through online interaction, users can learn relevant operation technologies of virtual reality online and immerse themselves in interactive experience, which can better stimulate users’ interest and make more effective use of relevant services [14]. In the development process of multisource information fusion interactive virtual reality technology, the technical feature of virtual reality is gradually integrated into the user and application object, becoming more closely connected and finally causing the user of the technology the feeling of “immersion” in the use experience [15].

The innovative contribution of this paper lies in the analysis of holographic projection based on multisource information fusion interactive virtual technology. The

approximate index of transformation matrix calculated by SIFT+MSA matching algorithm is 1.7073, which is less than 4.0017 obtained by SIFT matching algorithm. It can be seen that the error of CCD image registration parameters calculated by this algorithm is small. Holographic projection technology breaks through the limitations of traditional sound, light, and electricity; brings beautiful pictures to the audience; and gives people a dual world feeling of virtual and reality.

2. Related Work

2.1. Research Status at Home and Abroad. Liu et al. proposed that in the visual narration of holographic images, the identity of the viewer is also changing, from the object of being informed to the subject who can directly participate in the narration and even construct the narrative content. The narrative subject can not only be used as a narrative tool but also become the purpose and means of narration [16]. Wei proposed that exploring the unknown is the driving force for human progress, and expanding human living space is one of the best choices. However, for people living on earth, the application of holographic virtual technology to life will greatly expand our living space [17]. Che et al. proposed that holographic projection can not only be used alone but also be used together with other multimedia devices. The purpose of its application is to immerse people in a visual enjoyment different from print media with a convenient, cheap, and novel technology [18]. Li et al. proposed that holographic visual narrative can not only use the traditional linear time structure but also present the narrative form of coexistence of linear and nonlinear. Linear time can be deconstructed into narrative fragments, providing viewers with a variety of narrative elements of their choice when immersed in the same time. After selecting and entering a separate narrative element, the narrative is carried out in a linear time structure [19]. Pei et al. proposed that holographic virtual display is also an art form, and each art form has a certain connection with reality. Therefore, the immersion feeling brought by the display of the same content is different [20]. Pan et al. put forward that holographic virtual technology has been silent since its emergence in the last century, but with the continuous development of computer science, holographic virtual technology has only a broader application and development space. In this process, there is continuous communication and exchange between Si Gan Dou scientists and artists. The use of computer technology and network technology makes the communication between works of art and the audience more smooth and deepens the public’s understanding of art and science and technology [21]. Lei et al. proposed that holographic projection is highly technical, involving many professional requirements such as accurate photography technology, transformation algorithm for generating pure phase hologram, nanotechnology, materials science, and optics. In practical application, laser sensing technology, touch control technology, and voice control technology are also needed to realize perfect immersive visual effect and interactive experience [22]. Huang et al. proposed that the three-dimensional visual effect displayed

by holographic image is more direct and realistic, which is easier to stimulate the resonance of the viewer in visual perception, forming the viewer's immersed visual reality and space-time reality [23]. Wang et al. proposed that holographic projection transforms the traditional two-dimensional plane image into a dynamic, three-dimensional, and multidimensional viewing image. Consumers can place the holographic projection photo in the holographic projection frame and even add the holographic projection function to the smart phone in the future, which can be displayed directly by the mobile phone. Holographic projection photos break through the static state and can coexist acoustically and dynamically [24]. Ma et al. proposed that from the perspective of behavior, immersion is inseparable from the perception of presence; that is, the viewer's body and consciousness are naturally placed directly in the situation. Immersion is to make people focus on the current goal (created by the designer) and feel happy and satisfied but forget the real world situation. The direct participation of the body makes the viewer closer to the narrative content, understand the narrative in their own way, and obtain cognition [25].

2.2. Research Status of Holographic Immersion Chamber Based on Virtual Reality Technology. Based on multisource information fusion interactive virtual reality technology, this paper studies the holographic immersion room. The multisource information fusion interactive virtual reality technology can not only make the designer intuitively understand the intention but also make the construction party directly understand the design intention and design effect. It is also very helpful for customers to have an immersive experience. The integration of multisource information fusion and interactive virtual reality technology enhances the visibility and representativeness of the design model. Under the function of virtual display, users have a feeling of interactive design and visual design, which greatly enhances the marketing effect. Holographic image technology appears in front of the public with the development of human social science and technology. In terms of its historical time, holographic image technology is still a new thing. From the perspective of the application of holographic image technology in practice, it still has great uncertainty. Researchers and practitioners in relevant fields need to further develop the effective function of this technology. The narrative subject presented by holographic image breaks through the visual limitations of two-dimensional narrative subject or pure virtual environment in traditional visual narrative expression methods and brings tension and real three-dimensional visual experience to the viewer. Based on multisource information fusion, interactive virtual reality technology can achieve the following effects: perceptual systems such as visual, auditory, and tactile immersion and behavioral systems, such as direction, language immersion, and expression system immersion. Through the recognition of interactive virtual reality technology based on multisource information fusion and the sign language recognition system based on vision, virtual reality can restore nonverbal information such as gestures, facial expressions, and posture. In the field of information fusion, many scholars often refer to terms sim-

ilar to information fusion, such as data fusion, information fusion, and multisensor fusion. These concepts are both different and closely related. Virtual display technology is a new type of science and technology, which has just been popularized, so it is more attractive to most people. In this sense, virtual inch display has novel and unique characteristics, which is also one of the main reasons for the audience to come to the exhibition. The emergence of multisource information fusion interactive virtual reality technology has brought significant innovation to the imaging problem in the field of new media. This emerging technology not only makes outstanding contributions to virtual imaging but also has a far-reaching impact on promoting the development of other industries, including film and television industry, construction industry, exhibition planning industry, and tourism and leisure industry.

3. Principle and Model of Multisource Information Fusion Virtual Reality Technology

“Virtual reality technology” was put forward in the early 1980s. As a simulation technology and a challenging fashion frontier interdisciplinary subject, it combines simulation technology with computer graphics, man-machine interface technology, sensing technology, and multimedia technology to generate a virtual situation. This virtual three-dimensional dynamic situation integrating multisource information makes people feel like the real world. Data fusion is mainly aimed at all kinds of information fusion expressed in the form of data. When the information needed is the measured data of sensors, data fusion is called sensor fusion. Information fusion includes data fusion and sensor fusion, but the scope of information fusion is broader. In addition to data, the fused information can also be extended to images, notes, symbols, knowledge, and intelligence. At present, there is no clear distinction in the field. Virtual reality technology is the integration of electronic science and technology and information technology. In essence, it is a new communication medium and communication tool. From the perspective of media, virtual reality, with its new way of information exchange and unique use experience, produces a subversive way of communication-immersive communication. It has become a super media that surpasses popular media such as newspapers, magazines, radio, film, television; and Internet, realizes virtual crosstemporal interaction; and creates realistic on-the-spot communication. Using the similar method, the best matching position in the first level resolution can be found step by step, and the amount of search calculation decreases in geometric order. The flow chart of multisource information fusion algorithm is shown in Figure 1.

Virtual reality systems are usually equipped with a variety of sensing devices, including visual, auditory, and tactile devices. In the future, taste and olfactory sensing devices may be developed. In addition to official sensing devices, there are kinesthetic sensing devices and reaction devices. These devices enable the virtual reality system to have

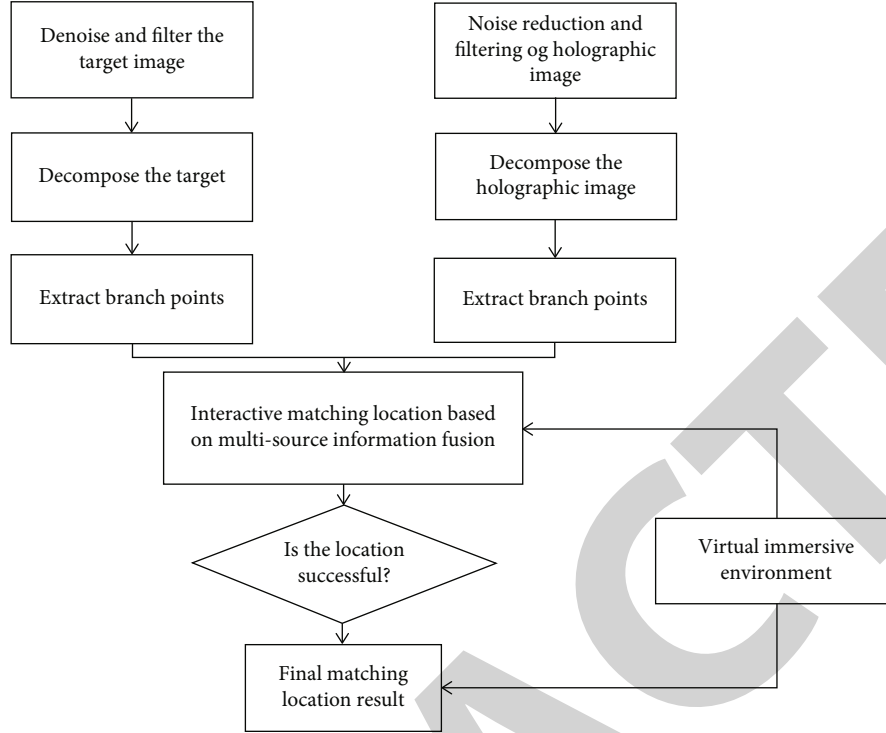


FIGURE 1: Flow chart of multisource information fusion algorithm in holographic immersion.

multisensory functions and enable users to obtain a variety of senses in the virtual environment. The organization of multisource information fusion is the data fusion joint command Laboratory of the U.S. Department of Defense. After continuous revision and practical expansion, this model has been determined as the actual standard of the U.S. defense information fusion system. A multisource information fusion virtual reality technology model is proposed, as shown in Figure 2.

The imaging band is $0.33 \sim 0.77 \mu\text{m}$, and the radiation of the ground object is composed of its own light radiation and the radiation of the surrounding environment reflected by it. Its radiation brightness can be expressed as

$$L_{\text{sum}} = L_0(\theta_r, \psi_r) + L_{rE}(\theta_r, \psi_r). \quad (1)$$

Among them, (θ_r, ψ_r) represents the radiation direction, its own radiation brightness is represented by $L_0(\theta_r, \psi_r)$, and its reflection brightness to surrounding objects is represented by $L_{rE}(\theta_r, \psi_r)$. In fact, in the $0.33 \sim 0.77 \mu\text{m}$ band, most ground objects can be regarded as diffuse reflectors, so the (θ_r, ψ_r) influence of radiation direction can be ignored, and the formula becomes

$$L_{\text{sum}} = L_0 + L_{rE}. \quad (2)$$

$L_0 = \varepsilon L_{h0}$ can be obtained according to Planck formula, where ε is the surface emissivity of diffuse reflector and L_{h0} is the radiance of absolute blackbody at the same temperature as the ground object. L_{rE} is composed of solar radiation and atmospheric scattering, which can be expressed as

$L_{rE} = (\rho/\pi)E_{\text{sum}}$, where $\rho = 1 - \varepsilon$ and E_{sum} represent the irradiance of solar radiation and atmospheric scattering, so it can be expressed as

$$L_{\text{sum}} = (1 - \rho) \cdot L_{h0} + \rho \cdot \frac{E_{\text{sum}}}{\pi}. \quad (3)$$

Generally, we think that the ground object is in the normal temperature state. According to Planck's law, L_{h0} is negligible. It can be seen that the main parameters affecting the visible light characteristics of the ground object are ρ and E_{sum} , which shows that the factors affecting the imaging quality of the visible light image are lighting conditions and meteorological factor, and explains the reason why the CCD sensor cannot work all day.

The reflectivity ρ represents the reflection ability of the ground object to the surrounding radiation, which can be described by the bidirectional reflection distribution function. For the opaque gray body, its transmittance $\tau = 0$, so it meets the requirements.

$$\varepsilon = \alpha = 1 - \rho. \quad (4)$$

The image is mainly disturbed by additive system noise and multiplicative speckle noise.

$$I(x, y) = R(x, y)\eta_0(x, y) + \eta_1(x, y), \quad (5)$$

where $I(x, y)$ represents the output image of the imaging system, $R(x, y)$ represents the scattering characteristics of ground objects, $\eta_0(x, y)$ represents multiplicative speckle noise, and $\eta_1(x, y)$ represents additive system noise.

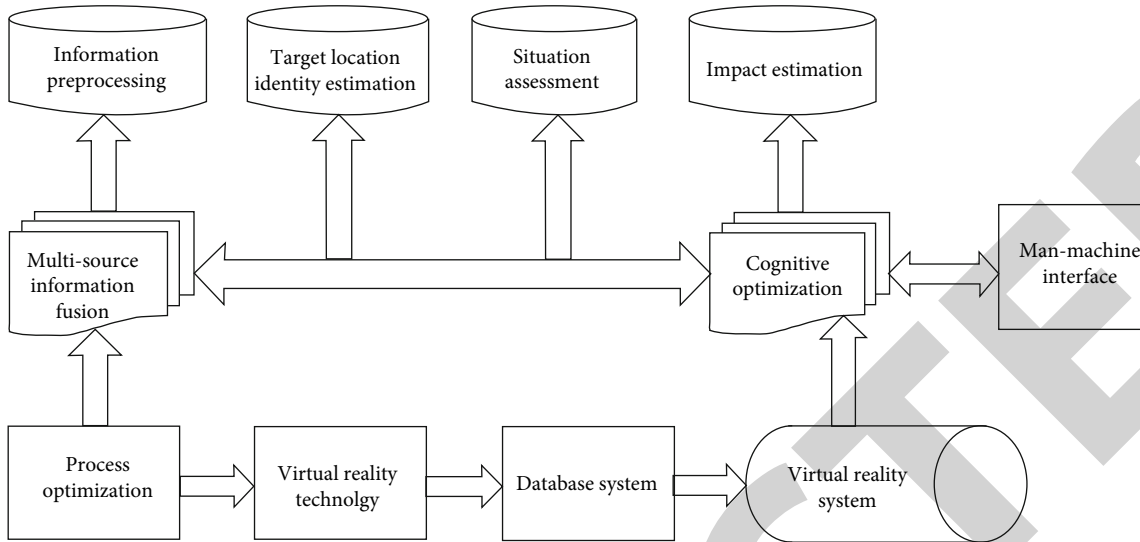


FIGURE 2: Multisource information fusion virtual reality technology model.

Gaussian filter is a commonly used low-pass filter, which can filter the noise of high-frequency part. The expression of Gaussian filter is

$$f'(x, y) = f(x, y) * g(x, y). \quad (6)$$

$f(x, y), f'(x, y)$ represents the pixel value before and after processing, $g(x, y)$ represents the Gaussian kernel function, and its expression is

$$g(x, y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (7)$$

The σ value determines the smoothing effect of Gaussian filter. The smoothing result of the image using the filter window with the size of t can be expressed as

$$f'(x, y) = \sum_{i=-t}^t \sum_{j=-t}^t f(x-i, y-j) \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (8)$$

This method has a good effect on removing Gaussian white noise, but at the same time, it will destroy some other details in the image and affect the extraction of point, line, and other types of features. Therefore, it is necessary to adjust the parameters of the filter to retain other features as much as possible while suppressing noise interference.

The mean filtering adopts the “windowed average” method to divide the gray level of the noise point (x, y) equally to the pixels with the size of $m \times n$ around it. The mean filtering expression of the image size is

$$f'(x, y) = \frac{1}{m \cdot n} \sum_{i,j \in m \times n} f(i, j), \quad (9)$$

where $f'(x, y)$ represents the noise point obtained by mean filtering, and the neighborhood range of point (x, y) is taken as $m \times n$. Obviously, the disadvantage of this filtering method is that the same processing is adopted for noise points and nonnoise points. While smoothing the noise interference, it will inevitably lose the detailed information in the image and still blur the edge.

For digital images, the filtering process in spatial domain can be expressed as the convolution of the original image and the filter kernel function.

$$\begin{aligned} f'(x, y) &= f(x, y) * h(x, y) \\ &= \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n)h(x-m, y-n). \end{aligned} \quad (10)$$

According to the relationship between spatial domain and frequency domain, the convolution in SD can be converted into product in FD. Virtual reality has developed rapidly in the past two years. Facebook, Microsoft, Google, and others have joined the game. Virtual reality creates an immersive interaction in which the machine construction environment is infinitely close to the reality perception environment, and conveys to users a complete and comprehensive physical and psychological feeling like the real situation. Human-computer interaction tends to return to people’s original state more and more. Information fusion is a general framework for simultaneous interpreting or combining methods and tools from different sensor data. The purpose of information fusion is to obtain higher quality information. Famous Chinese scholars Han Chongzhao and Zhu Hongyan believe that multisource information fusion is actually a functional simulation of the human brain’s comprehensive processing of complex problems. Multisource information fusion interactive virtual reality technology not only spreads and communicates across the limitations of time and space but also creates a technical

system that can make the virtual natural environment act on the human body and produce various senses and limb reactions, that is, using human body perception to construct a second nature constructed by optoelectronic materials and create an optoelectronic interface environment that can make it difficult for users to distinguish between true and false.

4. Realization of Holographic Immersion Chamber

4.1. Construction of Holographic Immersion Chamber System Based on Virtual Reality Technology. The expansion of space and the three-dimensional presentation of holographic images make the visual narrative no longer stick to the limited picture frame and construct the content and meaning of the narrative. The spatial dimension of visual narrative has the possibility of more diversified extension. The narrative elements in different picture frames are released from the space restrictions artificially cut and split, and more abundant narrative content is transmitted through more free arrangement and combination. Holographic projection immersion system based on multisource information fusion interactive virtual reality technology is mostly used in 360° booth and 270° booth of small products in the commercial application of exhibition cabinet. Most of the contents are relatively simple rotating animation. Of course, some are also used to display characters, but the action of characters is simple. In the commercial application of the stage, in order to meet the viewing angle of the stage, the single-chip holographic screen of 180° is mostly used. The application modes of multisource information fusion interactive virtual reality technology include virtual performance, virtual and real person interaction, real person performance, and holographic special effects. Compared with the large amount of equipment and space occupied by equipment in traditional dance beauty design, the requirements of new media technology for equipment space are greatly reduced, and the transformation of various effects can be easily realized.

The holographic immersion room system of multisource information fusion interactive virtual reality technology is an extension and innovation of traditional visual narration. It builds a bridge between the viewer and narration. Through the immersive on-site narrative experience, it increases the viewer's emotional input and active participation, shortens the distance between the viewer and narration, endows visual narration with new vitality, and transmits more profound and unforgettable narrative connotation. Holographic immersion room system is a dynamic display mode. Its operable and interactive form shortens the distance between consumers and products, and consumers directly experience all the functional features of products in their use and participation. Multisource information fusion interactive virtual reality technology this way not only improves the added value of products but also fully mobilizes consumers' interest, so as to stimulate consumer behavior and directly bring considerable economic benefits. The holographic immersion room system

uses holographic projection skills to build a scientific, innovative, and interactive environment. The details of goods are displayed through holographic imaging, which can attract the attention of visitors. It can also be applied in the fields of cars, large forklifts, watches, jewelry display, and so on, which can not only achieve the intention of showing to the audience but also reduce the consumption of booths and exhibits, holographic projection museum, holographic projection dance, holographic projection telephone, etc.

4.2. Experimental Results and Analysis. In this experiment, the image is registered with the same target image; the parameters of the transformation matrix are calculated, compared with the affine transformation matrix imposed by human; and the approximation index e of the transformation matrix is calculated. SIFT feature is a local feature of an image, which maintains invariance to rotation, scaling, and brightness changes and also maintains a certain degree of stability to angle changes, affine transformations, and noise. It also maintains good matching for the factors such as object motion, occlusion, and noise, so that the feature matching between two images with large differences can be realized. At the same time, taking the registration algorithm that separately uses sift as the matching feature and other matching strategies consistent with the algorithm in this section as a reference, the matching accuracy of the two algorithms under affine transformation is directly compared. The exact registration affine transformation parameters of images are shown in Table 1.

It can be seen from Table 1 that the approximation index of the transformation matrix calculated by the registration algorithm based on SIFT+MSA is 0.8646, which is less than 1.2697 obtained by the SIFT matching algorithm. It can be seen that the parameters calculated by the algorithm in this section are closer to the real value and have higher accuracy. The original image is too large and has been reduced to 83% of the display.

The matching points are obtained according to the two algorithms. The comparison between the affine transformation parameters fitted by the least square method and the real affine parameters is shown in Table 2.

The distortion of the scene in the imaging process will make the image out of proportion. Affine transformation can be used to correct various distortions. The parameters of affine transformation can be estimated by the least square method. It can be seen from Table 2 that the approximation index of the transformation matrix calculated by the registration algorithm based on SIFT+MSA is 1.7073, which is less than 4.0017 obtained by the SIFT matching algorithm. It can be seen that the error of CCD image registration parameters calculated by the algorithm in this section is smaller.

The matching points are obtained according to the two algorithms. The comparison between the affine transformation parameters fitted by the least square method and the real affine parameters is shown in Table 3.

It can be seen from Table 3 that the approximation index of transformation matrix calculated by SIFT+MSA

TABLE 1: Affine transformation parameters for accurate image registration.

Image	a	b	c	d	e	Approximation index
Real parameters	1.2	0.2	-2	0	1	0
SIFT	1.2082	0.4012	-4.0751	0.1013	1.1051	1.2697
SIFT+MSA	1.0771	0.2871	-2.1358	-0.0061	1.0035	0.8646

TABLE 2: Affine transformation parameters for accurate registration of visible images.

Visible image	a	b	c	d	e	Approximation index
Real parameters	0.7	0.2	51	0.2	0.6	0
SIFT	0.8895	0.1657	53.4828	0.2955	0.7693	4.0017
SIFT+MSA	0.7898	0.2751	50.9406	0.2974	0.6802	1.7073

TABLE 3: Affine transformation parameters for accurate image registration.

Infrared image	a	b	c	d	e	Approximation index
Real parameters	0.7	0.1	2	-0.2	1.1	0
SIFT	0.8991	0.1995	3.3548	-0.1988	1.2981	2.1686
SIFT+MSA	0.7925	0.1015	3.2677	-0.1996	1.2025	1.0037

TABLE 4: Comparison of registration results of real images.

Registration algorithm	Feature points extracted from real image and target image	Rough matching feature points	Fine matching feature points	Match score	Matching rate	RMSE	Subjective evaluation of local details
SIFT	(257, 1505)	121	30	46.32%	24%	1.4472	Dislocation
SIFT+MSA	(111, 693)	62	41	56.24%	66.66%	0.6424	High accuracy

registration algorithm is 1.0037, which is less than 2.1686 obtained by the SIFT matching algorithm, and the error is smaller.

The evaluation system of image registration algorithm compares the two algorithms, and the calculated evaluation indexes are shown in Table 4.

It can be seen from Table 4 that the feature points extracted by the SIFT algorithm from the real image and the target image are 257 and 1505, respectively, and the logarithm of feature points obtained by rough matching is 121. Based on this, the matching score is 46.32%. After further screening the matching points by RANSAC method, 30 pairs of fine matching points can be obtained, and the matching rate is 24%. After matching, the RMSE of the two images is 1.4472.

In this experiment, different numbers of recommended list items are set in the data set to compare the recommended diversity results of pUCP algorithm, phui growth algorithm, and this algorithm. Three experiments were carried out to compare the holographic image immersion. The experimental results are shown in Figures 3–5.

It can be seen from Figure 3 to Figure 5 that when the number of recommendation lists is small, the difference between the experimental results of the three algorithms is very small, because the fewer the items in the list, the greater

the possibility of approximation. However, with the increase of the number of lists, phui growth loses more item sets, and the resulting association rule base is inaccurate, resulting in similar items in the recommendation list, and the recommended diversity index will be significantly higher than the other two algorithms. Because the accuracy of this algorithm is higher than the other two algorithms, the holographic immersion chamber based on multisource information fusion interactive virtual reality technology is better than the other two algorithms in diversity.

In this experiment, holographic image technology creates a virtual immersive extraction method as follows: during the operation of the algorithm, insert multiple test points. In each test point, first, forcibly delete the garbage nodes generated during the operation of the algorithm, and then, extract the current memory consumption value, and take the maximum of these values as the memory consumption value of the current algorithm. Set the number of different recommended list items, and compare the recommended diversity results of pUCP algorithm, phui growth algorithm, and this algorithm. Conduct two experiments, respectively, for comparison, and the operation results are shown in Figures 6 and 7.

It can be seen from Figure 6 to Figure 7 that the memory consumption of the algorithm up growth increases with the

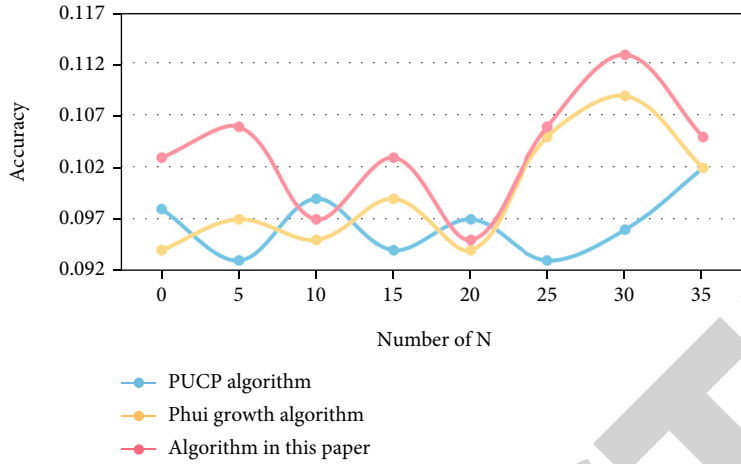


FIGURE 3: Diversity analysis of holographic immersion chambers recommended by different algorithms.

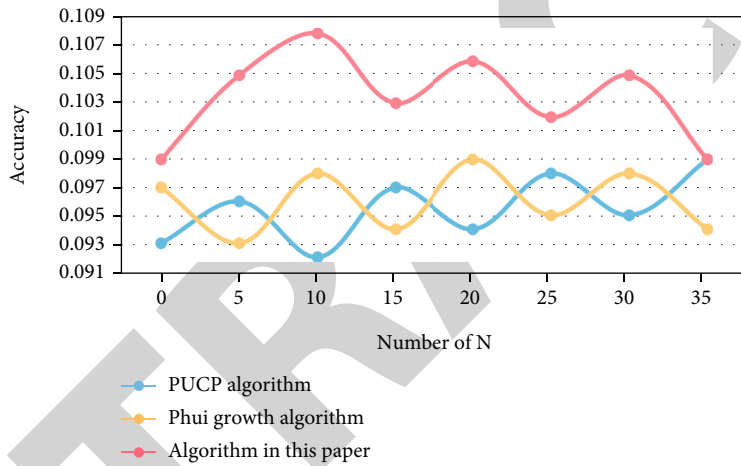


FIGURE 4: Diversity analysis of holographic immersion chamber recommended by different algorithms.

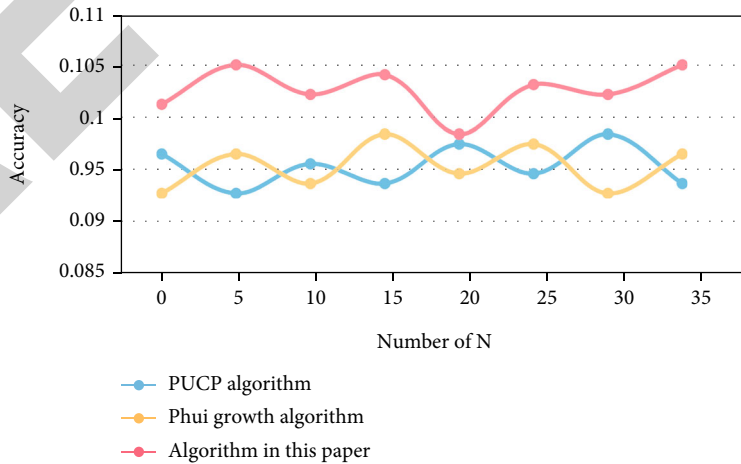


FIGURE 5: Diversity analysis of holographic immersion chamber recommended by different algorithms.

decrease of the minimum utility threshold. This is because the nodes in the utility mode tree increase with the decrease of the minimum utility threshold, but the memory con-

sumption of the algorithm in this paper and phui growth algorithm basically does not change much. This is because the two algorithms adopt a parallel strategy. The algorithm

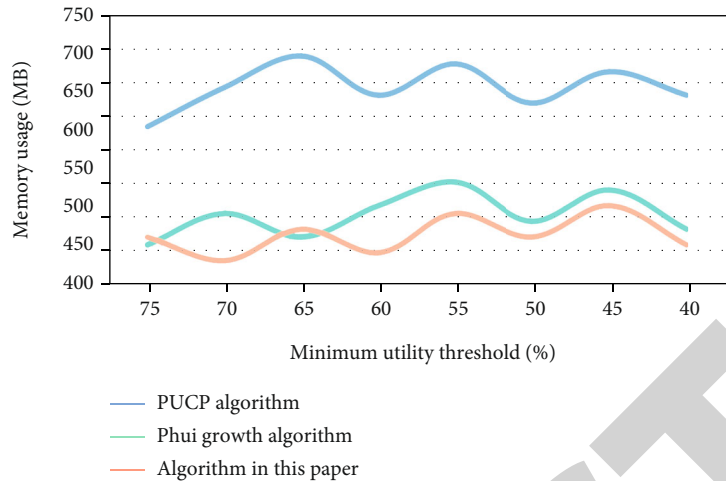


FIGURE 6: Comparison of virtual immersion consumption of holographic image technology under different algorithms.

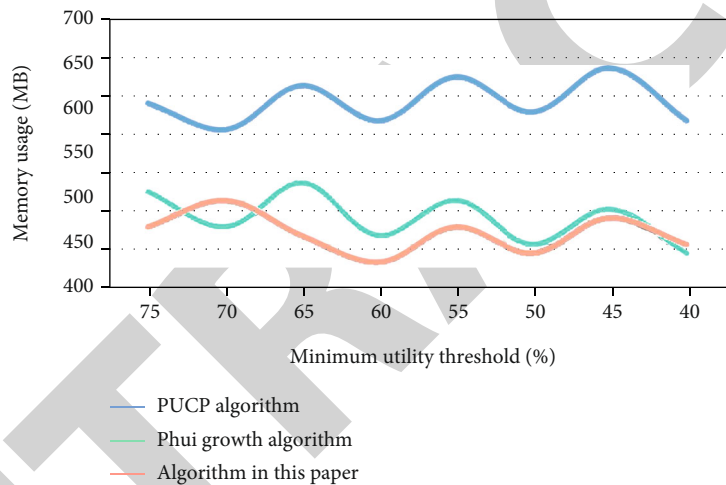


FIGURE 7: Comparison of virtual immersion consumption of holographic image technology under different algorithms.

in this paper uses the clustering partition method to build similar transactions into a utility pattern tree. The tree structure has fewer branches and uses the mode of pattern growth for mining, which reduces the number of candidate item sets. It adopts the tree structure and does not store noncandidate item sets. Therefore, it occupies less memory than the phui growth algorithm.

5. Conclusions

Multisource information fusion interactive virtual reality technology is an interdisciplinary subject integrating bionic technology, electronic technology, communication technology, and computer technology. It is also a cutting-edge subject prioritized by most countries. With the emergence of various new technologies, especially computer technology and communication technology, it has promoted the development of multisource information fusion interactive virtual reality technology. All kinds of new equipment relying on multisource information fusion interactive virtual reality

technology are emerging. Holographic projection will be diversified in the development, but with its gradual improvement, it will be widely used in the performance exhibition space. Once the holographic projection technology is popularized, the cost is reduced and the convenience is improved; its role cannot be underestimated. This paper studies the holographic immersion chamber based on multisource information fusion interactive virtual reality technology. The approximation index of the transformation matrix calculated by the registration algorithm based on SIFT+MSA is 1.7073, which is less than 4.0017 obtained by the SIFT matching algorithm. It can be seen that the error of CCD image registration parameters calculated by this algorithm is smaller. Holographic projection technology breaks through the limitations of traditional sound, light, and electricity; brings beautiful pictures to the audience; and gives people a double world feeling of coexistence of virtual and reality. In the future, the market development potential of holographic projection technology will be immeasurable. In the future, with the continuous maturity of technology,

Retraction

Retracted: Decompositions of Circulant-Balanced Complete Multipartite Graphs Based on a Novel Labelling Approach

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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] A. El-Mesady and O. Bazighifan, “Decompositions of Circulant-Balanced Complete Multipartite Graphs Based on a Novel Labelling Approach,” *Journal of Function Spaces*, vol. 2022, Article ID 2017936, 17 pages, 2022.

Research Article

Decompositions of Circulant-Balanced Complete Multipartite Graphs Based on a Novel Labelling Approach

A. El-Mesady ¹ and Omar Bazighifan ^{2,3}

¹Department of Physics and Engineering Mathematics, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

²Department of Mathematics, Faculty of Science, Hadhramaut University, Hadhramaut, Al Mukalla 50512, Yemen

³Department of Mathematics, Faculty of Education, Seiyun University, Hadhramout 50512, Yemen

Correspondence should be addressed to Omar Bazighifan; o.bazighifan@gmail.com

Received 27 May 2022; Accepted 24 June 2022; Published 18 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 A. El-Mesady and Omar Bazighifan. 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 applied scientists and engineers, graph theory is a strong and vital tool for evaluating and inventing solutions for a variety of issues. Graph theory is extremely important in complex systems, particularly in computer science. Many scientific areas use graph theory, including biological sciences, engineering, coding, and operational research. A strategy for the orthogonal labelling of a bipartite graph G with n edges has been proposed in the literature, yielding cyclic decompositions of balanced complete bipartite graphs $K_{n,n}$ by the graph G . A generalization to circulant-balanced complete multipartite graphs $K_{\underbrace{n,n,\dots,n}_m}$; $m, n \geq 2$,

is our objective here. In this paper, we expand the orthogonal labelling approach used to generate cyclic decompositions for $K_{n,n}$ to a generalized orthogonal labelling approach that may be used for decomposing $K_{\underbrace{n,n,\dots,n}_m}$. We can decompose

$K_{\underbrace{n,n,\dots,n}_m}$ into distinct graph classes based on the proposed generalized orthogonal labelling approach.

1. Introduction

As is well known, discrete mathematics is a field of mathematics that deals with countable processes and components. One of the most significant and intriguing disciplines in discrete mathematics is graph theory [1–3]. Graph theory is the study of structural models called graphs, which are made up of a collection of vertices and edges. Graph theory is extremely important in complex systems, particularly in computer science. Many scientific areas use graph theory, including engineering, coding [4, 5], operational research, biological sciences, and management sciences. For applied scientists and engineers, graph theory is a strong and vital science for evaluating and inventing solutions for a variety of issues. Graphs have recently been utilized as structural models for characterizing World Wide Web

connections and the number of links necessary to move between web pages [6].

Circulant graphs are a significant category of graphs [7–10]. Circulant graphs have gained a lot of attention in recent decades. The circulant graphs class includes complete graphs and classic rings topologies. The algebraic properties of circulant graphs have been studied in thousands of publications. Circulant graphs have been handled in a variety of graph applications, including wide area communication graphs, local area computer graphs, parallel processing architectures, very large-scale integrated circuit design, and distributed computing [11–13].

Several traditional parallel and distributed systems were built on the foundation of circulant graphs [14–16]. Circulant graphs have a wide range of practical uses, such as a structure in chemical reaction models [17], multiprocessor cluster

systems [18], small-world graph models [19], discrete cellular neural graphs [20], and as a basic structure for optical graphs [21], and so on.

The study of circulant graphs, including their characterization, analysis, and applications, is currently a popular issue in research. Several papers have been published that deal with graph decompositions by simpler graphs [22–24]. Decompositions of circulant graphs have several excellent contributions. For Cayley graphs labelled with Abelian groups, the Hamilton decomposition was investigated in [25]. The circulant graph is a particular case of the Cayley graph. It has been demonstrated that two Hamilton cycles may be used to decompose four-regular connected Cayley graphs [26].

For a certain recursive circulant graph, the Hamilton decompositions have been proven [27]. Every circulant graph has a corresponding circulant matrix [28]. Excellent descriptions of circulant matrices have been published in [28].

Definition 1. A circulant-balanced complete multipartite graph $K_{\underbrace{n,n,\dots,n}_m}$ is a simple graph having $mn = \sum_{l=1}^m n$ vertices.

The vertices of $K_{\underbrace{n,n,\dots,n}_m}$ are divided into m partitions

of cardinality n ; two vertices are said to be adjacent if they are found in two different partitions. The graph $K_{\underbrace{n,n,\dots,n}_m}$

has a degree equal to $(m-1)n$. The circulant graph $K_{\underbrace{n,n,\dots,n}_m}$

can be divided into $\delta K_{n,n}$, $\delta = \binom{m}{2}$.

Definition 2. A caterpillar graph $C_a(b_1, b_2, \dots, b_a)$ is a tree formed by the path $P_a = y_1 y_2 \dots y_a$ by linking a vertex y_i to b_i new vertices where $a \geq 1, b_1, b_2, \dots, b_a$ are integers greater than zero, $b_1, b_a \geq 1$ and $b_i \geq 0$ for $i \in \{2, 3, \dots, a-1\}$.

El-Mesady et al. have proposed an orthogonal labelling approach to decompose a certain circulant graph class with $2n$ vertices and n degree [29]. Circulant-balanced complete bipartite graphs are the name for this type of graph which is denoted by $K_{n,n}$. In cognitive radio graphs and cloud computing, bipartite circulant graphs can address a variety of challenges. For a good survey on several decompositions of circulant graphs, see [30–34].

In this study, we generalize the orthogonal labelling approach proposed in [29] to create edge decompositions of the graphs $K_{\underbrace{n,n,\dots,n}_m}; m, n \geq 2$ which are considered a

generalization to the graphs $K_{n,n}$. The following sections make up the current paper: The second section deals with the proposed novel orthogonal labelling approach. In the third section, the graph $K_{\underbrace{n,n,\dots,n}_m}$ is decomposed by infinite

classes of graphs. We generate many decompositions of $K_{\underbrace{n,n,\dots,n}_m}$ by connected caterpillars in the fourth section.

The fifth section introduces concluding remarks and future work.

2. A Novel Labelling Approach

Consider now the circulant-balanced complete multipartite graph with vertex set $V = \bigcup_{l=0}^{m-1} V_l$, where $V, l \in \{0, 1, \dots, m-1\}$ are m independent sets of vertices. There are bijective mappings $\varphi_l : V_l \rightarrow \mathbb{Z}_n \times \{l\}, l \in \{0, 1, \dots, m-1\}$ where the vertices in V_l are labelled by $\mathbb{Z}_n \times \{l\}$, see Figure 1.

The distance between two vertices $x_i \in \{0, 1, \dots, (n-1)_i\}$ and $y_j \in \{0, 1, \dots, (n-1)_j\}, 0 \leq i < j \leq m-1$ is the usual circular distance defined by $d\{x_i, y_j\} = \min\{|x_i - y_j|, n - |x_i - y_j|\}$. The edge $\{x_i, y_j\}$ is said to have length $d\{x_i, y_j\}$. Suppose $G = (V, E)$ is a subgraph with mn vertices and $\binom{m}{2}n$ edges, a labelling

$$\psi_k : V(G_k^{i,j}) \rightarrow \mathbb{Z}_n \times \{i, j\}, 0 \leq i < j \leq m-1, k$$

$$= \begin{cases} j & \text{if } i = 0, \\ mi + j(\text{mod } (i+1)) & \text{if } i > 0. \end{cases} \quad (1)$$

is considered an orthogonal labelling of $G \cong \bigcup_{k=1}^w G_k^{i,j}, w$

$$= \binom{m}{2}, 0 \leq i < j \leq m-1 \text{ if,}$$

(i) Each graph $G_k^{i,j}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $G_k^{i,j}$, and the length $n/2$ is found once in $G_k^{i,j}$ if n is even

(ii) For every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ ,

(iii) The length 0 is found $\binom{m}{2}$ times in G ,

(iv) The length $n/2$ is found $\binom{m}{2}$ times in G if n is even

Example 1. An orthogonal labelling of $K_{1,3}^{0,1} \cup P_4^{0,2} \cup K_{1,3}^{1,2}$ is shown in Figure 2.

Definition 3. Suppose G is a subgraph of $K_{\underbrace{n,n,\dots,n}_m}, x \in \mathbb{Z}_n$.

Then $G+x$ with $E(G+x) = \{\{a+x, b+x\} : \{a, b\} \in E(G)\}$ is called the x -translate of G .

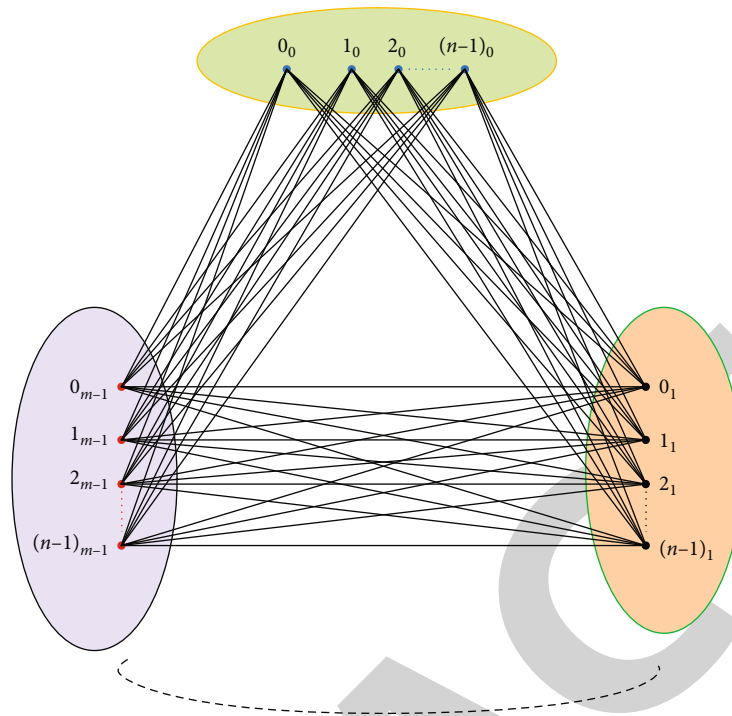


FIGURE 1: The labelling for $K_{\underbrace{n,n,\dots,n}_m}$.

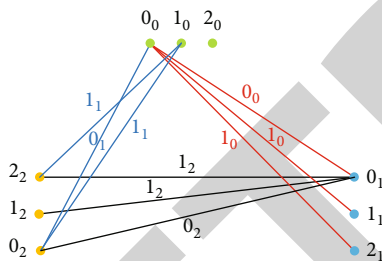


FIGURE 2: An orthogonal labelling of $K_{1,3}^{0,1} \cup P_4^{0,2} \cup K_{1,3}^{1,2}$.

The edge decomposition of circulant-balanced complete multipartite graphs and orthogonal labelling are linked in the next proposition.

Proposition 4. *If and only if there is an orthogonal labelling of $G \cong \bigcup_{k=1}^w G_k^{i,j}, 0 \leq i < j \leq m-1$, an edge decomposition of $K_{\underbrace{n,n,\dots,n}_m}$ can be constructed by G .*

Proof. Our goal is to show that $E(G^{i,j} + \omega) \cap E(G^{i,j} + \sigma) = \emptyset$ for all $\omega, \sigma \in \mathbb{Z}_n$. We assume, by way of contradiction, that $|E(G^{i,j} + \omega) \cap E(G^{i,j} + \sigma)| \geq 1$ for $\omega, \sigma \in \mathbb{Z}_n$ with $\omega \neq \sigma$. For the lengths $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, which are repeated twice in $G^{i,j}$, let $\{a, b\}$ and $\{c, d\}$ be two edges of $E(G^{i,j} + \omega) \cap E(G^{i,j} + \sigma)$ with length λ , then $\{a - \omega, b - \omega\}, \{c - \omega, d - \omega\}$ and $\{a - \sigma, b - \sigma\}, \{c - \sigma, d - \sigma\}$ are various edges with length λ in $E(G^{i,j})$. However, this is a contradiction because $G^{i,j}$ verifies the orthogonal labelling requirement (i). Let $\{a, b\}$

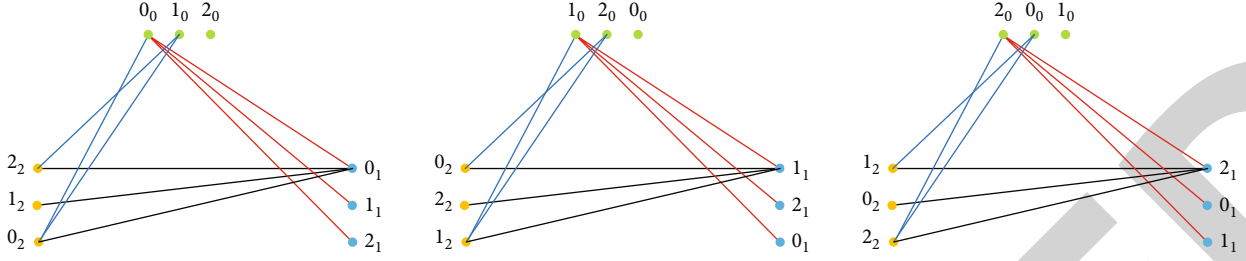
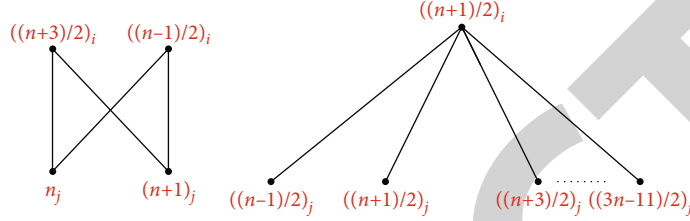
belong to $E(G^{i,j} + \omega) \cap E(G^{i,j} + \sigma)$ with length $l \in \{0, n/2\}$, n is even, then $\{a - \omega, b - \omega\}$ and $\{a - \sigma, b - \sigma\}$ are distinct edges in $E(G^{i,j})$, both with length l . However, this is a contradiction because $G^{i,j}$ verifies the orthogonal labelling requirement (i). Hence, $\bigcap_{x \in \mathbb{Z}_n} E(G + x) = \emptyset$. Also, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges with length λ , the length 0 is only found $\binom{m}{2}$ times in G , the length $n/2$ is only found $\binom{m}{2}$ times in G if n is even. Consequently,

$$\bigcup_{x \in \mathbb{Z}_n} E(G + x) = E\left(K_{\underbrace{n,n,\dots,n}_m}\right). \quad (2)$$

□

Example 2. An example of edge decomposition of $K_{3,3,3}$ by $K_{1,3}^{0,1} \cup P_4^{0,2} \cup K_{1,3}^{1,2}$ is shown in Figure 3.

In what follows, based on the aforementioned orthogonal labelling approach, we will decompose the circulant-balanced complete multipartite graph $K_{\underbrace{n,n,\dots,n}_m}$ by the $G \cong \bigcup_{k=1}^w G_k^{i,j}$, where the graphs $G_k^{i,j}, k \in \{1, 2, \dots, w\}, w = \binom{m}{2}, i \neq j \in \{0, 1, \dots, m-1\}$ are isomorphic. Also, we will consider

FIGURE 3: An edge decomposition of $K_{3,3}$ by $K_{1,3}^{0,1} \cup P_4^{0,2} \cup K_{1,3}^{1,2}$.FIGURE 4: The labelling for $(K_{2,2} \cup K_{1,n-4})^{i,j}$.

$$k = \begin{cases} j & \text{if } i = 0, \\ im + j \pmod{(i+1)} & \text{if } i > 0. \end{cases} \quad (3)$$

Proof. Suppose $V((K_{2,2} \cup K_{1,n-4})^{i,j}) = \{v_s : s \in \{0, 1, 2, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_1 , which can be defined by $\psi_k : V((K_{2,2} \cup K_{1,n-4})^{i,j}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by $\psi_k(v_0) = ((n+3)/2)_i$, $\psi_k(v_1) = ((n-1)/2)_i$,

$\psi_k(v_2) = ((n+1)/2)_i$, $\psi_k(v_{s+3}) = (((n-1)/2) + s)_j$, $s \in \{0, \dots, n-5\}$, and the edge set of $(K_{2,2} \cup K_{1,n-4})^{i,j}$ is

$$E((K_{2,2} \cup K_{1,n-4})^{i,j}) = \left\{ \left\{ \left(\frac{n+3}{2} \right)_i, n_j \right\}, \left\{ \left(\frac{n+3}{2} \right)_i, (n+1)_j \right\}, \left\{ \left(\frac{n-1}{2} \right)_i, n_j \right\}, \left\{ \left(\frac{n-1}{2} \right)_i, (n+1)_j \right\}, \left\{ \left(\frac{n+1}{2} \right)_i, \left(\frac{n-1}{2} \right)_j \right\} \right\} \cup \left\{ \left\{ \left(\frac{n+1}{2} \right)_i, \left(\frac{n+1}{2} + s \right)_j \right\} : s \in \{0, 1, \dots, n-6\} \right\} \quad (4)$$

see Figure 4. From the edge set of G_1 , the following conditions are verified: Each graph $(K_{2,2} \cup K_{1,n-4})^{i,j}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(K_{2,2} \cup K_{1,n-4})^{i,j}$, the length $n/2$ is found once in $(K_{2,2} \cup K_{1,n-4})^{i,j}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_1 has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_1 , and the length $n/2$ is found $\binom{m}{2}$ times in G_1 if n is even. Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_1 . \square

Theorem 6. Let $n > 1, m \geq 2$ be integers. Then, there is an orthogonal labelling for $G_2 \cong \bigcup_{0 \leq i < j \leq m-1} (K_{2,n})^{i,j}$.

Proof. Suppose $V((K_{2,n})^{i,j})$ is $V((K_{2,n})^{i,j}) = \{v_s : s \in \{0, 1, 2, \dots, n+1\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_2 , which can be defined by $\psi_k : V((K_{2,n})^{i,j}) \rightarrow \mathbb{Z}_{2n} \times \{i, j\}$ which is defined by

$$\psi_k(v_s) = s_i, s \in \{0, 1\}, \psi_k(v_{s+2}) = ((2(s-1)) \pmod{2n})_j, s \in \{1, \dots, n\}, \quad (5)$$

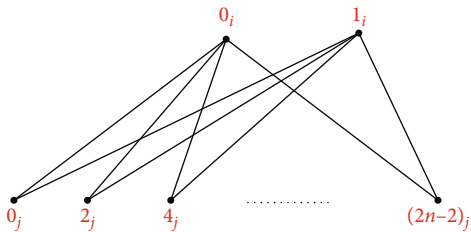


FIGURE 5: The labelling for $(K_{2,n})^{ij}$.

and the edge set of $(K_{2,n})^{ij}$ is

$$E((K_{2,n})^{ij}) = \left\{ \left\{ 0_i, (2s)_j \right\} : s \in \{0, 1, \dots, n-1\} \right\} \cup \left\{ \left\{ 1_i, ((2s) \bmod 2n)_j \right\} : s \in \{1, \dots, n\} \right\}, \tag{6}$$

see Figure 5. From the edge set of G_2 , the following conditions are verified: Each graph $(K_{2,n})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (2n-1)/2 \rfloor\}$, the length 0 is found once in $(K_{2,n})^{ij}$, the length n is found once in $(K_{2,n})^{ij}$, for every $\lambda \in \{1, 2, \dots, \lfloor (2n-1)/2 \rfloor\}$, G_2 has precisely 2.

$\binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_2 , and the length n is found $\binom{m}{2}$ times in G_2 .

Hence, $K_{\underbrace{2n, 2n, \dots, 2n}_m}$ can be decomposed by G_2 . \square

Theorem 7. Let $n \equiv 2 \pmod 6$ or $n \equiv 4 \pmod 6$, $m \geq 2$. Then, there is an orthogonal labelling for

$$G_3 \cong \bigcup_{0 \leq i < j \leq m-1} \left(\frac{n}{2} K_{1,2} \right)^{ij}. \tag{7}$$

$$E((C_8 \cup K_{1,n-8})^{ij}) = \{ \{0, 2_j\}, \{0, 4_j\}, \{4, 2_j\}, \{4, 3_j\}, \{2, 3_j\}, \{2, 5_j\}, \{8, 4_j\}, \{8, 5_j\}, \{1, 1_j\} \} \tag{10}$$

$\cup \{ \{1_i, s_j\} : s \in \{6, 7, \dots, n-4\} \}$, see Figure 7. From the edge set of G_4 , the following conditions are verified: Each graph $(C_8 \cup K_{1,n-8})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(C_8 \cup K_{1,n-8})^{ij}$, the length $n/2$ is found once in $(C_8 \cup K_{1,n-8})^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_4 has precisely 2. $\binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$

Proof. Suppose $V(((n/2)K_{1,2})^{ij}) = \{v_s : s \in \{0, 1, 2, \dots, 2(n-1)\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_3 , which can be defined by $\psi_k : V(((n/2)K_{1,2})^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by $\psi_k(v_s) = s_i, s \in \{0, 1, \dots, n-1\}$, $\psi_k(v_{n+s}) = ((2s) \bmod n)_j, s \in \{0, 1, \dots, n-1\}$, and the edge set of $((n/2)K_{1,2})^{ij}$ is $E(((n/2)K_{1,2})^{ij}) = \{ \{s_i, ((2s) \bmod n)_j\} : s \in \{0, 1, 2, \dots, n-1\} \}$, see Figure 6. From the edge set of G_3 , the following conditions are verified: Each graph $((n/2)K_{1,2})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $((n/2)K_{1,2})^{ij}$, the length $n/2$ is found once in $((n/2)K_{1,2})^{ij}$, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_3 has precisely 2. $\binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_3 , and the length $n/2$ is found $\binom{m}{2}$ times in G_3 . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_3 . \square

Theorem 8. Let $n \geq 9$, $m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_4 \cong \bigcup_{0 \leq i < j \leq m-1} (C_8 \cup K_{1,n-8})^{ij}. \tag{8}$$

Proof. Suppose $V((C_8 \cup K_{1,n-8})^{ij}) = \{v_s : s \in \{0, 1, 2, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_4 , which can be defined by $\psi_k : V((C_8 \cup K_{1,n-8})^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

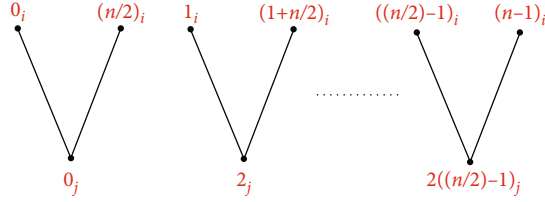
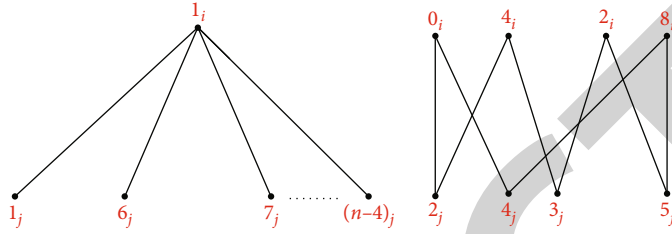
$$\begin{aligned} \psi_k(v_0) &= 0_0, \psi_k(v_1) = 1_0, \psi_k(v_2) = 2_0, \psi_k(v_3) = 4_0, \psi_k(v_4) \\ &= 8_0, \psi_k(v_s) = (s-4)_1, s \in \{5, \dots, n\}, \end{aligned} \tag{9}$$

and the edge set of $(C_8 \cup K_{1,n-8})^{ij}$ is

times in G_4 , and the length $n/2$ is found $\binom{m}{2}$ times in G_4 if n is even. Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_4 . \square

Theorem 9. Let $n \geq 7$, $m \geq 2$ be integers. Then, there is an orthogonal labelling for $G_5 \cong \bigcup_{0 \leq i < j \leq m-1} (C_6 \cup K_{1,1} \cup K_{1,n-7})^{ij}$.

Proof. Suppose $V((K_{1,1} \cup C_6 \cup K_{1,n-7})^{ij}) = \{v_s : s \in \{0, 1, 2, \dots, n\}\}$. The mapping ψ_k can be used to define an

FIGURE 6: The labelling for $((n/2)K_{1,2})^{ij}$.FIGURE 7: The labelling for $(C_8 \cup K_{1,n-8})^{ij}$.

orthogonal labelling for the subgraph G_5 , which can be defined by $\psi_k : V(K_{1,1} \cup C_6 \cup K_{1,n-7}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

$$\psi_k(v_0) = 0_i, \psi_k(v_1) = 1_i, \psi_k(v_2) = 3_i, \psi_k(v_3) = 4_i, \psi_k(v_4) = 6_i, \psi_k(v_5) = 1_j, \psi_k(v_6) = 2_j, \psi_k(v_7) = 3_j, \quad (11)$$

$\psi_k(v_8) = 5_j, \psi_k(v_s) = (s-2)_j, s \in \{9, \dots, n+1\}$, and the edge set of $(K_{1,1} \cup C_6 \cup K_{1,n-7})^{ij}$ is

$$E((K_{1,1} \cup C_6 \cup K_{1,n-7})^{ij}) = \{\{1_i, 1_j\}, \{0_i, 2_j\}, \{0_i, 3_j\}, \{4_i, 2_j\}, \{4_i, 5_j\}, \{6_i, 3_j\}, \{6_i, 5_j\}\} \cup \{\{3_i, s_j\} : s \in \{7, \dots, n-1\}\}, \quad (12)$$

see Figure 8. From the edge set of G_5 , the following conditions are verified: Each graph $(K_{1,1} \cup C_6 \cup K_{1,n-7})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(K_{1,1} \cup C_6 \cup K_{1,n-7})^{ij}$, the length $n/2$ is found once in $(K_{1,1} \cup C_6 \cup K_{1,n-7})^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_5 has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_5 , and the length $n/2$ is found $\binom{m}{2}$ times in G_5 if n is even. Hence, $\underbrace{K_{n,n,\dots,n}}_m$ can be decomposed by G_5 . \square

Theorem 10. Let $n \geq 5, m \geq 2$ be integers. Then, there is an orthogonal labelling for $G_6 \cong \bigcup_{0 \leq i < j \leq m-1} (2K_2 \cup K_{1,n-2})^{ij}$.

Proof. Suppose $V((2K_{1,1} \cup K_{1,n-2})^{ij}) = \{v_s : s \in \{0, 1, 2, \dots, n+2\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_6 , which can be defined by $\psi_k : V((2K_{1,1} \cup K_{1,n-2})^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

$$\psi_k(v_0) = 0_i, \psi_k(v_1) = 1_i, \psi_k(v_2) = (n-1)_i, \psi_k(v_{s+3}) = (s)_j, s \in \{0, \dots, n-1\}, \quad (13)$$

and the edge set of $(2K_{1,1} \cup K_{1,n-2})^{ij}$ is

$$E((2K_{1,1} \cup K_{1,n-2})^{ij}) = \{\{0_i, s_j\} : s \in \{0, 1, \dots, n-3\}\} \cup \{\{1_i, (n-1)_j\}, \{(n-1)_i, (n-2)_j\}\}, \quad (14)$$

see Figure 9. From the edge set of G_6 , the following conditions are verified: Each graph $(2K_{1,1} \cup K_{1,n-2})^{ij}$ has

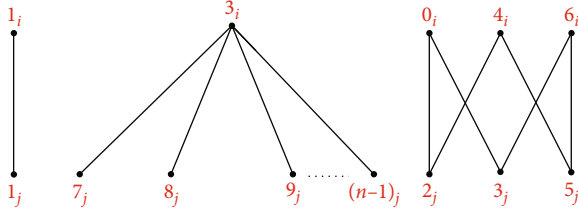


FIGURE 8: The labelling for $(K_{1,1} \cup C_6 \cup K_{1,n-7})^{ij}$.

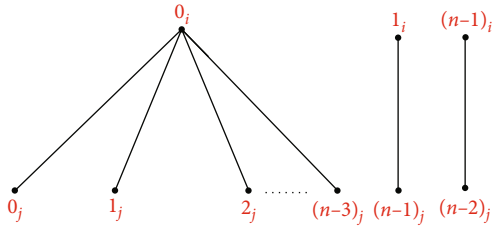


FIGURE 9: The labelling for $(2K_{1,1} \cup K_{1,n-2})^{ij}$.

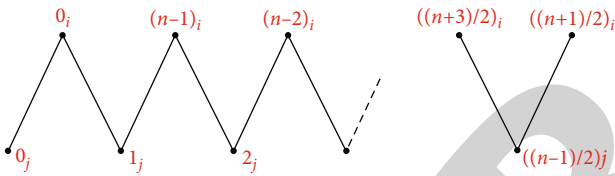


FIGURE 10: The labelling for $(P_{n+1})^{ij}$.

precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(2K_{1,1} \cup K_{1,n-2})^{ij}$, the length $n/2$ is found once in $(2K_{1,1} \cup K_{1,n-2})^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_6 has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_6 , and the length $n/2$ is found $\binom{m}{2}$ times in G_6 if n is even. Hence, $\underbrace{K_{n,n,\dots,n}}_m$ can be decomposed by G_6 . \square

Theorem 11. Let $n > 1, m \geq 2$ be integers. Then, there is an orthogonal labelling for $G_7 \cong \bigcup_{0 \leq i < j \leq m-1} (P_{n+1})^{ij}$.

Proof. Suppose $V((P_{n+1})^{ij}) = \{v_s : s \in \{0, 1, 2, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_7 , which can be defined by $\psi_k : V((P_{n+1})^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

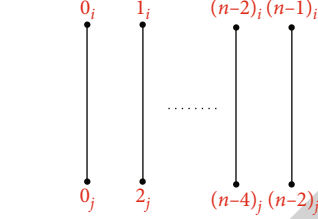


FIGURE 11: The labelling for $(nK_{1,1})^{ij}$.

$$\begin{aligned} \psi_k(v_s) &= ((n-s) \pmod n)_i, s \in \left\{0, 1, \dots, \frac{n-3}{2}\right\}, \psi_k(v_{n-1/2}) \\ &= \left(\frac{n+1}{2}\right)_i, \psi_k(v_{n-3/2+s+2}) = s_j, \end{aligned} \tag{15}$$

$s \in \{0, 1, \dots, (n-1)/2\}$, and the edge set of $(P_{n+1})^{ij}$ is $E((P_{n+1})^{ij}) = \{((n+1)/2)_i, ((n-1)/2)_j\} \cup$

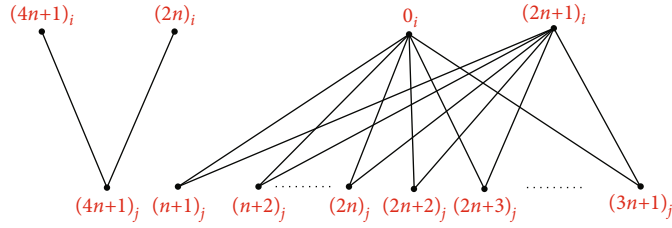
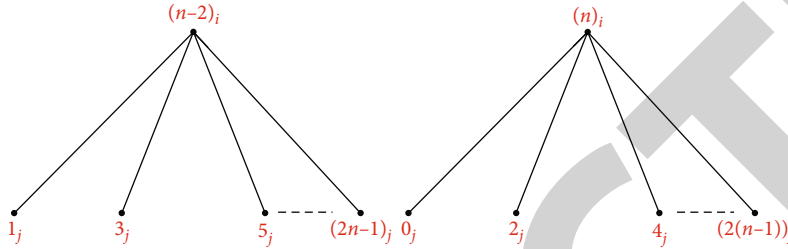
$\{((n-s) \pmod n)_i, (s+\alpha)_j\} : s \in \{0, 1, \dots, (n-3)/2\}, \alpha \in \{0, 1\}\}$, see Figure 10. From the edge set of G_7 , the following conditions are verified: Each graph $(P_{n+1})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is only present once in $(P_{n+1})^{ij}$, the length $n/2$ is found once in $(P_{n+1})^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_7 has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_7 , and the length $n/2$ is

found $\binom{m}{2}$ times in G_7 if n is even. Hence, $\underbrace{K_{n,n,\dots,n}}_m$ can be decomposed by G_7 . \square

Theorem 12. Let $n \equiv 1 \pmod 6, n \equiv 5 \pmod 6, m \geq 2$ be an integer. Then, there is an orthogonal labelling for $\underbrace{K_{n,n,\dots,n}}_m$ by

$$G_8 \cong \bigcup_{0 \leq i < j \leq m-1} (nK_{1,1})^{ij}.$$

Proof. Suppose $V((nK_{1,1})^{ij})$ is $V(nK_{1,1})^{ij} = \{v_s : s \in \{0, 1, 2, \dots, 2n-1\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_8 , which can be defined by $\psi_k : V((nK_{1,1})^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by $\psi_k(v_s) = s_i, s \in \{0, 1, \dots, n-1\}, \psi_k(v_{n+s-1}) = ((2(s-1)) \pmod n)_j, s \in \{1, 2, \dots, n\}$, and the edge set of $(nK_{1,1})^{ij}$ is $E((nK_{1,1})^{ij}) = \{s_i, ((2s) \pmod n)_j\} : s \in \{0, 1, \dots, n-1\}$, see Figure 11. From the edge set of G_8 , the following conditions are verified: Each graph $(nK_{1,1})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(nK_{1,1})^{ij}$, the length $n/2$ is found once in $(nK_{1,1})^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_8 has precisely $2 \cdot \binom{m}{2} = m(m$

FIGURE 12: The labelling for $(K_{1,2} \cup K_{2,2n})^{i,j}$.FIGURE 13: The labelling for $(2K_{1,n})^{i,j}$.

– 1) edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_8 , and the length $n/2$ is found $\binom{m}{2}$ times in G_8 if n is even. Hence, $\underbrace{K_{n,n,\dots,n}}_m$ can be decomposed by G_8 . \square

Theorem 13. Let $n \geq 1, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_9 \cong \bigcup_{0 \leq i < j \leq m-1} (K_{1,2} \cup K_{2,2n})^{i,j}. \quad (16)$$

Proof. Suppose $V((K_{1,2} \cup K_{2,2n})^{i,j}) = \{v_s : s \in \{0, 1, 2, \dots, 2n+4\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_9 , which can be defined by $\psi_k : V((K_{1,2} \cup K_{2,2n})^{i,j}) \rightarrow \mathbb{Z}_{4n+2} \times \{i, j\}$ which is defined by

$$\psi_k(v_0) = (4n+1)_i, \psi_k(v_1) = (2n)_i, \psi_k(v_2) = 0_i, \psi_k(v_3) = (2n+1)_i, \psi_k(v_4) = (4n+1)_j, G_{10} \cong \bigcup_{0 \leq i < j \leq m-1} (2K_{1,n})^{i,j}. \quad (19)$$

$\psi_k(v_{s+4}) = (n+s-4)_j, s \in \{5, \dots, n+4\}$, $\psi_k(v_{n+s}) = (2n+s-3)_j, s \in \{5, \dots, n+4\}$, and the edge set of $(K_{1,2} \cup K_{2,2n})^{i,j}$ is

$$E((K_{1,2} \cup K_{2,2n})^{i,j}) = \left\{ \left\{ (4n+1)_i, (4n+1)_j \right\}, \left\{ (2n)_i, (4n+1)_j \right\} \right. \\ \cup \left\{ \left\{ 0_i, s_j \right\}, \left\{ (2n+1)_i, s_j \right\} : s \in \{n+1, \dots, 2n\} \right\} \\ \left. \cup \left\{ \left\{ 0_i, s_j \right\}, \left\{ (2n+1)_i, s_j \right\}, s \in \{2n+2, \dots, 3n+1\} \right\} \right\}, \quad (18)$$

see Figure 12. From the edge set of G_9 , the following conditions are verified: Each graph $(K_{1,2} \cup K_{2,2n})^{i,j}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (4n+1)/2 \rfloor\}$, the length 0 is found once in $(K_{1,2} \cup K_{2,2n})^{i,j}$, the length $2n+1$ is found once in $(K_{1,2} \cup K_{2,2n})^{i,j}$, for every $\lambda \in \{1, 2, \dots, \lfloor (4n+1)/2 \rfloor\}$, G_9 has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_9 , and the length $2n+1$ is found $\binom{m}{2}$ times in G_9 . Hence, $\underbrace{K_{(4n+2), (4n+2), \dots, (4n+2)}}_m$ can be decomposed by G_9 . \square

Theorem 14. Let $n \geq 2, m \geq 2$ be integers. Then, there is an orthogonal labelling for

Proof. Suppose $V((2K_{1,n})^{i,j}) = \{v_s : s \in \{0, 1, 2, \dots, 2n+1\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{10} , which can be defined by $\psi_k : V((2K_{1,n})^{i,j}) \rightarrow \mathbb{Z}_{2n} \times \{i, j\}$ which is defined by $\psi_k(v_0) = (n-2)_i, \psi_k(v_1) = n_i, \psi_k(v_{s+2}) = s_j, s \in \{0, \dots, 2n-1\}$, and the edge set of $(2K_{1,n})^{i,j}$ is $E((2K_{1,n})^{i,j}) = \left\{ \left\{ n_i, (2s+1)_j \right\}, \left\{ (n-2)_i, (2s)_j \right\} : s \in \{0, 1, \dots, n-1\} \right\}$, see Figure 13. From the edge set of G_{10} , the following conditions are verified: Each graph $(2K_{1,n})^{i,j}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (2n-1)/2 \rfloor\}$, the length 0 is found once in

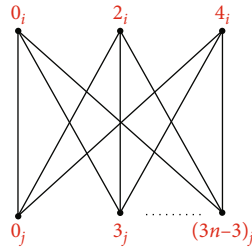


FIGURE 14: The labelling for $(K_{3,n})^{ij}$.

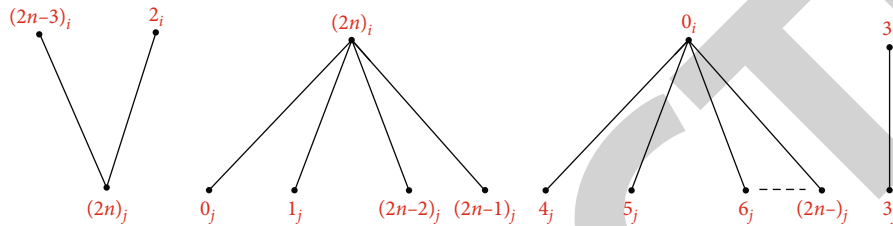


FIGURE 15: The labelling for $(K_{1,1} \cup K_{1,2} \cup K_{1,4} \cup K_{1,2n-6})^{ij}$.

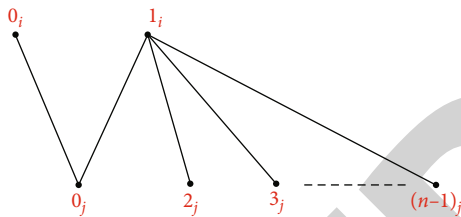


FIGURE 16: The labelling for $(C_2(1, n-2))^{ij}$.

$(2K_{1,n})^{ij}$, the length n is found once in $(2K_{1,n})^{ij}$, for every $\lambda \in \{1, 2, \dots, \lfloor (2n-1)/2 \rfloor\}$, G_{10} has precisely 2.

$\binom{m}{2} = m(m-1)$ edges of length λ , the length n is found $\binom{m}{2}$ times in G_{10} , and the length 0 is found $\binom{m}{2}$ times in G_{10} . Hence, $\underbrace{K_{2n, 2n, \dots, 2n}}_m$ can be decomposed by G_{10} . \square

Theorem 15. For all positive integers n with $\gcd(n, 3) = 1$, $m \geq 2$. Then, there is an orthogonal labelling for

$$G_{11} \cong \bigcup_{0 \leq i < j \leq m-1} (nK_{2,2})^{ij}. \quad (20)$$

Proof. Suppose $V((nK_{2,2})^{ij})$ is $V((nK_{2,2})^{ij}) = \{v_s : s \in \{0, 1, 2, \dots, 4n-1\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{11} , which can be defined by $\psi_k : V((nK_{2,2})^{ij}) \rightarrow \mathbb{Z}_{4n} \times \{i, j\}$ which is defined by

$$\begin{aligned} \psi_k(v_s) &= s_i, s \in \{0, 1, \dots, 2n-1\}, \psi_k(v_{2n+s}) \\ &= ((2s) \pmod{4n})_j, s \in \{0, 1, \dots, 2n-1\}, \end{aligned} \quad (21)$$

and the edge set of $(nK_{2,2})^{ij}$ is

$$\begin{aligned} E((nK_{2,2})^{ij}) &= \left\{ \{s_i, (2s)_j\} : s \in \{0, 1, \dots, 2n-1\} \right\} \\ &\cup \left\{ \{(s-2n)_i, ((2s-2n) \pmod{4n})_j\} : s \in \{2n, \dots, 4n-1\} \right\}. \end{aligned} \quad (22)$$

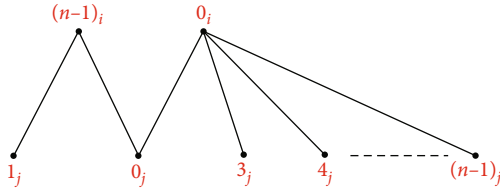
From the edge set of G_{11} , the following conditions are verified: Each graph $(nK_{2,2})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (4n-1)/2 \rfloor\}$, the length 0 is found once in $(nK_{2,2})^{ij}$, the length $2n$ is found once in $(nK_{2,2})^{ij}$, for every $\lambda \in \{1, 2, \dots, \lfloor (4n-1)/2 \rfloor\}$, G_{11} has precisely 2.

$\binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_{11} , and the length $2n$ is found $\binom{m}{2}$ times in G_{11} . Hence, $\underbrace{K_{4n, 4n, \dots, 4n}}_m$ can be decomposed by G_{11} . \square

Theorem 16. Let $n \geq 3, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{12} \cong \bigcup_{0 \leq i < j \leq m-1} (K_{3,n})^{ij}. \quad (23)$$

Proof. Suppose $V((K_{3,n})^{ij}) = \{v_s : s \in \{0, 1, 2, \dots, 2n+4\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{12} , which can be defined by $\psi_k : V((K_{3,n})^{ij}) \rightarrow \mathbb{Z}_{3n} \times \{i, j\}$ which is defined by $\psi_k(v_0) = 0_i$, $\psi_k(v_1) = 2_i$, $\psi_k(v_2) = 4_i$, $\psi_k(v_s) = (3(s-3))_j, s \in \{3, \dots, n+2\}$, and the edge set of $(K_{3,n})^{ij}$ is $E((K_{3,n})^{ij}) = \{\{a_i, b_j\} : a \in \{0, 2, 4\}, b \in \{0, 3, 6, \dots, 3n-3\}\}$, see Figure 14. From the

FIGURE 17: The labelling for $(C_3(1, 0, n-3))^{ij}$.

edge set of G_{12} , the following conditions are verified: Each graph $(K_{3,n})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (3n-1)/2 \rfloor\}$, the length 0 is only present once in $(K_{3,n})^{ij}$, the length $3n/2$ is found once in $(K_{3,n})^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (3n-1)/2 \rfloor\}$, G_{12} has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length 0 is found $\binom{m}{2}$ times in G_{12} , and the length $3n/2$ is found $\binom{m}{2}$ times in G_{12} if n is even. Hence, $\underbrace{K_{3n, 3n, \dots, 3n}}_m$ can be decomposed by G_{12} .

□ and the edge set of $(K_{1,1} \cup K_{1,2} \cup K_{1,4} \cup K_{1,2n-6})^{ij}$ is

$$E((K_{1,1} \cup K_{1,2} \cup K_{1,4} \cup K_{1,2n-6})^{ij}) = \left\{ \{3_i, 3_j\}, \{(2n-3)_i, (2n)_j\}, \{2_i, (2n)_j\}, \{(2n)_i, 0_j\}, \{(2n)_i, 1_j\}, \{(2n)_i, (2n-2)_j\}, \{(2n)_i, (2n-1)_j\} \right\} \cup \{ \{0_i, a_j\} : a \in \{4, 5, \dots, 2n-3\} \}, \quad (26)$$

see Figure 15. From the edge set of G_{13} , the following conditions are verified: Each graph $(K_{1,1} \cup K_{1,2} \cup K_{1,4} \cup K_{1,2n-6})^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, n\}$, the length 0 is found once in $(K_{1,1} \cup K_{1,2} \cup K_{1,4} \cup K_{1,2n-6})^{ij}$, for every $\lambda \in \{1, 2, \dots, n\}$, G_{13} has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , and the length 0 is found $\binom{m}{2}$ times in G_{13} . Hence, $\underbrace{K_{(2n+1)n, (2n+1)n, \dots, (2n+1)n}}_m$ can be decomposed by G_{13} .

□

4. Decompositions of $\underbrace{K_{n,n,\dots,n}}_m$ by Connected Caterpillars

Theorem 18. Let $n \geq 2, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{14} \cong \bigcup_{0 \leq i < j \leq m-1} (C_2(1, n-2))^{ij}, \quad (27)$$

Proof. Suppose $V((C_2(1, n-2))^{ij}) = \{v_s : s \in \{0, 1, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{14} , which can be defined by

Theorem 17. Let $n \geq 4, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{13} \cong \bigcup_{0 \leq i < j \leq m-1} (K_{1,1} \cup K_{1,2} \cup K_{1,4} \cup K_{1,2n-6})^{ij}. \quad (24)$$

Proof. Suppose $V((K_{1,1} \cup K_{1,2} \cup K_{1,4} \cup K_{1,2n-6})^{ij}) = \{v_s : s \in \{0, 1, 2, \dots, 2n+4\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{13} , which can be defined by $\psi_k : V((K_{1,1} \cup K_{1,2} \cup K_{1,4} \cup K_{1,2n-6})^{ij}) \rightarrow \mathbb{Z}_{2n+1} \times \{i, j\}$ which is defined by

$$\begin{aligned} \psi_k(v_0) &= 3_i, \psi_k(v_1) = (2n-3)_i, \psi_k(v_2) = 2_i, \psi_k(v_3) \\ &= 0_i, \psi_k(v_4) = (2n)_i, \psi_k(v_5) = 3_j, \psi_k(v_6) \\ &= (2n)_j, \psi_k(v_7) = 0_j, \psi_k(v_8) = 1_j, \psi_k(v_9) \\ &= (2n-2)_j, \psi_k(v_{10}) = (2n-1)_j, \psi_k(v_{s+1}) \\ &= (s-6)_j, s \in \{10, \dots, 2n+3\}, \end{aligned} \quad (25)$$

$\psi_k : V((C_2(1, n-2))^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

$$\begin{aligned} \psi_k(v_0) &= 0_i, \psi_k(v_1) = 1_i, \psi_k(v_2) = 0_j, \psi_k(v_s) \\ &= (s-1)_j, s \in \{3, 4, \dots, n\}, \end{aligned} \quad (28)$$

and the edge set of $(C_2(1, n-2))^{ij}$ is

$$E((C_2(1, n-2))^{ij}) = \left\{ \{0_i, 0_j\}, \{1_i, 0_j\} \right\} \cup \{ \{1_i, s_j\} : s \in \{2, 3, \dots, n-1\} \}, \quad (29)$$

see Figure 16. From the edge set of G_{14} , the following conditions are verified: Each graph $(C_2(1, n-2))^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(C_2(1, n-2))^{ij}$, the length $n/2$ is found once in $(C_2(1, n-2))^{ij}$, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_{14} has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times in G_{14} , and the length 0 is found $\binom{m}{2}$ times in G_{14} . Hence, $\underbrace{K_{n,n,\dots,n}}_m$ can be decomposed by G_{14} . □

Theorem 19. Let $n \geq 3, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{15} \cong \bigcup_{0 \leq i < j \leq m-1} (C_3(1, 0, n-3))^{ij}. \quad (30)$$

Proof. Suppose $V((C_3(1, 0, n-3))^{ij}) = \{v_s : s \in \{0, 1, \dots, n\}\}$.

$$E((C_3(1, 0, n-3))^{ij}) = \left\{ \{0_i, 0_j\}, \{(n-1)_i, 0_j\}, \{(n-1)_i, 1_j\} \right\} \cup \left\{ \{0_i, s_j\} : s \in \{3, 4, \dots, n-1\} \right\}, \quad (31)$$

see Figure 17. From the edge set of G_{15} , the following conditions are verified: Each graph $(C_3(1, 0, n-3))^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(C_3(1, 0, n-3))^{ij}$, the length $n/2$ is found once in $(C_3(1, 0, n-3))^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_{15} has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times in G_{15} if n is even, and the length 0 is found $\binom{m}{2}$ times in G_{15} . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_{15} . \square

$$\psi_k(v_0) = 0_i, \psi_k(v_1) = (n-1)_i, \psi_k(v_2) = 0_j, \psi_k(v_3) = 1_j, \psi_k(v_4) = (n-2)_j, \psi_k(v_{s+3}) = s_j, \quad (33)$$

$s \in \{2, 3, \dots, n-3\}$, and the edge set of $(C_4(1, 0, 0, n-4))^{ij}$ is

$$E((C_4(1, 0, 0, n-4))^{ij}) = \left\{ \{0_i, 0_j\}, \{0_i, 1_j\}, \{(n-1)_i, (n-2)_j\}, \{(n-1)_i, 1_j\} \right\} \cup \left\{ \{s_i, 0_j\} : s \in \{2, 3, \dots, n-3\} \right\}, \quad (34)$$

see Figure 18. From the edge set of G_{16} , the following conditions are verified: Each graph $(C_4(1, 0, 0, n-4))^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(C_4(1, 0, 0, n-4))^{ij}$, the length $n/2$ is found once in $(C_4(1, 0, 0, n-4))^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_{16} has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times

The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{15} , which can be defined by $\psi_k : V((C_3(1, 0, n-3))^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by $\psi_k(v_0) = 0_i, \psi_k(v_1) = (n-1)_i, \psi_k(v_2) = 0_j, \psi_k(v_3) = 1_j, \psi_k(v_s) = (s-1)_j, s \in \{4, 5, \dots, n\}$, and the edge set of $(C_3(1, 0, n-3))^{ij}$ is

Theorem 20. Let $n \geq 4, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{16} \cong \bigcup_{0 \leq i < j \leq m-1} (C_4(1, 0, 0, n-4))^{ij}. \quad (32)$$

Proof. Suppose $V((C_4(1, 0, 0, n-4))^{ij}) = \{v_s : s \in \{0, 1, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{16} , which can be defined by $\psi_k : V((C_4(1, 0, 0, n-4))^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

in G_{16} if n is even, and the length 0 is found $\binom{m}{2}$ times in G_{16} . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_{16} . \square

Theorem 21. Let $n \geq 5, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{17} \cong \bigcup_{0 \leq i < j \leq m-1} (C_5(1, 0, 0, 0, n-5))^{ij}. \quad (35)$$

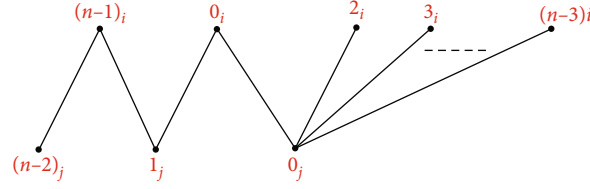


FIGURE 18: The labelling for $(C_4(1, 0, 0, n - 4))^{ij}$.

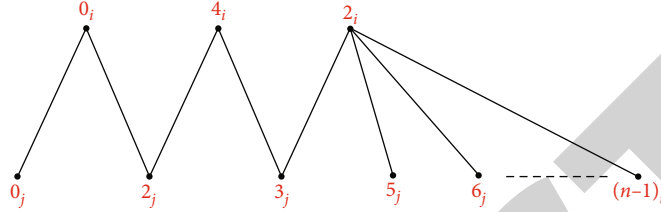


FIGURE 19: The labelling for $(C_5(1, 0, 0, 0, n - 5))^{ij}$.

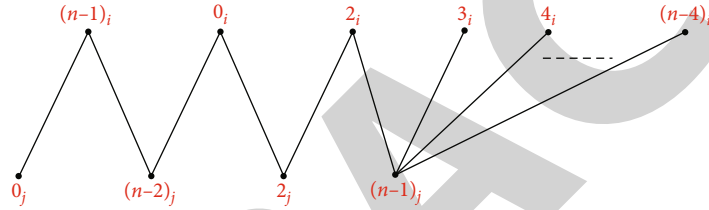


FIGURE 20: The labelling for $(C_6(1, 0, 0, 0, 0, n - 6))^{ij}$.

Proof. Suppose $V((C_5(1, 0, 0, 0, n - 5))^{ij}) = \{v_s : s \in \{0, 1, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{17} , which can be defined by $\psi_k : V((C_5(1, 0, 0, 0, n - 5))^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

$$\begin{aligned} \psi_k(v_0) = 0_j, \psi_k(v_1) = 0_i, \psi_k(v_2) = 2_j, \psi_k(v_3) = 4_i, \psi_k(v_4) = 3_j, \\ \psi_k(v_5) = 2_i, \psi_k(v_{s+1}) = s_j, s \in \{5, 6, \dots, n - 1\}, \end{aligned} \tag{36}$$

and the edge set of $(C_5(1, 0, 0, 0, n - 5))^{ij}$ is

$$E((C_5(1, 0, 0, 0, n - 5))^{ij}) = \{\{0_i, 0_j\}, \{2_i, 3_j\}, \{0_i, 2_j\}, \{4_i, 2_j\}, \{4_i, 3_j\}\} \cup \{\{2_i, s_j\} : s \in \{5, 6, \dots, n - 1\}\}, \tag{37}$$

see Figure 19. From the edge set of G_{17} , the following conditions are verified: Each graph $(C_5(1, 0, 0, 0, n - 5))^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n - 1)/2 \rfloor\}$, the length 0 is found once in $(C_5(1, 0, 0, 0, n - 5))^{ij}$, the length $n/2$ is found once in $(C_5(1, 0, 0, 0, n - 5))^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n - 1)/2 \rfloor\}$, G_{17} has precisely $2 \cdot \binom{m}{2} = m(m - 1)$

edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times in G_{17} if n is even, and the length 0 is found $\binom{m}{2}$ times in G_{17} . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_{17} . \square

Theorem 22. Let $n \geq 6, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{18} \cong \bigcup_{0 \leq i < j \leq m-1} (C_6(1, 0, 0, 0, 0, n-6))^{ij}. \quad (38)$$

Proof. Suppose $V((C_6(1, 0, 0, 0, 0, n-6))^{ij}) = \{v_s : s \in \{0, 1,$

$$\begin{aligned} \psi_k(v_0) &= 0_j, \psi_k(v_1) = (n-1)_i, \psi_k(v_2) = (n-2)_j, \psi_k(v_3) \\ &= 0_i, \psi_k(v_4) = 2_j, \psi_k(v_5) = 2_i. \end{aligned} \quad (39)$$

$\psi_k(v_6) = (n-1)_j, \psi_k(v_{s+4}) = s_i, s \in \{3, 4, \dots, n-4\}$, and the edge set of $(C_6(1, 0, 0, 0, 0, n-6))^{ij}$ is

$$E((C_6(1, 0, 0, 0, 0, n-6))^{ij}) = \left\{ \{(n-1)_i, 0_j\}, \{(n-1)_i, (n-2)_j\}, \{0_i, (n-2)_j\}, \{0_i, 2_j\}, \{2_i, 2_j\}, \{2_i, (n-1)_j\} \right\} \cup \left\{ \{s_i, (n-1)_j\} : s \in \{3, 4, \dots, n-4\} \right\}, \quad (40)$$

see Figure 20. From the edge set of G_{18} , the following conditions are verified: Each graph $(C_6(1, 0, 0, 0, 0, n-6))^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is only present once in $(C_6(1, 0, 0, 0, 0, n-6))^{ij}$, the length $n/2$ is found once in $(C_6(1, 0, 0, 0, 0, n-6))^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_{18} has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times in G_{18} if n is even, and the length 0 is found $\binom{m}{2}$ times in G_{18} . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_{18} . \square

$$E((C_7(1, 0, 0, 0, 0, 0, n-7))^{ij}) = \left\{ \{1_i, 2_j\}, \{1_i, 0_j\}, \{0_i, 0_j\}, \{0_i, 3_j\}, \{6_i, 3_j\}, \{6_i, 4_j\}, \{2_i, 4_j\} \right\} \cup \left\{ \{2_i, s_j\} : s \in \{6, 7, \dots, n-2\} \right\}, \quad (42)$$

see Figure 21. From the edge set of G_{19} , the following conditions are verified: Each graph $(C_7(1, 0, 0, 0, 0, 0, n-7))^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is only present once in $(C_7(1, 0, 0, 0, 0, 0, n-7))^{ij}$, the length $n/2$ is found once in $(C_7(1, 0, 0, 0, 0, 0, n-7))^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_{19} has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times in G_{19} if n is even, and the length 0 is

$\dots, n\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{18} , which can be defined by $\psi_k : V((C_6(1, 0, 0, 0, 0, n-6))^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

Theorem 23. Let $n \geq 7, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{19} \cong \bigcup_{0 \leq i < j \leq m-1} (C_7(1, 0, 0, 0, 0, 0, n-7))^{ij}. \quad (41)$$

Proof. Suppose $V((C_7(1, 0, 0, 0, 0, 0, n-7))^{ij}) = \{v_s : s \in \{0, 1, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{19} , which can be defined by $\psi_k : V((C_7(1, 0, 0, 0, 0, 0, n-7))^{ij}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by $\psi_k(v_0) = 2_j, \psi_k(v_1) = 1_i, \psi_k(v_2) = 0_j, \psi_k(v_3) = 0_i, \psi_k(v_4) = 3_j, \psi_k(v_5) = 6_i, \psi_k(v_6) = 4_j, \psi_k(v_7) = 2_i, \psi_k(v_{s+2}) = s_j, s \in \{6, 7, \dots, n-2\}$, and the edge set of $(C_7(1, 0, 0, 0, 0, 0, n-7))^{ij}$ is

found $\binom{m}{2}$ times in G_{19} . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_{19} . \square

Theorem 24. Let $n \geq 8, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{20} \cong \bigcup_{0 \leq i < j \leq m-1} (C_8(1, 0, 0, 0, 0, 0, 0, n-8))^{ij}. \quad (43)$$

Proof. Suppose $V((C_8(1, 0, 0, 0, 0, 0, 0, n-8))^{ij}) = \{v_s : s \in \{0, 1, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal

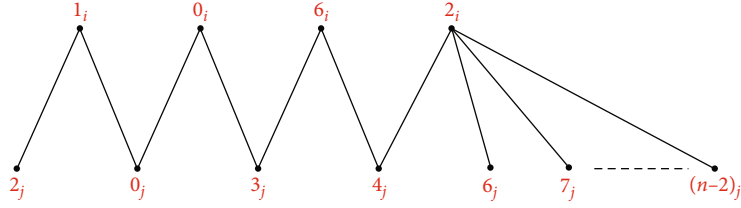


FIGURE 21: The labelling for $(C_7(1, 0, 0, 0, 0, 0, n - 7))^{i,j}$.

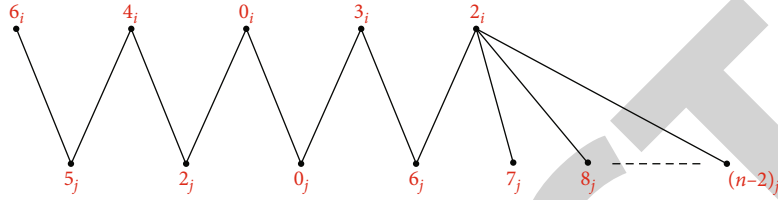


FIGURE 22: The labelling for $(C_8(1, 0, 0, 0, 0, 0, 0, n - 8))^{i,j}$.

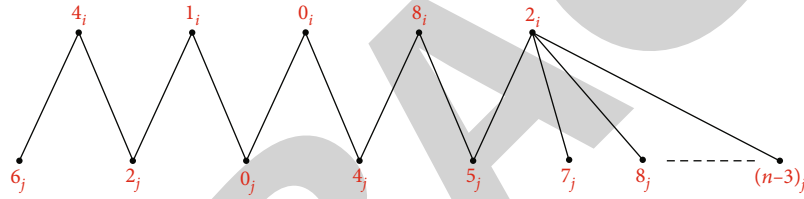


FIGURE 23: The labelling for $(C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}$.

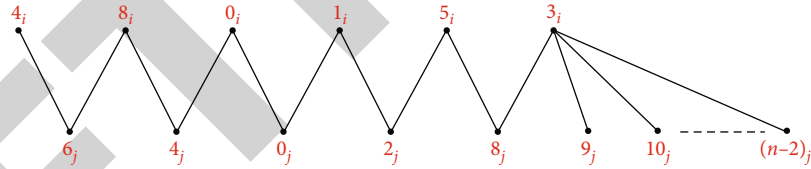


FIGURE 24: The labelling for $(C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n - 10))^{i,j}$.

labelling for the subgraph G_{20} , which can be defined by $\psi_k : V((C_8(1, 0, 0, 0, 0, 0, 0, n - 8))^{i,j}) \rightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by-

$\psi_k(v_0) = 6_i, \psi_k(v_1) = 5_j, \psi_k(v_2) = 4_i, \psi_k(v_3) = 2_j, \psi_k(v_4) = 0_i,$
 $\psi_k(v_5) = 0_j, \psi_k(v_6) = 3_i, \psi_k(v_7) = 6_j,$

$\psi_k(v_8) = 2_i, \varphi(v_{i+2}) = i_1, i \in \{7, 8, \dots, n - 2\}$, and the edge set of $(C_8(1, 0, 0, 0, 0, 0, 0, n - 8))^{i,j}$ is

$$E((C_8(1, 0, 0, 0, 0, 0, 0, n - 8))^{i,j}) = \{\{6_i, 5_j\}, \{4_i, 2_j\}, \{0_i, 2_j\}, \{0_i, 0_j\}, \{3_i, 0_j\}, \{3_i, 6_j\}, \{2_i, 6_j\}\} \cup \{\{2_i, s_j\} : s \in \{7, 8, \dots, n - 2\}\}, \quad (44)$$

see Figure 22. From the edge set of G_{20} , the following conditions are verified: Each graph $(C_8(1, 0, 0, 0, 0, 0, 0, n - 8))^{i,j}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n - 1)/2 \rfloor\}$, the length 0 is found once in $(C_8(1, 0, 0, 0, 0, 0, 0, n - 8))^{i,j}$, the length $n/2$ is found once in $(C_8(1, 0, 0, 0, 0, 0, 0, n - 8))^{i,j}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n - 1)/2 \rfloor\}$, G_{20} has precisely $2 \cdot \binom{m}{2} = m(m - 1)$ edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times in G_{20} if n is even, and the length 0 is found $\binom{m}{2}$ times in G_{20} . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_{20} . \square

$$E((C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}) = \{\{4_i, 6_j\}, \{4_i, 2_j\}, \{1_i, 2_j\}, \{1_i, 0_j\}, \{0_i, 0_j\}, \{0_i, 4_j\}, \{8_i, 4_j\}, \{8_i, 5_j\}, \{2_i, 5_j\}\} \cup \{\{2_i, s_j\} : s \in \{7, 8, \dots, n - 3\}\}, \tag{46}$$

see Figure 23. From the edge set of G_{21} , the following conditions are verified: Each graph $(C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n - 1)/2 \rfloor\}$, the length 0 is found once in $(C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}$, the length $n/2$ is found once in $(C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n - 1)/2 \rfloor\}$, G_{21} has precisely $2 \cdot \binom{m}{2} = m(m - 1)$ edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times in G_{21} if n is even, and the length 0 is found $\binom{m}{2}$ times in G_{21} . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_{21} . \square

$$\begin{aligned} \psi_k(v_0) &= 4_i, \\ \psi_k(v_1) &= 6_j, \psi_k(v_2) = 8_i, \psi_k(v_3) = 4_j, \psi_k(v_4) = 0_i, \psi_k(v_5) = 0_j, \psi_k(v_6) = 1_i, \psi_k(v_7) \\ &= 2_j, \psi_k(v_8) = 5_i, \psi_k(v_9) = 8_j, \psi_k(v_{10}) = 3_i, \psi_k(v_{s+2}) = s_j, s \in \{9, 10, \dots, n - 2\}, \end{aligned} \tag{48}$$

and the edge set of $(C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n - 10))^{i,j}$ is

$$E((C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n - 10))^{i,j}) = \{\{4_i, 6_j\}, \{8_i, 6_j\}, \{8_i, 4_j\}, \{0_i, 4_j\}, \{0_i, 0_j\}, \{1_i, 0_j\}, \{1_i, 2_j\}, \{5_i, 2_j\}, \{5_i, 8_j\}\} \cup \{\{3_i, s_j\} : s \in \{8, 9, \dots, n - 2\}\}, \tag{49}$$

Theorem 25. Let $n \geq 9, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{21} \cong \bigcup_{0 \leq i < j \leq m-1} (C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}. \tag{45}$$

Proof. Suppose $V((C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}) = \{v_s : s \in \{0, 1, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{21} , which can be defined by $\psi_k : V((C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}) \longrightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by-
 $\psi_k(v_0) = 6_j, \psi_k(v_1) = 4_i, \psi_k(v_2) = 2_j, \psi_k(v_3) = 1_i, \psi_k(v_4) = 0_j,$
 $\psi_k(v_5) = 0_i, \psi_k(v_6) = 4_j, \psi_k(v_7) = 8_i,$
 $\psi_k(v_8) = 5_j, \psi_k(v_9) = 2_i, \psi_k(v_{s+3}) = s_j, s \in \{7, 8, \dots, n - 3\},$ and the edge set of $(C_9(1, 0, 0, 0, 0, 0, 0, 0, n - 9))^{i,j}$ is

Theorem 26. Let $n \geq 10, m \geq 2$ be integers. Then, there is an orthogonal labelling for

$$G_{22} \cong \bigcup_{0 \leq i < j \leq m-1} (C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n - 10))^{i,j}. \tag{47}$$

Proof. Suppose $V((C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n - 10))^{i,j}) = \{v_s : s \in \{0, 1, \dots, n\}\}$. The mapping ψ_k can be used to define an orthogonal labelling for the subgraph G_{22} , which can be defined by $\psi_k : V((C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n - 10))^{i,j}) \longrightarrow \mathbb{Z}_n \times \{i, j\}$ which is defined by

see Figure 24. From the edge set of G_{22} , the following conditions are verified: Each graph $(C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n-10))^{ij}$ has precisely two edges of length $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, the length 0 is found once in $(C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n-10))^{ij}$, the length $n/2$ is found once in $(C_{10}(1, 0, 0, 0, 0, 0, 0, 0, 0, n-10))^{ij}$ if n is even, for every $\lambda \in \{1, 2, \dots, \lfloor (n-1)/2 \rfloor\}$, G_{22} has precisely $2 \cdot \binom{m}{2} = m(m-1)$ edges of length λ , the length $n/2$ is found $\binom{m}{2}$ times in G_{22} if n is even, and the length 0 is found $\binom{m}{2}$ times in G_{22} . Hence, $K_{\underbrace{n, n, \dots, n}_m}$ can be decomposed by G_{22} . \square

5. Conclusion

As known, there are several types of graphs labelling. Herein, we are concerned with orthogonal labelling notion. As a generalization to the orthogonal labelling approach provided in the literature for finding the decomposition of circulant-balanced complete bipartite graphs $K_{n,n}$, we have developed a generalized orthogonal labelling approach for decomposing the circulant-balanced complete multipartite graphs $K_{\underbrace{n, n, \dots, n}_m}$; $m, n \geq 2$, in this study. In the future, we will work to improve the orthogonal labelling approach so that it may be used with all types of circulant graphs.

Nomenclatures

K_m : Complete graph having m vertices
 kH : k disjoint unions of graph H
 $K_{m,n}$: Complete bipartite graph with size $m+n$, where the vertex set is divided into two sets with sizes m and n
 C_x : Cycle graph on x vertices
 P_m : Path graph on m vertices
 $V(G)$: Vertex set of graph G
 $E(G)$: Edge set of graph G
 $G \cup H$: Disjoint union of graphs G and H .

Data Availability

The data used to support the findings of this study are available from the corresponding author on request.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] A. Raheem, R. Hasni, M. Javaid, and M. A. Umar, "On cordial related labeling of isomorphic copies of paths and subdivision of star," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 23, no. 7, pp. 1381–1390, 2020.
- [2] H. M. A. Siddiqui, S. Hayat, A. Khan, M. Imran, A. Razzaq, and J.-B. Liu, "Resolvability and fault-tolerant resolvability structures of convex polytopes," *Theoretical Computer Science*, vol. 796, pp. 114–128, 2019.
- [3] J.-B. Liu, M. F. Nadeem, H. M. A. Siddiqui, and W. Nazir, "Computing metric dimension of certain families of Toeplitz graphs," *IEEE Access*, vol. 7, pp. 126734–126741, 2019.
- [4] A. El-Mesady, Y. S. Hamed, M. S. Mohamed, and H. Shabana, "Partially balanced network designs and graph codes generation," *AIMS Mathematics*, vol. 7, no. 2, pp. 2393–2412, 2022.
- [5] A. El-Mesady and Y. S. Hamed, "A novel application on mutually orthogonal graph squares and graph-orthogonal arrays," *AIMS Mathematics*, vol. 7, no. 5, pp. 7349–7373, 2022.
- [6] A. Schenker, M. Last, H. Bunke, and A. Kandel, "Clustering of Web documents using a graph model," in *Web Document Analysis: Challenges and Opportunities*, vol. 55, pp. 3–18, World Scientific, Singapore, 2003.
- [7] J.-C. Bermond, F. Comellas, and D. F. Hsu, "Distributed loop computer-networks: a survey," *Journal of Parallel and Distributed Computing*, vol. 24, no. 1, pp. 2–10, 1995.
- [8] F. K. Hwang, "A complementary survey on double-loop networks," *Theoretical Computer Science*, vol. 263, no. 1-2, pp. 211–229, 2001.
- [9] F. K. Hwang, "A survey on multi-loop networks," *Theoretical Computer Science*, vol. 299, no. 1-3, pp. 107–121, 2003.
- [10] C. S. Raghavendra and J. A. Sylvester, "A survey of multi-connected loop topologies for local computer networks," *Computer Networks and ISDN Systems*, vol. 11, no. 1, pp. 29–42, 1986.
- [11] O. G. Monakhov, E. A. Monakhova, A. Y. Romanov, A. M. Sukhov, and E. V. Lezhnev, "Adaptive dynamic shortest path search algorithm in networks-on-chip based on circulant topologies," *IEEE Access*, vol. 9, pp. 160836–160846, 2021.
- [12] Y. Aleksandr, E. V. Romanov, A. Y. Lezhnev, and A. A. Glukhikh, "Development of routing algorithms in networks-on-chip based on two-dimensional optimal circulant topologies," *Heliyon*, vol. 6, no. 1, article e03183, 2020.
- [13] Y. Aleksandr and V. A. Romanov, "Routing in triple loop circulants: a case of networks-on-chip," *Heliyon*, vol. 6, no. 7, article e04427, 2020.
- [14] R. Wilkov, "Analysis and design of reliable computer networks," *IEEE Transactions on Communications*, vol. 20, no. 3, pp. 660–678, 1972.
- [15] B. Bose, B. Broeg, Y. Kwon, and Y. Ashir, "Lee distance and topological properties of k-ary n-cubes," *IEEE Transactions on Computers*, vol. 44, no. 8, pp. 1021–1030, 1995.
- [16] W. J. Bouknight, S. A. Denenberg, M. I. De, J. M. Randall, A. H. Sameh, and D. L. Slotnick, "The illiac IV system," *Proceedings of the IEEE*, vol. 60, no. 4, pp. 369–388, 1972.
- [17] A. T. Balaban, "Reaction Graphs," in *Graph Theoretical Approaches to Chemical Reactivity*, D. Bonchev and O. Mekenyan, Eds., pp. 137–180, Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.
- [18] F. P. Muga and W. E. S. Yu, "A Proposed Topology for a 192-Processor Symmetric Cluster with a Single-Switch Delay," in *First Philippine Computing Science Congress*, p. 10, Manila, Philippines, 2000.
- [19] F. Comellas, M. Mitjana, and J. G. Peters, "Broadcasting in small-world communication graphs, in 9th Int. Coll. Structural Information and Communication Complexity

Retraction

Retracted: Exploration of Flood Prediction in Watersheds Based on the Fusion Analysis of Remote Sensing Big Data with Multiple Physical Fields

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Xu and Y. Ouyang, "Exploration of Flood Prediction in Watersheds Based on the Fusion Analysis of Remote Sensing Big Data with Multiple Physical Fields," *Journal of Function Spaces*, vol. 2022, Article ID 9422553, 10 pages, 2022.

Research Article

Exploration of Flood Prediction in Watersheds Based on the Fusion Analysis of Remote Sensing Big Data with Multiple Physical Fields

Minming Xu  and Yun Ouyang 

School of Mathematics and Physics, Hechi University, Guangxi, Yizhou 546300, China

Correspondence should be addressed to Yun Ouyang; 04013@hcnu.edu.cn

Received 16 May 2022; Revised 2 July 2022; Accepted 6 July 2022; Published 16 July 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Minming Xu and Yun Ouyang. 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.

It is well known that flooding brings great losses to people's production and life, just because it is unsuspected, is extremely extensive, and has high frequency. More than half of the people in China live in flood-prone areas, and their lives and properties are threatened. Flood risk assessment is one of the measures of flood management, and it is economically and socially important to assess flood risk. The scope of the traditional monitoring and forecasting early warning system is mainly limited to the area affected by flash floods, while the objective reality of real-time monitoring, forecasting, and early warning of flood disasters in the area affected by river floods and by the scheduling of riverine terrace power stations is not possible. The multiphysics remote sensing big data fusion analysis can divide the data into grids according to the grid method in the sliding time window, filter the normal data by information entropy in each grid, judge the remaining data that may be abnormal by using local abnormality factor, and eliminate the abnormal data according to the judgment result. This paper introduces the application of remote sensing in flood control field, proposes the framework of basin flood prediction based on multiphysics field remote sensing big data fusion analysis, and designs each functional module of the system according to the object-oriented idea to realize the functions of data management, image processing, spatial analysis, and simulation output. The method can change the problems of cumbersome data processing and basic parameter rate determination in traditional hydrological methods and can find certain regularity through the connection between all related factors. Meanwhile, the use of artificial intelligence and other technical means makes the calculation speed faster and the obtained results are closer to the actual measured values, which is beneficial to guide the practical work.

1. Introduction

North of the Qinling and Huaihe rivers and downstream of the Yellow River basin in China is a relatively typical semi-humid and arid region under temperate monsoon climate. Every July and August, under the influence of the western Pacific subtropical monsoon, the region experiences a steep increase in rainfall resulting in floods along the river basin, posing a serious threat to people's property and life safety [1]. Flooding has brought about a bad impact on human production and life and has become one of the biggest losses from natural disasters [2]. A real-time flash flood disaster monitoring and forecasting early warning system covering

the whole basin has been basically built, and the existing basin flash flood disaster early warning system cannot meet the actual needs of flood control departments in terms of joint river-wide flood scheduling and joint warning [3]. However, how to integrate, mine and analyze, and expand the application of the survey and evaluation results of flash flood disaster prevention and control is a major issue facing the watershed and even the whole Guangxi water conservancy industry.

In the process of struggling with floods, people have concluded the experience of regulating floods with water conservancy projects and have controlled floods to some extent by constructing a large number of flood control projects such as

embankments and reservoirs [4]. After the reservoirs were built, some water conservancy projects were implemented later in their upstream basins, resulting in changing the runoff process, and if the flood prediction model did not reflect this change in a timely manner, the accuracy of the prediction model would not be up to the requirements [5]. However, flood control engineering treats the symptoms but not the root cause, and consumables are labor intensive, so the concept of flood treatment has changed from control to management, from relying on flood control engineering to combining flood control and management [6]. With the development of satellite remote sensing technology, especially the increase in the number of domestic remote sensing satellites launched and the significant reduction in the cost of use, the operational use of multiphysics field remote sensing technology means for regular flood monitoring has become possible [7]. Multiphysics remote sensing, in a broad sense, is a detection technique that senses a target object or natural phenomenon at a long distance without direct contact [8]. Flood risk uncertainty addresses the factors in flood risk, which include the gestating environment, the causative factors, and the hazard-bearing bodies [9]. With the development of society and human changes to the environment, the gestation environment has become more and more complex, including man-made factors in addition to natural ones, such as expansion of fields, occupation of river channels, and global warming [10]. Hydrological forecasting provides information for flood control and a basis for making decisions, which can forecast flood processes accurately and timely and can provide a scientific basis for the development of flood control plans.

The main problems can be divided into the following: the pollution of rivers caused by human waste and sewage and the prevention and control of natural disasters such as river water allocation and flooding. By analyzing the hydrological situation in the river network through the numerical simulation model of the river network, we can provide the law of the evolution of each hydraulic element of the river with time and space and provide an important database for the flood control planning of the river network. However, the inherent drawback of single-point rainfall measurement determines that it cannot fully represent the real spatial and temporal distribution of precipitation in a watershed or region, while in most cases, the observed values of runoff usually can truly reflect the changes of precipitation in a watershed. Therefore, through the use of multiphysics field remote sensing satellite data, we provide regular monitoring of water body dynamics during the flood warning period for flood monitoring needs and combine with ground monitoring data and flood forecasting data to carry out water body inundation extent, impact evaluation, and impact forecasting.

The innovation of this paper is to adopt the ordered weighted average method of multiphysical field remote sensing big data fusion analysis. By changing the decision-making risk coefficient, we can change the importance of evaluation indicators and get the evaluation results of different attitudes. The technical process of flood remote sensing monitoring and the establishment of flood prediction model are discussed. The idea of watershed flood forecasting based

on multiphysical field remote sensing big data fusion analysis is to divide uncertainty into objective world uncertainty and human cognitive uncertainty. In this paper, a flood forecasting model is proposed, which uses the input of control quantity to replace the impact of human activities in each river section and subtly examines these methods from other angles.

2. Flood Prediction of River Basin Based on Multiphysical Field Remote Sensing Big Data Fusion Analysis

2.1. Technical Process of Flood Remote Sensing Monitoring. Permanent flood multiphysics field remote sensing monitoring needs to follow an operational technical process to guarantee the monitoring procedures and result specification. To ensure the standardization of monitoring procedures and results, the level of monitoring technology should be improved, and the professional training of environmental monitoring technicians should be strengthened [11]. It is necessary to provide monitoring personnel with the opportunity to visit and learn and broaden their horizons. Experts with rich monitoring knowledge and years of practical monitoring experience are invited to provide business guidance and operational technical skills training. Strive to improve the level of monitoring point distribution, on-site sampling, laboratory analysis, and preparation of monitoring reports. Strengthening the construction of environmental monitoring quality management system is to ensure the accuracy of environmental monitoring data [12]. Especially in the case of multiple people and groups involved, it is possible to ensure the convergence of data and product results between the processes. The steps of the flood risk level evaluation in the study area are broadly based on six steps for summary analysis and evaluation, as shown in Figure 1 below.

First is the emergency monitoring preparation. Emergency monitoring includes rainfall monitoring information, meeting information, local disaster reporting information, reaching emergency monitoring forecast setting monitoring start indicators, or receiving monitoring instructions from relevant departments to start monitoring procedures. Internet technology was applied to collect information on meteorological cloud maps, radiation, radar, etc. on a global scale and high-precision multiphysics field remote sensing of impact data, historical hydrology, geographic data, etc. The Masking equation for storage and discharge can be written as

$$W = K[xI + (1 - x)O] = KQ', \quad (1)$$

where K is storage parameter, x is specific gravity factor, and Q shows discharge.

The data are cleaned automatically or manually, the database is designed to unify the data format and requirements, a refined global hydrograph is generated based on the DEM data, and the hydrological parameters of the corresponding area are automatically analyzed based on

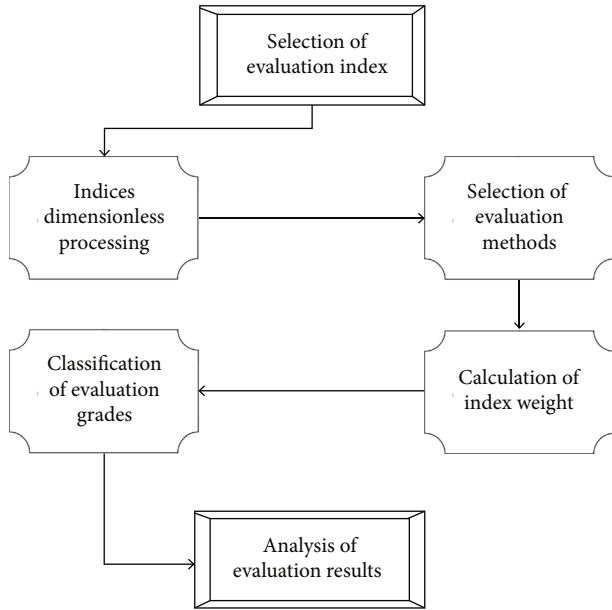


FIGURE 1: Flood risk assessment steps.

geomorphological data, land use data, global image data, etc. The threshold value T of the method is calculated by averaging E , squared difference between pixels P , and root mean square value Q between pixels for this window together, and the parametric equations for performing the threshold calculation are as follows:

$$T = a * E + b * p + c * Q. \quad (2)$$

For different factors in flood risk with different weight priorities, we obtained a flood risk hierarchical analysis structure using hierarchical analysis. The test focuses on how clouds interfere with the acquisition of water surface information, because a large number of samples are required to construct the correlations. However, there is not much multiphysics field remote sensing information that can be relied on, so this study uses all the collected multiphysics field remote sensing data and then removes the reservoir water surface information under cloud disturbance. This requires both real-time flood computation and real-time model computation, which is a complex project. The commonly adopted approach is shown in Figure 2.

The second is satellite image processing. The biggest difference between satellite image processing based on emergency flood monitoring and conventional situation processing is to complete the processing as soon as possible to buy time and improve timeliness. The parameters such as socioeconomic development data, river topography and geomorphology, reservoir capacity and hydropower station generation scale, irrigation area crop type, and irrigation scale within the basin are collected in conjunction with the basin flood measurement and reporting needs. If a stochastic process is a Markov decision process, then for that stochastic process, there must be a strategy that yields better results than the other strategies in all cases, which is also called the optimal strategy, then

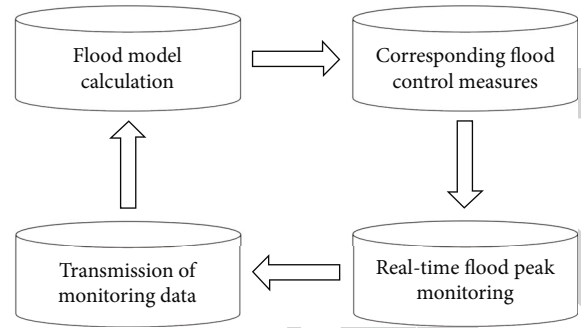


FIGURE 2: Real-time flood routing process.

$$\pi^*(s) = \arg \max Q(s, a), \quad (3)$$

where s is time, $\pi^*(s)$ is optimal strategy, and a is action.

Then, the optimal value return function can be obtained:

$$Q^*(s, a) = \sum_{x_s \in S} T(s, a)(R(s, a) + \gamma V^*(s^*)), \quad (4)$$

where γ is impact factor.

Through the judgment matrix obtained by two-by-two comparison of each factor in the hierarchy, so as to obtain the maximum characteristic root of the judgment matrix and the corresponding eigenvector, we get the weight ranking of the corresponding factors, which provides reference and basis for decision-making. The following equation is often used to calculate the entropy value between samples, the smaller the entropy value indicates the greater the similarity between samples. For an n -dimensional space, the Minkowski distance between point X and point Y can be expressed as

$$\text{dis}(X, Y) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}, \quad i = 1, 2, \dots, n, \quad (5)$$

where x_i is the i dimension characteristic data of point X and y_i is the i dimension characteristic data of point Y .

The analysis uses historical disaster data and historical hydrological data to correct hydrological parameters for precipitation forecasting results, which are gridded at $3 \text{ km}^2 \times 3 \text{ km}^2$. Compared with large and medium-sized reservoirs, small reservoirs have larger gaps in management level and system due to regional differences, which can reduce the accuracy of actual survey and operation information. Moreover, some small reservoirs lack complete information on the change of reservoir volume even during the field floods. If the gray value of the background can be reasonably seen as constant throughout the image and all objects have almost the same contrast with the background, then as long as the correct Min value is chosen, using a fixed global closed value will generally have a better segmentation effect.

Finally, the last step is water body extraction. After the preprocessing of multiphysics field remote sensing images is completed, water body extraction is carried out to obtain water body extent information, through specific technical

means to open up the meteorological data collection channel between the flood forecasting end of the basin and the National Weather Bureau, Weather Forecast Center, and Meteorological Agency. The weather data collection channels with the National Weather Bureau of China, the National Weather Bureau of the United States, the European Weather Forecast Center, and the Japan Meteorological Agency are established, where the minute precipitation data from the Public Weather Service Center of the China Meteorological Administration are used for 0~1.5 h short prognosis. At the same time, historical disaster data are collected and the model parameters are rechecked, so as to form a closed loop and continuously improve the forecast accuracy. Later, the detected data with higher accuracy can be applied when calculating the correlation between reservoir capacity and area of various small reservoirs. A threshold that works well in one area of the image may work poorly in other areas. Based on the weight ranking and expert experience evaluation, we obtained the basic probability distribution function of the evidence-based inference theory, and the risk level of the study area was obtained by fusing the flood factors with the evidence-based theory. In this case, it is appropriate to take the grayscale threshold as a function value that varies slightly with the position in the image.

2.2. Establish Flood Forecasting Model. Due to the delayed nature of storm floods and the need for real-time forecasting, a direct “rainfall-water level” model needs to be considered. The river model and the flood algorithm are combined into one, specifically for storm flood prediction. In the early warning response tracking system, the staff of flood control office will start the corresponding flash flood prevention and control plan by the level of warning information and carry out the flood control work according to the preparation of the plan; the business process of the early warning system is shown in Figure 3 below.

First of all, the forecast accuracy has to meet or exceed the results obtained by following the steps of determining the model and then performing the flood evolution, including the accuracy of both the flood arrival moment and the maximum flood flow. A more objective evaluation of the forecast error of the model is made using the Nash efficiency factor:

$$E_{NS} = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O}_i)^2}, \quad (6)$$

where O_i is the i th measured value and P_i is the i th analog value.

There are uncertainties in estimating economic losses linked to human lives, uncertainties in economic accounting of flood control projects, and uncertainties in estimating discount rates and other economic parameters. As the incoming runoff is back-propagated from the water level, there are also often missing data or abnormal data values in the collected data. After the soil moisture satisfies the field water holding capacity, all rainfall produces flow, which corresponds to the hyperinfiltration produced flow. The structure of the storage full produced flow is

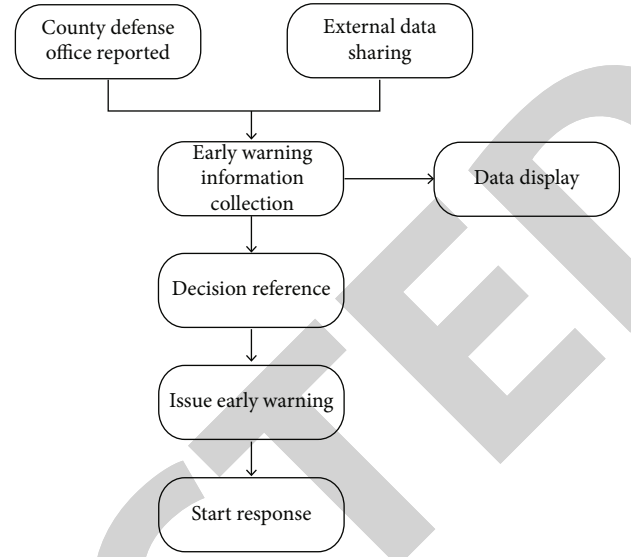


FIGURE 3: Business flow chart of the early warning system.

$$P - E - R = WM - W_1, \quad (7)$$

where B is rainfall period, E is time period evapotranspiration, W_1 is soil moisture at the beginning of the period, R is time period discharge, and WM is field capacity.

Therefore, the derivative term in the differential equation of water flow motion is approximated by a difference equation using the Tairau series expansion as a tool, so that a system of difference equations can be obtained at each calculation time. While assessing the importance of the indicators, the importance of the indicators is reranked by considering the influence of the magnitude of the indicator attribute values, and the decision risk factor is changed to adjust the participation of the indicators in the evaluation. Combine indicators between “with” and “or” and adjusting “with” and “or” by changing the magnitude of the decision risk factor “ratio to provide a comprehensive decision strategy. If a system of difference equations solves a disaster, i.e., each equation can be solved independently, it is called explicit format; conversely, if a joint solution is required, it is called implicit format. The weights of the feature terms obtained from the calculation are normalized by the following formula:

$$W(t, d) = \frac{1 + \log_2 tf(t, d) \times \log_2(N/n_t)}{\sqrt{-\sum_{t \in d} [(1 + \log_2 tf(t, d) \times \log_2(N/n_t))]}}, \quad (8)$$

where $W(t, d)$ is the weight, $tf(t, d)$ is the frequency, N is the total number of all videos in the training set, and n_t is the number of documents in which the word t appears.

Secondly, the established model is used to calculate the flooding algorithm based on the monitoring data and to calculate the areas that are vulnerable to danger; flood warnings are communicated to the area, and flood control measures such as dike reinforcement or early reservoir release are carried out for the area; with the different Tairau spreading and parallel formats used, the differential formats can be divided

into first order, second order, and even higher order according to the order of approximation accuracy and also into central and upwind (or eccentric) formats according to the nature of the formats. For a century sequencex with n samples, construct an order column.

$$S_k = \sum_{i=1}^k r_i, \quad k = 2, 3, \dots, n. \quad (9)$$

Among them,

$$r_i = \begin{cases} +1, & \text{when } x_i > x_j \\ 0, & \text{when } x_i \leq x_j \end{cases} \quad j = 1, 2, \dots, i. \quad (10)$$

Before the simulation calculation, the data for outlier rejection, missing value interpolation, and other work, after processing the formation of the data series, in the calculation can improve the accuracy of the prediction results, to get more desirable results. For optical images using a simple water body index method to extract the water body boundary, the basic principle is to use the band ratio operation method to achieve the purpose of highlighting the water body, weakening the non-water body. No time parameter is provided, time parameter in year-month-day format is provided; weighted average of 10 rainfall stations is provided; rainfall data of 10 stations are provided, respectively; runoff data of the previous day is provided; runoff data of the previous day is not provided. Using the finite element assembly algorithm, the river water level equation and the Hanpoint water level equation are superimposed to solve the overall equation, which improves the stability of the algorithm and fast convergence of the model because the overall equation takes into account the strong scour effect of water level distribution and change in the river network. Let $x \in R^n$ be an unknown parameter vector, the measure y is a m -dimensional random vector, and a set of samples of y with capacity N is $\{y_1, y_2, y_3\}$, for which the statistic is

$$\hat{x}^{(N)} = \phi\{y_1, y_2, \dots, y_n\}. \quad (11)$$

Finally, waiting for the flood water to flow through the area, the flood water that travels in the downstream channel is evolved in real time based on the actual monitoring data transmitted back, and so on. To make the differential equation correctly reflect the physical mechanism of water flow, such as using the central format to calculate rapids, it is physically inappropriate to use only the continuity of the solution. Control differential equations express the physical laws of conservation of mass and conservation of momentum, while differential equations sometimes do not strictly maintain the nature of conservation; the numerical solution will appear in the amount of water and momentum imbalance of conservation errors. For the satellite image data such as HSPA-1, ENVISION 1A/1B, RESOURCE-3, and RESOURCE-1 02C, which only have blue, green, red, and near-red bands, the normalized difference water index (NDWI) is used to automatically extract water bodies using

the contrast between absorption and reflection of water bodies in the green and near-red bands. The data structure of finite element is used for river network description similar to the rod system structure, and the data preparation is simple for the problem of adding river cross sections in the program study phase. By means of 3S technology and Internet technology, while relying on the water and rainfall data information that has been collected into the reservoir, rapid response and unified command and dispatch of flash flood-related events are realized.

3. Application Analysis of Big Data Fusion Analysis in Basin Flood Prediction

3.1. Calibration Analysis of Water Source Parameters. Due to the complexity of hydrological elements, describing the relationship between systems by physical quantities is inherently error-prone, coupled with the fact that many actual measurements are not observable, or are missed, or the observations are not representative. Energy losses are mainly concentrated in the data transmission process. In order to compress the amount of data and reduce energy consumption while improving the accuracy of data fusion, a water source model parameter rate determination method is proposed. This method often uses the empirical method to determine the initial values of the parameters and then uses the hydrological model calculation to perform the calculation of the production and convergence processes. Then, the hydrological model calculations are used to calculate the production and sink processes, and the process is compared with the measured process, and the optimization and debugging are continuously carried out to determine the optimal values of the parameters empirically using the principle of minimum error. Next, to verify whether the control law is effective, the performance indexes are selected as 10, 20, and 30, and the simulation curves are shown in Figure 4.

First, continuous data for a certain period are selected for the calculation, and the selected time series data information including the abundance and exhaustion years gives better results. In the distributed big data fusion principle, the distributed data set is first obtained and the data set is divided into several subsets. Then, by drawing Tyson polygons based on the location of watershed stations, the area weights of Tyson polygons are used to obtain the weighted sum of watershed surface rainfall. The data collection and caching mechanism of the complex heterogeneous multilayer network is centered on real-time sensing of data and accurate processing of decision information, eliminating the quality constraints of network data by redundant information and various types of interference from the external environment. For individual phase to be analyzed, the selected data can be selected based on the geographic base map and remote sensing base map through interactive methods and exported. Based on the operational data of the reservoir and the multi-physical field remote sensing data, the relationship between the reservoir volume and the watershed area is described; then, what changes in the volume of the small reservoir ponds and dams are derived based on the evaporation information of the precipitation, and the flood interception

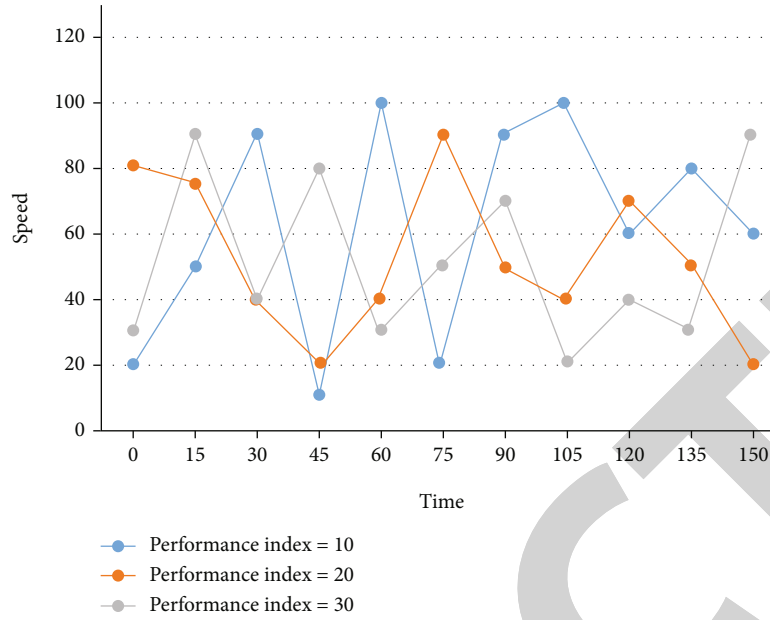


FIGURE 4: Simulation curve.

volume of the corresponding flood period is derived. By visualizing the rainfall map with histograms according to time, users can clearly grasp the local rainfall changes. The comparison curves of water level and storage volume at different times are shown in Figure 5 below.

Secondly, the calculated results are compared with the measured results to calculate the error, and the optimal values of the parameters are obtained by repeating the calculation until the final error is minimized. Then, several subsets are clustered at the same time to get several clustering centers. In order to well fuse the detection data of the wireless sensor network and make the obtained estimated values match the actual values more closely to achieve the purpose of high accuracy and high reliability state estimation, the Mann-Kendall test is introduced to compress the data in order to reduce the data communication. Using the Mann-Kendall test, the annual rainfall in the watershed was tested for abrupt changes and the annual rainfall Mann-Kendall curve was plotted. Kendall's statistic curve is shown in Figure 6.

However, the data collection and transmission probability control for multilayer networks and the decision mechanism to maintain the stability of the system become the bottleneck problem of network big data integration. So the flash flood disaster prediction and early warning system carries out simulation prediction and comprehensive research and judgment by establishing mathematical models for various types of disasters. It uses 2D and 3D visualization to present the results and determine the current warning level and hazard level of the event. And according to the simulation prediction and comprehensive research and judgment results to predict the secondary and derived from the disaster, the scope of influence, consequences, and risk of flash flood disaster are assessed. When there is a large variation of water flow along the course (such as the existence of bottom slope), the calculation of the nonflush term is more

complicated and may bring considerable errors. Therefore, the amount of regulation of each flood control means should be allocated and the prediction results should be obtained by computer algorithm. Other objective functions can also be used as criteria to evaluate the error, such as annual runoff and surface water flow process, although not as objective as the flow process line, but its advantage is that it only reflects local factors, such as only reflecting the production of flow and has nothing to do with the confluence. The higher the elevation of the region and the larger the standard deviation of elevation represent, the easier the drainage of the region and the less likely to have waterlogging and flooding, according to which the influence degree distribution under the combined effect of elevation and standard deviation of elevation is obtained, as shown in Table 1 below.

Finally, the process line is the total result of the whole process of flow production and convergence, which can be used as an optimization criterion and can also be called overall optimization. To determine whether the sum of the cluster centers is less than the scale threshold of the data fusion problem, if the result is yes, then the number of data is the value obtained from the sum of the data clustering centers is clustered and the resulting categories are fused to complete the distributed data fusion. In the case that the sum of squared residuals of all measured values is minimized, the arithmetic mean is considered as the most reliable value to complete the estimation of the actual value. If the rapid flow is influenced by the upstream flow pattern only, while the slow flow is influenced by both upstream and downstream flow patterns, this can be achieved by taking the difference backward along the feature. When solving the criterion weights of the influencing factors relative to the total study content, it is necessary to solve the relative weights of the same level factor judgment matrix for the factors at the previous level first.

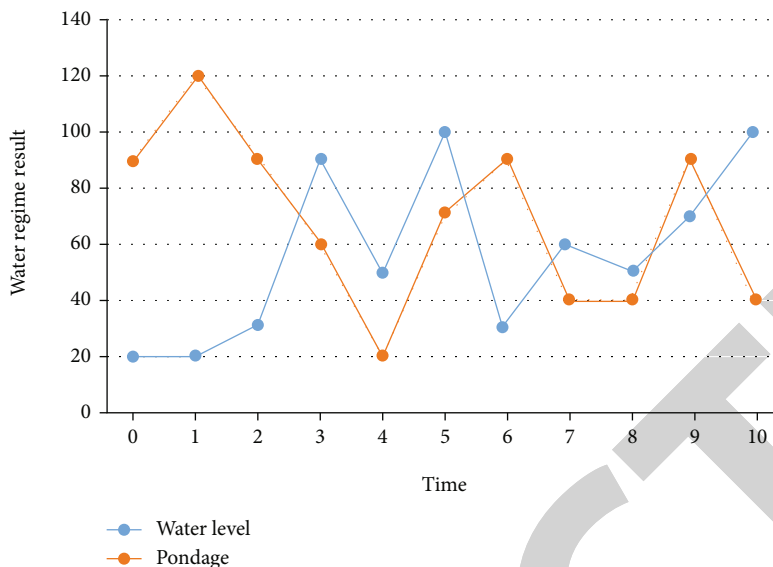


FIGURE 5: Comparison curve of water level and water storage capacity.

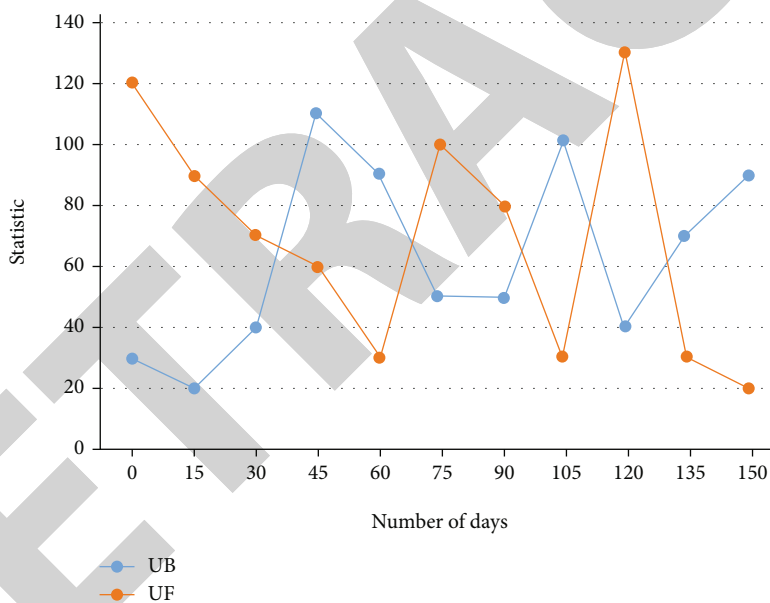


FIGURE 6: Annual rainfall Mann Kendall statistics curve.

TABLE 1: Distribution table of comprehensive influence degree.

Elevation and elevation standard deviation	0-5	5-10	>10
0-20	0.772	0.567	0.417
20-40	0.651	0.437	0.393
40-60	0.578	0.329	0.377

3.2. Calculation and Analysis of Order Weight. In the study of flood risk evaluation based on ordered weighted averaging, when decision-makers are more pessimistic, the lower the importance of the sequential layers, the higher the weight they are assigned, and the evaluation results show the least important attributes of the indicators. A two-by-

two comparison of factors at each level compares their importance to the study content and is represented by numerical values. In order to study the change process of incoming flow in flood and nonflood periods, the information of incoming flow in flood and nonflood periods from 2012 to 2021 was used to draw Figure 7.

First, the problem is analyzed and studied. For the study of flood risk evaluation, it is necessary to classify and stratify the analysis of the influencing factors of flooding. Since the fluxes transported across the interface between control bodies are equal in size and opposite in direction for adjacent control bodies, the fluxes along all internal boundaries cancel each other for the whole calculation domain. Assuming the river as a channel, the relationship between the water

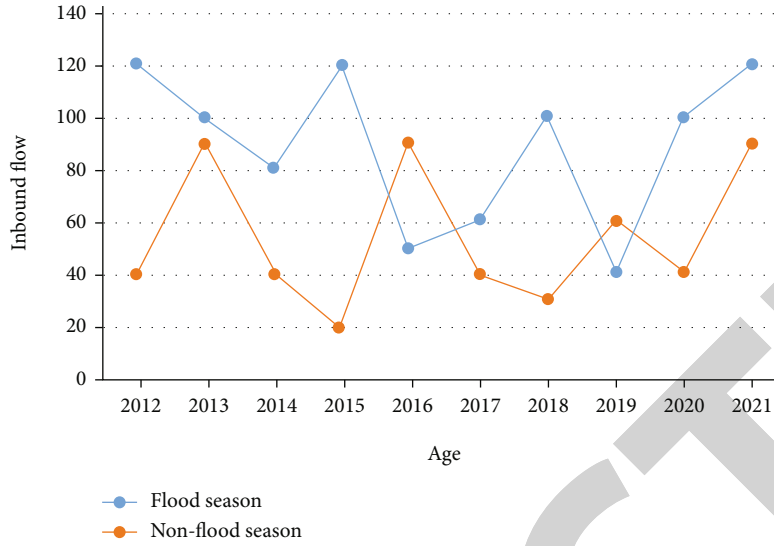


FIGURE 7: Change process diagram of inflow in flood season and nonflood season.

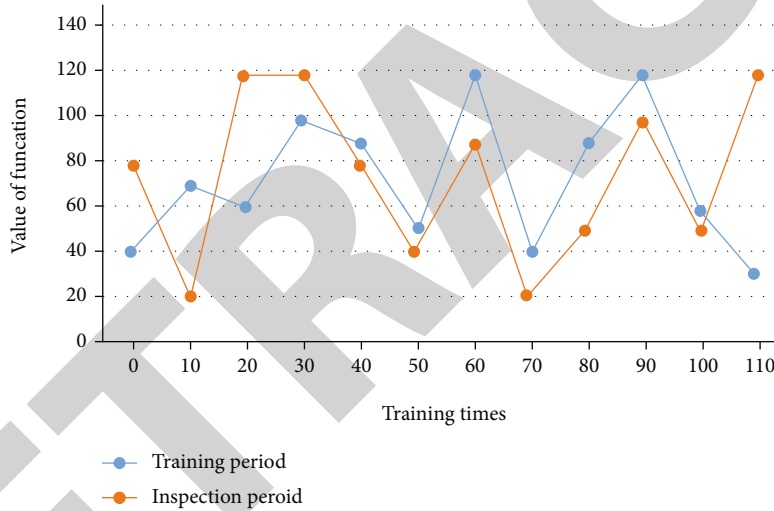


FIGURE 8: Calculation effect of time series plus rainfall runoff module.

level and the storage volume of each river section is found, and the change in water level and the rate of advance of the flood are obtained according to the relationship between the difference between the inlet and outlet of the tank storage theory and the constant value of the increase and decrease of the tank storage volume. The data is normalized and passed to Keras, a deep learning framework, in a NumPy data structure, and the optimal clustering result is used to limit the region for generating cluster heads. Next, the LSTM framework structure is selected to determine the dimensionality of the input vector in terms of the eigenvalue dimension. Using the system field statistics function, the flooded area is counted and the total flooded area is output or recorded, and the results of the loss calculation corresponding to different eigenvalues are shown in Figure 8.

Secondly, the dominant or subordinate relationship between factors at the same level is analyzed, and the lower level factors are subordinate to and influence the lower level

factors, and according to this rule, the recursive hierarchical structure model is established. The process of flow at the mouth is often high peaks, steep shapes, rapid and slow flow conversion, and water depth disparity, resulting in the classical format that often fails, requiring the use of a format based on completely shallow water equations and high resolution of discontinuity and the need to use a small time step; the explicit format is usually more favorable than the implicit format. It generalizes, abstracts, and simplifies the objective phenomena of nonconstant flow in open channels based on the basic principles of mass conservation law, Newton's second law, and energy conservation to obtain this nonlinear hyperbolic system of partial differential equations. The output dimension of the LSTM cell window is selected to be 100, and the dense function is connected to compress the 100-dimensional output into 1-dimensional, and a suitable excitation function is selected to convert the final forecast results. After finishing the inundation state extraction,

TABLE 2: Order weight under decision risk coefficient.

Decision risk coefficient	V1	V2	V3
$\alpha = 0.1$	0.0001	0.0002	0.0003
$\alpha = 0.2$	0.0263	0.372	0.441
$\alpha = 0.3$	0.0341	0.382	0.458

the area field is added to the vector file, the projection is set to equal area projection, and the area field is recalculated by using the system calculation geometry function. Calculate the presoil water content, deduce the real-time soil infiltration rate, derive the net rainfall process from the precipitation process, and analyze the influencing factors of river network storage in the basin. The data collected in the abnormal data rejection results are transmitted to the cluster head node to complete the distributed data intelligent fusion.

Finally, the eigenvectors of the maximum eigenvalue of the judgment matrix are normalized to the weights, after which the relative weights of the elements of the lowest level are multiplied by the relative weights of the elements of the upper level, which is the criterion weight of the final required elements. There are often some extremely shallow water areas in the flow field, the flow friction loss should be specially treated (because the water depth is zero when the friction ratio drop calculated by Manning's formula tends to infinity), and the numerical solution is required not to produce false oscillations, so as to avoid calculation instability. Adjusting the values of decision risk coefficients, we can get the order weights under other decision risk coefficients, which are summarized in Table 2.

The basic equations of nonconstant flow in open channels are formed by the continuous and momentum equations, and their corresponding functions elucidate the relationship between the flow elements and the process coordinates and time. The loss functions are selected as mean square error, mean absolute error, and root mean square error (MSE), followed by the Adam optimizer, and the flood process and flood volume in the study area are simulated using Python language programming, and the simulation results are compared with the measured information. Because the cluster head selection is random, the number of member nodes within the cluster varies from cluster to cluster, which will make the network of cluster head nodes with more member nodes consume more energy and lead to unbalanced network load and shorter survival period of network nodes after data fusion. Therefore, if you need to calculate the inundation depth and water volume, you can extract the boundary range of the water body that needs to calculate the water volume through the system and turn the water body boundary surface to line and line to scatter point.

4. Conclusions

With the rapid advancement of computer technology, the application of big data is getting more and more widespread attention. Big data analysis methods are new technologies and new methods based on the original calculation methods

as well as mathematical models, which are constantly improved and updated. Flood risk uncertainty and multi-source information fusion are a hot topic in current research; however, the causes affecting flood occurrence are complex, and the uncertainty of the objective world and human cognition reinforces the uncertainty in flood risk assessment. Big data fusion analysis can significantly improve data utilization and address the problems of high network energy consumption and short survival period of network nodes in current data fusion algorithms. The application of multiphysics field remote sensing big data fusion analysis to flood risk evaluation research better reflects the differences in risk attitudes of different decision-makers, because different decision-makers have different criteria for considering the importance of indicators, which is more in line with the actual situation of flood risk evaluation. The application of basin flood forecasting based on multiphysics field remote sensing big data fusion analysis to carry out operational normal monitoring of floods can complement the ground monitoring, forming a three-dimensional monitoring, which can provide data support for flood and drought disaster defense. It can also be connected to the flood control command system, as a part of the flood control command system, through cooperation with meteorology, hydrology, rivers, Skynet, and other departments to centralize the data through the server and the system to facilitate the relevant command decision-makers to grasp the data, timely planning, and timely release, to protect everyone's life and property safety. On the other hand, the study did not analyze the relationship between the relevant factors. It is necessary to use artificial intelligence and other technical means to find a certain regularity. It makes the calculation speed faster, and the results obtained are closer to the measured values, which is of great benefit to guide the practical work.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Hechi University Curriculum Model Reform Project: Research and Practice on Classroom Teaching Reform of "Advanced Mathematics" in Science and Engineering under the Training Mode of Applied Talents in Local Universities (No. 2015KTJY20); Guangxi Higher Education Reform Project: Research and Practice on Value Guidance of Infiltration Teaching Mode of "Mechanics" Course (No. 2021JGA302); National Natural Science Foundation of China: Boundedness and asymptotic behavior of stochastic, non-smooth and other complex

Retraction

Retracted: Modeling and Analysis of Multifocus Picture Division Algorithm Based on Deep Learning

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. You, "Modeling and Analysis of Multifocus Picture Division Algorithm Based on Deep Learning," *Journal of Function Spaces*, vol. 2022, Article ID 8326626, 10 pages, 2022.

Research Article

Modeling and Analysis of Multifocus Picture Division Algorithm Based on Deep Learning

Hongxia You 

Department of Liberal Studies, Jiangsu College of Engineering and Technology, Nantong 226007, China

Correspondence should be addressed to Hongxia You; youhx@jcet.edu.cn

Received 16 May 2022; Revised 27 June 2022; Accepted 1 July 2022; Published 16 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Hongxia You. 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 complex machine learning algorithm, deep learning can extract object shape information and more complex and advanced information in images by using a deep learning model. In order to solve some problems of deep learning in image feature extraction and classification, this paper designs a modeling method of multifocus image segmentation algorithm based on deep learning. The acceleration effect of FPGA (field programmable gate array) on deep learning and weight sharing is analyzed. By introducing deep learning, the trouble of determining the weight coefficient is eliminated, and the energy function is simplified. Therefore, the relevant parameters of multifocus image segmentation can be easily set, and better results can be obtained. The multifocus image segmentation algorithm based on deep learning can not only obtain closed and smooth segmentation curves but also adaptively deal with topology changes due to high segmentation accuracy and stable algorithm. The results show that the model effectively combines the local and global information of the image, so that the model has good robustness. The depth learning algorithm is used to calculate the average value of local inner and outer pixels of an image. Even for complex images, relatively simple contour curves can be obtained.

1. Introduction

Images are able to describe what people see with their eyes; however, it is quite difficult for computers to process human visual objects [1]. The multifocus picture division is a key step from picture handling to analysis of pictures, which is aimed at dividing the image into several specific regions with unique properties and extract the target of interest [2]. The advent of computers has provided the possibility to study images in greater depth and with the help of computer technology, image engineering has gradually evolved. Deep learning is a large-scale nonlinear circuit with real-time signal processing capability; like cellular automata, it consists of a large number of cellular elements and allows direct communication between only the closest cells [3]. Multifocus image segmentation is a crucial step from image processing to image analysis and is a fundamental computer vision technique [4]. In image analysis, entropy, as a statistical feature of an image, captures the magnitude of the quantity of data contained in the image [5]. This is due to the fact that image segmentation, separation of objectives, extraction of

characteristics, and determination of parameters transform the raw picture into a more extracted and compressed shape, which makes higher layer of analysis and comprehension feasible [6].

A multifocus picture division is a difficult and hot point in picture handling, which makes the basic pavement for the analysis and reconstruction of the image later [7]. The picture handling system can be expressed in three levels, and the combination of these three levels is also called image engineering low-level processing techniques which mainly include image filtering, image enhancement, image restoration, image restoration, and image coding and compression. The multifocused picture division is a class of picture handling between the bottom-level picture handling and the middle-level analysis of pictures. Picture division is an important element in analysis of pictures, which is the process of dividing the original image into several subregions that are not connected to each other and extracting the part of interest from these regions [8]. Traditional picture division methods are mainly based on the underlying image features that can be observed by the human eye, such as color,

texture, and edges [9]. When there is a large difference between the target and the background in the image, it can be segmented completely. However, when the features such as grayscale and color of the target and the background are closer or the background in the image is complex, the segmentation results obtained by traditional picture division methods often have more missplitting and the segmented targets are incomplete and lack of detail information [10].

CNN (convolutional neural network) is a locally connected network in which each unit is interconnected only with its neighboring neurons, and the influence of other neurons in the neighborhood is realized by passing information from unit to unit. Simply put, it is a machine learning algorithm that gives computers the ability to learn potential patterns and features from a large amount of existing data to be used for intelligent recognition of new samples or to anticipate the likelihood of something in the future. The advantage is that its multilayer structure can automatically extract different levels of distinguishing features of an image from a large number of samples, and these features are more effective and robust compared to the underlying features designed and extracted manually. The multifocus picture division is the key process from analysis of pictures to image understanding, separating the region of interest from the image, which simplifies the complexity of picture handling on the one hand and provides the basis for subsequent extraction of image features and quantitative analysis on the other.

The innovative points of this paper are shown below:

- (1) In order to assist the complete segmentation of the region, the representation of the target shape is particularly important. In this paper, the shape is characterized by building a model, and subsequently, the a priori shape information is used to assist the multifocus picture division
- (2) With the introduction of deep learning, the trouble of determining the weighting coefficients is eliminated and the energy function is simplified; thus, it is easy to set the relevant parameters for the multifocus picture division and obtain better results
- (3) The multifocus picture division algorithm based on deep learning can not only obtain closed and smooth segmentation curves but also handle topological changes adaptively due to the high segmentation accuracy and stable algorithm

2. Modeling Idea of Multifocus Picture Division Algorithm Based on Deep Learning

2.1. Construction Method of Conditional Random Field Model. The conditional random field is a discriminative model for modeling the probability distribution between the marker X and the observation Y . In the multifocus picture division, if the kind of similarity present in the same region is particularly valued and the uniformity of the segmented region is required too much, it will result in many blank regions or irregular edges. If there is a difference in

the gray value between the target and the background in a gray image, the target and the background in the image can be separated by threshold segmentation. If only one threshold is selected, the segmentation is called single-threshold segmentation, which will divide the image into two parts: the target and the background. The segmentation with multiple thresholds is called multithreshold segmentation, which will divide the graph into multiple regions.

The network objective function is then measured in terms of the sum of squared total errors.

$$J_p(t) = \frac{1}{2} \sum_R \left\| d_{Rp} - y_{Rp} \right\|^2, \quad (1)$$

where $J_p(t)$ is the p group input objective function.

The conditional random field model can combine knowledge and experience of the target to be segmented based on the underlying visual properties of the image itself such as edge, texture, grayscale, and color, and the target to be segmented, such as the shape, brightness, color, and other empirical knowledge of the target, in an organic way. Depending on the level of abstraction and the research method, there are three characteristic levels: picture handling, analysis of pictures, and image understanding. This is shown in Figure 1 below.

First, the conditional probability is defined using the DBN network and the marker variables of pixel points on the graph structure and the marker variables of neighboring pixel points. The basic unit of DBN is the cell: it contains linear and nonlinear circuit elements, typical elements are linear capacitance, linear resistance, and linear and nonlinear controlled and independent current sources. The structure of DBN is no connection within the top two layers, full connection between each layer, and the remaining other layers are The DBN model structure is shown in Figure 2.

Before segmentation, a suitable threshold is determined and then all pixels are compared with the threshold, and pixels greater than or equal to the threshold are classified as the target class and those smaller are classified as the background class. The neuron nodes on each characteristic image of the convolution layer are obtained by convolving various zones of the input image using the same convolution kernel. This operation ensures that all neuron nodes on the feature map share a set of convolutional kernel weights. In deep neural networks, the neuron nodes between layers are no longer fully connected, and the local spatial correlation between layers is used to connect the neuron nodes of each adjacent layer only to the neuron nodes of the upper layer that are close to it, i.e., locally connected. A multilayer forward network is built with the function to calculate the output M of the hidden layer.

$$M_j = f \left(\sum_{i=1}^n W_{ij} X_i - b_j \right), j = 1, 2, \dots, m. \quad (2)$$

For the actual noisy image, the edge points detected by the differential operator alone do not form a closed boundary to separate the target from the background. The

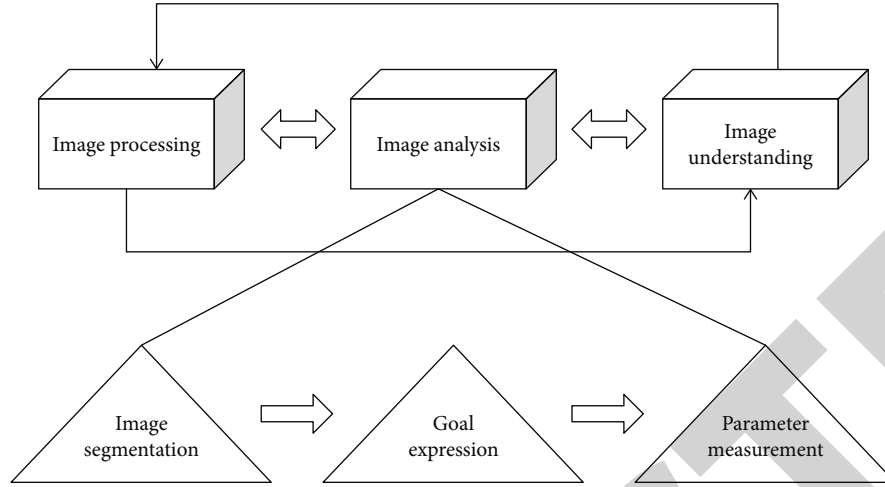


FIGURE 1: Position of the multifocus picture division in image engineering.

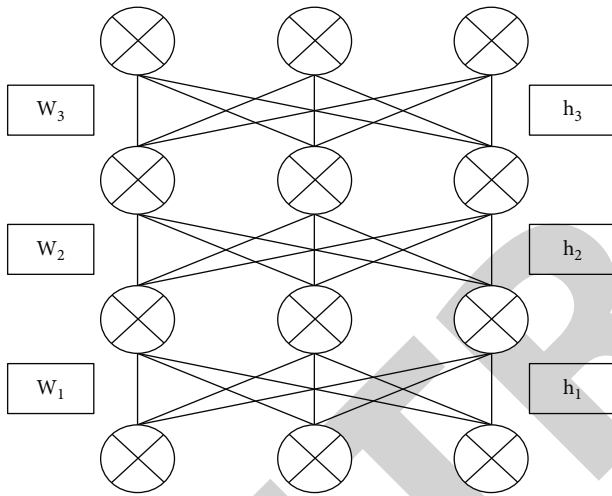


FIGURE 2: DBN network structure.

multiregion segmentation of this conditional random field model has only one unsigned distance level set function, and then, the model is solved using the VIIM algorithm to achieve the evolution of the boundary curves in each region. The large scaling factor a is used at low frequencies, with a large observation range on the time axis, while the frequency domain is equivalent to using low-frequency wavelets for a generalized observation. a is the scale factor, reflecting the frequency information of the signal; b is the translation factor, reflecting the time information of the signal. The signal power spectrum of the wavelet transform can be defined as

$$S(\omega) = \frac{2\pi}{N} \sum_{i=1}^N \frac{|W_s(a, iT_g)|^2}{\Delta\omega}. \quad (3)$$

There is always a discontinuity between two adjacent regions with different gray values, and this discontinuity is often easily detected by finding the first- or second-

order derivatives. After continuous processing, the correlation coefficient matrix of the evaluation index is calculated:

$$R = (r_{ij})_{p \times p} \quad (4)$$

where r_{ij} is the correlation coefficient of the i th evaluation sample with the j th indicator.

Second, the labeling of each pixel considers only its color characteristics, which often have complex color probability distributions for natural images. To obtain better modeling results, a Gaussian mixture model can be used to model the probability density function. GMM and its variant k -means (K -means) algorithm are clustering tools often used in industrial practice. Because GMM introduces the concept of implicit variables in modeling, we cannot directly use MLE (maximum likelihood estimate) for parameter estimation. Then, the EM (expectation maximization) algorithm is introduced to train the model with hidden variables. To be able to obtain the model results quickly, AMG and SFM algorithms can be used to increase the time step and reduce the area to be computed, respectively. Multiple forward propagation and backward propagation can be used after which the model learns more distinguishing features of the input image and obtains a higher correct classification rate. The input samples are normalized to ensure that the data are in the specified range $[1, 2]$ this interval, which facilitates the processing of data and improves the network efficiency. The input sample normalization process is given by

$$X = \frac{T - T_{\min}}{T_{\max} - T_{\min}}, \quad (5)$$

where T is the unprocessed input value, T_{\max} is the maximum value of neural network input, and T_{\min} is the minimum value of neural network input.

Since SFM can accept input images of arbitrary size, the feature map of the last convolutional layer is upsampled using a deconvolution layer to restore it to the same size as the input image. Thus, a prediction can be generated for

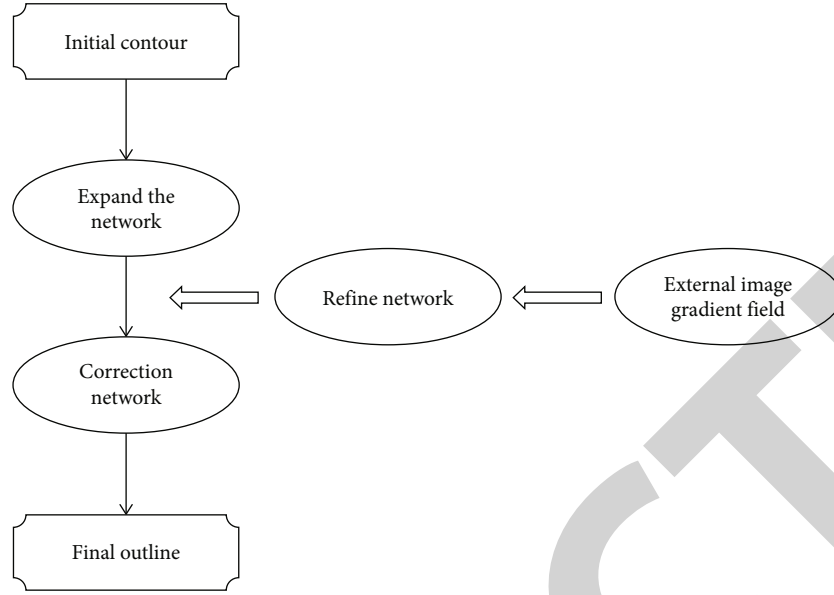


FIGURE 3: Processing block diagram of multifocus picture division algorithm.

each pixel while preserving the spatial information in the original input image. After the division of pixel points is completed, the two types of pixel points form two image regions representing the target and the background, respectively, and the segmentation is completed.

Finally, the identification of the markers in the image is done by two potential functions and the characteristic observed data of the whole image. The best marker Y , which is the marker corresponding to the maximum a posteriori probability, can be converted into the minimization of the energy function. Therefore, it is most convenient to use threshold segmentation when the difference between target and background grayscale in the original image is more obvious. Threshold segmentation accomplishes the work of obtaining a closed contour of a given shape from the (discontinuous) edge points in the image space by an accumulation operation in the parameter space. When using the region extraction method, the pixels are assigned to individual objects or regions; in the boundary method, only the boundary between the presence and the region needs to be determined; in the edge method, the edge pixels are first determined and they are connected to form the desired boundary. Meanwhile, the high resolution of downsampling is fused with the features of upsampling to further improve the accuracy of segmentation localization. Most importantly, deep learning can be loaded on smart mobile terminals, which in turn makes it easier and faster to receive data from remote servers and upload and download the data.

2.2. Model Structure of Multifocus Picture Division Algorithm. The multifocus picture division algorithm model uses the SegNet model as the backbone network for segmentation, which contains two components: an encoder and a decoder. The multifocus picture division algorithm is done by the DTCNN that finds the image gradients one at a time and is invariant throughout the multilayer iterations. The

processing module for each direction is composed of 3 DTCNN networks: the contour curve expansion network, the refinement network, and the correction network, as shown in Figure 3.

First, the last layer of convolutional features in Conv1 and Conv2 stages, and all convolutional layer features in Conv3, Conv4, and Conv5 stages are selected. For a given image operation, local rules are those features that reflect the essential and structurally invariant nature of the picture handling process, usually a set of if-then statements. The development of the model requires a server design via Nginx, which in turn facilitates the loading of third-party libraries needed for the service invocation process on mobile terminals. This project also requires the implementation of a Python based web microframework using Flask. When the net input to the output node is too large, the output layer, for example, is normalized by the above sample normalization formula to obtain

$$\delta_j^l = (d_j - y_j) u_j^l \cdot (1 - u_j^l). \quad (6)$$

Therefore, when the variance of foreground and background gray values is the largest, the probability of foreground and background misclassification is the smallest and the segmentation result is the most accurate. A properly designed serial boundary algorithm can remove the false edge points and directly obtain a single pixel wide and continuous boundary. A properly designed serial boundary algorithm selects one or more values and compares them with the grayscale values of the image to segment the target and background areas. The number of targets can be determined by the number of values selected. The initial contour curve is artificially given on the target region of interest, i.e., the initial value is one or more closed curves, and an energy function is minimized so that the contour curve is deformed

by the motion in the image and finally approximates the target boundary of the region. The class of forward neural network functions N is dense in the space of vector-valued continuous functions, in a consistent parametric sense, where the class of forward network functions is defined as a set of the following type.

$$N = \left\{ Y \in R^m, Y = W^{(K+1)} N_k \{ N_{k-1} [\dots N_1 (X)] \} \right\}, K = 1, 2, 3, \dots \quad (7)$$

Next, all features of the last three stages (Conv3-D, Conv2-D, and Conv1-D) are selected and do layer-by-layer, pixel-by-pixel summation fusion for them as well. Using the least squares derivation, the algorithm is less complex and the operation time is shorter compared to other threshold segmentation methods and is used more frequently. The grayscale value corresponding to the position of the valley of the histogram is taken as the thresholding value for the original image and use this as the boundary to divide all pixels of the image into two parts the one located on one side of the valley point is the target and marked with a target marker. The learning method of the template is essentially an optimization-seeking process, and the obtained template coefficients are the optimal solutions that meet the requirements, so the template can be designed by the mathematical optimization-seeking method. The features are expanded sequentially, where the aggregated features at each moment are obtained by splicing the feature words with different granularity at that moment, and the results are shown in the following equation:

$$c_i = [c_i^1; c_i^2; \dots; c_i^k], \quad (8)$$

where c_i^k is the convolution window of the i word which is the characteristic representation of k .

The cloud computing layer uploads the scene images and performs multifocus picture division using the FCN8-VGG16 network when the network is open. The trained U-Net model is downloaded, and the segmented image and U-Net model at the remote cloud server are uploaded to the smart mobile terminal via a 5G mobile network. The mechanism that drives the contour motion lies in minimizing the energy function when making the region that satisfies the consistency expand to the maximum or contract to the minimum, i.e., the boundary of the consistency region.

Finally, all the feature graphs above are stitched together in a way that they are cascaded and then fed into the fresh convolution layer and Softmax layer for studying and getting the final result graph of division. This is the metric space of the image, which determines the position corresponding to the histogram valley points, i.e., the threshold, to obtain the convolutional layer clustering of the image. The convolution layer uses the ReLU function, and no data is filled in while the convolution operation is performed, so the convolution operation affects the pixels of the feature map. Both the prior probability P (observed data|label) and the poste-

rior probability P can be evaluated, while the traditional feedback neural network only evaluates the posterior probability, that is, the network weights and radial basis function parameters are determined, and the output of DL can be expressed as

$$y_j = \sum_{i=1}^h \omega_{ij} \exp \left(-\frac{\|x_p - c_i\|^2}{2\sigma^2} \right), j = 1, 2, 3, \dots, n, \quad (9)$$

where ω_{ij} is the weight of hidden layer and output.

Suppose the targets with uniform grayscale in the image show a clustering phenomenon on the metric space, i.e., there are obvious pluralities in the grayscale histogram waveform and valley points. Then, the grayscale corresponding to the grayscale histogram valley points are used as sub-points, and the interval is divided into N segments, and each segment is used as a different marker. The radial basis function and dynamic weights are the key of interval segmentation, and the expression of radial basis function is

$$R = (x_p - c_i) = \exp \left(-\frac{\|x_p - c_i\|^2}{2\sigma^2} \right), \quad (10)$$

where $\|x_p - c_i\|$ is the norm of $x_p - c_i$, x_p is the sample data input by the input layer, and c_i and σ re the center and width of radial basis function.

The principle is to design the fitting target with a set of curves to be selected and define the energy function associated with each curve in the set to be selected. The energy function is designed with the principle that favorable properties lead to energy reduction, including continuity and smoothness of the curves. The remote service component layer serves remote procedure calls through a web server in the background and tensor flow on the GPU (graphics processing unit) to perform image transfer and data processing operations. Therefore, the model can be used to process large image databases, and the picture division model in this paper is highly efficient and robust compared to the commonly used picture division models.

3. Application Analysis of Deep Learning in Multifocus Picture Division Algorithm Modeling

3.1. FPGA Accelerated Analysis of Deep Learning. This section explores FPGA (field programmable gate array) acceleration schemes for deep learning to enable fast and efficient processing of data for deep learning. For those images with strong contrast between target and background, the method can quickly and accurately segment multi-focused images into the images we want. The value of FPGA acceleration is that it provides a unified solution to a range of computer vision problems and has been successfully applied to computer vision fields such as boundary extraction, picture division and classification, motion tracking, 3D reconstruction, and stereo visual matching. The mean square error of deep learning after FPGA acceleration is also consistently lower

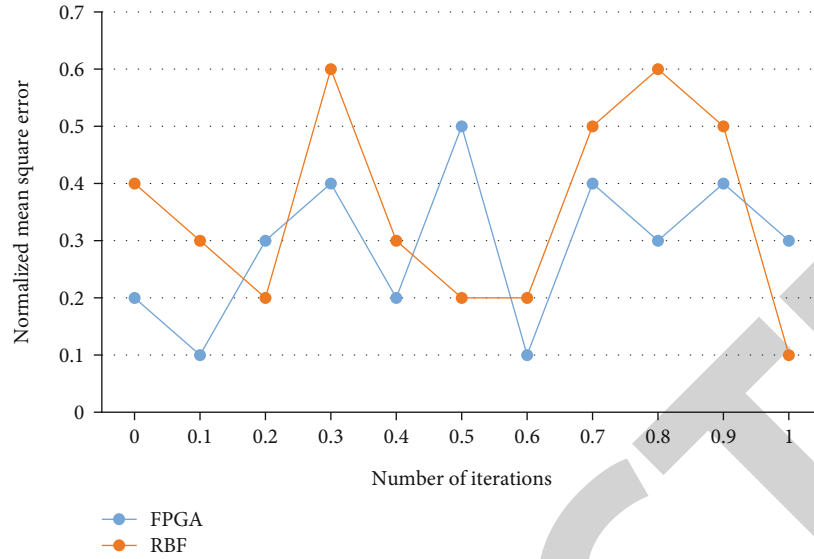


FIGURE 4: Mean square error of model.

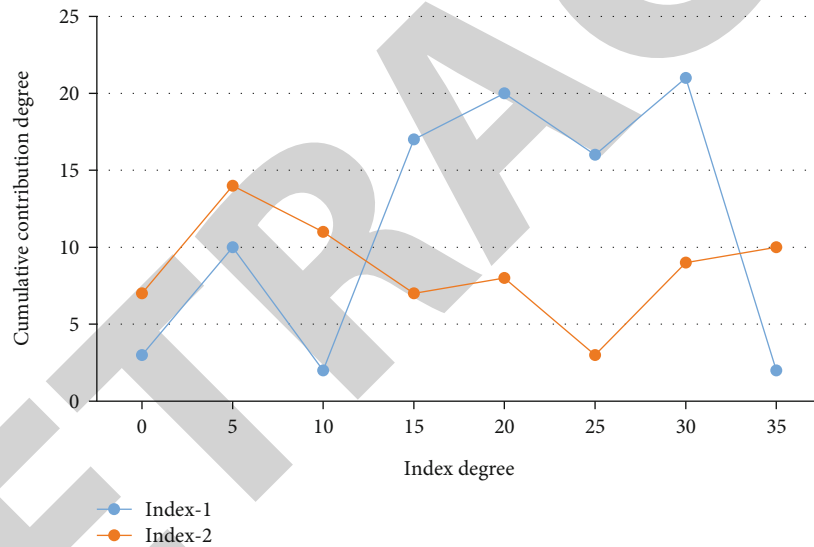


FIGURE 5: Processing results of evaluation indexes by principal component analysis.

than that of RBF neural network, indicating that FPGA acceleration can improve the prediction accuracy of the model, and the model mean square error comparison is shown in Figure 4 below.

First, the process of computing the product accumulation of the input feature map and the weight matrix needs to be performed $K * K * N$ times the product accumulation of the weights and the input, and the N input feature map is convolved to obtain the M size $R * C$ output feature map. In order to obtain the color feature probability distribution of target and background, the initial region information of manual marker needs to be extracted separately. The convolutional characteristics of the appropriate phase in the encoder and decoder are coupled in a cascade by channel splicing and then fed into a newly convolutional level to be further learned and complete the ultimate classification fore-

cast. The proposed deep learning is used to analyze the multifocus picture division evaluation metrics, and the metrics that contribute the most to the picture division are filtered out, and the results of the principal component analysis are shown in Figure 5 below.

When the value of this pixel is +1, if the sum of the pixel values in its neighborhood is greater than the threshold, then this pixel is not an edge pixel and the value of this pixel tends to 0. Conversely, if the sum of the pixel values in its neighborhood is less than this threshold, then this pixel is an edge pixel. As for the position of the boundary, since the serial approach takes fuller consideration and relies on the previously obtained results when making the current decision, the accuracy of the position of the whole boundary is easily guaranteed to be high if the starting point is chosen correctly.

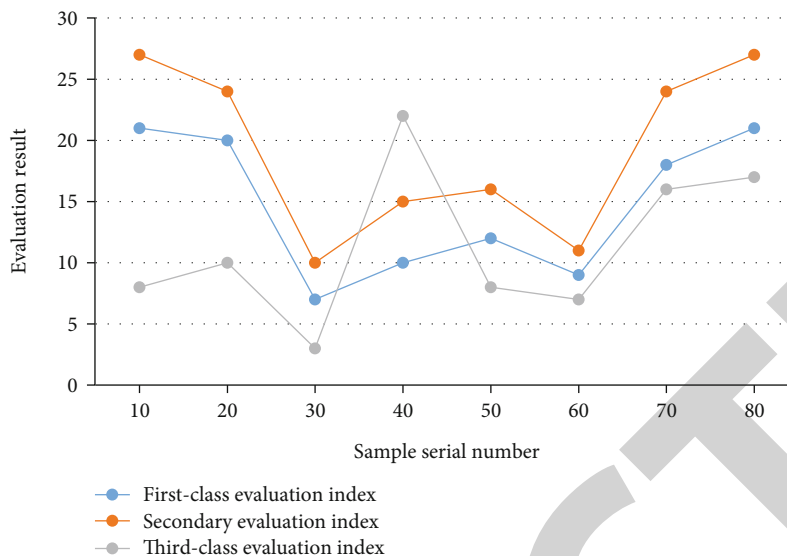


FIGURE 6: Evaluation results of different samples.

Secondly, when two matrices are matrix multiplied, each output result needs to take one row and one column of data of the two matrices, respectively, for product accumulation to get the final result. The manually extracted target and background pixels both have the same pattern information, so the color feature space can be described by pattern points. Prior to coupling the cascade of convolutional characteristics in each of the latter three phases in the encoder and decoder, the sophistication of the model after cascading multiple levels of features can be reduced to simplify the training of the model, and the level dependence of the characteristics extracted from the deeper convolutional layers is considered to maintain the local consistency of the features and improve the performance of division. Boundary tracking finds the neighboring points with the highest gray level from the original image and finds the neighboring points with the maximum gray level as the boundary points. In the specific implementation, it is necessary to first find a raw pixels for each image zone to be divided as the growing starting position, then take the raw pixels and compare them with the pixels in the surrounding neighborhood, and if there are pixels with the identical or comparable properties to the raw pixels, they are incorporated into the zone where the raw pixels are situated. The neural network is able to automatically learn previous experience from the provided data samples without the tedious process of searching and representation and is able to automatically approximate those functions that best portray the laws of the sample data, and the evaluation results under different sample order numbers are shown in Figure 6.

Finally, to reduce the number of repeated accesses from off-chip memory DDR, on-chip storage is used to cache the input data and the weight matrix. The initial marker region is smoothed and filtered using the same pattern points that many pixels have to characterize the local feature points corresponding to some pixels. The use of pattern points to describe the color feature space allows for a signif-

icant reduction in algorithm complexity when estimating whether a pixel belongs to the target or the background. Since pixels in the same region of the original image have similar features, these pixels are combined to form different regions by some predefined rules (growth rules). Thus, in a noise-free monotonic point-like image, this algorithm will trace the maximum gradient boundary, but even a small amount of noise will make the tracking temporarily or permanently off-boundary. For feature fusion, the CMLFDNN model amplifies the last three phases of features in the encoder and decoder, respectively, level-by-level and does pixel-by-pixel summation to fuse them prior to coupling them in cascade, which takes into account the dependency of the levels of accumulated characteristics retrieved in different convolution phases and maintains the local consistency of the active contours. In this way, the active contour can converge to the local boundary under the guidance of the energy function as it moves within the range of values taken, and the continuity and smoothness of the curve itself are maintained.

3.2. Weight Sharing Analysis of Deep Learning. In the case of the same spectral difference, if the spatial distance or image gradient is smaller, the weight is larger; on the contrary, the weight is smaller. For example, the neighboring pairs of points on both sides of the potential boundary have a small distance but a large gradient. The traditional weight sharing model is generally based on the Gaussian mixture model, which is also very simple and effective in multifocus picture division algorithms. The Gaussian mixture model is a kind of highly global optimization-seeking model, which is widely used in the fields of signal processing, image processing, and pattern recognition. The accuracy curve of the model evaluation is shown in Figure 7.

However, there is noise in segmentation and there are many segmentation algorithms based on a small number of

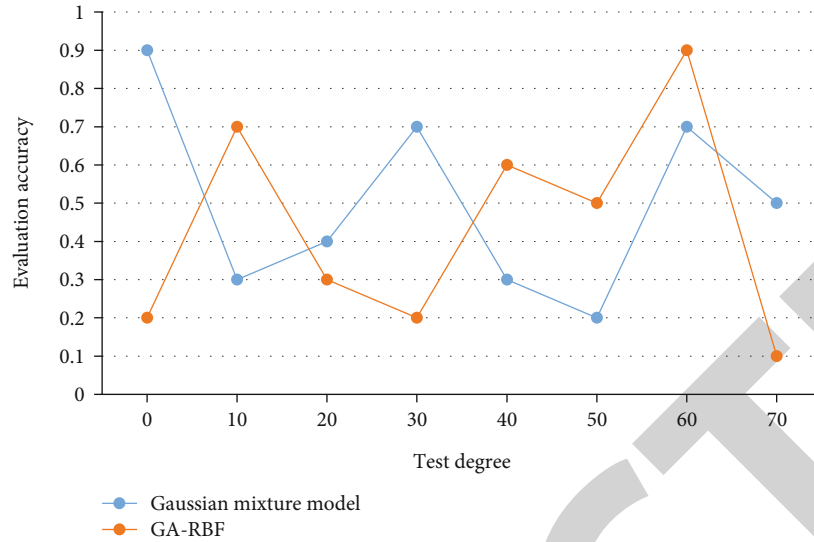


FIGURE 7: Accuracy curve of model evaluation.

local data points. Therefore, weight sharing based on deep learning has the interactive ability to represent the shape of anatomical structures concisely and analytically and segment the target boundary precisely. In deep learning, each convolution filter of the convolution layer is repeatedly applied to the whole perceptual field to convolve the input image, and the convolution result constitutes the feature map of the input image to extract the local features of the image.

First of all, in deep learning, it is assumed that layer $m - 1$ is the input layer and all neuron nodes in layer m are connected with all neuron nodes in layer $m - 1$. The process of convolutional computation to matrix multiplication is actually merging the input loop and the convolutional kernel loop, and merging the output row loop, thus reducing the number of loops to 2. In other words, take a point first and then compare the adjacent points. If they are similar, they will be merged together. If they are not similar, the calculation will be stopped immediately. Then, the gradient of the energy generalization function is calculated and the evolution equation of the level set function is obtained by using the continuous gradient descent method, and the final zero level set is determined as the multifocus picture division result. Here, it is assumed that the subsampling layer in the network is in the middle of the two convolutional layers, and if the postlevel sampling layer is a fully connected layer, we can solve the sensitivity map for subsampling directly. That is, an unsupervised codec network is trained to extract the deep features of the input image, and then, these features and spatial frequencies are used to measure the image pixel activity and obtain the decision map. Under the condition of limited training sample data, increasing the number of convolutional filters in the network model will make the number of weight parameters to be learned increase, which may cause the network to be difficult to reach a steady state during the training process. The experimental results of the four network models are compared as shown in Table 1 below.

TABLE 1: Comparison of experimental results of four network models.

Classifier	C1	C2	C3	Misclassification rate
Conv-net-1	3	17	26	8%
Conv-net-2	4	24	32	17%
Conv-net-3	7	27	41	27%

Second, the m layer feature image contains neurons, and the weight parameters are shared among different connections, so we can still use gradient descent to learn the shared weight parameters. The equivalent mapping of convolution calculation into matrix multiplication simplifies the loop structure, which is more suitable for FPGA design and more convenient for subsequent design optimization. In the reverse update phase of deep learning, the cost function will change very slowly at first, and then, the learning process of the network will become very slow. Therefore, the strength of the final gradient at each location will be enhanced by the nearby gradients by weighting the gradient at each location in the local window. That is, the loss function is constructed based on the deviation of the output result from the desired result, and the weights of the ω are continuously adjusted using the BP algorithm to minimize the result of the loss function. Since both edges and noise are grayscale discontinuities, it is difficult to overcome the effect of noise by directly using differential operations. Therefore, the image should be smoothed and filtered before detecting the edges with the differential operator and then apply the consistency check method to adjust the decision map to get the final decision map. The performance of the deep learning-based multifocus picture division algorithm model in this paper is compared with the GA-BPNN evaluation model and SVM-based evaluation model. The simulation data remain unchanged, showing that the constructed genetic algorithm optimized RBF neural network model has high evaluation accuracy and computational efficiency, as shown in Figure 8 below.

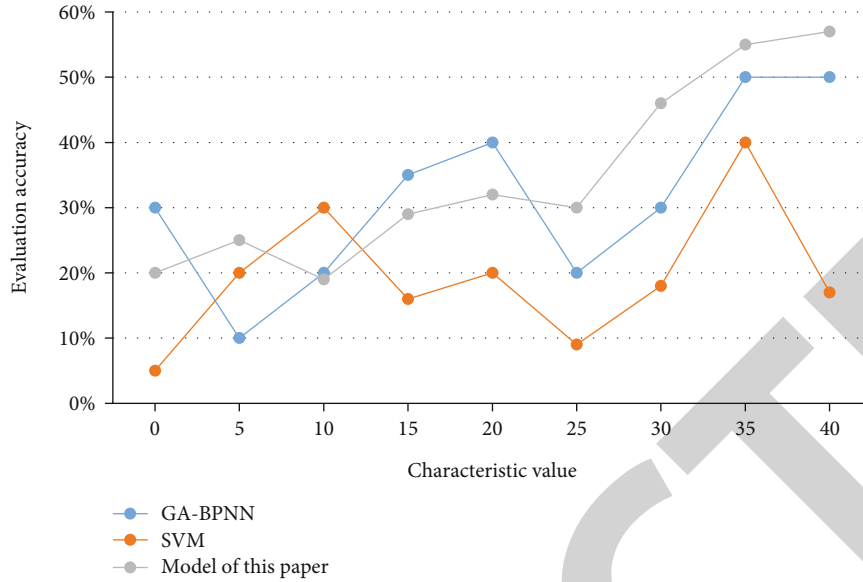


FIGURE 8: Comparison results of evaluation performance.

TABLE 2: Resource consumption table.

	Available	256-node language model	428-node language model	639-node language model
ALMs	23189	65%	49%	37%
RAMs	18924	38%	28%	19%
DSPs	9956	42%	39%	26%

Finally, an improvement to the original gradient descent method is needed to make the gradient of the shared weights the sum of the gradients of the shared connection parameters without considering the location of the local features when feature extraction is performed on the image. In the process of converting the convolution calculation into matrix multiplication, the input feature map does not need to be changed, but the size of the weight matrix and the order of data arrangement need to be rearranged. Considering the resource limitation of the FPGA, $N=5$ is set, in which case the resource consumption of the three LSTM models is shown in Table 2.

The cost function is redefined and the partial derivatives of the cost function with respect to the weights and biases are positively correlated with the errors. When the error is large, the bias derivatives are guaranteed to be relatively large even if the derivative values of the nodes are small at this time, thus making the learning rate relatively large. This makes it suitable for detecting changes in focus and has the best overall detection performance under normal imaging conditions. The image can be divided into regions, and the convolution kernel of the first network layer calculates these regions separately. After the convolution kernel gets the feature responses for these regions, the responses are used as input data for the second network layer, and so on. Training deep learning through supervised learning enables the network to learn the complementary relationships of different focused regions in the original image to achieve a globally clear image.

4. Conclusions

Image target segmentation refers to the assignment of a category label to each pixel in an image. It plays a forward and backward role in the whole study, both as a test of the effectiveness of all image preprocessing and as a basis for subsequent analysis of pictures and interpretation. However, for the variational level set model where the internal energy is a length generalized function, the conventional picture division calculating the Sobolev gradient of the whole energy generalized function does not show higher computational efficiency. And deep learning algorithms are widely used for processing various types of data such as images, videos, sounds, and texts. It is capable of various tasks such as image recognition, target detection, speech recognition, and machine translation and has achieved great success in various fields such as security, medical, entertainment, and finance. A convolutional neural network is a classical deep learning algorithm that combines feature extraction and classification problems, omitting the work of manual rule-based feature extraction. In this paper, we design a modeling method for the multifocus picture division algorithm based on deep learning and analyze FPGA acceleration for deep learning and weight sharing. The model effectively combines the local and global information of the image, which makes the model have better robustness. The method of deep learning computation is used in calculating the local internal and external pixel averages of the image, and a relatively simple contour curve can be obtained even for complex images.

Retraction

Retracted: Prediction of Urban Scale Expansion Based on Genetic Algorithm Optimized Neural Network Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Kuang, "Prediction of Urban Scale Expansion Based on Genetic Algorithm Optimized Neural Network Model," *Journal of Function Spaces*, vol. 2022, Article ID 5407319, 11 pages, 2022.

Research Article

Prediction of Urban Scale Expansion Based on Genetic Algorithm Optimized Neural Network Model

Hewu Kuang ^{1,2}

¹*School of Economics & Management, South China Normal University, Guangzhou 510631, China*

²*School of Insurance, Guangdong University of Finance, Guangzhou 510521, China*

Correspondence should be addressed to Hewu Kuang; 2017010013@m.scnu.edu.cn

Received 30 April 2022; Revised 14 June 2022; Accepted 27 June 2022; Published 14 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Hewu Kuang. 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 urbanization, the urban population is becoming more and more dense, and the demand for land is becoming more and more tense. Urban expansion has become an indispensable part of urban development. This paper studies the optimization of neural network structure by genetic algorithm, puts forward the prediction model of urban scale expansion based on a genetic algorithm optimization neural network, and compares the performance of the model with the basic model. A genetic algorithm BP neural network (GA-BP) optimized by the genetic algorithm is used to shorten the running time of the algorithm and improve the prediction accuracy, but it is easy to fall into local solution. The genetic algorithm is improved by immune cloning algorithm, and the CGA-BP neural network model is established to obtain the global optimal solution. Compared with the BP neural network model and GA-BP neural network model, the CGA-BP neural network model converges faster, and the training times reach the error condition after 79 times, while the BP neural network model and GA-BP neural network model need 117 times and 100 times, respectively, and the fitness value corresponding to the number of iterations of the model is larger. Therefore, the CGA-BP neural network algorithm can make prediction more accurately and quickly and predict the expansion of urban scale through urban conditions.

1. Introduction

Urban spatial expansion is the most direct manifestation of urban land use change. Urban spatial expansion is a comprehensive reflection of the changes of spatial layout and structure in the process of urbanization. With the advancement of global urbanization, urban spatial expansion has become a hot field in urban development research at home and abroad. The rapid development of urbanization in China has brought the unprecedented characteristics of urbanization. The continuous expansion and disorderly spread of cities in space have brought a series of problems. The contradiction between supply and demand of urban land is becoming more and more acute, which has become one of the main challenges that China will face now and in the next decade. Under the background that the concept of urban space management has been widely recognized, the rational and orderly expansion of urban space has more strategic significance in China. Therefore, the research on the simulation

and prediction of urban spatial expansion has important and urgent social and economic needs.

With the development of urbanization, urban expansion has become an inevitable trend. Predicting its scale is conducive to the direction of social development and the proportion of economic investment [1]. A neural network model is one of the common prediction means. The neural network has the characteristics of direct input to direct output. It adjusts its parameters to make the output results meet the conditions. Its high accuracy is one of the necessary conditions for prediction [2]. The genetic algorithm is to make up for the defect of neural network falling into local solution. The common genetic algorithms are the niche genetic algorithm and immune clonal genetic algorithm. The niche genetic algorithm is to preserve the population, reduce the fitness of the population, increase the probability of elimination, and increase the diversity of the population. It is a more accurate model [3]. The research is using the genetic algorithm to optimize the model constructed by the neural

network, which is more accurate and scientific for prediction. In order to reduce the redundant connection and unnecessary computational cost of neural network, quantum immune cloning algorithm is applied to the optimization process of neural network. The neural network structure is optimized by generating weights with sparsity. The algorithm can effectively delete the redundant connections and hidden nodes in the neural network and improve the learning efficiency, function approximation accuracy, and generalization ability of the neural network at the same time.

Immune clonal genetic algorithm is used to optimize neural network. Immune clonal genetic algorithm essentially adopts proportional replication operator and proportional mutation operator. Select the antibody with the highest affinity before antigen and antibody for replication, get a new set, and generate the global optimal antibody through high mutation probability. This makes the immune clonal algorithm obtain the optimal solution quickly on the basis of maintaining the diversity of the population, which is the advantage of the model, such as more adaptability and efficiency.

The innovation of this research is to solve the defect that the neural network model is easy to fall into the local optimal solution by combining the characteristics of introducing new species and increasing species and group diversity into the genetic algorithm. The urbanization expansion is gradually mature this year. In this field, the prediction is made by combining the genetic algorithm and neural network for the first time, and the nonlinear function characteristics of the neural network are used to make the prediction of urban expansion more accurate and provide a scientific basis for urban development.

The research is mainly divided into four parts. Section 2 mainly analyzes and summarizes the current application and effect of the genetic algorithm and neural network. Section 3 introduces the factors affecting the development of urban scale and constructs the prediction model of the genetic algorithm optimization neural network. Section 4 analyzes and compares the new capabilities of the optimization model and the traditional model. Section 5 evaluates the analysis and research results and puts forward the deficiencies in the research.

2. Related Work

Dawid's team applied the genetic algorithm to the spider web model and explained the results of different settings by simulating different coding schemes and fitness functions with state dependence. It found the relationship between the coding and convergence attributes of the genetic algorithm, the adaptability of the genetic algorithm in the economic system and the learning behavior of bounded rational groups [4]. Researchers such as Domashova et al. proposed a genetic algorithm to build the best framework of multilayer perceptron, create a random population, confirm the selection method of fitness function and parent sample selection, and propose the modification of crossover and mutation operators to ensure the operability of the algorithm for large and small individuals and solve the classification problem with minimal error [5]. Shukla and other scholars use the

model established by the convolutional neural network to diagnose COVID-19 patients with chest X-ray images and use the multiobjective genetic algorithm to adjust the recognition parameters. The experiment shows that 20x cross-validation can overcome the overfitting dilemma, and the recognition of the COVID-19 model has achieved obvious results, providing technical support for the process of diagnosis and screening of diseases [6]. Feng and other researchers proposed a network structure based on genetic algorithm, which generates a basic model dominated by an integrated algorithm by adjusting parameters and takes the original personal identification data as the numerical value. The results show that the prediction accuracy of the integrated model is better than the basic model, making personal identification occupy an important position in biological feature recognition [7]. Nikbakht's team applied the genetic algorithm to the neural network to find the optimal solution and optimized the hidden layer, integration point, and neuron number of each layer to achieve the highest accuracy, so as to predict the stress distribution of the structure. The results show that the prediction accuracy of the neural network has been significantly improved after the optimization of the genetic algorithm [8]. Han and other researchers used the genetic algorithm to optimize the BP neural network and build a GA-BP hybrid model. 13 of the 16 selected product design schemes were added to the hybrid model as parameters. After training, the predicted and actual values were obtained. Finally, the remaining schemes were verified. The results showed that the errors of several groups of data were 3.4%, 9%, and 3.1%, reflecting that the hybrid GA-BP model can predict the design scheme quickly, conveniently and accurately [9]. Ni's team proposed a router optimization algorithm based on the genetic ant colony algorithm in an IPv6 environment. The algorithm integrates genetic algorithm and ant colony algorithm, rewards or punishes by comparing the smoothness of the search path and the best path, and trains the obtained optimal solution. Experimental results show that the algorithm effectively solves the network quality problem and improves the service quality to users [10].

Researchers such as Gamidi and Ke proposed an artificial neural network model to predict the properties of eutectic based only on pure component physical properties. In the model grouping, the proportion of training model and test model is 7:3. When the training stops, the total value is higher than 0.986. The corresponding dissolution properties of eutectic and ideal solubility of eutectic are successfully predicted [11]. Zhang and other scholars combined PM2.5. The prediction model based on the Bi-LSTM neural network is constructed based on the concentration and meteorological information data. The meteorological characteristics are introduced and compared with the neural network model based on Gru. The results show that the model is effective for PM2.5. The prediction error of the concentration is smaller [12]. Pattnaik's team developed a prediction model based on the artificial neural network and Taguchi to predict the response. By using a lower level Taguchi orthogonal array to conduct a small number of experiments in the process of wire laying out, the results show that the modified model can predict the response more accurately [13].

Kenanoğlu and other researchers established an artificial neural network model by using LM learning algorithm to adjust the weight in the forward network. The results show that the model has high-precision prediction of motor torque, motor power, and emissions and plays an important role in promoting low-carbon emissions [14]. Abuelezz's team proposed a model for predicting the vertical total electron content of the upper ionosphere based on neural network and compared it with the IRI2016 model. The results show that the IRI2016 model is inferior to this model in all cases [15]. Chen and other researchers established a film scoring prediction model based on the convolutional neural network. Through experiments, it is found that the performance of a ten-layer convolutional neural network is the best, and the accuracy is about 56%, and the predicted value is close to the actual value, which provides a new model for film scoring prediction [16]. Gade's team proposed a Newmark sliding displacement prediction model based on the neural network. The mixed effect algorithm was used to evaluate the event residual. It was found that the model had no deviation and was consistent with the observed displacement characteristics, proving the applicability of the model and effectively predicting the seismic slope displacement hazard [17]. Ma and other researchers put forward a network state prediction algorithm of intelligent production line, which calculates and predicts the operation of the network through the optimized BP neural network. The results show that the optimal data prediction cycle is obtained in a large number of network data prediction experiments, and the accuracy is more than 90%, which is of great significance to improve the quality of network communication [18].

Based on the above analysis, genetic algorithm and neural network structure play an important role in classification, screening, and prediction, improve the accuracy of results, and promote the development of various fields. However, the relevant algorithms are rarely used in urban scale expansion, and the model provides a new scientific basis for urban scale expansion prediction.

3. Application of CGA-BP Neural Network in Urban Scale Expansion Model

3.1. Prediction of Urban Scale Expansion and Construction of Neural Network. Urban construction land is directly related to the rational layout and optimal utilization of urbanization land and has an important impact on the competitiveness and development of the city. With the rapid development of social economy and the improvement of urbanization level, the demand for urban construction land will continue to increase. The acceleration of urban construction land expansion has become an inevitable trend. This paper plans and optimizes the layout to improve the utilization efficiency and rational utilization of urban construction land.

The expansion of urban scale is to meet the balance of supply and demand. Therefore, it is necessary to meet the rational layout and maximize the benefits. The feasibility of urban scale expansion should consider the following factors: demographic factors, economic factors, location factors, social factors, policy factors, etc. [19]. The traditional predic-

tion methods include the linear regression method and gradient descent method. The linear regression method takes the factors affecting the price as variable parameters, which are expressed by $x = [x_1, x_2, \dots, x_m]^T$ (T represents matrix transposition), d represents the sum of output variables, w_1, w_2, \dots, w_m represents unknown parameters, and the expected error of the model is expressed by ε , so the linear regression equation shown in formula (1) can be established.

$$d = \sum_{j=1}^m w_j x_j + \varepsilon. \quad (1)$$

The gradient descent method randomly initializes the factors affecting house prices to achieve the purpose of optimizing the function and obtains the optimized linear regression model, as shown in

$$h(\theta) = \sum_{j=0}^n \theta_j x_j. \quad (2)$$

In formula (2), m represents the number of training set samples, n represents the number of features, $h(x)$ represents the fitting function, $J(\theta)$ represents the loss function, and the loss function formula is shown in

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (y^i - h_{\theta}(x^i))^2. \quad (3)$$

Random gradient descent method and batch gradient descent method are commonly used gradient descent methods. In the random gradient descent method, the partial derivative of the error function $J(\theta)$ is shown in

$$\frac{\partial J(\theta)}{\partial \theta_j} = -\frac{1}{m} \sum_{i=1}^m (y^i - h_{\theta}(x^i)) x_j^i \quad (4)$$

After the update of each θ , the function of risk minimization is obtained, as shown in

$$\theta_j^i = \theta_j + \frac{1}{m} \sum_{i=1}^m (y^i - h_{\theta}(x^i)) x_j^i. \quad (5)$$

It can be seen from equations (4) and (5) that when the value of m is particularly large, the process of fitting the function will consume more time. In the batch gradient descent method, all sample sizes are trained and the partial derivatives of parameters are obtained. The calculation formula is shown in

$$\theta_j^i = \theta_j + (y^i - h_{\theta}(x^i)) x_j^i \quad (6)$$

After eliminating the characteristic function, the batch gradient descent method can obtain the optimal solution of more data with fewer samples and determine which gradient descent method to choose according to the actual demand (brindle k m et al. 2021) [20]. During data analysis, some

original data have problems such as duplication or deletion. Therefore, the original data needs to be preprocessed so that the data can be applied to the model. The data processing method is to remove the uniqueness. The data that does not belong to the characteristics of the data sample and has no impact on the modeling can be deleted directly. The missing data can be supplemented or deleted by the average. The data often has orders of magnitude differences, so the data of different orders of magnitude and different quantity levels are standardized and regularized to ensure the experimental reliability. The common ones are min max standardization and Z-score standardization [21].

Min max standardization is to map the value x of attribute A into the $[0, 1]$ interval through standardization, and the formula is shown in

$$x' = \frac{x - \min x}{\max x - \min x}. \quad (7)$$

Z-score standardization is to regularize the average and standard deviation of the original data x . The obtained formula is shown in

$$x' = \frac{x - x_{\text{mean}}}{x_{\text{std}}}. \quad (8)$$

In equation (8), mean represents the average value and std represents the standard deviation. Set the data set as $D = \{(\vec{x}_1, y_1), (\vec{x}_2, y_2), \dots, (\vec{x}_N, y_N)\}$, $\vec{x}_i = (x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(d)})^T$ and calculate the norm L_p as shown in

$$L_p(\vec{x}_i) = \left(\|x_i^{(1)}\|^p + \|x_i^{(2)}\|^p + \dots + \|x_i^{(d)}\|^p \right)^{1/p}. \quad (9)$$

The result of regularization is shown in

$$\vec{x}_i = \left(\frac{x_i^{(1)}}{L_p(\vec{x}_i)}, \frac{x_i^{(2)}}{L_p(\vec{x}_i)}, \dots, \frac{x_i^{(d)}}{L_p(\vec{x}_i)} \right)^T. \quad (10)$$

The specific process of BP neural network model is shown in Figure 1.

3.2. Neural Network Optimized by Immune Clonal Genetic Algorithm. The good genes suitable for the environment will be selected by the environment to survive, and the unsuitable genes will be gradually eliminated. This natural law idea of survival of the fittest is applied to the process of finding the optimal solution to obtain the genetic algorithm (GA) [22]. Genetic algorithm can adjust the search algorithm through its own adaptation, without formulating certain rules. It can also find the optimal solution with probability when the objective function is not clear and can process multiple sample data at the same time, so as to improve the efficiency of the algorithm without avoiding the local optimal solution [23].

At the initial stage of the algorithm, data samples will be randomly generated. The samples have the attributes to solve the problem. After imitating the way of biological

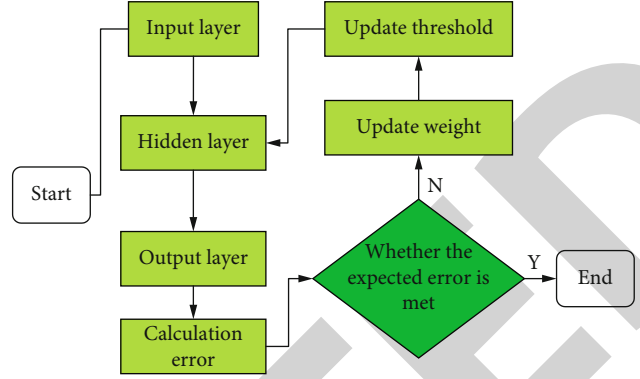


FIGURE 1: BP neural network algorithm flow.

selection, cross-reproduction, and obtaining mutated genes, the algorithm will generate a sample set more suitable for solving the problem and finally obtain the optimal solution through repetition. Then, the algorithm is accelerated by adjusting the parameters to obtain the optimal solution without falling into the local optimal solution [24]. The main steps are to code with specified numbers, letters, and other methods, generate random initial samples, evaluate the fitness in the initial sample individuals, and judge whether the operation meets the termination conditions through selection, crossover, and mutation. The flow chart is shown in Figure 2.

The chromosome coding of genetic algorithm is very important to the evolutionary efficiency. The commonly used coding includes binary coding, gray coding, and floating point coding. Although binary coding has many processing features and high accuracy in genetic algorithm, it is not universal. The commonly used coding is floating-point coding. Floating-point coding uses real numerical values to realize algorithm operation, which is direct, simple, and widely used. The fitness function of genetic algorithm is used to evaluate the fitness value of sample individuals. The design of fitness function needs to avoid the problems that large individual differences in the early stage of evolution affect the search ability and small differences in the later stage of evolution weaken the performance of the algorithm. The selection of fitness function directly affects the convergence speed of genetic algorithm and whether it can find the optimal solution. Because genetic algorithm basically does not use external information in evolutionary search, it only uses the fitness of each individual of the population to search based on the fitness function. Because the complexity of fitness function is the main component of the complexity of genetic algorithm, the design of fitness function should be as simple as possible to minimize the time complexity of calculation. The transformation method of fitness function value is shown in

$$\text{Fit}(f(x)) = \frac{1}{1 + c + f(x)}, c \geq 0, c + f(x) \geq 0, \quad (11)$$

where c represents the bound estimation of $f(x)$, and equation (11) is used to find the minimum value problem.

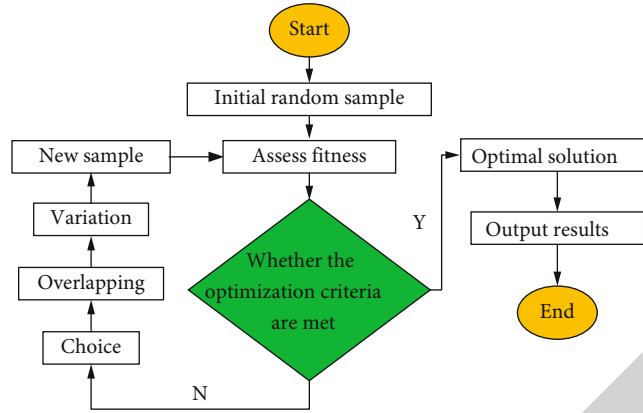


FIGURE 2: Chart of genetic algorithm.

Similarly, the problem of finding the maximum value can be obtained, and the function is shown in

$$\text{Fit}(f(x)) = \frac{1}{1 + c - f(x)}, c \geq 0, c - f(x) \geq 0. \quad (12)$$

Genetic operation includes the process of selection, crossover, and mutation. Individuals with problem-solving attributes are selected from large samples to form new samples. The quality of individuals directly affects the results of the algorithm. In order to meet the randomness of sampling, the sampling mode of roulette can be selected. The probability formula of Roulette is shown in equation (13). Genetic operation includes the process of selection, crossover, and mutation. Individuals with problem-solving attributes are selected from large samples to form new samples. The quality of individuals directly affects the results of the algorithm. In order to meet the randomness of sampling, the sampling mode of roulette can be selected. The probability formula of roulette is shown in

$$p_k = \frac{f(x_k)}{F} = \frac{f(x_k)}{\sum_{k=1}^n f(x_k)}, k = 1, 2, \dots, n., \quad (13)$$

where p_k represents the probability of being selected, x_k represents any individual, and $f(x_k)$ represents the fitness value of any individual. After the selection operation is completed through the above function, the crossover operation is carried out. First, the parent is selected. The parent is a combination obtained by random pairing of two individuals, and the individual selection is random sampling from the sample with the highest fitness value. The individual attributes are exchanged through one-point crossing and multipoint crossing to obtain the offspring. The one-point crossing diagram is shown in Figure 3.

Randomly select a point on the chromosome, divide the chromosome into two parts, and exchange the parts before or after the intersection to obtain new offspring. This process is cross-operation [25]. Now, swap the 5678 alleles of the parent in Figure 2, and the chromosome number is encoded in binary. Then, the fitness values f of the parent before

crossover are $(2^4 + 2^2 + 2^0)^2 = 441$ and $(2^7 + 2^4 + 2^1)^2 = 21316$, respectively. After exchange, the fitness values of the offspring are $(2^4 + 2^1)^2 = 324$ and $(2^7 + 2^4 + 2^2 + 2^0)^2 = 22201$. The results show that the fitness value of the offspring may evolve more than that of parents through the recombination of parents. Mutation operation is based on the theory of small probability. Some genes of chromosome mutate to form new individuals. Binary coding is to change "1" into "0" or "0" into "1." When such mutation operator is used in GA, it can not only prevent the immature phenomenon of the algorithm in the training process but also achieve the effect of optimizing GA. The parameters of the genetic algorithm depend on the population size M , and according to the actual situation, the value range is 10 – 200, the crossover probability P_c and the general value range of crossover probability 0.4 – 0.99, and the mutation probability P_m and the general value range 0.0001 – 0.1.

Due to the great limitations of the BP neural network, the convergence speed of function is slow and the accuracy of prediction value is relatively low. Using Hamming distance method to measure the common points or similarities between different individuals, the penalty function is applied to increase the elimination probability of individuals with low fitness and maintain individuals with high fitness. This method is niche technology. Adjusting the fitness of all individuals is to use the mechanism of sharing function to express the sharing degree with the sum of the sharing values among all individuals, as shown in

$$S_i = \sum_{j=1}^n S(d_{ij}). \quad (14)$$

The sharing value of individual i and individual j is represented by d_{ij} , S represents the sharing function, and the sum of sharing degrees of all individuals is represented by S_i . The adjustment formula of individual fitness is shown in

$$\text{Fit}(f_s(i)) = \frac{f(i)}{S_i}. \quad (15)$$

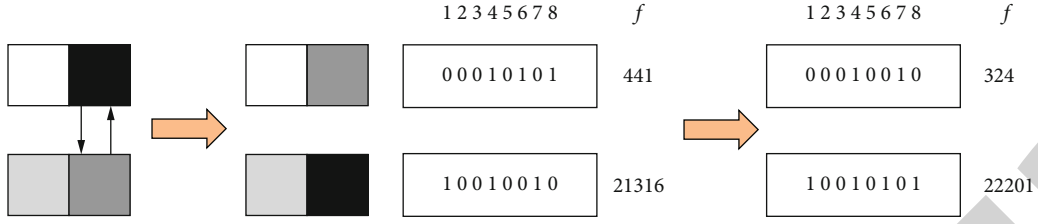


FIGURE 3: One-point intersection example.

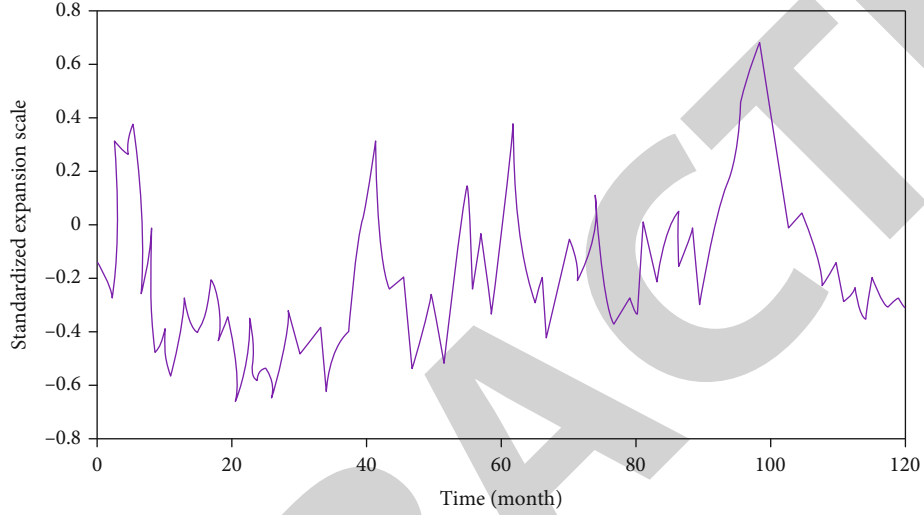


FIGURE 4: Normalized urban expansion curve.

The step of immune cloning algorithm is to randomly generate a set H , which contains antibodies, and the size of the antibody is expressed in N . Then, the affinity between each antibody in the set is calculated, and the n antibodies with the highest affinity value are reconstituted into a new set. Copy the antibodies in the new set. Each antibody copies a certain number to temporarily form a clone set. The cloneable set is given a certain mutation probability, and the mutated set is represented by C_n . Calculate the antibody affinity in the mutated set and then select the first few antibodies with the highest affinity value. These antibodies form a memory cell set and replace the antibodies with low affinity value in the initial set H , so as to maintain the diversity of antibody population.

4. Experimental Design and Analysis

Too much sample data selection will lead to slow operation speed. On the contrary, too little sample data will affect the accuracy of prediction. Here, 500 groups of data samples are selected, of which 400 groups are randomly selected as input data, and the remaining 100 groups are used as test samples to test the effect of the model through error comparison. The columns contained in a dimension that represent the dimension are called dimension attributes. Get dimension properties from the report. Or talk to business personnel to get. When the attribute hierarchy is instantiated as a series of dimensions instead of a single dimension, it is called

TABLE 1: Contribution rate of different dimensions.

Dimension	Contribution rate (%)	Dimension	Contribution rate (%)
1	47.1767	8	93.0468
2	58.2130	9	95.1792
3	67.7811	10	96.8751
4	74.3844	11	97.5912
5	80.8125	12	98.3073
6	85.8746	13	99.6110
7	89.9968	14	100.0000

a snowflake pattern. The operation of merging the attribute hierarchy of a dimension into a single dimension is called denormalization. Using snowflake mode, users need a lot of correlation operations in the process of statistical analysis, which has high complexity and poor query performance. It is convenient, easy to use, and has good performance. Due to different dimensions between data sources, effective comparison cannot be carried out. It needs to be normalized to convert dimensional data into dimensionless data and map the data of statistical probability between $[0, 1]$, and the mapping of data is in the range of $[-1, 1]$. The normalization formula of the original data through the premnmx function is shown in

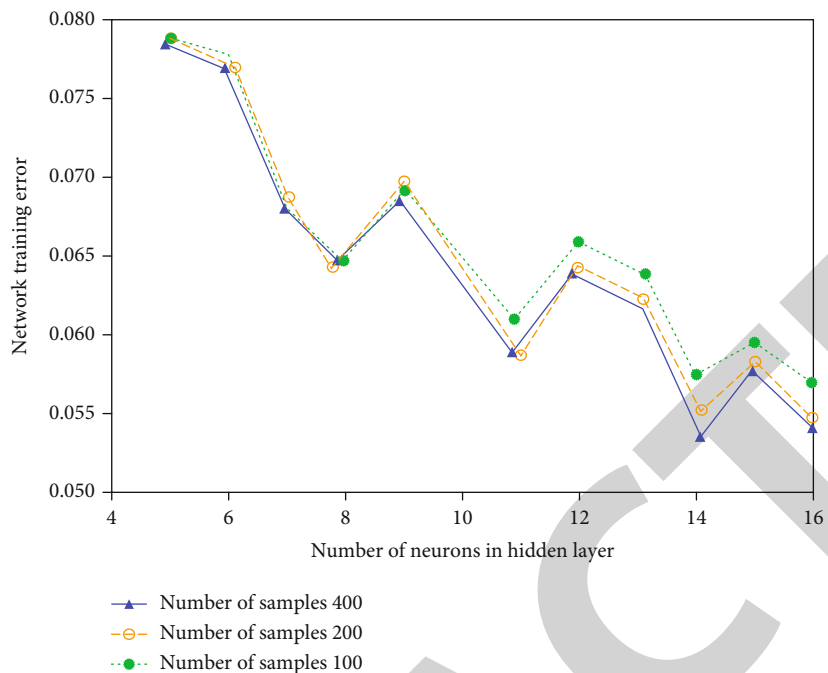


FIGURE 5: Network training error corresponding to the number of hidden layers.

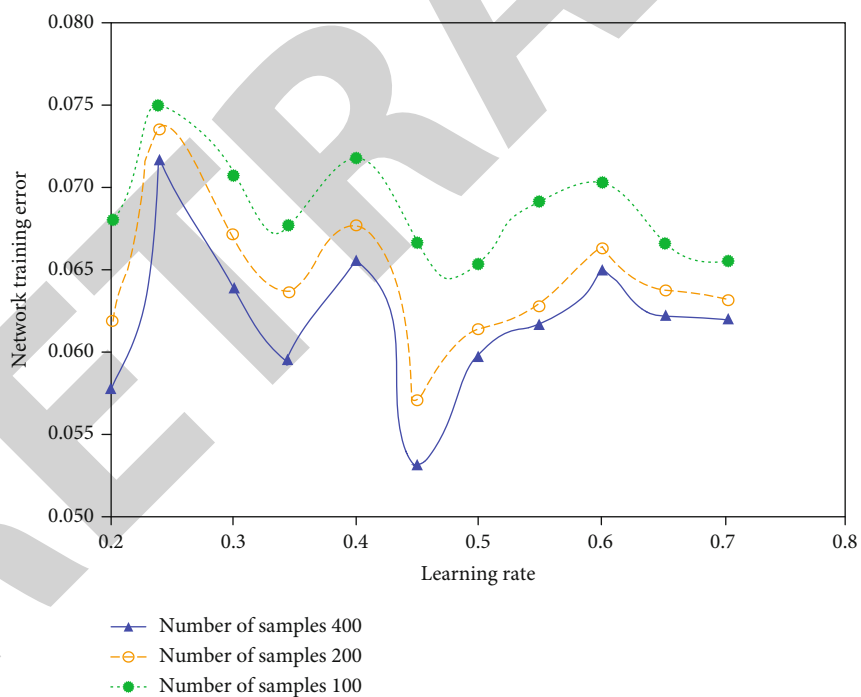


FIGURE 6: Network training error corresponding to different learning rates.

$$\begin{aligned}
 \text{PN} &= \frac{2(p - \min p)}{\max p - \min p} - 1, \\
 \text{TN} &= \frac{2(t - \min t)}{\max t - \min t} - 1.
 \end{aligned}
 \tag{16}$$

In equation (16), p is the input sample of the original data, t is the output sample of the original data, $\min p$ and

$\min t$ represent the minimum value of the original input sample and output sample, respectively, and $\max p$ and $\max t$ represent the maximum value of the original input sample and output sample, respectively. PN and TN obtained by normalization represent the factors affecting the expansion of urban scale and the value of expansion scale, respectively. The data obtained after fitting by the BP neural network is still the normalized data, and the inverse

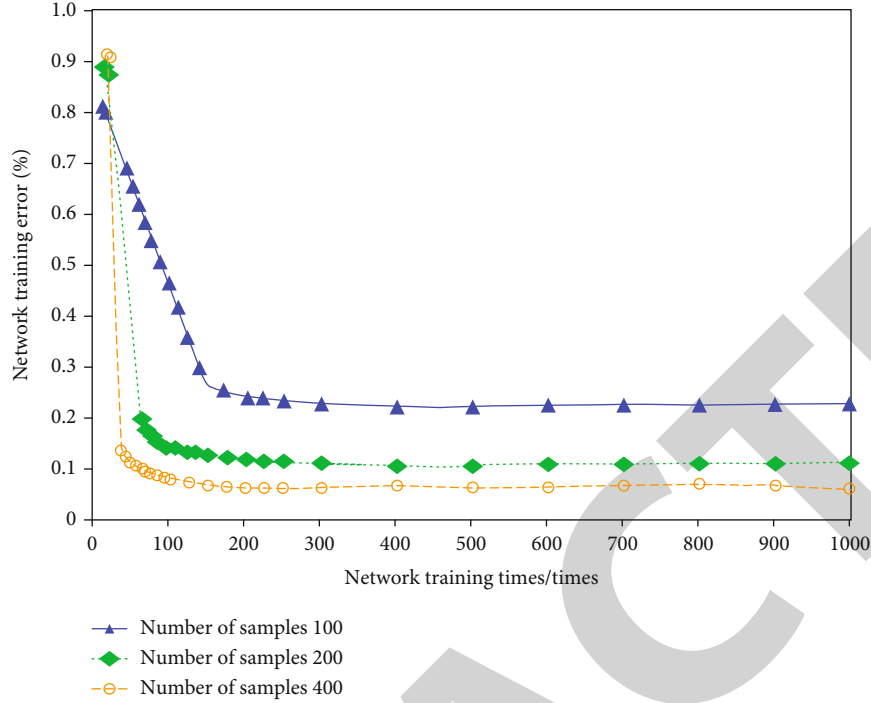


FIGURE 7: Network training error corresponding to different training times.

normalization function is also required to obtain the normal value. After normalizing the original data output samples, the numerical diagram is obtained, as shown in Figure 4.

The dimensionality of data samples is different and needs to be reduced. PCA dimensionality reduction method can be used. Based on linear transformation, the high-dimensional data is projected into the low-level space, that is, the dimensions with a large amount of information in the data are saved and the unimportant information is deleted, which can effectively reduce the running time of the algorithm. Firstly, the sample features are averaged, that is, the difference between the sample features and the feature mean of the rigid sample, the eigenvalues and eigenvectors in the covariance matrix of the sample are calculated, and the eigenvectors corresponding to the previous maximum eigenvalues are selected and recombined into a new matrix to obtain a new sample set. According to the experimental data, the contribution rates corresponding to different dimensions are obtained according to PCA dimensionality reduction, as shown in Table 1.

According to the dimension reduction experiment, the convergence curve of contribution rate after 9-dimensional class is gradually slow, so 9-dimensional data are selected for experimental simulation. In the construction of the BP neural network model, a three-layer network structure is selected to meet the requirements of the operation rate. Nine-dimensional data is used as the number of neurons in the input layer of neural network, and the number of neurons in the output layer is the result of urban expansion prediction. The number of neurons in the hidden layer needs to be determined by trial and error. Too few or too many neurons in the hidden layer will affect the output error. There-

TABLE 2: Algorithm and simulation time corresponding to different cross-probabilities.

Crossover probability	Algorithm error	Simulation time (seconds)
0.4040	0.0438	0.2679
0.4888	0.0396	0.2679
0.5737	0.0815	0.1733
0.6585	0.0374	0.2836
0.7434	0.0754	0.1576
0.8282	0.0840	0.1733
0.9130	0.0281	0.6618
0.9979	0.0399	0.8193

fore, the number of neurons in the hidden layer is determined to be 5-16 according to the common empirical formula. When the number of sample sets is different and other conditions are the same, the comparison of network errors corresponding to different numbers of nodes in the hidden layer is shown in Figure 5.

According to the simulation results in the figure, when the training error reaches the minimum value of 0.053645, the number of neurons in the hidden layer is 14, so the number of neuron nodes is 14. To sum up, the number of neurons in the input layer of the experimental model is 9, the number of neurons in the hidden layer is 14, and the number of neurons in the output layer is 1.

If the learning rate of BP neural network is set too high, it may lead to the error of optimal solution. If the learning rate is too small, it will lead to the long running time of the algorithm. The appropriate learning rate can make the

TABLE 3: Parameters of neural network.

Parameter name	Parameter value	Parameter name	Parameter value
Number of network layers	3	Training times	150
Enter the number of layers	9	Population size	50
Number of hidden layers	14	Number of iterations	80
Number of output layers	1	Crossover probability	0.652
Learning rate	0.45	Variation probability	0.01

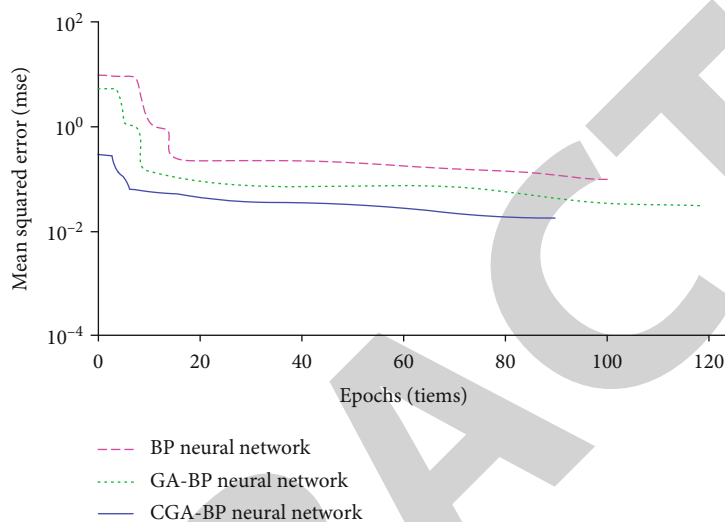


FIGURE 8: Comparison of BP neural network training under different models.

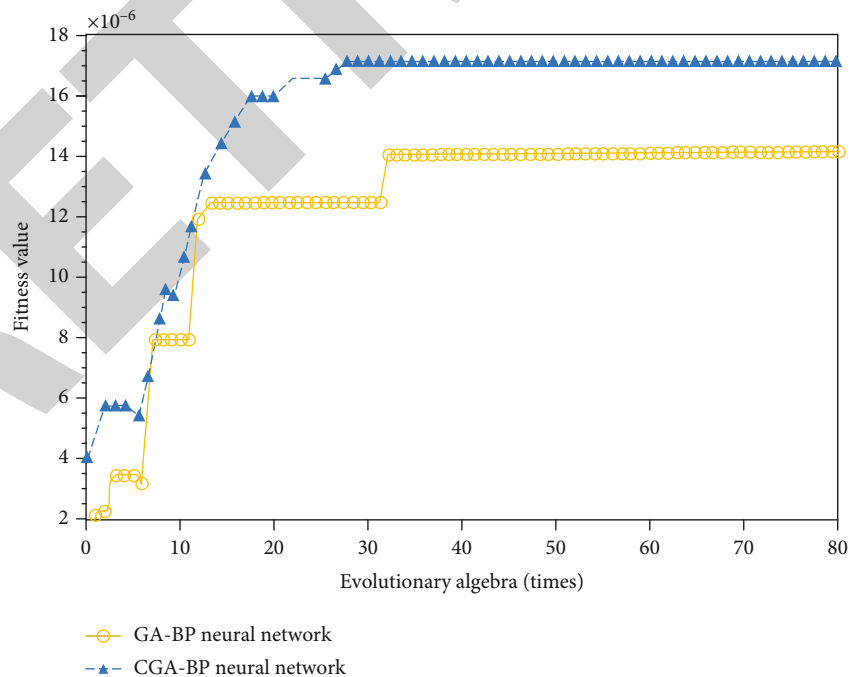


FIGURE 9: Comparison of iteration times of genetic algorithm under different algorithms.

algorithm have better operation performance. Set the learning rate to 0.2 to 0.7. The results are shown in Figure 6.

Different training times also have a certain impact on the prediction accuracy. Therefore, gradually increase the training times based on the determination of the neural network framework structure and analyze its correlation with the accuracy prediction. The results are shown in Figure 7. The population size can be determined by the optimized performance of the algorithm. After comparing the prediction accuracy and simulation time, it can be determined that the population size is the best of 50.

The convergence accuracy of the genetic algorithm will be affected by the number of iterations. It is necessary to find a balance between efficiency and accuracy when the algorithm converges. When the crossover probability is large, the offspring will produce unstable changes and destroy the excellent model of the population. On the contrary, when the crossover probability is small, the evolution process of the algorithm will slow down and spend a lot of time in the genetic search stage. Therefore, the crossover probability is generally limited to 0.4-0.99. The error size and simulation time results corresponding to different crossover probabilities are shown in Table 2. When the algorithm error is small and the simulation time is appropriate, the crossover probability is 0.652.

Because the mutation probability will affect the generation of new individuals, the mutation probability is too high, and the generation probability of new individuals will also increase, which may also destroy the existing excellent pattern structure. On the contrary, the mutation probability is too small, there is no change of new individuals, and it is not conducive to the convergence of the algorithm. Therefore, the range of mutation probability is limited to 0.0001-0.1, and the initial value is 0.01. Through the above experiments, the parameters of the prediction model for predicting urban scale expansion are obtained, which are summarized as shown in Table 3.

After the model parameters are determined, the performance of the model is reflected through the model results. The relationship between the training times under different models and the target error is shown in Figure 8.

The mean square error levels of the three neural networks can be seen from Figure 8. Among them, the mean square error level of the BP neural network is the highest, and GA-BP neural network. The lowest error is the CGA-BP neural network. The difference between the estimated quantity and the estimated quantity is the smallest. After training the GA-BP neural network and CGA-BP neural network model, the relationship between iteration times and fitness value is shown in Figure 9.

5. Conclusion

For urban development, a prediction model of urban scale expansion based on the BP neural network optimized by Xiaosheng genetic algorithm is proposed and compared with other models. The results show that the neural network parameters have suitable values, the number of model dimensions is about 9, the contribution rate of the model

is more than 95% and tends to be stable, the network training error is the smallest, the number of hidden layer nodes is 14, and the model learning rate is 0.45. The number of data samples will also affect the network training error, and with the increase of training times, the error change is no longer significant. When the number of samples is 400, the number of training times is 50, and the error tends to be flat. Considering the optimization accuracy and simulation time, the population size is the best value of 50. When the number of iterations of the model is 150, the genetic algorithm begins to converge and the running time of the algorithm is less. The crossover probability can be seen from Table 2. When the crossover probability is 0.652, the error is small on the basis of moderate simulation time, and the mutation probability can neither be large nor small, so it is set to 0.01. When the parameter conditions are the same, the BP neural network model, GA-BP neural network model, and CGA-BP neural network model are compared. The CGA-BP neural network model achieves the target error when the number of training is 79, which is better than the other two models. The fitness value corresponding to the number of iterations also shows that the best fitness value of the CGA-BP neural network when the number of iterations is the least is also higher than other models. Therefore, the neural network model optimized by genetic algorithm is more suitable for the prediction of urban scale expansion. However, there are still some limitations in this paper. Limited by various conditions, the research still has deficiencies, and the parameter setting is not comprehensive enough. Genetic algorithm is insufficient to optimize the neural network structure model. In the follow-up work, it is necessary to further improve the quality of the research model and obtain the prediction results more accurately.

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 he/she has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] B. Alharthi and T. A. El-Damaty, "Study the urban expansion of Taif City using remote sensing and GIS techniques for decision support system," *Advances in Remote Sensing*, vol. 11, no. 1, pp. 1-15, 2022.
- [2] Y. Wang, L. Wang, F. Yang, W. di, and Q. Chang, "Advantages of direct input-to-output connections in neural networks: the Elman network for stock index forecasting," *Information Sciences*, vol. 547, pp. 1066-1079, 2021.
- [3] F. I. Chou, W. H. Ho, and C. H. Chen, "Niche genetic algorithm for solving multiplicity problems in genetic association studies," *Intelligent Automation and Soft Computing*, vol. 26, no. 3, pp. 501-512, 2020.

Retraction

Retracted: Application of Internet of Things Based on Big Data Ecosystem in Factory Energy Consumption Analysis Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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] A. Li, C. Zhang, and L. Li, "Application of Internet of Things Based on Big Data Ecosystem in Factory Energy Consumption Analysis Model," *Journal of Function Spaces*, vol. 2022, Article ID 7631323, 10 pages, 2022.

Research Article

Application of Internet of Things Based on Big Data Ecosystem in Factory Energy Consumption Analysis Model

Ang Li ^{1,2}, Chen Zhang ¹, and Lei Li ²

¹Institute of Electronic Engineering and Optoelectronic Technology, Nanjing University of Science and Technology Zijin College, Nanjing 210023, China

²College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Correspondence should be addressed to Ang Li; liang878@njust.edu.cn

Received 10 May 2022; Revised 24 June 2022; Accepted 28 June 2022; Published 12 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Ang 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 demand for energy in factories is huge. With abundant energy supply, factories can raise their production capacity to a new height. Energy is the material basis of domestic economic development. As the largest developing country in the world, energy shortage has become an urgent problem for China. In this paper, an Internet of Things based on big data ecosystem is proposed to analyze the energy consumption of the factory and build a model. The Internet of Things technology of big data ecosystem can be summarized as a technology that uses information sensing devices to complete the transaction and network connection according to the protocol content. A total of 853,000 power distribution operations were carried out in the power grid. In 2019, the average ratio of decision tree algorithm, machine learning algorithm, and machine learning algorithm was 36.8%, 37.4%, and 43.5%, respectively. Compared with the three methods, the method in this paper increased by 37.9% year-on-year and reduced the power outage by 2.63 million households, which is equivalent to a corresponding reduction of 35 users per operation. The functional requirements of the IoT energy consumption analysis system in factories based on big data ecosystem are reflected in three aspects: energy consumption monitoring and management, power control and management, and energy consumption supervision and analysis. Based on the management of energy consumption monitoring and power control through the software platform, the functional requirements of the system are analyzed.

1. Introduction

After the founding of New China, especially since the reform and opening up, China's manufacturing industry has developed rapidly and continuously, and a complete and independent industrial system has been built, which has effectively promoted the process of industrialization and modernization and significantly enhanced the overall national strength. Panfactory electric power pays more attention to how to mine the value of collected data; that is, through the analysis and mining of data, it can achieve the optimal management of the whole power system of the factory and then promote energy conservation and consumption reduction on the energy consumption side and use energy intelligently, thus

providing important support for the transformation and upgrading of the energy system [1, 2]. At present, factories are playing an extremely important role. Uninterrupted operation refers to the operation in which factory personnel directly contact live lines or equipment, or factory personnel use special tools, equipment, or devices to work on live lines or equipment, so as to carry out maintenance and testing on uninterrupted power lines or equipment [3, 4]. The demand for energy in factories is huge. With abundant energy supply, factories can raise their production capacity to a new height. Energy is the material basis of domestic economic development. As the largest developing country in the world, energy shortage has become an urgent problem to be solved in China [5]. In this case, it is necessary to supervise the

energy and further help the factory to reduce costs and increase efficiency, so as to build a more perfect smart factory.

This paper proposes an Internet of Things based on big data ecosystem, analyzes the energy consumption of factories, and constructs a model. The Internet of Things technology of big data ecosystem can be summarized as a technology that uses information sensing equipment to complete the connection between transactions and networks according to the protocol content [6, 7]. In this technology, things mainly use information sensing media to promote information data exchange, sorting, adjustment, and optimization and then complete the optimization, identification, positioning, tracking, supervision, and other functions of intelligent system [8]. The Internet of Things of big data ecosystem is the application extension and business expansion based on the Internet. The composition and operation of the Internet of Things can achieve the connection between things and networks. The Internet of Things is regarded as a great opportunity for development and change in the information field. The Internet of Things of big data ecosystem is a network with self-configuration capability. Similar to the Internet, the Internet of Things must also operate according to the communication protocol [9, 10]. Internet of Things technology has been extended to all aspects of people's life. It can be said that driven by Internet of Things technology, it has played an important role in both social and economic development and the improvement of people's quality of life [11]. The rapid development of the Internet has led to a sharp increase in the number of users and data of various Internet applications. Only increasing the storage capacity of single point devices cannot solve the massive data processing needs of users and enterprises. At the same time, the big data ecosystem gradually developed on the basis of distributed file system to meet the storage and calculation of massive data has been gradually improved [12].

Based on the Internet of Things technology of big data ecosystem and the sensor nodes of Internet of Things that monitor the factory energy consumption analysis data of actual production equipment in factories as specific data sources, this paper puts forward an innovative design concept for the overall architecture of the factory energy consumption analysis system based on the Internet of Things of big data ecosystem, adapts the characteristics of big data of factory energy consumption analysis, designs and implements a complete factory big data processing system, and puts forward energy-saving strategies according to the analysis of big data of factory energy consumption by this system [13, 14]. The research and development of processing technology and system architecture of massive data generated in the analysis of energy consumption in factories are still in its infancy. Because of its many differences with Internet big data, a set of processing methods and tools that directly copy Internet big data cannot meet the demand of Wang Ye production for big data processing [15]. However, due to insufficient investment in information technology in traditional industries, relatively backward technology, and lack of rational use of data, the big data ecosystem of factory energy consumption analysis is still far from perfect.

The innovation of this paper is to propose an Internet of Things based on big data ecosystem, analyze the energy consumption of factories, and establish a model. Compared with the traditional three methods, the outage time of the proposed method is reduced by 15.3% year-on-year, which is equivalent to reducing the downtime of 35 users by 2.33 h each time. In the future, the Internet of Things must face the realistic problem of the proliferation of Internet of Things devices. At the same time, mobile communication itself has high requirements for the reliability and immediacy of communication. Therefore, on the premise of meeting the needs of all parties, mature dynamic spectrum management in the Internet of Things can be realized. And correctly use the spectrum sensing technology to realize the content of spectrum sharing in the Internet of Things.

Based on the big data ecosystem, the application of the Internet of Things in the factory energy consumption analysis model is analyzed. The Section 1 describes the background of industrialization and modernization. The Section 2 analyzes the research status of related work. The energy consumption of the Internet of Things in the factory is analyzed. Section 3 analyzes the principle and model of the Internet of Things big data ecosystem. Section 4 implements the plant energy consumption analysis model and designs the IoT plant energy consumption analysis system. Section 5 summarizes the full text. The research will gradually increase data types and design more targeted algorithms to process and analyze data, so as to further optimize energy-saving strategies and verify the energy-saving effects of relevant strategies.

2. Related Work

2.1. Research Status at Home and Abroad. Ulusoy et al. proposed that in most factories, due to the failure to coordinate the energy distribution among various production departments, the comprehensive utilization rate of factory energy is still high, and it is difficult to achieve the energy-saving goal, so it is impossible to achieve the statistical analysis of energy consumption of the whole factory, and there is no data support for whether the energy consumption is saved [16]. Gutierrez-Osorio et al. proposed that the energy consumption system uses information technology to promote energy conservation and consumption reduction, which is an important measure to improve the fine management level of enterprises, and provides technical support for enterprises to accurately grasp and analyze the development trend of energy conservation and consumption reduction and make scientific use of energy [17]. Li et al. put forward that the factory still records and manages energy information in a manual way, which wastes a lot of labor costs. The data stays on paper, which cannot guarantee the accuracy of the data and provide data basis for leaders' decision-making quickly and accurately. At the same time, there is a lack of statistics on the operation of equipment startup and shutdown, resulting in low efficiency of energy information processing and lack of decision-making mechanism to effectively use energy consumption data and deeply analyze the operation status of energy system, and the control of energy consumption

cost always lags behind the occurrence of cost [18]. Mumtaz et al. Proposed the acquisition method of node preprocessing, which means that the acquisition node preprocesses the sensor data to a certain extent. For example, aggregate multiple data to a certain extent, and then transmit. After preprocessing, the bandwidth required for the transmission of collected data will be greatly reduced, which will lead to the corresponding reduction of energy consumption [19]. Orenga-Rogla et al. proposed that for factories, in order to directly or indirectly reduce the total amount of greenhouse gas emissions and achieve “zero emission” of carbon dioxide, it is necessary to implement staggered peak power consumption; that is, according to the load characteristics of the power grid, transfer part of the load during the peak period of power grid to the low period of power consumption through administrative, technical, and economic means, so as to reduce the peak and valley load difference of the power grid and optimize the allocation of resources and improve the security and economy of power grid [20]. Ahad and Biswas proposed to establish a perfect energy consumption monitoring and energy management system; realize the informationization, visualization, and controllability of the dynamic process of energy consumption; and monitor and manage the energy consumed in the factory production process, which can greatly improve the energy efficiency [21]. Sánchez et al. put forward that nearly 50% of the energy saving in California comes from the energy efficiency services of public utilities. Applying accurate electricity metering data can directly provide users with the basis for energy-saving decisions and, at the same time, improve and change the current electricity consumption mode to improve the energy efficiency. As electricity involves all users, the energy saving generated by energy efficiency services will account for 50% of the total energy saving in the whole state [22]. Broring et al. put forward that in order to ensure that the carbon emissions reach the standard without affecting the realization of the enterprise’s production objectives, the factory should realize automatic meter reading by the Internet of Things, remote control, and other technologies, and the computer should monitor the energy consumption data in real time and make corresponding adjustments to help enterprises save costs, reduce carbon emissions, and achieve green production [23]. Asch et al. put forward a model-based approach combining the advantages of preprocessing, further abstracting the data collection model, selecting a reasonable algorithm according to the characteristics of data, calculating the collected data to a certain degree, and finally sending information such as a certain coefficient or weight of the relevant data, thus reducing the traffic volume and greatly reducing the energy consumption [24]. The central air conditioner produced by Jan et al. is the key energy-consuming equipment, which usually accounts for more than 50% of the total energy consumption of the enterprise. Compared with other energy-consuming equipment, the cost of energy consumption control and renovation of central air-conditioning is small, and it will not affect production [25].

2.2. Research Status of Internet of Things in Factory Energy Consumption Analysis. The above open-source technologies

and frameworks have been widely used in the field of Internet, but there are still many improvements and adaptations for plant energy consumption analysis. Therefore, the Internet of Things based on big data ecosystem is studied in the factory energy consumption analysis in this paper. The factory energy consumption analysis platform designed in this paper is also designed and implemented on the basis mentioned above. Energy consumption is divided into two parts: energy consumption online monitoring system and energy consumption management system. It can provide some energy consumption statistics for enterprises and summarize them into the local total energy. The other is to help enterprises save energy fundamentally. According to the annual and quarterly comprehensive energy consumption of the products, the energy consumption at the plant level is managed in a multiangle and multilatitude manner. From the aspects of energy use type, monitoring area, production process/section time, sub items, etc., the energy consumption statistics of enterprises, year-on-year and month-on-month analysis of energy consumption per unit product, and energy consumption per unit output value are carried out by means of curves and digital tables. Find out the loopholes and unreasonable places in the process of energy use, so as to adjust the energy distribution strategy and reduce the waste in the process of energy use. In the field of Internet of Things of big data ecosystem, embedded system is also widely used. It is a place for embedded system technology in the system development of all kinds of Internet of Things terminals. Based on the big data ecosystem, the Internet of Things faces the total power consumption, total water consumption, total incoming line, and other elements in the process of plant energy consumption analysis. The enterprise integrated energy consumption management system is an integrated management and control system. It is aimed at the energy instruments such as water, electricity, gas, and steam in industrial enterprises and carries out remote data acquisition and control through networking, integrating wired and wireless measurement and control and computer LAN. Form a network system with multiple transmission media to monitor the operation status of on-site energy media in real time. The parameters of each instrument can be remotely collected and set through the network, and the running status of the instrument can be monitored in real time through the computer terminal. The results of energy consumption monitoring are reflected in the form of histogram and data, in which the histogram faces each cycle such as day, month, and year. In the process of querying factory energy consumption analysis, you can freely filter indicators and set the energy consumption data range, time range, and indicator type required in the query.

3. Internet of Things Principle and Model of Big Data Ecosystem

At present, the related technologies of big data are mainly used in the Internet environment, which is used to process the device data and massive media data generated in the network. The Internet of Things technology of big data ecosystem mainly uses communication radio frequency

identification technology and data communication technology to form an integrated network. Its main feature is to realize information sharing and provide convenient conditions for information exchange and development in various industries. With the maturity of technologies in the field of big data processing, the demand for data mining has become more and more obvious, and related machine learning algorithms have been applied to big data ecosystems, and a series of open-source machine learning frameworks for big data ecosystems, represented by Mahout and MLlib, have emerged. The function realization of the factory energy consumption analysis model is based on the Internet of Things of big data ecosystem. Using the query tool script of the Internet of Things of big data ecosystem can enable non-technical personnel to query and count the data in the massive data scene like relational database, which is a very reasonable solution. The database language provided by the plant energy consumption analysis model and the data statistics requirements of the platform realized in this paper that are not strong in real-time requirements use the plant energy consumption analysis model. The IoT flow chart of big data ecosystem is shown in Figure 1.

The real-time data query processing of Internet of Things based on big data ecosystem can meet the needs of users for fast data query and can quickly locate energy consumption information and equipment information in massive data. Making use of the advantages of the Internet of Things of the pan-big data ecosystem in comprehensive state perception and efficient information processing to carry out pioneering changes in the analysis and management of factory energy consumption can effectively change the defects of long energy consumption and low efficiency. This paper presents an energy consumption analysis model of Internet of Things in factories based on big data ecosystem. The overall structure of the model is shown in Figure 2.

The model is mainly based on the data acquisition system, and the function analysis module realizes the establishment and management of uninterrupted power operation scheme. The data acquisition system mainly collects relevant data from the electricity acquisition system, PMS2.0 system, marketing system, GIS system, and SCADA system. The PMS2.0 system takes asset life cycle management as the main line, condition maintenance as the core, realizes the integration of drawings and numbers, and has the linkage mechanism of "account-card-object." The real condition of equipment can be collected through the PMS system.

The application of energy detection method in the local detection of secondary users is the most common, because its implementation difficulty is relatively low, the calculation cost is low, and the judgment method is direct. A binary detection model can be used to simulate whether the received main user signal exists. The sampling signal $e(t)$ of the secondary user at time t is as follows:

$$\begin{cases} H_1 : e(t) = s(t)h(t) + w(t), \\ H_0 : e(t) = w(t), \end{cases} \quad (1)$$

where H_1 represents the busy channel, that is, the primary user is using the channel; H_0 is the opposite; $s(t)$ represents the signal of the primary user; $w(t)$ is the additive white Gaussian noise (AWGN) existing in the channel; and $h(t)$ is the channel gain between the secondary user and the primary user. Assuming $s(t)$ and $w(t)$ are independent of each other, using the energy detection method, when the secondary user samples S times on a channel, the energy calculated by the secondary user for the received signal is

$$E = \sum_{n=1}^S \frac{|e(n)|^2}{S}. \quad (2)$$

Because when n changes, E follows the standard normal distribution, E constitutes a chi-square distribution. γ is the instantaneous signal-to-noise ratio of the received signal. When the main user does not exist, E follows the central chi-square distribution with degree of freedom $2S$. When the main user exists, E follows the noncentral chi-square distribution with $2S$ degree of freedom and 2γ noncentral parameter

$$\begin{cases} H_1 : E \sim \chi_{2S}^2(2\gamma), \\ H_0 : E \sim \chi_{2S}^2. \end{cases} \quad (3)$$

Assuming that the number of sampling points S is large enough and that the signal and noise are zero-mean circularly symmetric Gaussian random distributions and they are independent of each other, according to the central limit theorem, E conforms to Gaussian distribution.

$$\begin{cases} H_1 : E \sim N\left(\sigma_w^2(1+\gamma), \frac{2\sigma_w^2(1+\gamma)^2}{S}\right), \\ H_0 : E \sim N\left(\sigma_w^2, \frac{2\sigma_w^4}{S}\right). \end{cases} \quad (4)$$

The detection probability and false alarm probability in single energy detection are

$$P_d = P\{E \geq \lambda | H_1\} = Q\left(\frac{\lambda}{\sqrt{(2/S)\sigma_w^2(\gamma+1)}} - \sqrt{\frac{S}{2}}\right), \quad (5)$$

$$P_f = P\{E \geq \lambda | H_0\} = Q\left(\frac{\lambda - \sigma_w^2}{\sqrt{(2/S)\sigma_w^4}}\right), \quad (6)$$

$Q(x) = (1/\sqrt{2\pi}) \int_x^{+\infty} e^{-(t^2/2)} dt$ is the right tail function of the standard normal distribution, which decreases with the increase of x , and is a decreasing function. Therefore, it can be seen from formula (5) that the higher the signal-to-noise ratio, the greater the detection probability. It can be seen from formula (6) that the larger the sampling point S is, the smaller the false alarm probability is.

By analyzing the characteristics of the noise energy of the signal, the high and low thresholds λ_H and λ_L are set. The noise uncertainty parameters of wireless environment are the ratio of actual noise power to Gaussian white noise

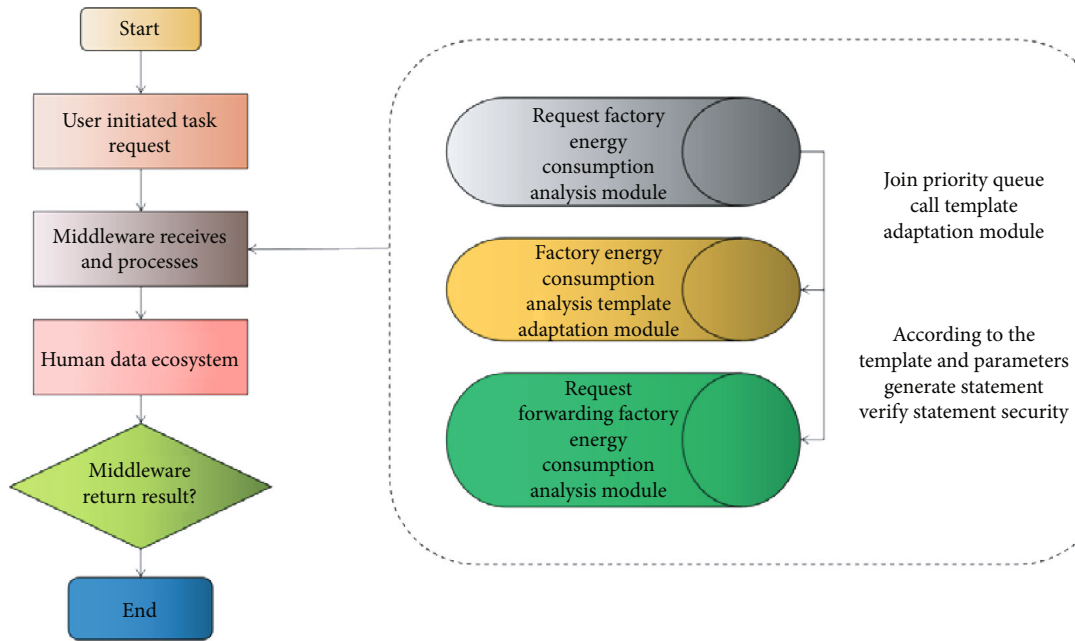


FIGURE 1: Internet of Things flow chart of big data ecosystem.

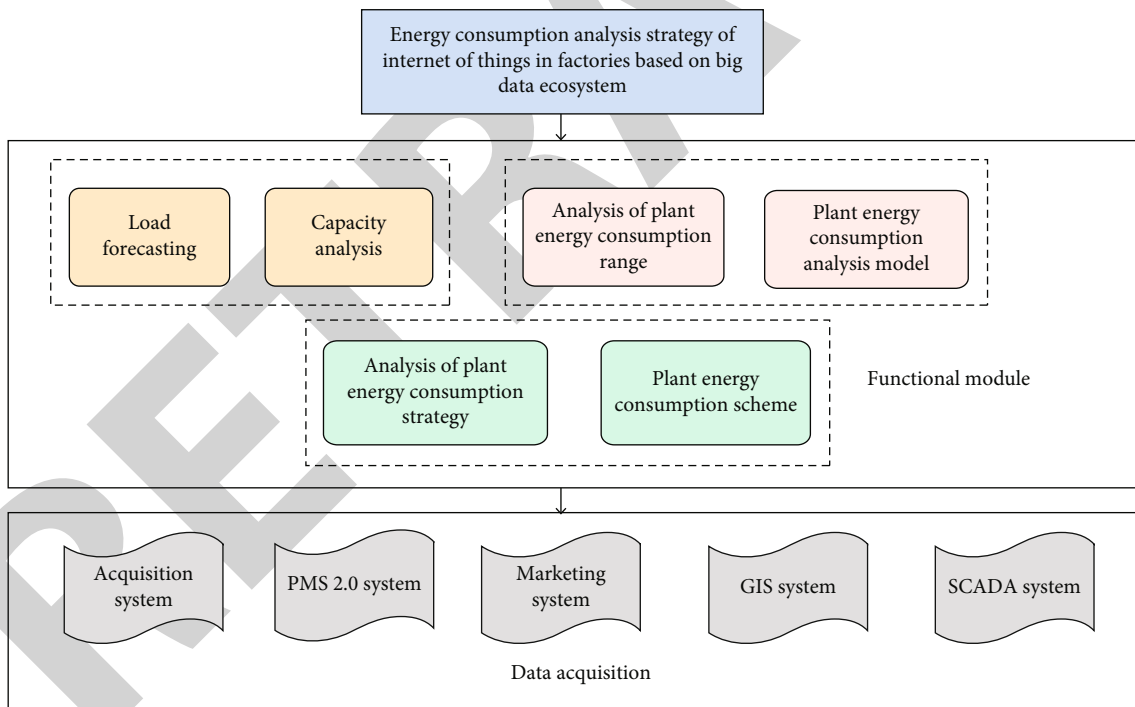


FIGURE 2: Analysis model diagram of IoT energy consumption in factory based on big data ecosystem.

power ρ , $\rho = \sigma_w^2 / \sigma_w^2$. The high and low thresholds of double-threshold energy detection method are

$$\begin{aligned} \lambda_H &= \rho\lambda = \rho\sigma_w^2 \left(\sqrt{\frac{2}{S} Q^{-1}(P_f)} + 1 \right), \\ \lambda_L &= \frac{\lambda}{\rho} = \frac{\sigma_w^2}{\rho} \left(\sqrt{\frac{2}{S} Q^{-1}(P_f)} + 1 \right). \end{aligned} \tag{7}$$

When $E \geq \lambda_H$, it is determined that the main user exists. If $E \leq \lambda_L$, it is determined that the main user does not exist. If it is between two threshold values, the result is considered unreliable and no definite conclusion is given for the time being. The high and low thresholds are related to the channel state and can be adjusted adaptively according to the channel state. When the noise power in the channel is large, the difference between the high and low thresholds can be increased when the value increases. When the channel

quality is good, the difference between high and low thresholds can be reduced by reducing the value to make the determination more accurate. If the ρ value in the channel is 1, the high and low thresholds of $\lambda_L = \lambda_H = \lambda$ are equal, which is equivalent to single threshold detection.

When the energy observation value E_i is in the middle of the high and low thresholds λ_H and λ_L , the secondary user thinks that the subchannel result is unreliable and will temporarily store the data locally. If FC wants to make a second round of judgment, it will upload the value to FC. Secondary user's decision D_i for i subchannel can be expressed as

$$\begin{cases} 0 \leq E_i \leq \lambda_L : D_i = 0, \\ E_i \geq \lambda_H : D_i = 1. \end{cases} \quad (8)$$

In the local decision of secondary users, among the K subchannels detected by secondary users, K_1 get the decision result 1 and K_2 get the decision result 0. In addition, the subchannels with no definite conclusion of $(K - K_1 - K_2)$ are considered unreliable, and the energy data of these "unreliable channels" are stored locally. Then, the secondary user will conclude the channel state according to the " m -out-of- k " rule above, and the conclusion D_i is

$$\begin{cases} K_1 < A : D_i = 0, \\ K_1 \geq A : D_i = 1. \end{cases} \quad (9)$$

In this formula, $l \in [1, N_A]$ represents the l secondary user, and $A \in [1, K]$ is the threshold value when the secondary user locally fuses the results of each subchannel; that is, the secondary user will judge that there is a primary user signal on the measured subchannel when the judgment result of at least A subchannel in K subchannel is 1. Then, the secondary user whose local judgment D_i is 1 will send the local judgment result to FC, and the secondary user whose local judgment is 0 will not send the local judgment.

As the Internet big data-related technology has experienced a long period of development, the technology has become mature, and many of its methods have been applied to many industries and environments outside the Internet. As a medium, the Internet of Things integrates everything in the world with the virtual Internet to form a unified integrated network. The operation of the world will carry out social and economic activities based on the integrated network.

4. Implementation of Plant Energy Consumption Analysis Model

4.1. Design of Energy Consumption Analysis System for Internet of Things Factory. With the popularization of the Internet of things infrastructure and the diversification of sensor information collection, the amount of data generated by the bottom information collection nodes of the Internet of Things based on the big data ecosystem is increasing exponentially. How to reasonably and effectively process and utilize these data and make them become the basis of

TABLE 1: Real-time display page loading time of energy consumption test group.

Test group	Average loading time
Group 1	5.62 s
Group 2	5.25 s
Group 3	5.36 s
Group 4	5.57 s

TABLE 2: Cluster statistics query function select statement test.

Test group	Average loading time
Group 1	180.502 s
Group 2	182.608 s
Group 3	179.004 s
Group 4	188.553 s

TABLE 3: Test of count statement of statistical query function.

Test group	Average loading time
Group 1	30.677 s
Group 2	31.677 s
Group 3	31.522 s
Group 4	31.687 s

TABLE 4: KL divergence value when k is 3.

Clustering dimension	Cluster 1	Cluster 2	Cluster 3
Devld	0.448	0.156	0.024
Date	0.217	0.422	0.367
Time	0.271	0.487	0.517
Period	0.324	0.713	0.571

TABLE 5: KL divergence value when k is taken as 5.

Clustering dimension	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Devld	0.725	0.131	0.612	0.598	0.758
Date	0.138	1.078	0.467	0.312	0.603
Time	0.222	1.395	0.648	0.352	0.976
Period	0.391	2.402	0.815	0.538	1.067

industrial production and management is an urgent problem to be solved at present. The Internet of Things based on big data ecosystem has self-organization ability and generally requires low power consumption of nodes. Therefore, the ability of the Internet of Things equipment of the plant energy consumption analysis system is limited. For various optimization algorithms in cognitive radio, the main goal of optimization is to improve the accuracy of spectrum sensing. Once spectrum sensing is applied to the Internet of

TABLE 6: KL divergence value when k is 6.

Clustering dimension	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Devld	0.597	0.832	0.456	0.604	0.844	0.607
Date	0.367	0.354	0.365	0.412	0.154	0.402
Time	0.454	0.575	0.388	0.495	0.264	0.595
Period	0.688	0.835	0.687	0.686	0.542	10.95

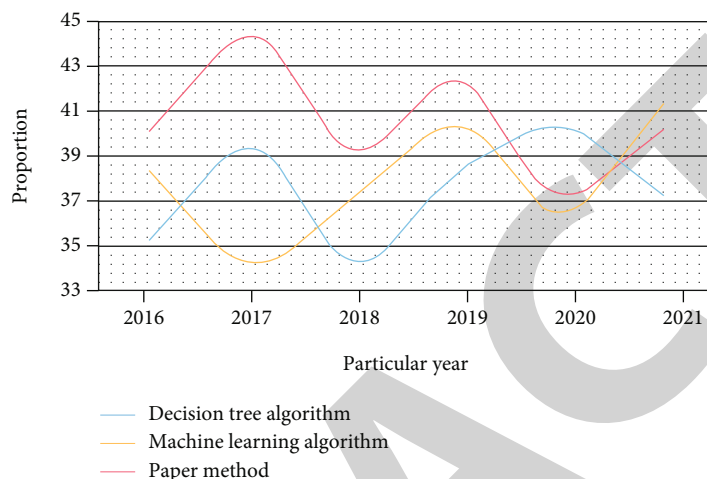


FIGURE 3: Number of times of live work in distribution network of power grid factory.

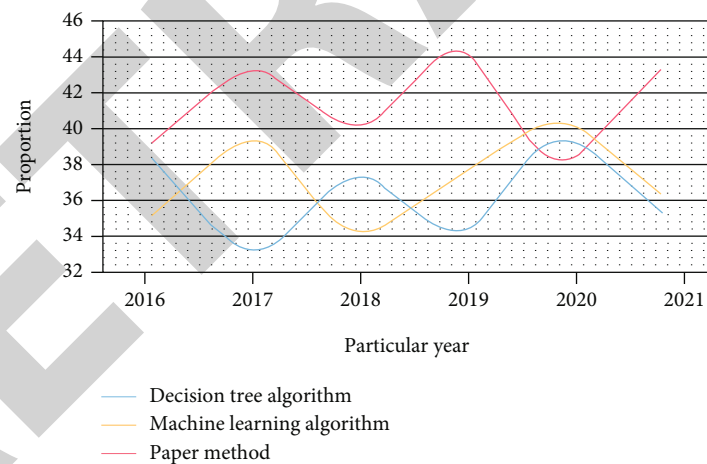


FIGURE 4: Number of times of live operation in distribution network of power grid factory.

Things, it is necessary to consider the actual energy consumption factors. The functional requirements of the plant energy consumption analysis system of the Internet of Things based on the big data ecosystem are reflected in three aspects: energy consumption monitoring and management, power control management, and energy consumption supervision and analysis. The energy consumption supervision is analyzed on the basis of managing energy consumption monitoring and power control through the software platform, and the functional requirements of the system are analyzed here.

4.2. *Experimental Results and Analysis.* Aiming at the demand of real-time energy consumption information, this experiment uses real-time power information to represent real-time energy consumption information. The data display is accurate and has certain reference value. At the same time, the webpage loading time has been tested in 4 groups, 10 times in each group, and the average loading time of each group is shown in Table 1.

For the application scenario of statistical query, select the energy consumption data in a certain period of time. Import the energy consumption data into HDFS and create a table.

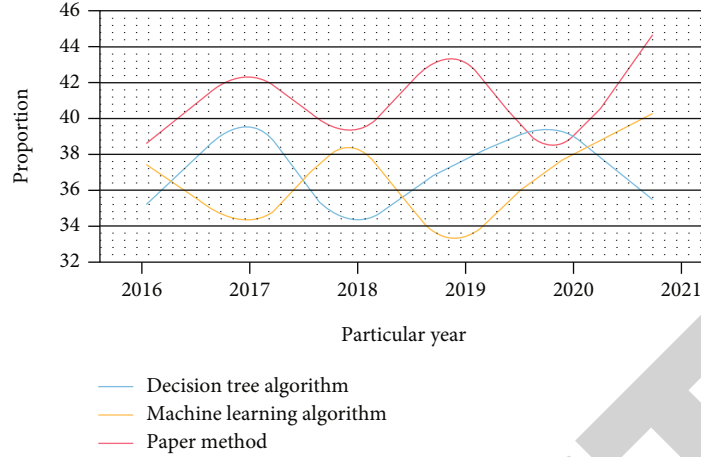


FIGURE 5: Number of times of live operation in distribution network of power grid factory.

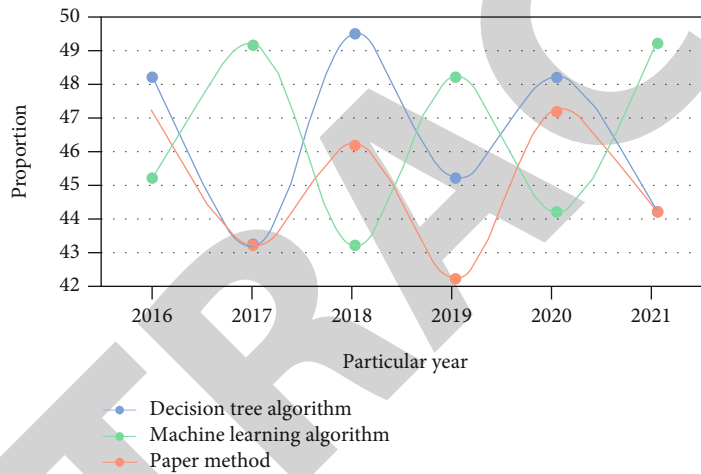


FIGURE 6: Average power outage time of plant users.

Here, the basic select and count statements are selected for testing. Each statement is divided into five groups for testing, and each group is tested ten times. The average query time is shown in Tables 2 and 3.

In cluster analysis, the selection of the number of clusters K has great influence on the results of cluster analysis. After many experiments, too many K values are selected, and only the most representative K values are selected into the following description; that is, K values of 3, 5, and 7 are selected for clustering. When the number of clusters K is 3, the results obtained according to the cluster analysis are shown in Table 4.

From the comparison in Table 4, it can be seen that the sensor network nodes represented by Devld have the highest correlation among clusters 1 and 3, and the sensor network nodes imply working areas. Different working areas in the actual factory are responsible for different groups, so the working groups have great influence on clustering. In cluster 2, the working time KL value of equipment period is the largest, which indicates that its correlation is high in cluster 2. Continue to analyze the KL value of the dimension of period

employees' length of service which is in the second largest position in cluster 3, so the correlation of employees' length of service is larger in cluster 3, which is also the key index.

When the value of K is 5, the same experiment is carried out on the data, as shown in Table 5.

When the value of K is 6, repeat the above clustering process, as shown in Table 6.

It can be seen from Tables 5 and 6 that there is little difference in KL values of Devld in each cluster except cluster 2, and the KL values of each dimension in other clusters are relatively large in the two dimensions of period and time, indicating that they are highly correlated in each cluster. It can be concluded that when k is 6, the KL value of each dimension is basically similar to that when k is 5, and the dimension with the highest correlation in each cluster is the same. After many tests of taking a larger value of k , the distribution of KL value basically does not change much, so it is more appropriate to take a value of k of about 3 to 4.

With the rapid development of social economy, the traditional uninterrupted operation can no longer meet the power supply reliability requirements of people's demand

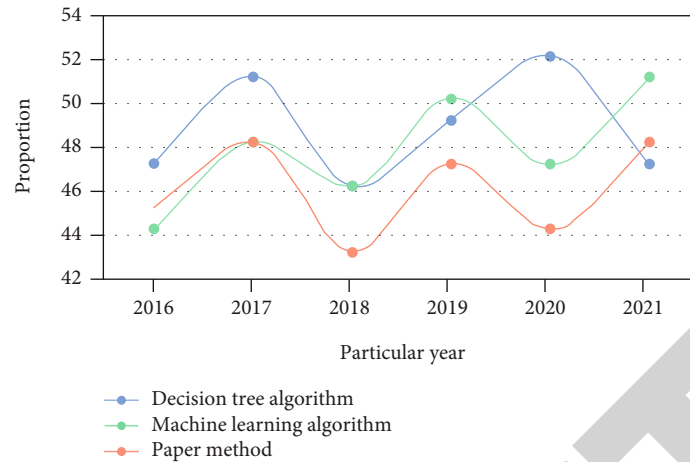


FIGURE 7: Average power outage time of plant users.

for a better life because of its long time-consuming and low efficiency. In this experiment, for the analysis of the effectiveness of live work in distribution network of a factory from 2016 to 2021, the decision tree algorithm, machine learning algorithm, and the method in this paper, including the completion times of live work, are used for three experimental comparisons. The experimental results are shown in Figures 3–5.

As can be seen from Figures 3–5, there are 853,000 times of uninterrupted power distribution in power grid factories. When the current year is 2019, the average percentage of decision tree algorithm, machine learning algorithm, and machine learning algorithm is 36.8%, 37.4%, and 43.5%, respectively. Compared with the three methods, the method in this paper increases by 37.9% year-on-year and reduces the power outage by 2.63 million h households. The excess power supply is about 748.59 million kW-h, which is equivalent to the power generation of 1.36 million kW installed power plant for about 2 years.

Similarly, this experiment is aimed at the analysis of the effect of reducing the average outage time of power outage users in power grid factories from 2016 to 2021 and adopts decision tree algorithm, respectively. The machine learning algorithm and the method in this paper, including the number of live-line operations, are compared twice. The experimental results are shown in Figures 6 and 7.

From Figures 6 and 7, it can be seen that the power grid factory reduces the average outage time of users. When the year is 2020, the average proportion of decision tree algorithm is 50.1%, the average proportion of machine learning algorithm is 45.5%, and the average proportion of this method is 30.5%. Compared with the three methods, the method in this paper reduces 15.3% year-on-year, which is equivalent to reducing the outage time of 35 users by 2.33 h each time. The future Internet of Things must face the realistic problem of the proliferation of Internet of Things devices. At the same time, mobile communication itself has high requirements for the reliability and immediacy of communication. Therefore, how to realize mature dynamic spectrum management in the Internet of Things and properly use spectrum sensing technology under the

condition of meeting the requirements of all parties is the key to realize spectrum sharing in the Internet of Things.

5. Conclusions

With the development of large-scale data, modern personal data protection technology continues to develop. In principle, it cannot prevent the disclosure of personal data. At the same time, the current laws are imperfect and lack strong technical support. The IoT of big data ecosystem is studied in the factory energy consumption analysis model. The power grid plant has 853,000 times of uninterrupted power distribution. When this year is 2019, the average percentages of decision tree algorithm, machine learning algorithm, and machine learning algorithm are 36.8%, 37.4%, and 43.5%, respectively. Compared with the three methods, this method has a year-on-year increase of 37.9% and a reduction of 2.63 million hours of power failure. The energy consumption data collected by the system is not rich enough, and the analysis and mining of energy consumption data need to be further deepened. On the premise of meeting the needs of all parties, this paper can realize the mature dynamic spectrum management in the Internet of Things. And correctly use the spectrum sensing technology to realize the content of spectrum sharing in the Internet of Things. In the future work, data types will be gradually increased, and more targeted algorithms will be designed to process and analyze the data, so as to further optimize the energy-saving strategies and verify the energy-saving effects of relevant strategies.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Retraction

Retracted: Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Zhang, "Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data," *Journal of Function Spaces*, vol. 2022, Article ID 3367256, 11 pages, 2022.

Research Article

Evaluation of Mathematical Models for Teaching Precision in the Framework of Big Data

Ming Zhang 

School of Applied Technology, China University of Labor Relations, Beijing 100048, China

Correspondence should be addressed to Ming Zhang; zhangm@cullr.edu.cn

Received 27 April 2022; Revised 14 June 2022; Accepted 16 June 2022; Published 12 July 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Ming 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.

With the continuous development and application of data computing in the whole society in the construction of digital campus and intelligent campus of each higher education institution. In the environment of education, universities use these data well, which not only affects the orderly operation of higher education but also will become an inexhaustible power to help higher education promote the reform and innovation of education and teaching system. In this paper, we focus on the teaching operation and students' independent learning by taking students' evaluation data and students' online learning data of Y school as the research objects. We conducted a preliminary analysis and transformation of students' evaluation data of a university, eliminated the abnormal evaluation data by using the improved cosine phase dissimilarity algorithm, standardized the evaluation data by using the normalization method, and used the traditional *K*-modes algorithm. Based on these three problems, the traditional *K*-modes algorithm was improved in three aspects, including the determination of the number of clustering families, the determination of the measurement of clustering distances, and the experimental results showed that the improved algorithm was more reasonable and effective.

1. Introduction

Since the McKinsey report first introduced the concept of "big data," human society has been paying more and more attention to "big data," and people are widely aware. The potential value of "big data" has been widely recognized. Therefore, in view of the great potential, all countries in the world regard "big data" as important as oil, coal, minerals, natural gas, etc., and the research on "big data" at home and abroad has entered a brand new era. Research on "big data" has a crucial impact on various fields such as scientific progress, national security, and national education [1].

According to the study, the continuous expansion of data application will have a far-reaching impact on all fields of the whole society [2]. Big data is not only affecting people's learning, living, working, and thinking in an unprecedented way but also influencing the way of production, operation, organization, and management in various industries. It can be said that the era of big data is not only affecting the industry revolution but also the big changes in education [3]. Therefore, many experts and scholars have

conducted a series of fruitful researches and discourses based on this, which have reform and innovation of university education and teaching system.

On March 29, 2012, BDRI was launched in the United States, elevating big data to a national strategy [4]. At the same time, some U.S. teaching fields have made all-round reports on intelligent computing, which has accelerated the change of the whole field. [5]. In the report, the basic situation of online education big data application in the United States is analyzed with the case of online education big data, especially the case of adaptive personalized learning system discusses more comprehensively the development situation and a series of challenges and problems that will be faced by online education in the future [6]. Desire2Learn, a Canadian company, develops systems for student-directed learning, academic alert, and intervention services, and so on [7].

Its technology is no longer a traditional information technology, but provides the needs of intelligent transformation, integrating technologies, services, and solutions such as artificial intelligence, 5G, cloud computing, and the Internet of Things. At present, from intelligent manufacturing to

intelligent transportation, remote diagnosis and treatment, and online education, all walks of life are experiencing the process of digital and intelligent transformation [8]. In particular, 5G not only accelerates the intelligent transformation of the industry but also makes the traditional “cloud pipe end” IT architecture including cloud computing center, network pipeline, and terminal ineffective, thus giving birth to a new architecture [9]. “New IT” can also support intelligent manufacturing and promote the high-quality development of manufacturing industry. With digital and intelligent transformation and upgrading, China’s huge manufacturing industry is transforming into high-quality high-end manufacturing and intelligent manufacturing [10].

It is very important to deeply explore the potential resources and application value of big data in higher education to guide higher education institutions to fully implement the fundamental task of moral education and the concept of student-centered education, to promote the deep integration and intelligent transformation of higher education teaching activities, to promote the scientificization of higher education teaching system and decision-making system, to help the transformation and upgrading of scientific research paradigm in higher education institutions, and to build a more scientific management system in higher education institutions [11]. It is very important to promote decision-making system of higher education, help the transformation and upgrading of scientific research paradigm of higher education, and build up a more scientific management system of higher education [11]. A large amount of dynamic data will be generated in the process of school teaching and learning. With the help of big data technology, people can easily collect and mine these data to improve school teaching and learning and promote school quality and efficiency. And decision-making can make the evaluation results with higher reliability and credibility and make the decision-making more scientific and accurate [12].

Taking the student evaluation data and student online learning data of Y school as the research object, this paper focuses on teaching operation and students’ autonomous learning. The traditional *K*-modes algorithm is innovatively improved in three aspects: the determination of cluster family number and the determination of cluster distance measurement. The experimental results show that the improved algorithm is more reasonable and effective. This paper makes a targeted exploration from the two aspects of teaching operation and students’ online learning and draws some preliminary conclusions with certain reference value. Student evaluation, expert evaluation, peer evaluation, supervision, and evaluation institutions, and daily teaching monitoring data can be formed into more comprehensive, accurate, popularized, and influential results. It will provide more targeted decision-making, reform, and innovation basis for education and teaching managers and front-line teachers [13].

This paper focuses on two aspects of new era of big data environment of university education as the research background, takes the real student evaluation data and student online learning behavior data of a university as the research object, and carries out the relevant application research on

the teaching operation and learning situation of a university by using the improved clustering algorithm and the neural network algorithm based on machine learning, and the experimental result method has good scientific, and the thesis is divided into five sections: Section 1 describes the research background of this paper and the main structure of this paper; Section 2 introduces the current situation of domestic and foreign research in related fields and summarizes the research significance of this paper; Section 3 takes the student evaluation data of a university as the research object and adopts to model and analyze the school teaching and learning situation based on the abnormal data elimination and specification of the student evaluation data. The operation is modeled and analyzed, and the teacher teaching status evaluation model is established. In Section 4, the proposed scheme is tested and analyzed. Section 5 summarizes the research content of this paper and provides an outlook on future research directions.

2. The Related Works

Scholars have different interpretations on the definition or connotation of precision teaching. Precision teaching is divided into four stages: precision teaching breeding, precision teaching creation, precision teaching expansion, and precision teaching informatization. Carefully analyzing the statements of scholars at home and abroad on precision teaching, we find that no matter at which stage of precision teaching, scholars’ statements on precision teaching have the following common characteristics: data-based teaching, evaluation based teaching, learner centered teaching, and teaching that emphasizes recording and analyzing students’ learning behavior and performance. Therefore, we believe that the essence of accurate teaching is the spirit of scientific reflection [14, 15].

Chen Shongyeh et al. pointed out that precision teaching is supported by many theories, such as teaching theory, learning theory, curriculum theory, and technology theory. There are also many idealized models in implementation. Its experience is that the theory must be localized, and school-based. Only by integrating theory with practice can it be more widely used and reflect the localization characteristics of precision teaching [16]. Zhang Junchao et al. pointed out that when carrying out precision teaching, many regions or schools will adopt a one size that fits all approach in operation, requiring teachers to fully adopt new technologies and subvert the original teaching methods or teaching habits. According to our practical experience, we believe that teachers’ habits should be changed step by step and students’ habits should not be changed as much as possible [17]. Zhang Yannan et al. believe that the arrival of the era of big data in education will certainly be a revolutionary change due to its many characteristics such as comprehensiveness, real-time, and potential, which make a series of problems that are difficult to solve in the traditional education field are greatly improved, such as the balance, science, and rationality of education, and then the education model, education implementation path, and scientific evaluation of education [18]. Yang Xianmin et al. believe that one of the important

reasons for students' heavy academic burden is that teachers do not grasp the learning situation. Teachers' teaching effect is positively correlated with their understanding of students. The greatest value of big data precision teaching lies in accuracy, which is to deeply and fully understand the learning situation by collecting student data. Therefore, if we want to collect students' data fully and completely, we should collect students' classroom homework, synchronous test, simulated test, and other data horizontally. The data of students from enrollment to graduation are collected vertically [19].

At present, many studies have comprehensively discussed the concept of educational evaluation and the context of big data in education and pointed out specific implementation paths, which have indicated the direction for the comprehensive reform and coordinated development of education teaching [20]. In the environment of big data in education, based on the accelerated evolution of the new round of technological revolution and industrial change, take the national strategy, new industry, new economic development, and the future development direction of industry as the guide, and take the cultivation of innovation and practical as the guide. With the cultivation of innovation spirit and practical ability of college students as the core, the universities use these data well, which not only affects the orderly operation of the whole education teaching system but also will become the inexhaustible power teaching system in higher education.

K -modes algorithm has many advantages, but there are also many shortcomings, so He, San, Ng et al. proposed the method of calculating the distance between sample data by the frequency of occurrence within a class; Hsu et al. calculated the distance between sample data based on hierarchy; Ganti et al. and Ahmad et al. used the degree of cooccurrence data; these studies have made some improvements to K -modes algorithm. These studies have improved the K -modes algorithm, and its clustering ability has been greatly improved in the corresponding applications, but there are still shortcomings in the problems discussed in this study, mainly because they all ignore the different attributes of the connections and differences between different objects, and for this reason, some improvements are made to the K -modes algorithm for the specific problems studied in this paper. On this basis, this paper takes the big data of education and teaching in a university as the background and establishes relevant mathematical models with the help of big data related theories and technologies to guarantee that the evaluation and prediction of school teaching are more scientifically based and convincing.

3. Application of Improved K -Modes Algorithm in the Evaluation of Teaching Status of University Teachers

3.1. Basic Structure of Neural Network. K -modes algorithm is a clustering algorithm used to classify attribute data in data mining. K -modes algorithm is an extension of K -means algorithm. It can only deal with numerical data, but not classified attribute data. However, the traditional K -

-means algorithm is only suitable for data sets with continuous attributes, and neural algorithm is needed to supplement the data sets with discrete attributes. In the process of practical application, a single neuron cannot fit too complex mapping relationship. We need to build more complex networks to approximate those more complex objective functions. Using multilayer networks can sometimes find a good convergence relationship after less iterative training, as shown in Figure 1.

Neurons, as the most basic building blocks of neural networks, are responsible for computational or processing functions, which are usually expressed in terms of computational functions called excitation functions, and their functions vary greatly for different applications, so they should be selected with great care. The biological structure of a neuron is shown in Figure 2.

3.2. Problem Formulation and Data Selection and Data Structure Analysis of Evaluation Data. The evaluation indicators are as follows:

(1) Quality literacy

They have high moral character and serious and responsible teaching, pay attention to the image of teachers, strictly abide by teaching discipline, and do not suspend and transfer classes at will, no late arrival, and early departure, profound professional knowledge, solid practical skills, fluent expression in Mandarin, and standardized board writing.

(2) Teaching attitude

The teaching attitude is correct, the class is well prepared, and the teaching schedule can be reasonably arranged according to the syllabus; the class is organized in an orderly and active atmosphere, and the teaching content is organized rigorously without reading from the text.

(3) Teaching skills

The lectures are clear, accurate, and focused on difficult points; they focus on the systematic, scientific, and advanced knowledge and are able to update the teaching content in a timely manner by combining the frontiers and dynamics of the field.

(4) Extracurricular sessions

He teaches and nurtures students, focuses on the cultivation of students' innovative and practical abilities, pays attention to the process management, reviews homework carefully, provides timely counseling, answers questions, often takes the initiative to communicate with students, pays attention to students' feedback, and continuously improves teaching.

The alternatives are the following: excellent, good, moderate, pass, and fail.

In order to illustrate the problem, improve the efficiency of the experiment, and focus the work on the improvement

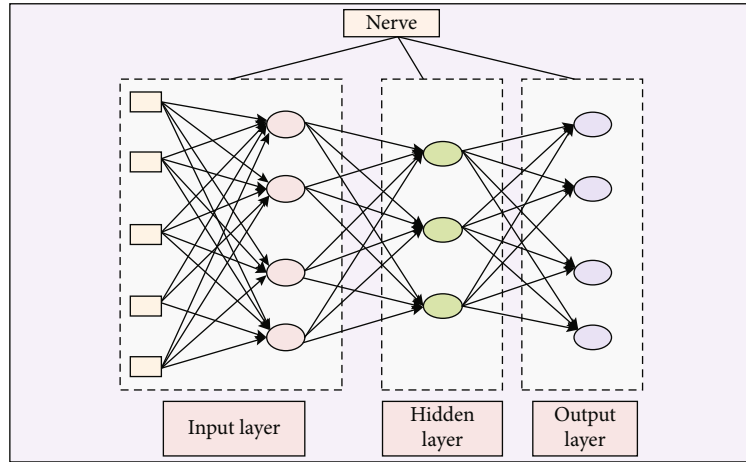


FIGURE 1: Schematic diagram of neural network structure.

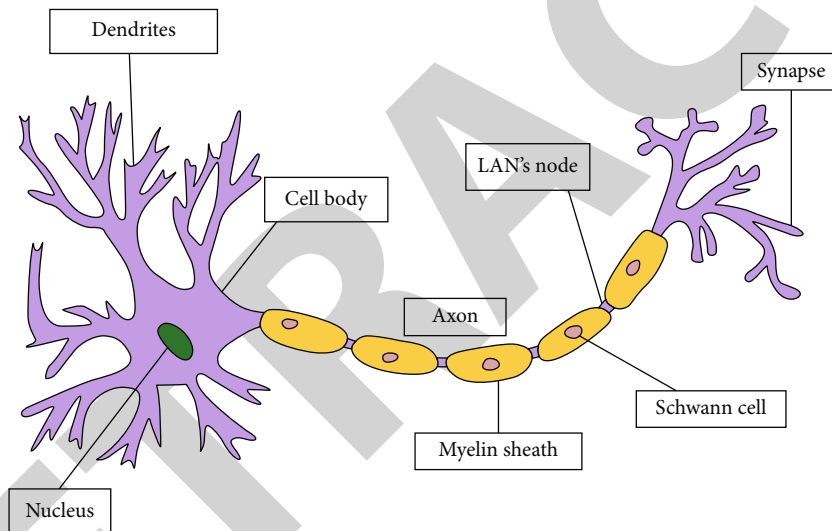


FIGURE 2: Schematic diagram of the biological structure of a neuron.

of the method, all the data of student evaluation in the follow-up of this paper are the data of one college of the university (referred to as College R for the convenience of the subsequent narrative) for the last ten academic years, i.e., from the second half of 2009 to the second half of 2019, with a total of 237,924 student evaluation records.

The data tables related to the evaluation data table in the Academic Affairs Management System of Campus Y are: Student Information Data Table (Student Number, Name,), Teacher Information Data Table (Employee Number, Name,), Course Information Data Table (Course Code, Course Name,), teaching task data sheet (academic year, semester, course code, course selection number, employee number,), course selection data sheet (academic number, course selection number,), and student evaluation. There are 7 tables of indicators (serial number, evaluation item, content, and evaluation level). The student evaluation data table is based on XN, XQ,

XH, XKKH, and ZGH as the main code, and the sex associated with other data is the external code, for example, for the student information table data, its external code is XH, to effectively ensure the integrity and consistency of all relevant data in the system.

3.3. Student Evaluation Data Preprocessing. In this paper, there are four evaluation indicators for student evaluation data, and each indicator has five values, which is a multidimensional subtype problem. Since the subtypes are different from the numerical data, it is difficult to measure the differences between the data in a hierarchical way, and also for the sake of presentation, without changing the interpretation, each evaluation value of student evaluation is transformed accordingly in this paper, i.e., each different subtype is assigned a rank that exactly matches its value. The specific transformation strategy is as follows:

$f(x)$ is the transformation function of the evaluation value;

x : for PJ_i original rating, $x \in \{\text{excellent, good, moderate, failing}\}$, $i = 1, 2, 3, 4$;

$$\begin{aligned} f(x) &= 5, x = \text{excellent}, \\ f(x) &= 4, x = \text{good}, \\ f(x) &= 3, x = \text{medium}, \\ f(x) &= 2, x = \text{pass}, \\ f(x) &= 1, x = \text{failed}. \end{aligned} \quad (1)$$

Anomalous data in student evaluation data are often due to the fact that individual students do not evaluate a course taught by a particular instructor objectively and fairly, but rather evaluate it with a strong personal touch, resulting in a large deviation between their evaluations and those of other students. There is a significant correlation between students' cognitive style and some dimensions of college students' social adaptability. Field-independent cognitive style is significantly correlated with learning adaptability, career choice adaptability, self-care adaptability, and physical and mental adaptability. Field-dependent cognitive style is significantly correlated with interpersonal adaptability and role adaptability. Therefore, these anomalous data need to be removed from the evaluation of teaching data.

For example, for two sample data x, y , their original evaluation values are (1, 1, 1, 1) and (5, 5, 5, 5), and the replacement values are (1-3, 1-3, 1-3, 1-3) and (5-3, 5-3, 5-3, 5-3), respectively.

The corresponding phase dissimilarity is calculated as

$$\cos(x, y) = \frac{(-1 \times 2) + (-1 \times 2) + (-1 \times 2) + (-1 \times 2)}{\sqrt{(-1)^2 + (-1)^2 + (-1)^2 + (-1)^2} \times \sqrt{2^2 + 2^2 + 2^2 + 2^2}} = -1. \quad (2)$$

Applying the method above, 237924 student evaluation records from 1326 categorical sample data files of College R were subjected to anomalous data elimination, and the results are shown in Figure 3.

3.4. Application of the Improved K-Modes Algorithm to the Evaluation of College Teachers' Teaching Ability. After eliminating and merging the abnormal evaluation data in the previous section, the K-modes algorithm will be applied to evaluate the teaching ability with the evaluation of 59 instructors in a semester in College R. As we can see from the discussion of K-modes algorithm in Section 2, K-modes algorithm is a 0-1 matching algorithm based on whether the values of each attribute of the sample data are the same or not. However, the algorithm also has three serious shortcomings in solving some specific problems with certain relationships among attributes.

Firstly, in the k-means algorithm, K is given in advance, and the selection of this k value is very difficult to estimate. In the initial K-means clustering, we need to determine an initial partition algorithm, and then, we need to optimize an initial partition algorithm. The selection of this initial clustering center has a great impact on the clustering results.

Finally, the algorithm needs to constantly adjust the sample classification and constantly calculate the adjusted new cluster center. Therefore, when the amount of data is very large, the time overhead of the algorithm is very large.

In this paper, each uses the method of minimizing sum of squared error (SSE) to determine the number of clustering families K .

(1) Definition of SSE

$$\text{SSE} = \sum_{l=1}^k \sum_{x \in L_l} \text{Dist}(x, Z_l)^2, \quad (3)$$

where k is the number of clustering families, Z_l is the clustering center of the l cluster L_l , and $\text{Dist}(x, Z_l)$ is the similarity between the sample data x and the clustering center Z_l . Different similarity calculations often bring large differences to the clustering results, since the frequency-based similarity calculation is measured by the frequency of each attribute of each sample data in the whole sample data set.

(2) Calculation of the frequency of sample data

In order to obtain the frequency of each attribute of the sample data, it is necessary to count the number of occurrences of each rating value in the whole sample data set. Let x_{ij} be the evaluation value of the attributes of the sample data, $f(x_{ij})$ be the frequency of x_{ij} , C be the set of allowed values of x_{ij} , for this example, C be {5 (excellent), 4 (good), 3 (moderate), 2 (pass), 1 (fail)}, C_p be the number of times the element of p in the set C appears in the sample data set, where $i = 1, 2, \dots, n$, n is the number of sample data, $j = 1, 2, \dots, m$, m is the dimension of the sample data, $p = 1, 2, \dots, P$, $P = |C|$ is the base of the set C , i.e., $f(x_{ij})$ can be calculated by the formula:

$$\begin{aligned} f(x_{ij}) &= C_1, x_{ij} = 5, \\ f(x_{ij}) &= f(x_{ij}) = C_2, x_{ij} = 4, \\ f(x_{ij}) &= f(x_{ij}) = C_3, x_{ij} = 3, \\ f(x_{ij}) &= f(x_{ij}) = C_4, x_{ij} = 2, \\ f(x_{ij}) &= f(x_{ij}) = C_5, x_{ij} = 1. \end{aligned} \quad (4)$$

(3) Frequency-based similarity calculation

The frequency of the sample data is calculated by the following formula:

$$\text{AVF}(x_i) = \frac{1}{m} \sum_{j=1}^m f(x_{ij}), \quad (5)$$

where x_i is the sample data and $f(x_{ij})$ is the frequency of the sample data x_i on the attribute j .

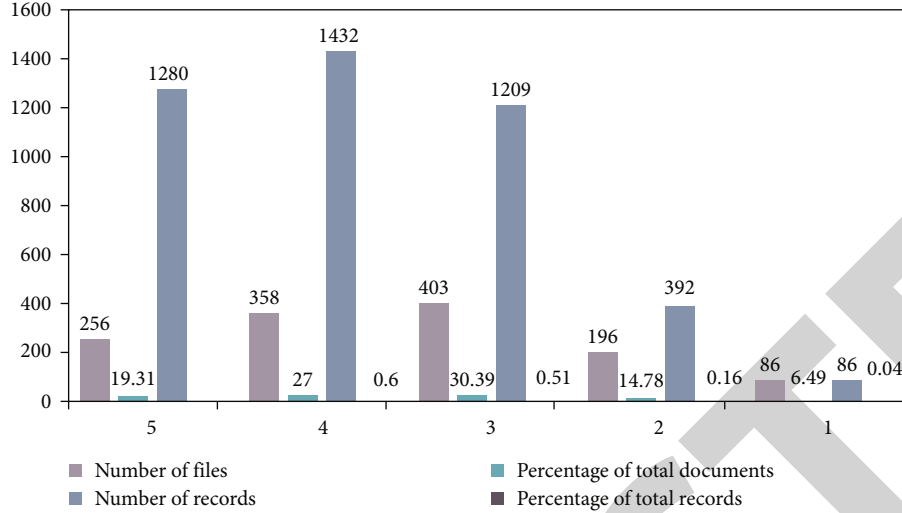


FIGURE 3: Number of records containing anomalous data versus number of files and number of records.

(4) Preclustering to determine the number of clusters K

The frequency-based similarity was used to pre-cluster the classes 1-7 with K -modes, and the minimum error sum of squares versus the number of clusters K was obtained as shown in Figure 4.

From Figure 5, it is obvious that the inflection point of the image appears at $K = 3$, so it is more appropriate to cluster 3 classes for a semester of student evaluation results in College R. In order to prove this point, the same method was used to precluster all teachers' evaluation data of school Y. The preclustering results showed that the inflection point of the image appeared at $K = 5$; when the dimensionality of the sample data was doubled, the preclustering results showed that the inflection point of the image appeared at $K = 7$. Therefore, the value of K varies for different types and scales of problems and should be treated with caution in specific problems.

From the principle of K -modes algorithm to achieve clustering, it is known that for the randomly selected K clustering centers, iterate and update the clustering centers continuously according to the closest distance principle, and the goal is to minimize the SSE. Inspired by this idea, the scheme of this paper is specified as follows.

Each sample data point other than the one already selected as a cluster center is used as the hypothetical initial cluster center at $l + 1$, and the SSE with $l + 1$ is calculated, and the data point with the fastest decreasing SSE value is selected as the true initial cluster center at. The data point with the fastest decreasing SSE value was selected as the true $l + 1$ initial cluster center, repeated, and terminated when K is the known number of classifications.

Based on the above idea, the following model is developed:

$$\begin{aligned}
 Z &= \{z_l | z_l \in X, l \leq K\}, \\
 z_l &= \max_{x \in X} \text{DJJSSE}(x), \\
 \text{DJJSSE}(x) &= \left(\sum_{j=1}^{l-1} \text{Dist}(x, x_j)^2 + \min_{x \in X} \left(\sum_{i=1}^n \text{Dist}(x, x_i)^2 \right) \right), \quad (6)
 \end{aligned}$$

where Z is the set of initial clustering centers, z_l is the sample data point that makes SEE fall fastest x , i.e., the l initial clustering center, x is the sample data, and n is the number of sample data.

The traditional K -modes algorithm of clustering distance calculation uses the similarity (0-1 match, 0 for same and 1 for different) of sample data and cluster center data for the metric; this method is feasible and effective for applications where the sample attributes take values independently of each other, but if there is a certain connection between different attributes of the sample, and different values under the same attribute also have a certain connection; the method's exposed deficiencies are very large. For example, in this case, if the students' evaluation of teacher 1 is (2, 2, 3, 2), and the evaluation of teacher 2 is (5, 4, 5, 4), and the cluster center is (3, 5, 4, 3), the distance of teacher 1 from the cluster center is (1, 1, 1, 1) = 4, and the distance of teacher 2 from the cluster center is (1, 1, 1, 1) = 4, using the traditional K -modes algorithm for similarity measurement. (1, 1, 1, 1) = 4, which is obviously far from the actual situation. Therefore, on the one hand, the results of clustering may vary greatly depending on the initial clustering centers, and the results are unstable, as can be seen from the table comparison; on the other hand, because the distance measure does not consider the connection between different attributes and different values of the same attribute and simply uses 0-1 matching to measure the distance, the clustering results are poor and do not match the actual situation. Therefore, in this paper, the distance measure of co-occurrence proposed by Ahmad et al. is used to improve the similarity of the traditional K -modes algorithm by using the similarity between different attribute values under the same attribute with other attributes.

The application of the improved K -modes algorithm for cluster analysis of student evaluation data to achieve the evaluation of teachers' teaching ability can be summarized in two stages and eight basic steps, as follows.

Phase I: data preprocessing, consisting of four basic steps

Step 1: sample data conversion, i.e., rating values from (excellent, good, moderate, passing, and failing)

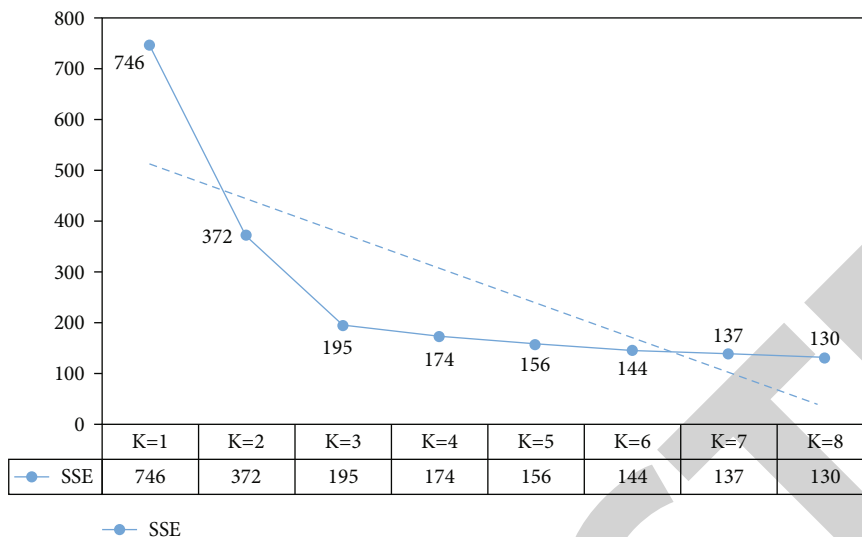


FIGURE 4: Plot of the minimum error sum of squares versus the number of clusters K .

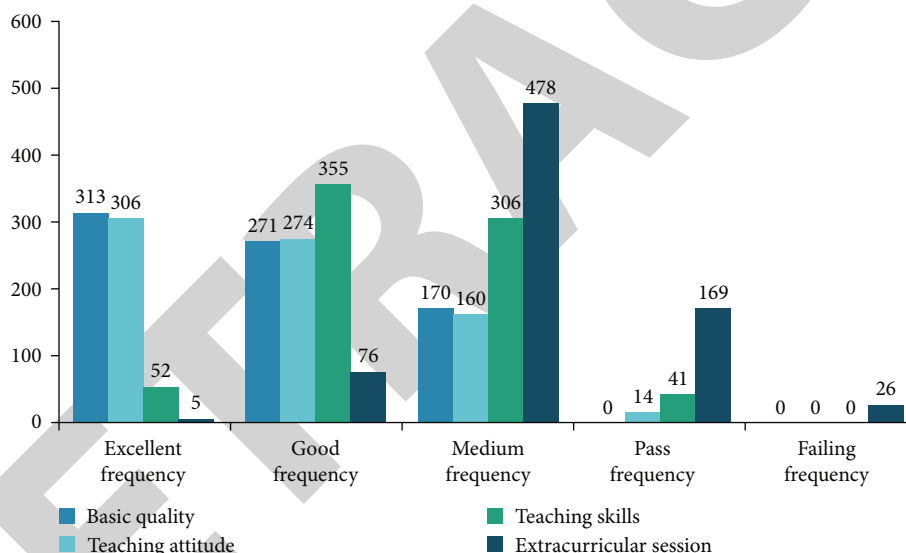


FIGURE 5: Statistics of students' evaluation of teaching in a semester in school Y.

Convert to (5, 4, 3, 2, 1)

Step 2: sample data splitting, i.e., all the evaluation data are split into multiple classification data files by “academic year + semester + course code + employee number” as the splitting code

Step 3: sample anomaly data elimination, i.e., using the cosine dissimilarity calculation model, the sample data whose evaluation values in each categorical data file are significantly different from other evaluation values are eliminated from the corresponding categorical data file

Step 4: merge and restore the sample classification data files, i.e., for the sample classification data files, merge each teacher's evaluation into one record according to the merge code of “academic year + semester + employee number”, and restore it

Phase II: clustering evaluation, consists of four basic steps

Step 1: a reasonable number of clusters K is determined by applying the frequency-based (AVF) similarity calculation algorithm for clusters that have not been given a specific number of clusters K or have been given a number of clusters K but have poor clustering results

Step 2: even poor usability of clustering results due to the selection of outlier data points or proximity points, K initial clustering centers are determined by using the sum of least errors squared (SSE) algorithm to avoid the occurrence of two types of situations

Step 3: traditional K -modes clustering algorithm does not consider the relationship between different attributes and different values of the same attribute in the 0-1 matching method when calculating the clustering distance, the

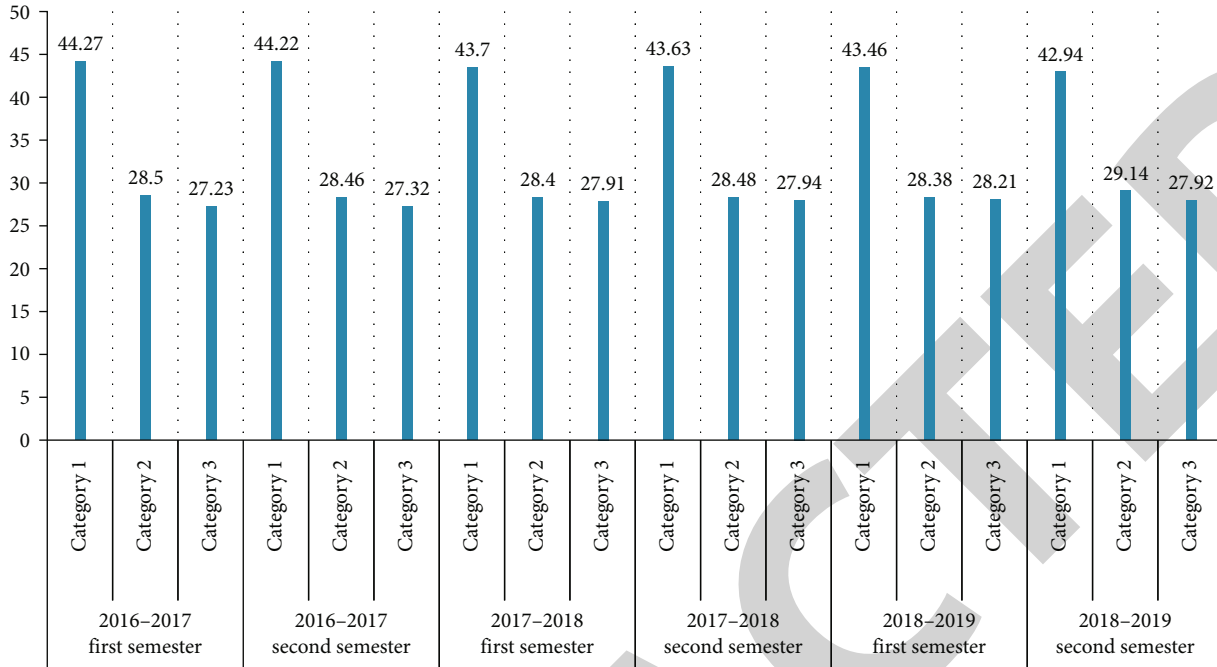


FIGURE 6: Histogram of the percentage of teachers in the three categories at school Y.

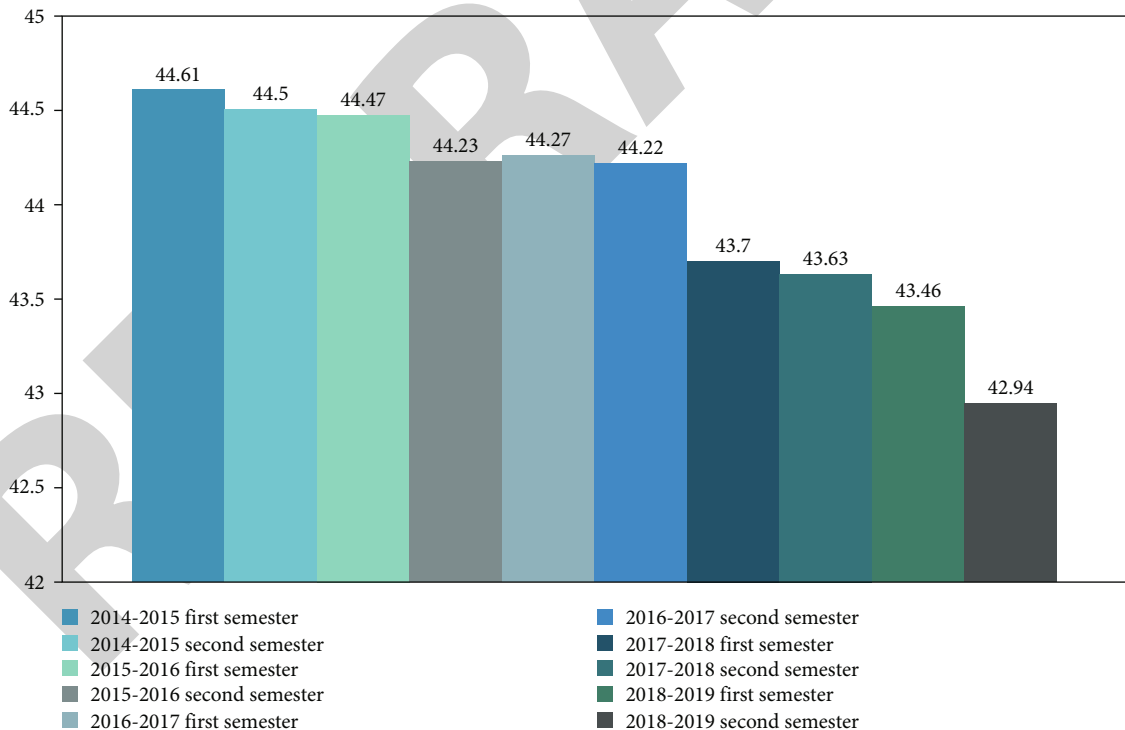


FIGURE 7: The proportion of first-class teachers has changed in the past five years.

traditional *K*-modes clustering algorithm is improved by using the cooccurrence of evaluation value as the distance measure

Step 4: based on the clustering results, we analyze the overall teaching situation of all teachers and the teaching sit-

uation of each teacher in each semester by semester, managers to grasp the overall teaching situation of all teachers, and teachers in each teaching unit in a timely manner and thus provide a scientific basis for correct decision-making and targeted measures; on the other hand, we provide

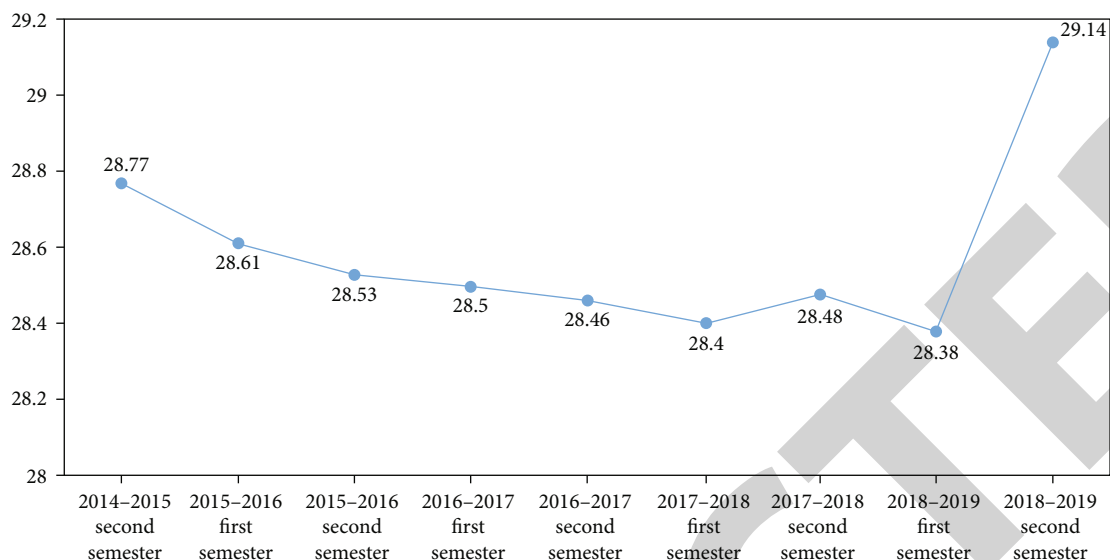


FIGURE 8: The percentage of category 2 teachers in school Y over the past five years.

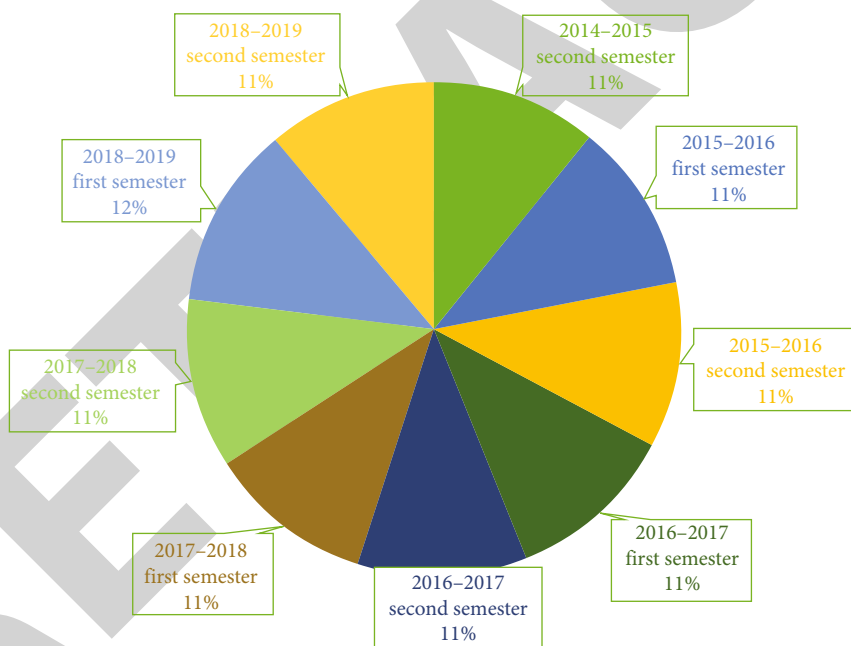


FIGURE 9: Changes in the proportion of three types of teachers in Y school in the past five years.

teachers with a timely understanding of their own teaching situation and that of other teachers, so that they can take targeted measures to improve their teaching. On the other hand, it provides teachers with a timely understanding of their own and other teachers' teaching conditions, so that they can take targeted measures to enhance the internal driving force for continuous improvement of teaching

4. Analysis of Simulation Results

The teaching status can reflect the teaching operation of a university as a whole, but this whole is also composed of

individual teachers, so it can also reflect the overall teaching ability of teachers in a university. Based on the improved *K*-modes clustering algorithm, the teaching situation of the university was analyzed by clustering the teachers of the whole university, the teachers of secondary teaching units, and the individual teachers at three levels based on the time period of academic year semester.

The change in the percentage of teachers in the three categories is shown in Figure 6. The percentages of teachers in the three categories in each semester were basically equal, with the first category (at least 3 out of 4 indicators rated "good" or higher) accounting for at least (43%, 45%), and

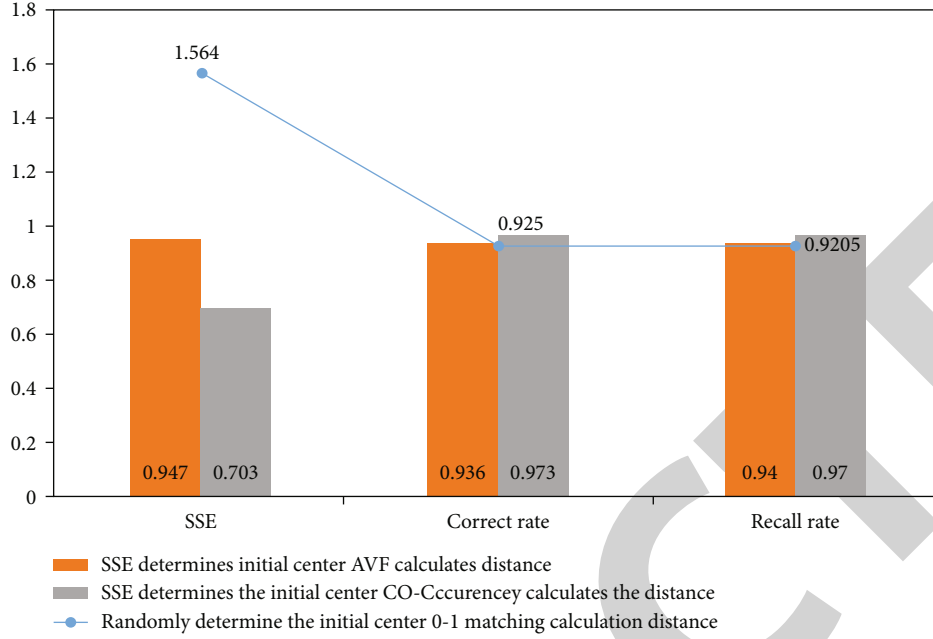


FIGURE 10: Comparison of three different clustering algorithms for each index in teaching evaluation data.

the second category (2-3 out of 4 indicators rated “good”), “good” grade. The second category (at least one of the four indicators is rated as “moderate”) is within the range of (26%, 28.5%); the third category (all four indicators are rated below “moderate”) is within the range of (26%, 28.5%). It can be seen that although students’ evaluation groups are changing, their perceptions of teachers’ teaching are basically the same, indicating that students’ evaluation is fair and just overall. Figure 7 shows the change in the proportion of first-class teachers over the past five years.

Figure 8 shows the proportion of class II teachers in school y in the past five years. Figure 9 shows the changes in the proportion of three types of teachers in school y in the past five years. From the clustering results, it is obvious that a little more than 40% (the first category of clustering) of teachers in the school are relatively qualified, nearly 30% (the second category of clustering) are in the basic qualified state, and nearly 30% (the third category of clustering) are not in the “basic quality” and “teaching.” This indicates that the overall teaching operation of the university is not optimistic and is far from the fundamental goal of establishing moral education and still needs to be reversed and improved with great efforts.

The results of three k -modes were compared in terms of cluster centers, number of clusters, percentages, sum of squares of least errors, correctness, and recall rates, using the evaluation of a semester at YR as an example.

$$\begin{aligned} \text{Correct rate} &= \frac{\left(\sum_{i=1}^k e_i\right)}{n}, \\ \text{Recall rate} &= \frac{\left(\sum_{i=1}^k (e_i/e_i + m_i)\right)}{k}. \end{aligned} \quad (7)$$

Figure 10 compares three different clustering algorithms for each index in the teaching evaluation data. From the experimental results of the three different algorithms, it is obvious that the improved K -modes outperform the first two algorithms in all metrics.

5. Summary and Outlook

Focusing on teaching operation and students’ autonomous learning, this paper uses the improved K -modes algorithm to cluster analyze the classroom teaching operation and preliminarily analyzes and converts the student evaluation data of a university. On this basis, the abnormal evaluation data is eliminated according to the improved cosine algorithm, and the normalization algorithm is used to normalize the data. The traditional K -modes algorithm is used to cluster the data, and the main problems are pointed out. Taking the real data of a university as the research object, this paper makes a targeted exploration from the two aspects of teaching operation and students’ online learning and draws some preliminary conclusions with certain reference value, which can form a more comprehensive, accurate, popularized, and influential result of students’ evaluation, expert evaluation, peer evaluation, supervision, and evaluation institutions, and daily teaching monitoring data. It will provide more targeted decision-making, reform, and innovation basis for education and teaching managers and front-line teachers. However, there are still some limitations in this paper. Research needs to determine the number of cluster family K . Determine the initial K cluster centers and then the distance measurement from the K cluster centers. In future research, K -modes algorithm can be improved in these three aspects to evaluate teachers’ teaching ability.

Retraction

Retracted: Optimization of College Students' Employment Prosperity Index System Based on Multiple Logit Models

Journal of Function Spaces

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 2023

Copyright © 2023 Journal of Function Spaces. 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. Xiang and W. Hu, "Optimization of College Students' Employment Prosperity Index System Based on Multiple Logit Models," *Journal of Function Spaces*, vol. 2022, Article ID 9367675, 8 pages, 2022.

Research Article

Optimization of College Students' Employment Prosperity Index System Based on Multiple Logit Models

Wei Xiang¹ and Weizhen Hu²

¹Zhejiang Normal University, Jinhua, Zhejiang 321004, China

²Jinhua Radio and Television University (Zhejiang Commerce and Trade School), Jinhua, Zhejiang 321004, China

Correspondence should be addressed to Wei Xiang; xiangwei@zjnu.edu.cn

Received 10 May 2022; Revised 22 June 2022; Accepted 24 June 2022; Published 12 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Wei Xiang and Weizhen Hu. 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 employment of university graduates has always been a matter of great concern to educational administrative departments and governments at all levels. EPI (employment prosperity index) is a weather vane that reflects the situation of the labor market, and it is also an important reference for the national economic situation. This paper constructs college students' EPI from different aspects, such as employment environment, employability, employment status, and public service, focuses on the calculation method of college students' EPI, and obtains the general formula for calculating college students' EPI. Based on the data obtained from a university, the EPI system is optimized by using MLM (multinomial logit model). Through empirical research, it is found that EPI, which is optimized by the system, has the function of evaluating the employment environment of college students, which not only provides macroemployment policies for the government but also provides scientific and effective employment service policies for universities.

1. Introduction

As an important part of the market economy, the indicators reflecting the changing trend of the employment market need to be timely, accurate, and representative, which is of great practical significance to the government's macromarket regulation, the recruitment of enterprises, and the employment of workers. China has entered the economic transition period; while abandoning the theory of only economic growth, employment data has also become the key indicator of governments at all levels. EPI (employment prosperity index) describes the overall and long-term employment situation in order to predict the future trend of employment and economic growth. With the improvement of human resource market and social security system in various places, designing and constructing EPI has become a realistic topic before us. The implementation of the family planning policy has led to the reduction of the birth population, which in turn has led to the weakening of social consumption, the increase of national savings, the decline of social consumption demand, the slow growth of employment, and employment difficulties.

EPI is a comprehensive description of the overall situation of labor and employment, so finding out the important factors that affect the employment boom is the first step in the compilation of the index, and many scholars have made a lot of research results [1]. Literature [2] shows that the trade surplus has obvious promoting effect on employment in China, but the promoting effect and contribution rate of consumption to employment growth are more obvious. Maintaining a moderate inflation level is conducive to the growth of employment. The promotion effect of investment on employment lags behind for about half a year. In literature [3] combined with the empirical data of the study area, the regional EPI was constructed by using AHP and factor analysis. Literature [4] uses the diffusion index method to compile the employment demand prosperity index. Literature [5, 6] holds that unemployment is the result of college graduates' unwillingness to reduce their retained wages. It is found that the research on college students' employment from different disciplines has different focuses, and the research on higher education employment from multidisciplinary and compound perspectives has made great progress. However, there are many shortcomings in the

theoretical deduction and problem-solving. Literature [7] points out that the adjustment of college students' employment policy should constantly improve the employment market of college students with interest mechanism, guide college students' employment with interest compensation mechanism, and ensure the public interest in college students' employment with interest balance mechanism. Literature [8] points out that the change of employment policy not only has its own development rules and characteristics but also has the characteristics of being influenced by social-related factors. In literature [9, 10] using the prosperity index compiling method, the diffusion index and composite index of employment demand cycle prosperity are compiled, and the index system of labor demand cycle prosperity analysis is formed.

At present, China's educational decision-making on college students' employment generally lacks scientific basis and quantitative criteria, largely because the quantitative research on college students' employment is not deep and the support of effective data is lacking [11]. However, only according to the "employment rate," it is difficult to reflect the actual situation of college students' employment, nor can it provide a scientific basis for solving the employment difficulties of college students. Therefore, the construction of EPI for college students can not only reflect the current social employment situation but also predict the current employment development trend, so as to help government departments make corresponding decisions and solve the upcoming problems on the premise of anticipating the employment development trend. The establishment of EPI model for college students can not only predict the quantity and types of talents in various sectors of society but also calculate the future employment prosperity of existing college students and measure the rationality of cultivating various talents in universities. This paper focuses on how to calculate EPI of college students.

2. Research Method

2.1. EPI System for College Students. EPI of college students mainly reflects the restriction and influence of economic growth on the number and structure of college students' employment, and it is reflected by quantitative and intuitive data. The main object of EPI for college students is contemporary college students, which is different from the general EPI. Analyzing the employment problem of college students from the perspective of employment prosperity greatly enriches the current index system for evaluating the employment of college students. Based on a large number of objective data, it is convincing and credible to evaluate the employment environment of college students.

When analyzing the degree of regional employment prosperity, an effective and reliable index is urgently needed to comprehensively and intuitively reflect the active degree of the local enterprise employment market, reveal the ups and downs of the human resource market within the monitoring scope, and provide guidance for employers and job seekers to realize benign interaction [12, 13]. In view of this, through investigation and analysis, our research group has determined the constituent factors that affect the degree of

regional employment prosperity, including the inflow of labor, the outflow of labor, the total number of employees in the enterprise, the demand for human resources, the supply of human resources, the recruitment completion rate of the enterprise, the recruitment ratio, the change rate of the long-term unemployed, and the reference unemployment rate.

Establishing a scientific index system is an important link of EPI analysis and synthesis for college students. The determination of the comprehensive quality evaluation index must conform to the objective and practical evaluation concept and play a guiding role in improving the comprehensive quality of college students. Therefore, it must be constructed with a rigorous and scientific attitude. Set up a special leading group in colleges and universities. The members of the group must be researchers or related personnel with certain professional qualities. A systematic index system is formed by investigating the opinions of college students, graduates, and relevant employment managers through procedural steps to form a relatively complete thinking. It provides a reliable index data source for comprehensive evaluation through its quotation. The following principles are followed in the selection of indicators: the principle of comprehensiveness: the index system should reflect the employment situation and changing trend of college students from all angles as far as possible; principle of importance: the change of indicators has an important impact on the employment situation and trend of college students; principle of representativeness: the selected indicators can reflect the employment situation and changing trend of college students and have certain representativeness; and principle of operability: the obtained data of indicators should be easy to collect and collate. Therefore, the index system of EPI established in this subject is shown in Figure 1.

The employment quality of college students is related to the development of education. Increasing investment in education will enhance the development level of higher education and provide a financial basis for cultivating high-quality college students, and the employability of high-quality talents will be correspondingly strong. The factors that affect college students' employment competitiveness include students' subjective ability to work. The objective aspects of college students' employment competitiveness are affected by social development, including development goals and school characteristics. In order to solve the employment problem of graduates, it is more difficult for enterprises to recruit. Improving the employment competitiveness of graduates is a key project to adapt to the reform of higher education and the development of market economy. It is also an important measure to ensure the survival and development of universities.

One of the functions of EPI for college students is to establish a contact intermediary between the market and students and to play the role of employment service for college students through this intermediary. For college students, it can help them broaden their professional caliber according to market needs and establish rational employment expectations [14]. For college graduates, EPI can be used to guide them to grasp employment opportunities reasonably.

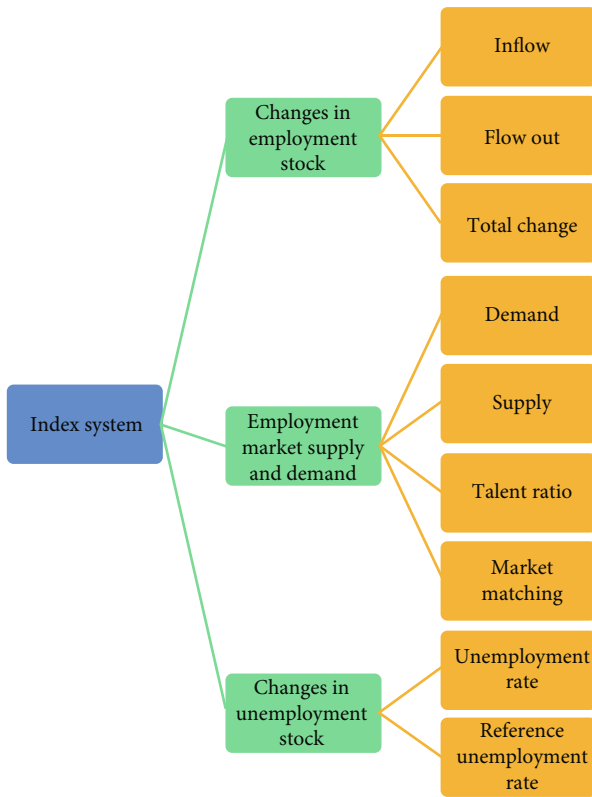


FIGURE 1: EPI system.

2.2. *The Calculation of College Students' EPI.* Employment is the greatest livelihood. The purpose of stabilizing growth and protecting market players is to protect employment, that is, to protect people's livelihood and social stability [15]. With the increase in the number of college graduates, the severe employment situation, and fierce competition, the number of postgraduate applicants is increasing year by year, and the number of postgraduate entrance examinations has surged. In literature [7], the data of talent activity in major cities are compared. Talent activity = the ratio of the number of people with strong job – seeking intention to the talent base. The higher the talent activity, the more talents flow in, and the stronger the attraction of the city to talents (Figure 2).

Due to the pressure of work, life, and environment in first-tier cities, talents are moving to new first-tier cities with more comfortable life and less competition. In addition, due to the shortage of land in first-tier cities and the high cost of living, in recent years, major Internet and technology companies have started to move many businesses to new first-tier and second-tier cities for the long-term development of their companies, which has brought a large number of jobs to these cities. There are more and more new opportunities, and at the same time, the housing prices are within the affordable range, so the level of talent activity in these areas is at a high level.

The role of human capital investment is not immediate [16, 17]. Its role can continue to play in a long period of time. Although the knowledge is updated quickly, the cumulative effect of human capital stock is long-lasting. The update is only a further development on the original basis,

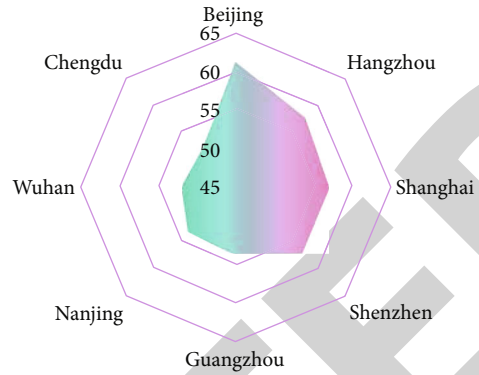


FIGURE 2: Talent activity.

and human capital investment just adapts to this way. The role of human capital investment in urban competitiveness is often indirect, difficult to quantify, and time-delayed. Its role is generally realized through labor force, scientific and technological progress, and other factors and links.

This paper puts forward the competitiveness model (Figure 3). In this model, the author puts forward four kinds of material resources and four kinds of human capital elements and divides the direct and indirect effects of environmental changes on material and human capital elements into four stages:

- (1) The city provides its own labor force and natural resources
- (2) The role and status of politicians and administrators are particularly important, and their thoughts and behaviors play a decisive role in the development of urban economy
- (3) Entrepreneurs have become the core role
- (4) It is no longer possible to maintain the competitive advantage for a long time by relying solely on entrepreneurs' adventurous spirit. In this dynamic process, the functions of the four types of human capital are also different, and each stage has a human capital that plays a major role

The direct factors, including enterprise development and industrial structure, support each other and restrict each other. Among them, scientific and technological innovation is the internal power of enterprise development, and its role in urban competitiveness should be realized by enterprises. Industrial structure grasps the development direction from a macroperspective, and enterprises are the core of the whole force. The activities of enterprises are related to all aspects of urban development.

Referring to the calculation method of economic prosperity index, the calculation method of college students' EPI first calculates the benchmark fluctuation index H_{index} :

$$H_{\text{index}}(t) = \frac{\sum_{i=1}^N d_i(t) W_i}{\sum_{i=1}^N W_i}. \tag{1}$$

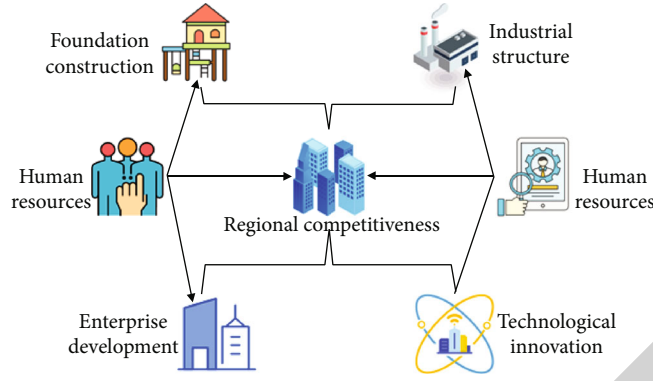


FIGURE 3: The function model of human capital on urban competitiveness.

$H_{\text{index}}(t)$ is the benchmark fluctuation index, which takes different values with time; $d_i(t)$ is the value of the i th index at time t ; W_i is the weight of the i th index, which is generally determined by correlation coefficient weight or expert scoring method; N is the number of indicators.

Employment growth rate is the ratio of the increase of employment population to the employment population in the previous period. Among them, the employment population refers to the population aged 16 or above, engaged in certain social labor or business activities, and obtained labor remuneration or business income. The employment growth rate reflects the actual utilization of all labor resources in a certain period. The index formula is

$$E_{\text{growth rate}} = \frac{E_{\text{employed population}} - E_{\text{Last employment population}}}{E_{\text{Last employment population}}} \times 100\%. \quad (2)$$

Among them, $E_{\text{growth rate}}$ is the growth rate of employed population; $E_{\text{employed population}}$ is the current employment population; $E_{\text{Last employment population}}$ is the population in permanent employment.

Long-term unemployment index is

$$U = \frac{F_i/P}{F_j/P} \times 100\%. \quad (3)$$

Among them, U is the long-term unemployment rate index; F_i, F_j is the number of people receiving unemployment insurance for six consecutive months in the base period and the number of people receiving unemployment insurance for six consecutive months in the current period; P is the total number of participants.

In order to facilitate the unity of calculation, all types of indicators are positively processed in this paper. The specific method is as follows:

For the index of reversibility, this paper takes the reciprocal to make it positive:

$$X_i = \frac{1}{X_t}, \quad (4)$$

where X_t is the actual value and X_i is the converted index quantity.

The following is the index conversion with moderate optimal criteria:

$$X_i = \frac{1}{X_t - X_\beta}, \quad (5)$$

where X_β is the median value of the group composed of the actual value and the optimal value, which is the converted index quantity.

In this paper, the Pearson correlation coefficient method is used to calculate the correlation matrix of each index. By screening the highly correlated variables and indicators, only one of the highly correlated indicators and variables is reserved. The calculation formula is

$$r = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N\sum x_i^2 - (\sum x_i)^2} \sqrt{N\sum y_i^2 - (\sum y_i)^2}}. \quad (6)$$

The value range of correlation coefficient r is $[-1, 1]$. When $r \rightarrow 1$ or -1 is used, it indicates that the correlation degree is stronger; on the contrary, the correlation coefficient $r \rightarrow 0$ is weaker.

2.3. Optimized Implementation of EPI System. EPI can not only objectively and truly reflect the present situation and trend of employment but also fully show the macroeconomic situation, the relationship between labor supply and demand, and the implementation effect of employment policy [18]. Because the labor transferred from the primary and secondary industries cannot meet the demand of talents in the tertiary industry from skills, knowledge, and other aspects, resulting in the friction of the structure of labor supply and demand, there may be a situation that the supply of labor in the primary and secondary industries exceeds demand, while the supply of labor in the tertiary industry is in short supply.

Employment elasticity is the percentage change of employment quantity corresponding to every percentage change of economic growth. The change of employment elasticity depends on factors such as economic structure

and labor cost. The capital investment and labor cost required for a certain amount of labor employment constitute the unit cost of employment. Employment elasticity is related to economic growth and employment growth. Employment elasticity is the ratio between the two growth rates, so a high employment elasticity does not mean that employment is more full. On the contrary, low employment elasticity does not mean that unemployment is more serious. Generally speaking, in the process of economic maturity, the elasticity of employment will gradually decrease. The decreasing employment elasticity indicates that the labor increment required to create an incremental value becomes smaller, which actually means the improvement of labor productivity. If the proportion of labor-intensive economy such as small- and medium-sized enterprises and service industry in the economic structure is large, the proportion of capital is low, the employment cost is relatively low, and the employment elasticity is high.

At present, it is more and more difficult for college graduates to find jobs, and the pressure is increasing. On the one hand, it is related to the rapid growth of the number of college graduates employed; on the other hand, the employment of college graduates is closely related to the socioeconomic situation, industrial structure, and graduates' own psychology. With the continuous development of social economy, the average per capita income of the people is constantly increasing, and the labor force will first move from the primary industry to the secondary industry. When the per capita national income is further increased, the labor force will move to the tertiary industry as a whole.

Logistic regression model is one of the most commonly used statistical analysis models when analyzing classified dependent variables. The logit form of logistic regression model is

$$\ln \left(\frac{p_i}{1-p_i} \right) = \alpha + \beta x_i. \quad (7)$$

When there are k independent variables, the model is extended to

$$\ln \left(\frac{p_i}{1-p_i} \right) = \alpha + \sum_{k=1}^k \beta_k x_{ki}. \quad (8)$$

Logistic regression in the general sense requires that the dependent variable y has only two values, namely, the binary variable. In fact, the logistic regression model is not limited to the application of binary response variables [19].

For multiclass response variables, that is, the response variables with three or more classes, logistic regression is also applicable as long as the model is slightly improved.

MLM (multinomial logit model) is based on logistic probability density function. If the variable t is a random variable of logistic function, its probability density function is

$$f(t) = \frac{e^{-1}}{(1+e^{-1})^2}, -\infty < \infty. \quad (9)$$

In MLM, the observed probability of a y value of 1 is

$$p_i = p[T \leq \alpha + \beta x_i] = F(\alpha + \beta x_i) = \frac{1}{1 + e^{-(\alpha + \beta x_i)}}. \quad (10)$$

In order to better study the influence of different industries on college students' employment, a linear regression analysis is adopted to establish a regression model between each industry and college students' employment, and a linear regression model between the output value of different industries and the number of college students' employment is preliminarily established:

$$y_i = a + bx_i + \varepsilon_i, \quad (11)$$

where y is the dependent variable, x is the independent variable, ε is the random error, a, b is the regression parameter to be estimated, b can represent the employment contribution rate per unit output value, and subscript i represents the i observed value.

Using the retail price index of urban commodities can not only reflect the changes of urban retail prices in different periods but also analyze the impact of retail price changes on the monetary expenditure and living consumption level of urban residents. The calculation of average price of representative products is

$$p_t = \sum \frac{p_i}{n}, \quad (12)$$

where p_t is the average price of the product with the t specification, p_i is the price of the i th survey of the product with the t specification in the survey period, and n is the number of surveys of the product with the t specification in the survey period.

Taking whether graduates are employed or not as dependent variables, and taking the above-mentioned different factors as independent variables, a regression model is established. Because there are only two possibilities for dependent variables, 0 and 1 (employed is 1 and unemployed is 0), which are dichotomous variables, and explanatory variables are classified variables, so they are not suitable for ordinary linear regression models. Therefore, this paper adopts binomial logit model for analysis. The model form is as follows:

$$\ln \left(\frac{p}{1-p} \right) = a + \sum_{t=1}^n b_t x_t + \varepsilon. \quad (13)$$

In the formula, p is the probability of graduates' employment, and $1-p$ is the probability of graduates' nonemployment; a is a constant term; x_t is the explanatory variable; b_t is the regression coefficient of each explanatory variable; ε is a random error term.

3. Result Analysis

The employment part of the original data in this example comes from the statistical data of the employment situation of fresh graduates of H University Employment Guidance Service Center. After obtaining the original data, strict coding input and verification have been carried out. Therefore, the data quality is reliably guaranteed.

Through the overall analysis of the above comprehensive score, we will further analyze the influence of each dimension on EPI score in detail and summarize the dimension indicators that have important influence on EPI score. Figure 4 shows the EPI scores of college students in 2013-2020 according to dimensional indicators.

We reflect the unfair employment phenomenon through the income gap; that is, as the income gap between urban and rural residents in H University area is gradually widening, the unfair employment phenomenon will adversely affect the EPI scores of college students. On the other hand, the economic development index shows a tortuous upward trend.

Under normal circumstances, with the development of economy, the GDP created by people should develop in the same direction as the regional economic growth. In terms of social insurance and labor supply and demand, the scores of both show a trend of first increasing and then decreasing. The possible reason for this phenomenon is that the government policy and the social market environment have different demands on the labor force.

The motivating factors that influence graduates' satisfaction include job satisfaction, job nature, and promotion opportunities. Policy and administration, interpersonal relationship, and personal life are the health care factors that affect graduates' employment satisfaction. With the expansion of the scope of responsibilities and working face of the talent market, the government should take active policies and measures to strengthen the service of the talent market for college graduates' job selection and employment, establish and improve the service window specifically for college graduates, collect the demand information of enterprises and institutions for college graduates extensively, and effectively combine it with the employment of college graduates. Figure 5 shows the scores of EPI secondary indicators of college students from 2013 to 2020.

It can be seen that the average rate of talent supply ratio is higher than the rate of demand. The main difference between the high-paid talent market and the labor market is that although the groups needed by the two are the same, the employment tendency of college students is, of course, directed at high-paid jobs. This will lead to the pressure that the supply of some posts far exceeds the demand of posts. However, some post personnel are seriously lacking. The policy comes from the demand of the society for this kind of policy. Therefore, if we want to have an idealized policy, we should not only ensure the process of policy implementation but also ensure the timely formulation of employment policy. In the process of policy implementation, all aspects of policy implementation should give timely information feedback to the problems arising from policy implementation and revise the employment policy. This is the idealized policy.

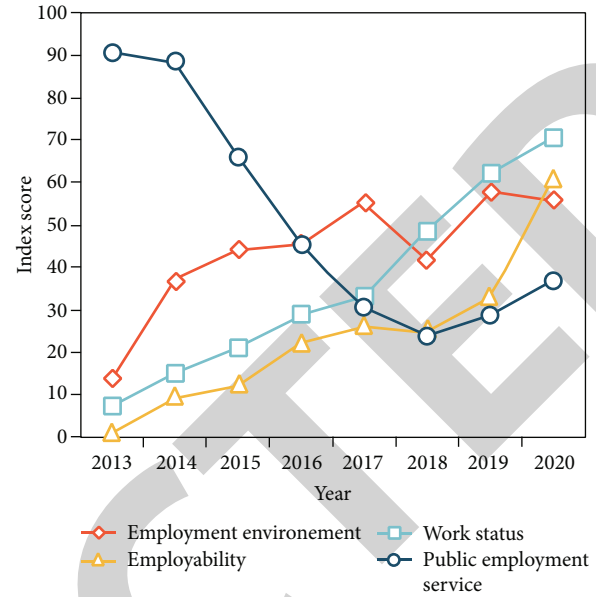


FIGURE 4: EPI scores of college students by dimensional indicators from 2013 to 2020.

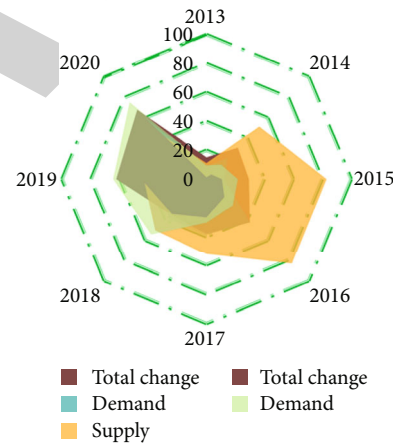


FIGURE 5: Scores of EPI secondary indicators of college students from 2013 to 2020.

The main purpose of this logistic regression model is to analyze the influencing factors of university graduates' reluctance to stay in H city. Therefore, when setting the dummy variable for the explained variable "whether to stay in H city," we take "stay in H city" as the benchmark variable and set the value to 0, take "not staying in H city" as the observation variable, and set the value to 1. Logistic regression is carried out by SPSS software, and three very insignificant factors, such as salary amount, consumption level, and corresponding position, are eliminated. The regression results are shown in Table 1.

Considering the significance level, among the influencing factors, salary level, educational background, talent policy, and the number of famous enterprises have the greatest influence on staying in H city. Among them, the salary level is the most important factor, and its chi-square test value reaches 60.1%. That is to say, the main factor that

TABLE 1: Results of econometric regression analysis after MLM correction.

Explanatory variable	Coefficient of regression	Significant level
Academic degree	-1.25	0.00
Gender	0.74	0.03
Salary level	2.24	0.00
Job prospect	1.89	0.07
Talent policy	1.04	0.001
Economic conditions	1.01	0.05

TABLE 2: Changes of EPI of enterprises of different scales from month to month.

Scale	Demand number	Number of applicants	EPI increase
Miniature	27.1%	28.3%	-0.01%
Small-sized	1.2%	26.1%	-0.22%
Medium-sized	-2.3%	25.5%	-0.23%
Large-scale	10.5%	21.1%	-0.15%

graduates do not stay in H city is that the salary level in H city is too low. The correlation coefficient of knowledge of talent policy is 1.04, that is to say, the lack of knowledge of talent policy in H city has caused graduates to be reluctant to stay in H city to some extent.

To some extent, it is also detrimental to career development. Many universities lack professional career guidance teachers, which makes career planning education useless. Most students do not really understand the exact meaning of career planning, and they do not really master the methods and skills of career planning. Most of the plans made are mere formality and lack of practicality, which leads to the failure of career planning to give full play to its due role. This may be the reason why the vocational qualification certificate and career planning have no significant impact on the employment of college graduates.

Table 2 lists the EPI and month-on-month changes of enterprises of different sizes in the first quarter of 2019.

In Table 2, from the month-on-month changes, in terms of demand numbers, the number of medium-sized enterprises decreased, while the number of small enterprises was basically the same, while the number of large and microenterprises increased significantly, with an increase of 10.5% and 27.1%, respectively. In terms of EPI, except for microenterprises, the EPI of enterprises of other sizes increased to a certain extent year-on-year. This may be due to the fact that microenterprises are in the “start-up stage,” and most enterprises have weak market competitiveness and short life cycle, so the employment demand has declined to some extent.

According to the calculation results of the factors of economic conditions, employment factors, and employment policies of large- and medium-sized cities, the scores of the employment monitoring indexes of eight cities in 2020 can be obtained, as shown in Figure 6:

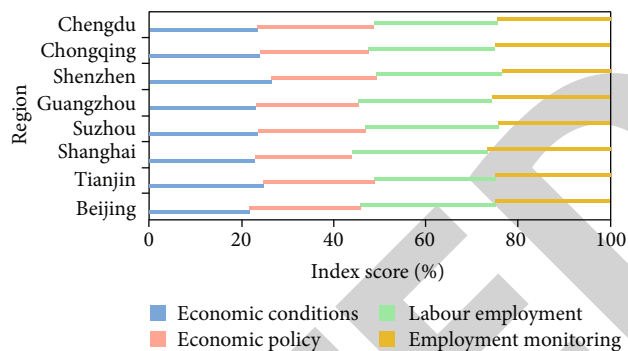


FIGURE 6: Scores of employment monitoring index of cities in 2020.

As can be seen from the above table, from the index score, the employment situation in economically developed areas is also better. The score of Shanghai is far lower than that of other first-tier cities, which is due to the low economic condition monitoring index and employment monitoring index of Shanghai. Shanghai's economic growth slows down in 2020. Due to industrial transformation and upgrading, the contradiction between supply and demand in the labor market is prominent. These factors affect the scores of employment monitoring index in large- and medium-sized cities.

4. Conclusion

Based on the data of a university, this paper uses MLM (multinomial logit model) to optimize the EPI system. Through empirical research, it is found that the system-optimized EPI has the function of evaluating the employment environment of college students. It not only provides the government with macroemployment policies but also provides scientific and effective employment service policies for colleges and universities. As an index system to evaluate and guide college students' employment, EPI mainly includes employment elasticity and employment effect of college students in different industries. EPI after system optimization can not only reflect the overall trend of the job market but also reflect the structural changes of the job market such as subregion, subindustry, suboccupation, and sub-enterprise scale. At the same time, we can also get the long-term trend, seasonal characteristics, and periodic changes of the prosperity degree of the job market through the decomposition and analysis of this index, so as to grasp the long-term, short-term, and medium-term prosperity degree changes of the job market more carefully. It provides guidance for employers and job seekers to realize benign interaction, provides intuitive graphic data support for government departments at all levels to understand the dynamics of human resource market, and provides scientific basis for labor and employment management service departments to grasp the overall employment situation and make targeted employment policy decisions.

However, the study still has some limitations. The conclusion and applicability of the research on the employment competitiveness of this paper are conditional. The research

Retraction

Retracted: Optimization of the Intelligent Sports Health Management System Based on Big Data Technology and Internet of Things

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Dang, "Optimization of the Intelligent Sports Health Management System Based on Big Data Technology and Internet of Things," *Journal of Function Spaces*, vol. 2022, Article ID 5053150, 19 pages, 2022.

Research Article

Optimization of the Intelligent Sports Health Management System Based on Big Data Technology and Internet of Things

Jian Dang 

Department of Sports, Xinxiang Medical University, Xinxiang 453003, China

Correspondence should be addressed to Jian Dang; 031007@xxmu.edu.cn

Received 26 April 2022; Revised 27 May 2022; Accepted 15 June 2022; Published 11 July 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Jian Dang. 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.

People often suffer from unpredictable injuries during physical exercise. One of the important reasons is the absence of a scientific sports health management system. Therefore, the construction of such a scientific and effective system has gradually attracted the attention of scholars, which is of great significance to realizing people's scientific and personalized physical fitness. An intelligent sports health management system based on big data analysis and the Internet of things (IoT) is constructed to solve this problem. The system consists of the user, IoT, cloud, system analysis, evaluation, and data layers. Firstly, a new multilabel feature selection algorithm is proposed in the system analysis layer. The suggested multilabel feature selection algorithm maps the sample space to the label space through the L_{21} norm. Then, the consistency of various topologies is guaranteed by combining with feature popularity so that the factors affecting user health can be better selected. Secondly, the experiment is compared with SCLS, SSFS, and six other multilabel feature selection algorithms in 6 classic medical multilabel datasets. Experimental results under five indexes show the effectiveness and superiority of the proposed feature selection algorithm. Finally, the feasibility of the proposed intelligent sports management system is analyzed.

1. Introduction

Nowadays, people will pay more attention to their health and beauty and more people will choose to experience a happy life, which requires the support of a fitness system tailored for themselves. Scientific and reasonable fitness behavior can meet people's needs for a better life and make people have a beautiful appearance and happy mood. Moreover, the unscientific fitness mode often brings all sorts of misfortune to people, for example, causing muscle injury, joint deformation, heart function injury, iron deficiency anemia, sports sex hematuria proteinuria, and other diseases. At present, people's cognition of how to reasonable fitness is still relatively narrow and most of them follow the trainer's years of experience to give a more rational way of wellness. However, despite this, there will still be many nonstandard, accessible to damage practices of physical fitness and damage user health. Combining the excellent experience of coaches with

people's situations and building a tailor-made sports health management system for people has become an urgent problem to be solved. In life, we should nip it in the bud. User health should be based on prevention rather than treatment. Scientific and reasonable health management and fitness are the main ways to maintain health.

In the era of rapid development of science and technology, user health is no longer limited to running, Tai Chi, long jump, and other ways. There are a variety of exercises, even for one part of the body, such as wrist kicks, which have become popular in recent years. It has become a headache for people to choose fitness methods for themselves. After all, the correct choice will get twice the result with half the effort. The wrong choice is likely to be ineffective or even counterproductive. This paper proposes an intelligent sports health management system based on big data analysis and the IoT to solve the abovementioned problems. The main contributions of this paper are as follows:

- (1) In order to make the intelligent sports health management system more targeted, combined with big data analysis technology, a new multilabel feature selection algorithm is suggested to ensure that the main factors affecting user health can be extracted effectively by this method
- (2) Combining the proposed multilabel feature selection algorithm with the IoT technology, the basic framework of the intelligent sports management system is constructed
- (3) A series of comparative experiments were carried out on six classical medical multilabel datasets, and the experimental results proved the effectiveness of the suggested algorithm. The feasibility of the suggested intelligent sports health management system is illustrated through the feasibility analysis

The structure of the paper is as follows: in the second section, the symbol description and related work are given. In the third section, the theoretical support, optimization solution technique, algorithm design, and convergence proof of the proposed feature selection model are given and the design principle of the proposed system is introduced. In the fourth section, the experimental settings are compared with SCLS and SSFS and the experimental results are analyzed. Finally, the summary and prospects of this paper are given in the fifth section.

2. Literature Review

In this section, we not only make a brief overview of feature selection in big data analysis but also make a brief overview of the health management system as follows:

There has been much research on feature selection algorithms in recent years. In many works, feature selection models are divided into three types: filter [1–3], wrapper [4, 5], and embedded [6, 7]. This paper mainly uses embedded feature selection. Literature [8] combines logistic regression, manifold learning, and sparse regularization to construct a feature selection algorithm. In the literature [9], a possible structure sharing (LSS) term is designed and the construction of the feature selection algorithm is completed in combination with spiritual learning.

Both [10, 11] used dynamic graphs to learn the basic manifold structure of samples or labels and then combined them with linear regression to build feature selection models. Reference [10] strengthens the local connection between samples and labels by combining with subspace to better special features. Reference [11] strengthened the correlation between the weight matrix and sample space and between the weight matrix and label space by comprehensively restricting the weight matrix, making the weight matrix more representative of the weight of features and more accessible to distinguish features.

Furthermore, several multilabel feature selection algorithms include a variety of factors. For instance, many scholars use a multilabel feature selection strategy characterized by mutual information:

Among them, modeling the feature selection process as a multicriteria decision process was proposed for the first time in [12]. This approach applies to multilabel data, using the TOPSIS (order preference technology similar to ideal solution) approach as the well-known MCDM algorithm to evaluate features based on relationships with multiple labels as different criteria.

The SCLS algorithm was proposed in [13], which is an effective feature selection algorithm. Still, the feature selection algorithm is easily affected due to the excessive combination of labels and features. Multilabel feature selection algorithms also use mutual information to evaluate features. They feature sorting algorithms to achieve feature selection, such as MDMR [14], PMU [15], and FIMF [16].

And there are a lot of studies on health management systems. The comprehensive health management service platform for the public is proposed in [17], which promotes the development of sports health management services through intelligent sensors and intelligent health system detection equipment. Aiming at the problems of a long cycle and high cost of existing intelligent sports health management systems, an improved particle swarm optimization algorithm is proposed in [18] to optimize the intelligent sports health management system. A new type of intelligent sports management system is proposed in [19], which is constructed by using information technology and human-computer interaction technology under artificial intelligence and combining it with deep learning technology. There is an organized review of healthcare management systems in [20] and enhancement of healthcare management systems through many of the latest IoT-oriented healthcare applications.

Firstly, most of the abovementioned multilabel feature selection algorithms are based on linear regression and mutual information. Among them, the multilabel feature selection algorithm based on linear regression has poor robustness due to the loss function (least square) of linear regression. Moreover, the multilabel feature selection algorithm based on mutual information has high algorithm complexity and is not suitable for the analysis of high-dimensional data. Therefore, in order to better analyze user data features, a multilabel feature selection algorithm based on L_{21} norm regression is constructed to promote the analysis ability of the system. In addition, the existing intelligent sports health management system seldom uses expert experience data to guide data analysis. Finally, a new type of intelligent sports health management system is constructed by adding expert experience data into system analysis and combining the proposed multitag feature selection algorithm.

3. Method

To address the problem identified in Section 2, in this section, the design principle of an intelligent sports health management system based on big data analysis and IoT is introduced in detail from three aspects: feature selection in big data analysis technology, the IoT technology, and intelligent sports health management system.

3.1. IoT Technology. The IoT technology refers to the use of the contractual network protocol, and the use of IoT information induction of all kinds of equipment (infrared image sensor, the world's first GPS satellite positioning system, radio and laser image scanner, etc.) will be connected to the Internet at the same time with any kind of objects. The IoT network can simultaneously achieve a variety of voice and image control, track objects' location, and make image monitoring and information management to facilitate the IoT information network exchange and data communication. Thus, the IoT has unique advantages in fitness data collection and information output.

About the IoT and its application in intelligent sports health management systems, the first aspect is data collection: a large number of actual and reliable data are the basis for the construction of the suggested model, including not only users' physical health status data but also medical index data, as well as the experience summary data of well-known coaches, etc. With these data, data mining technology can be used to analyze people's physical conditions and give a reasonable way of fitness. At the same time, the IoT technology can collect information through various intelligent fitness equipment, such as brilliant fitness bracelets, and innovative tracks.

Among them, wearable devices have been widely used in sports activities. The most common are smart bracelets, mobile phones, etc. Such devices can collect a user's daily energy intake and output and data such as the user's heart rate and exercise.

In terms of data output, the IoT technology can push the evaluation results of big data analysis and suggested fitness methods to users through various intelligent devices, such as wearable devices, mobile phones, computers, and VR. These intelligent devices typically have screens that allow users to read information pushed by the IoT.

Among them, VR equipment can help users learn standard movements. Users can achieve the maximum fitness effect in the shortest time. In addition, due to the novelty of VR, it can also enhance users' interest in fitness, develop more users, and lead the wave of reasonable fitness for all.

3.2. Symbol Description. This subsection will briefly introduce the symbols and meanings used in this article. The specific content is shown in Table 1. For any matrix M :

In addition, $X \in R^{n \times d}$ represents the sample matrix, $Y \in R^{n \times m}$ represents the real label matrix, and $W \in R^{d \times m}$ is the matrix of coefficients.

3.3. Big Data Analysis Technology (Feature Selection). Big data analysis technology is an essential component of knowledge discovery, using computer algorithms to analyze data. In many databases, the required data should be obtained and the data should be properly converted, mined, and utilized to obtain valuable information. Generally speaking, the objects of big data analysis are basically structured, semi-structured, or other structured data.

Feature selection is a kind of big data analysis technology that selects d most representative features from n features of the sample. Feature selection is used for data dimensionality

TABLE 1: Notations of this paper.

Symbol	Meaning
M_{ij}	The i th row and j th column of M
M_{i*}	The i th row vector of M
M_{*j}	The j th column vector of M
M^T	M transpose
$tr(M)$	Trace of M

reduction. The selected features are also the most representative, so we can analyze the main factors affecting user health through feature selection to provide more targeted scientific and reasonable fitness methods. Due to the complexity of data from users and their lives, multilabel feature selection is more of the system, for analyzing the main factors affecting user health. Therefore, we propose a new multilabel feature selection method as follows:

Given multilabel dataset $\mathbf{X} \in R^{n \times d}$, its corresponding label set $\mathbf{Y} \in R^{n \times m}$. According to the assumption of the problem, the loss function $L(\mathbf{W})$ of the model is constructed and the optimal solution is sought by minimizing $L(\mathbf{W})$. The formula is as follows, where $\mathbf{W} \in R^{d \times m}$ is the feature weight matrix, which can reflect the importance of features.

$$\Phi = \min L(\mathbf{W}). \quad (1)$$

In addition, the penalty function $R(\mathbf{W})$ is often imposed on the weight matrix \mathbf{W} and the property of constraint \mathbf{W} is that the model is further idealized:

$$\Phi = \min L(\mathbf{W}) + \alpha R(\mathbf{W}), \quad (2)$$

where α is the penalty factor.

Due to the least square loss function of linear regression, although the calculation is simple, it is not as robust as the L_{21} norm. Therefore, L_{21} norm regression is used to describe the relationship between the sample set and the label set:

$$L(\mathbf{W}) = \|\mathbf{XW} - \mathbf{Y}\|_{2,1}. \quad (3)$$

The basic framework of the suggested algorithm is constructed:

$$\Phi = \min \|\mathbf{XW} - \mathbf{Y}\|_{2,1} + \alpha R(\mathbf{W}). \quad (4)$$

In the suggested algorithm, we constrain the learning of feature weight matrix in the following ways:

Let $\mathbf{X}_{\cdot i}$ represent the i th-column vector of dataset \mathbf{X} and $\mathbf{W}_{i\cdot}$ represent the i th-row vector of the weight matrix \mathbf{W} ; if $\mathbf{X}_{\cdot i}$ and $\mathbf{X}_{\cdot j}$ are similar, then, $\mathbf{W}_{i\cdot}$ and $\mathbf{W}_{j\cdot}$ are similar. Thus, the similarity between features can be used to guide the learning of the weight matrix and build the feature manifold learning model:

Input: data matrix $X \in R^{n \times d}$, label matrix $Y \in R^{n \times m}$, regularization parameters α and β , select the number of features K .
 Output: feature selection result I .

- a) Calculate L_{X^T} according to Eq (5) and (6).
- b) The initialization matrices H and D are the identity matrices.

Repeat:

$$\text{Update } W: W = (D^{-1}X^T H X + \alpha D^{-1}L_{X^T} + \beta)^{-1} D^{-1}X^T H Y$$

$$\text{Update } D: D = 1/\|W_{i\bullet}\|_2$$

$$\text{Update } H: H = 1/\|(XW - Y)_{i\bullet}\|_2$$

Until it converges.

- c) Compute $\|W_{i\bullet}\|_2$; ($i = 1, 2, \dots, d$) and sort it to find the first k largest assigned to I

ALGORITHM 1: The suggested algorithm.

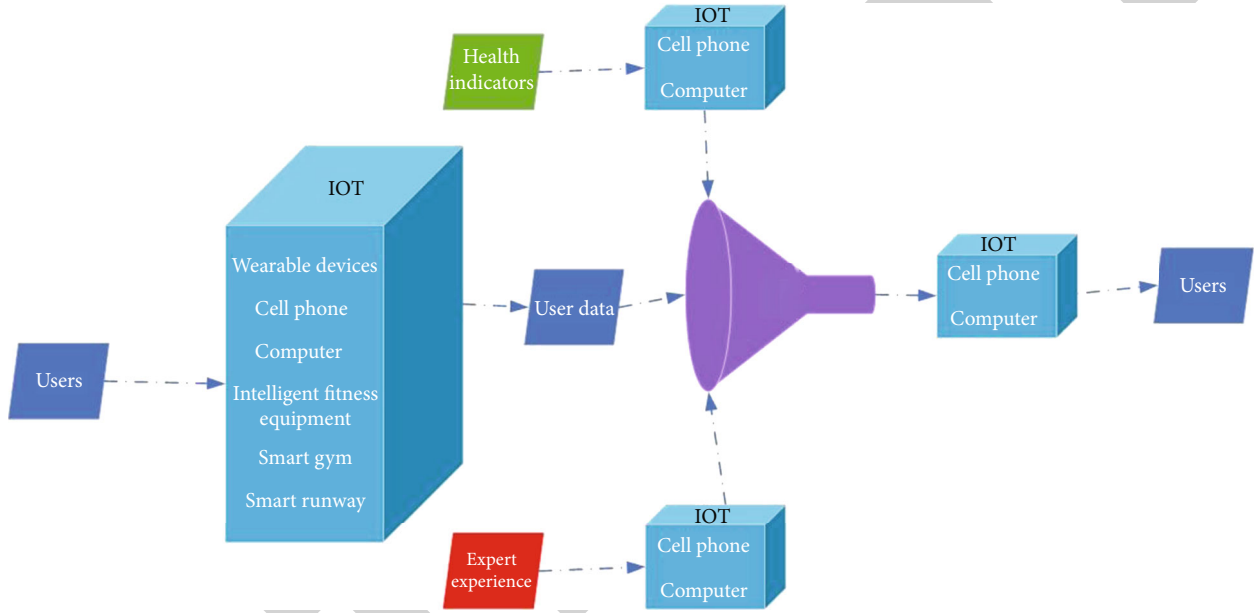


FIGURE 1: Structure diagram of smart sports health system.

TABLE 2: Dataset parameters.

No.	Dataset	Instances	Features	Labels	Cardinality	Training	Test
1	Abide	2417	103	14	4.237	1500	917
2	Oasis	593	72	6	1.869	391	202
3	Ddsm	645	260	19	1.014	322	323
4	Mias	2407	294	6	1.047	1211	1196
5	Mura	600	294	5	1.236	400	200
6	Luna16	662	1185	27	1.252	463	199

TABLE 3: Display of optimal experimental results under index hamming loss.

Algorithm	MFS_MCDM	SSFS	SCLS	MDMR	PMU	FIMF	Ours
Abide	0.2014	0.2137	0.2006	0.1999	0.2006	0.2021	0.1947
Oasis	0.2302	0.2418	0.2500	0.2409	0.2673	0.2252	0.2162
Ddsm	0.0471	0.0495	0.0499	0.0505	0.0504	0.0520	0.0468
Mias	0.1066	0.1290	0.1073	0.1348	0.1137	0.1587	0.1024
Mura	0.2050	0.2164	0.2110	0.2240	0.2270	0.2340	0.2000
Luna16	0.0342	0.0428	0.0326	0.0024	0.0056	0.0030	0.0020

TABLE 4: Display of optimal experimental results under index ranking loss.

Algorithm	MFS_MCDM	SSFS	SCLS	MDMR	PMU	FIMF	Ours
Abide	0.1742	0.1925	0.1745	0.1710	0.1723	0.1747	0.1673
Oasis	0.1864	0.2010	0.2056	0.1994	0.2570	0.2012	0.1700
Ddsm	0.1944	0.2138	0.2655	0.2591	0.2585	0.2586	0.2055
Mias	0.1201	0.1463	0.1129	0.1444	0.1290	0.1994	0.1027
Mura	0.2258	0.2477	0.2167	0.2550	0.2483	0.2662	0.1954
Luna16	0.0480	0.0869	0.0653	0.0066	0.0080	0.0078	0.0062

TABLE 5: Display of optimal experimental results under index one-error.

Algorithm	MFS_MCDM	SSFS	SCLS	MDMR	PMU	FIMF	Ours
Abide	0.2356	0.2508	0.2268	0.2366	0.2366	0.2366	0.2236
Oasis	0.3119	0.3455	0.3614	0.3564	0.3614	0.3515	0.2673
Ddsm	0.5465	0.5872	0.6454	0.6744	0.6395	0.7035	0.5291
Mias	0.3169	0.3661	0.2977	0.3905	0.3904	0.4983	0.2617
Mura	0.4200	0.4300	0.4000	0.4450	0.4700	0.5000	0.3650
Luna16	0.4221	0.5477	0.4472	0	0	0	0

TABLE 6: Display of optimal experimental results under index coverage.

Algorithm	MFS_MCDM	SSFS	SCLS	MDMR	PMU	FIMF	Ours
Abide	6.4569	6.6772	6.4482	6.3642	6.3708	6.3740	6.2814
Oasis	1.9851	2.0842	2.1139	2.0891	2.3614	2.0545	1.9158
Ddsm	2.2755	2.5542	3.2012	3.0495	3.0526	3.0526	2.3746
Mias	0.7032	0.8329	0.6681	0.8253	0.7492	1.0953	0.6154
Mura	1.1750	1.2520	1.1650	1.3200	1.2900	1.3550	1.0650
Luna16	1.6181	2.7236	2.1809	0.5930	0.6080	0.5930	0.5729

TABLE 7: Display of optimal experimental results under index average precision.

Algorithm	MFS_MCDM	SSFS	SCLS	MDMR	PMU	FIMF	Ours
Abide	0.7551	0.7312	0.7563	0.7579	0.7562	0.7552	0.7655
Oasis	0.7815	0.7584	0.7496	0.7551	0.7143	0.7510	0.8071
Ddsm	0.5302	0.5143	0.4435	0.4158	0.4435	0.4074	0.5162
Mias	0.8058	0.7727	0.8163	0.7633	0.8034	0.6906	0.8365
Mura	0.7288	0.7208	0.7437	0.7058	0.7002	0.6791	0.7655
Luna16	0.7071	0.6044	0.6882	0.9919	0.9907	0.9915	0.9939

$$\begin{aligned}
& \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \|\mathbf{W}_{i\bullet} - \mathbf{W}_{j\bullet}\|_2^2 \mathbf{S}_{X^T ij} \\
&= \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d (\mathbf{W}_{i\bullet} - \mathbf{W}_{j\bullet}) (\mathbf{W}_{i\bullet} - \mathbf{W}_{j\bullet})^T \mathbf{S}_{X^T ij} \\
&= \text{Tr}(\mathbf{W}^T (\mathbf{P}_{X^T} - \mathbf{S}_{X^T}) \mathbf{W}) = \text{Tr}(\mathbf{W}^T \mathbf{L}_{X^T} \mathbf{W}),
\end{aligned} \tag{5}$$

where $\mathbf{L}_{X^T} \in \mathbf{R}^{d \times d}$ is the Laplacian matrix of the Eigen similar matrix; $\mathbf{P}_{X^T} \in \mathbf{R}^{d \times d}$ is the diagonal matrix, $\mathbf{P}_{X^T ii} = \sum_{j=1}^d \mathbf{S}_{X^T ij}$ is the i th diagonal element of \mathbf{P}_{X^T} , \mathbf{S}_{X^T} is the feature similarity

matrix, and $\mathbf{S}_{X^T ij}$ represents the similarity between feature $\mathbf{X}_{\bullet i}$ and $\mathbf{X}_{\bullet j}$. In this study, \mathbf{S}_{X^T} is calculated by Pearson's correlation coefficient with square constraint and its formula is as follows:

$$\mathbf{S}_{X^T ij} = \begin{cases} \frac{\|\mathbf{X}_{\bullet i} - \mathbf{X}_{\bullet j}\|_2^2}{\sigma}, & \text{if } \mathbf{X}_{\bullet i} \in \mathbf{N}_K(\mathbf{X}_{\bullet j}) \text{ or } \mathbf{X}_{\bullet j} \in \mathbf{N}_K(\mathbf{X}_{\bullet i}), \\ 0, & \text{others,} \end{cases} \tag{6}$$

where $\mathbf{N}_K(\mathbf{X}_{\bullet i})$ is the K neighborhood of $\mathbf{X}_{\bullet i}$.

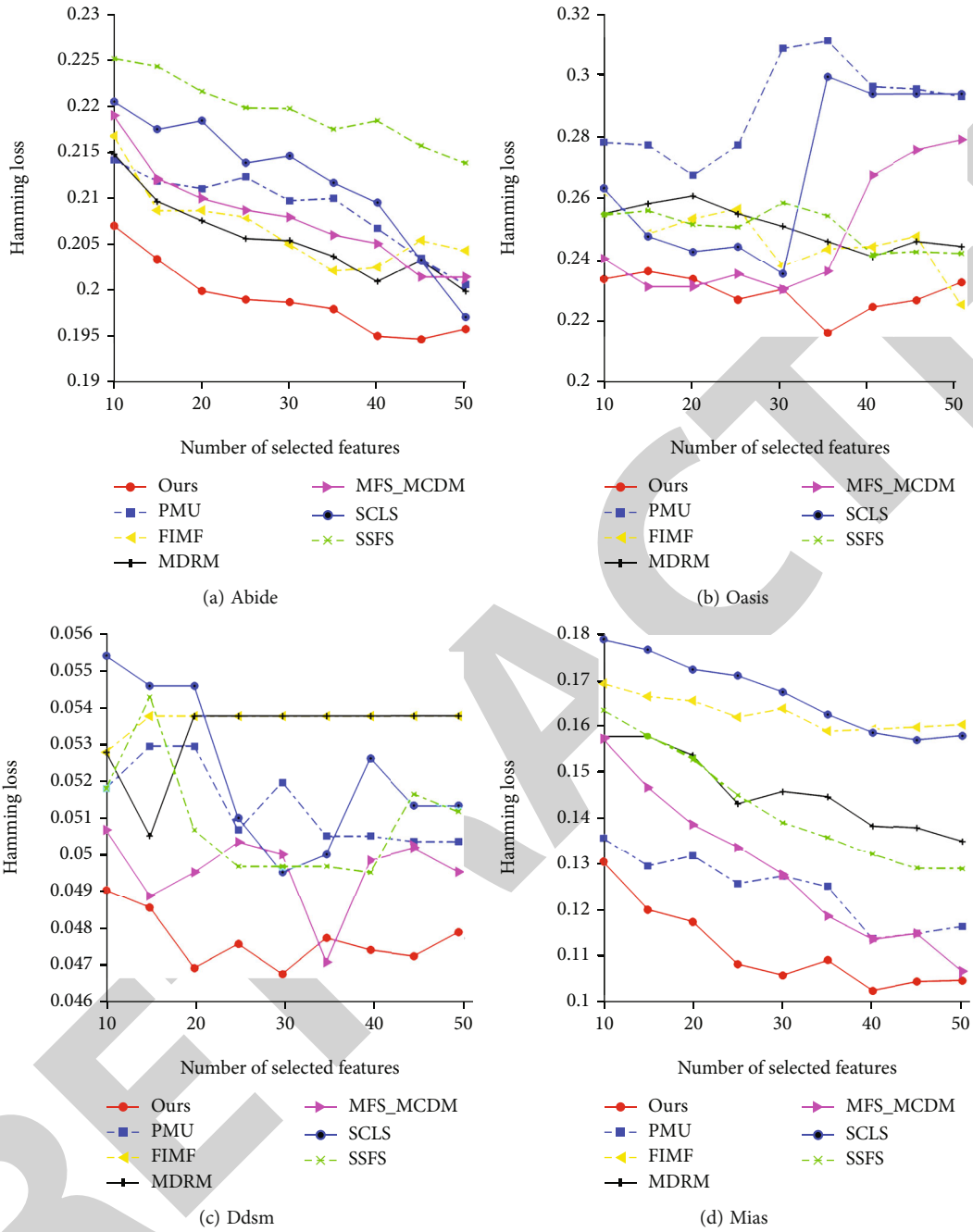


FIGURE 2: Continued.

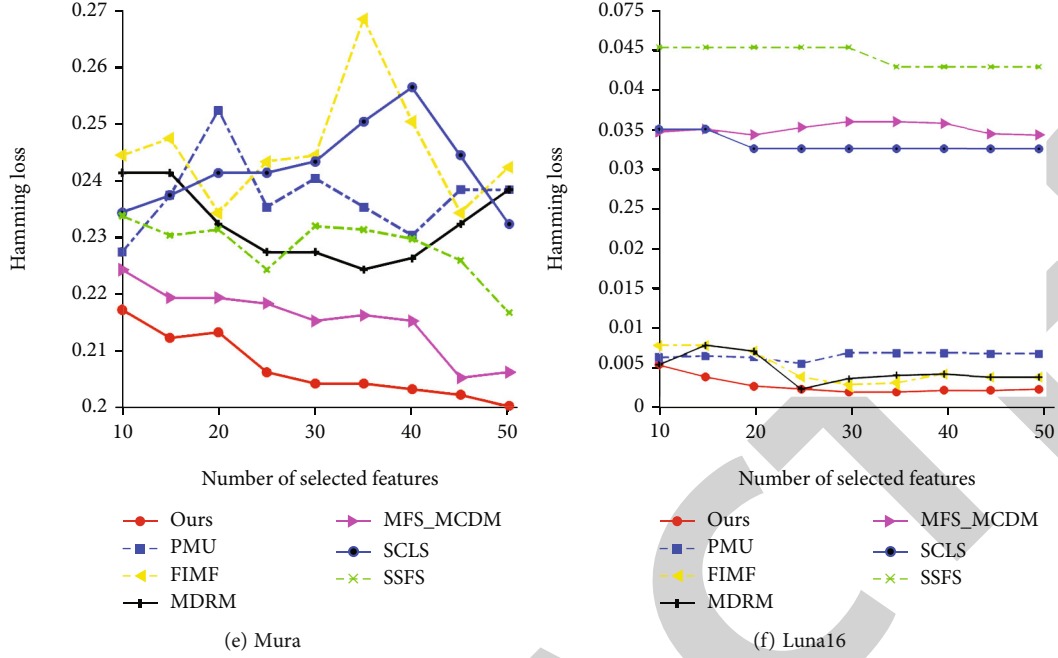


FIGURE 2: Hamming loss: the quality of each algorithm varies with the selected feature quantity.

Thus, a multilabel feature selection model is constructed:

$$\Phi = \min \|\mathbf{X}\mathbf{W} - \mathbf{Y}\|_{2,1} + \alpha \text{tr}(\mathbf{W}^T \mathbf{L}_X \mathbf{W}). \quad (7)$$

Meanwhile, in order to enable the suggested algorithm to perform feature selection better and faster, $L_{2,1}$ norm sparse constraint is applied to \mathbf{W} in equation (7):

$$\Phi = \min \|\mathbf{X}\mathbf{W} - \mathbf{Y}\|_{2,1} + \alpha \text{tr}(\mathbf{W}^T \mathbf{L}_X \mathbf{W}) + \beta \|\mathbf{W}\|_{2,1}. \quad (8)$$

In terms of the optimization solution of the abovementioned formula, due to the nonsmoothness of the $L_{2,1}$ norm, it can be transformed into the following:

$$\Phi = \min \text{tr} \left((\mathbf{X}\mathbf{W} - \mathbf{Y})^T \mathbf{H} (\mathbf{X}\mathbf{W} - \mathbf{Y}) \right) + \alpha \text{tr}(\mathbf{W}^T \mathbf{L}_X \mathbf{W}) + \beta \text{tr}(\mathbf{W}^T \mathbf{D} \mathbf{W}), \quad (9)$$

where $\mathbf{H} = 1/\|(\mathbf{X}\mathbf{W} - \mathbf{Y})_{i\cdot}\|_2$, $\mathbf{D} = 1/\|\mathbf{W}_{i\cdot}\|_2$.

For the solution of equation (9), the derivative function with respect to \mathbf{W} is required:

$$\frac{\partial \Phi}{\partial \mathbf{W}} = \mathbf{X}^T \mathbf{H} \mathbf{X} \mathbf{W} - \mathbf{X}^T \mathbf{H} \mathbf{Y} + \alpha \mathbf{L}_X \mathbf{W} + \beta \mathbf{D} \mathbf{W}. \quad (10)$$

Set the derivative to 0:

$$\mathbf{X}^T \mathbf{H} \mathbf{X} \mathbf{W} - \mathbf{X}^T \mathbf{H} \mathbf{Y} + \alpha \mathbf{L}_X \mathbf{W} + \beta \mathbf{D} \mathbf{W} = 0. \quad (11)$$

Thus, it can be concluded that the updating formula for \mathbf{W} is as follows:

$$\mathbf{W} = (\mathbf{D}^{-1} \mathbf{X}^T \mathbf{H} \mathbf{X} + \alpha \mathbf{D}^{-1} \mathbf{L}_X + \beta)^{-1} \mathbf{D}^{-1} \mathbf{X}^T \mathbf{H} \mathbf{Y}. \quad (12)$$

Through solving the abovementioned problems, the suggested multilabel feature selection algorithm is shown in Algorithm 1:

The big data analysis technology can conduct cluster analysis and collation of the massive sports fitness and health data of residents collected to evaluate the user health status and sports injury status. Applying big data processing technology in sports injury assessment can improve the accuracy and effectiveness of sports injury assessment and enhance the efficiency of sports injury data processing. Specifically, the predictive model predicts possible sports injuries and chronic diseases caused by excessive exercise for users; data mining technology is used to analyze people's fitness behaviors to guide sports training. The recommendation model is used to recommend personalized scientific fitness methods for users.

3.4. The Structure and Function of the Suggested System. The intelligent sports health management system based on big data analysis and the IoT comprises five layers (Figure 1): user, IoT, cloud, system analysis, evaluation, and data, respectively.

3.4.1. First, Basic Information Collection and Management Functions. This function is used to manage the basic health information of users. The specific role is to input and view primary personal exercise data (including resident number, name, region, age, gender, weight, item, duration of continuous exercise, time of each exercise, sports injury history, and past medical history). Data entry and viewing of the personal physical activity status (including basic personal information, maximum heart rate during exercise, exercise time, amount of exercise, calories consumed, and lactic acid accumulation). Introduction of personal fitness and health

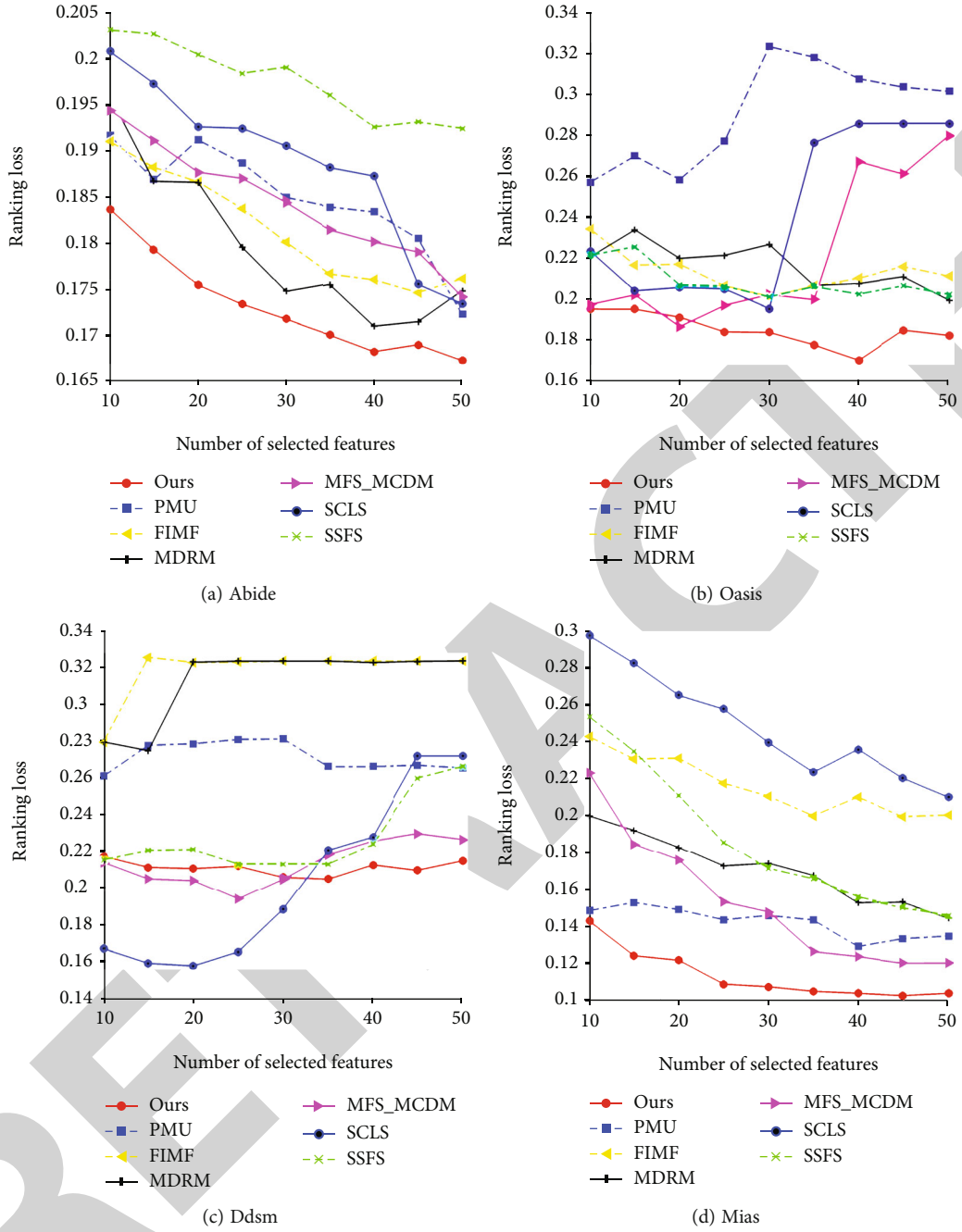


FIGURE 3: Continued.

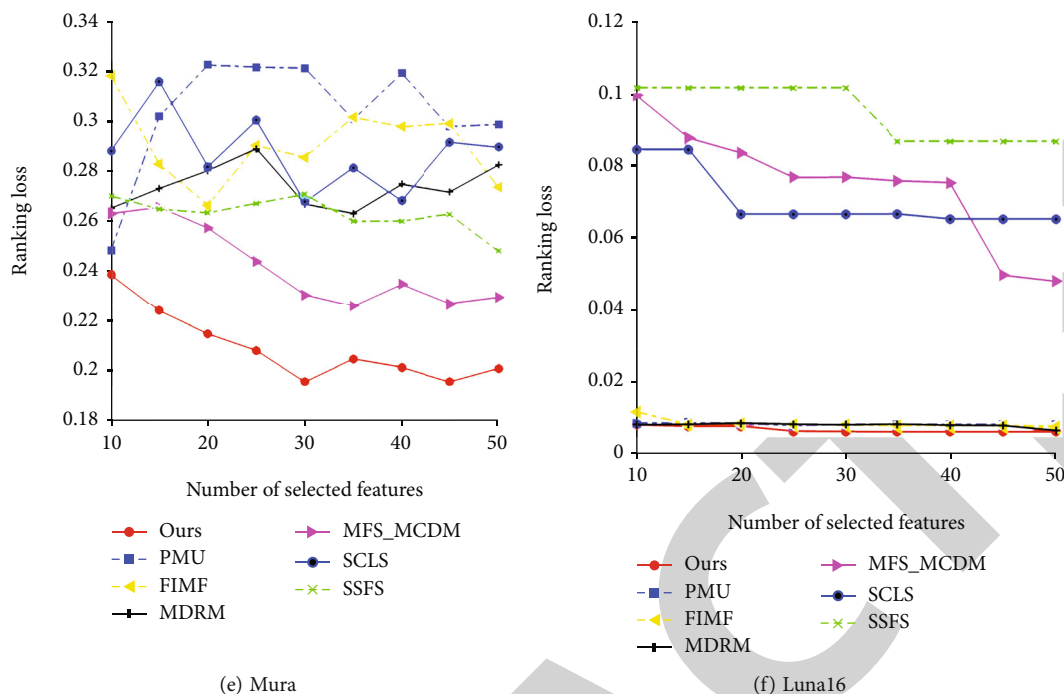


FIGURE 3: Ranking loss: the quality of each algorithm varies with the selected feature quantity.

information, including: residents will automatically import WeChat sports information, wearable smart device information, innovative fitness equipment information into the system, etc; Fitness and health information statistics, including fitness statistics chart, health status statistics chart, etc.

3.4.2. Secondly, Data Analysis, and Health Status Assessment Function. The feature can assess a user health at the right time based on the user input, information collected by wearable devices and intelligent fitness equipment, and big data analysis technology. This function uses a predetermined health assessment model to assess the user health at rest and during exercise according to the real-time collection of basic exercise information and exercise status information. These predetermined fitness assessment models are derived from fitness assessment models in sports medicine.

3.4.3. Thirdly, Exercise Reminder and Disease Prediction Function. Based on big data analysis technology and the sports medicine model, this function predicts sports injuries and diseases that may occur during users' fitness. Fuzzy clustering technology was used to cluster users into multiple user types according to the five indicators of age, weight, gender, exercise program, duration of continuous exercise, and medical history. Collaborative filtering technology and a fuzzy time series prediction model were used to predict possible sports injuries and diseases.

3.4.4. Finally, a Scientific and Reasonable Fitness Program Recommended Functions. The feature is based on collaborative filtering of content, data mining techniques such as association rules, and medical models of motion. This func-

tion can recommend scientific exercise and fitness methods for residents according to their age, gender, weight, medical history, exercise environment and other information, so as to correct the wrong actions of residents in the exercise, avoid sports injury, physical disease and other adverse conditions, and improve the health promotion effect of physical fitness.

4. Results and Discussion

To verify the effectiveness of the proposed feature selection algorithm and the feasibility of the proposed intelligent sports fitness system, six classical medical datasets were used to test the effectiveness of the suggested multilabel feature selection algorithm and the quality of the FSML algorithm was compared with SCLS [13], MDMR [14], PMU [15], FIMF [16], SSFS [9], and MFS_MCDM [12]. In addition, ML-KNN [21] is used for classification and evaluation.

4.1. Experimental Setup. Six classic datasets from the medical field were used: Abide, Oasis, Ddsm, Mias, Mura, and Luna16. The specific parameters of each dataset are shown in Table 2:

In terms of experimental settings, first, the experimental data are discretized [22]. Second, set the parameters in algorithms MDMR, ML-KNN, and FIMF to default values. Finally, because some parameters need to be randomly initialized in the MFS_MCDM algorithm and SSFS algorithm, the experimental results given by the MFS_MCDM algorithm and SSFS algorithm are the mean of 10 times running and the experimental results provided by all algorithms are the optimal results under the optimal parameters.

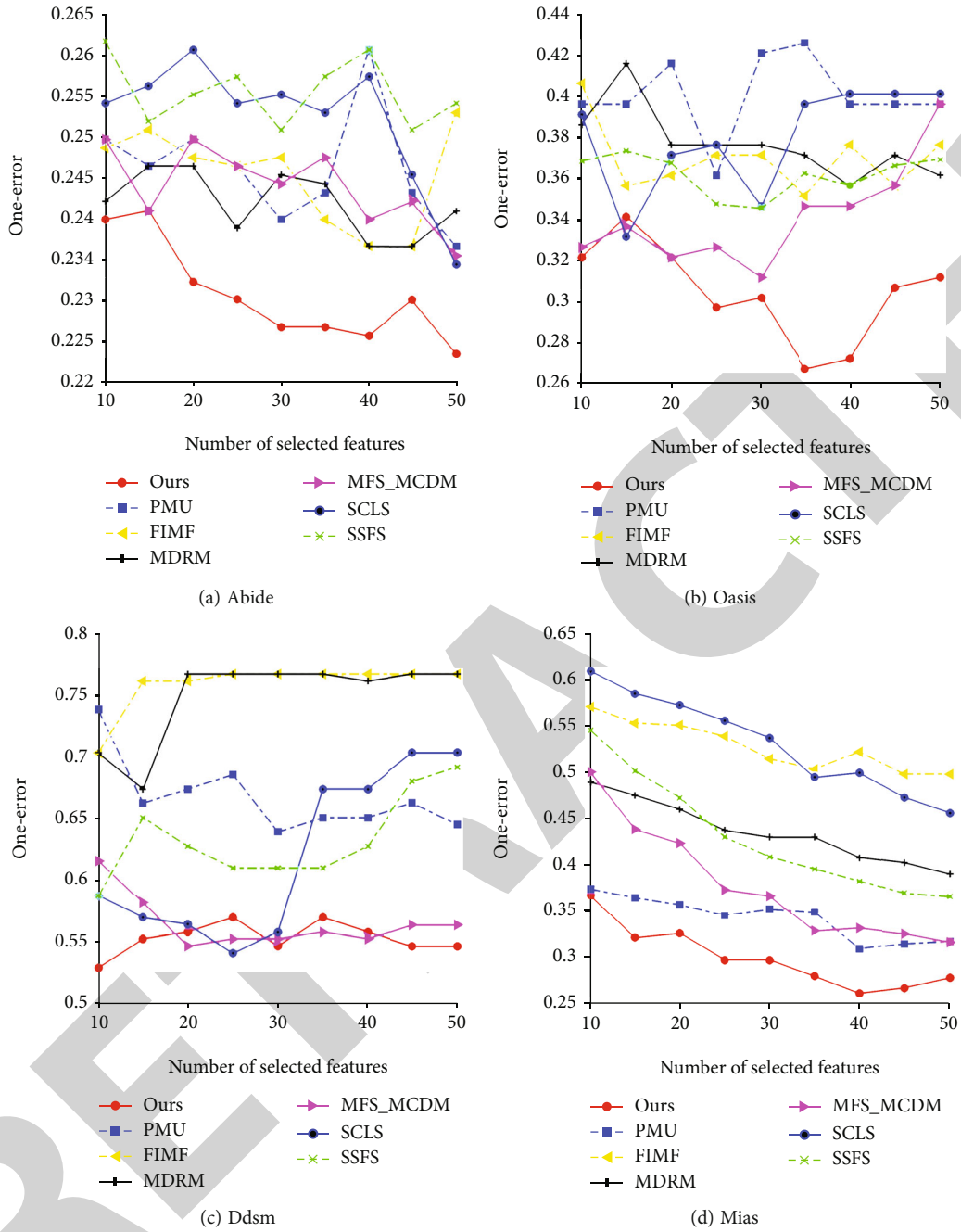


FIGURE 4: Continued.

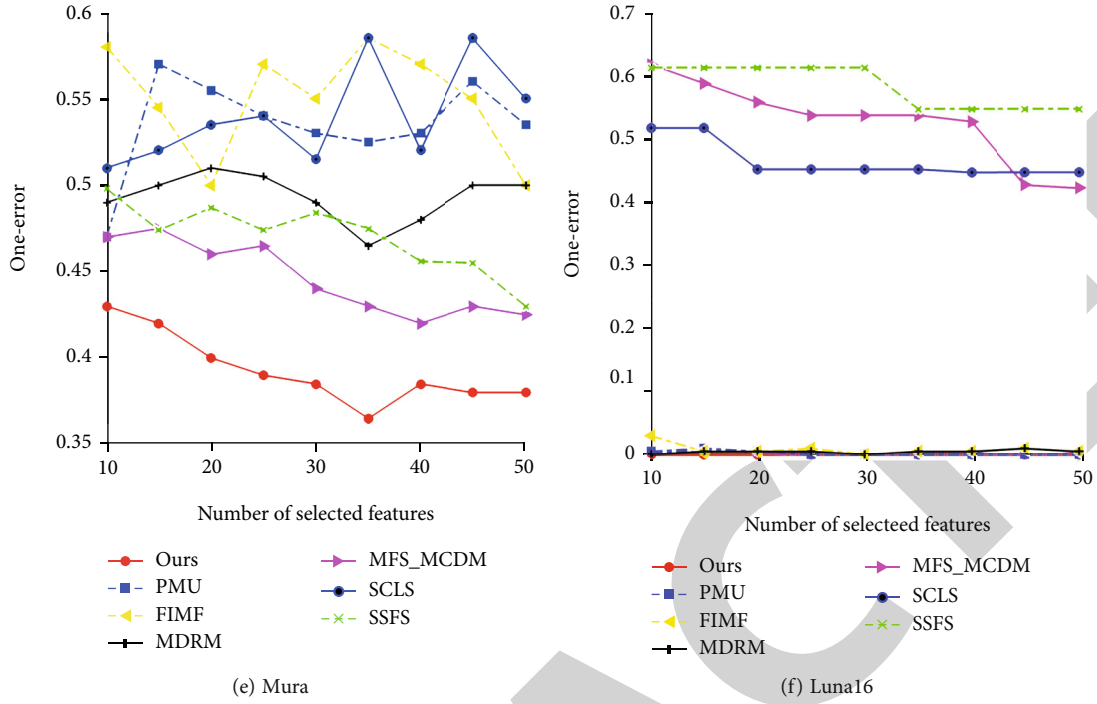


FIGURE 4: One-error: the quality of each algorithm varies with the selected feature quantity.

In terms of evaluation indicators, five evaluation indexes: hamming loss, one-error, coverage, average precision, and ranking loss, were selected to evaluate and compare the quality of each experimental algorithm comprehensively. The specific meanings of each indicator are as follows:

Let $\mathbf{D} \in \mathbf{R}^{n \times d}$ be the sample data of the training set and $\mathbf{Y} \in \mathbf{R}^{n \times m}$ be the corresponding label set data. $\mathbf{h}(\mathbf{D}_{i \bullet})$ represents the binary label vector, and $\text{rank}_{i \bullet}(\mathbf{q})$ represents the rank predicted by label $\mathbf{Y}_{\mathbf{q} \bullet}$.

- (1) Hamming loss indicates the percentage of misclassified labels

$$\text{HL} = \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \|\mathbf{h}(\mathbf{D}_{i \bullet}) \Delta \mathbf{Y}_{i \bullet}\|_1, \quad (13)$$

where Δ is the symbol of symmetry difference

- (2) Ranking loss is the ratio of labels to the reverse order

$$\text{RL} = \frac{1}{n} \sum_{i=1}^n \frac{1}{1_m^T \mathbf{Y}_{i \bullet} \cdot 1_m^T \widetilde{\mathbf{Y}}_{i \bullet}} \sum_{\mathbf{q}: \mathbf{Y}_{i \bullet}^{\mathbf{q}}=1} \sum_{\mathbf{q}': \mathbf{Y}_{i \bullet}^{\mathbf{q}'}=0} (\mathbf{P}), \quad (14)$$

where $\mathbf{P} = \delta(\text{rank}_{i \bullet}(\mathbf{q}) \geq \text{rank}_{i \bullet}(\mathbf{q}'))$, $\delta(\mathbf{z})$ is indicator functions and $\widetilde{\mathbf{Y}}_{i \bullet}$ is the complement of $\mathbf{Y}_{i \bullet}$ on \mathbf{Y}

- (3) One-error indicates the sample proportion of the “most relevant predicted label” that does not exist in the “real label.”

$$\text{OE} = \frac{1}{n} \sum_{i=1}^n \delta(\mathbf{Y}_{i \bullet}^{\mathbf{l}} = 0), \quad (15)$$

where $\mathbf{l} = \arg \min_{\mathbf{q} \in [1, m]} \text{rank}_{i \bullet}(\mathbf{q})$

- (4) Coverage is the average number of moves that the “sorted labels” need to make to cover the real label correlation set

$$\text{CV} = \frac{1}{n} \sum_{i=1}^n \arg \max_{\mathbf{q}: \mathbf{Y}_{i \bullet}^{\mathbf{q}}=1} \text{rank}_{i \bullet}(\mathbf{q}) - 1 \quad (16)$$

- (5) Average precision refers to the label whose correlation is higher than that of a particular label

$$\text{AP} = \frac{1}{n} \sum_{i=1}^n \frac{1}{1_m^T \mathbf{Y}_{i \bullet} \cdot \mathbf{q}: \mathbf{Y}_{i \bullet}^{\mathbf{q}}=1} \sum_{\mathbf{q}': \mathbf{Y}_{i \bullet}^{\mathbf{q}'}=1} \frac{\mathbf{P}}{\text{rank}_{i \bullet}(\mathbf{q})} \quad (17)$$

Among the five evaluation indexes, except coverage, the range of the other indexes is $[0, 1]$. Among these five evaluation indexes, the larger the value of the average precision index

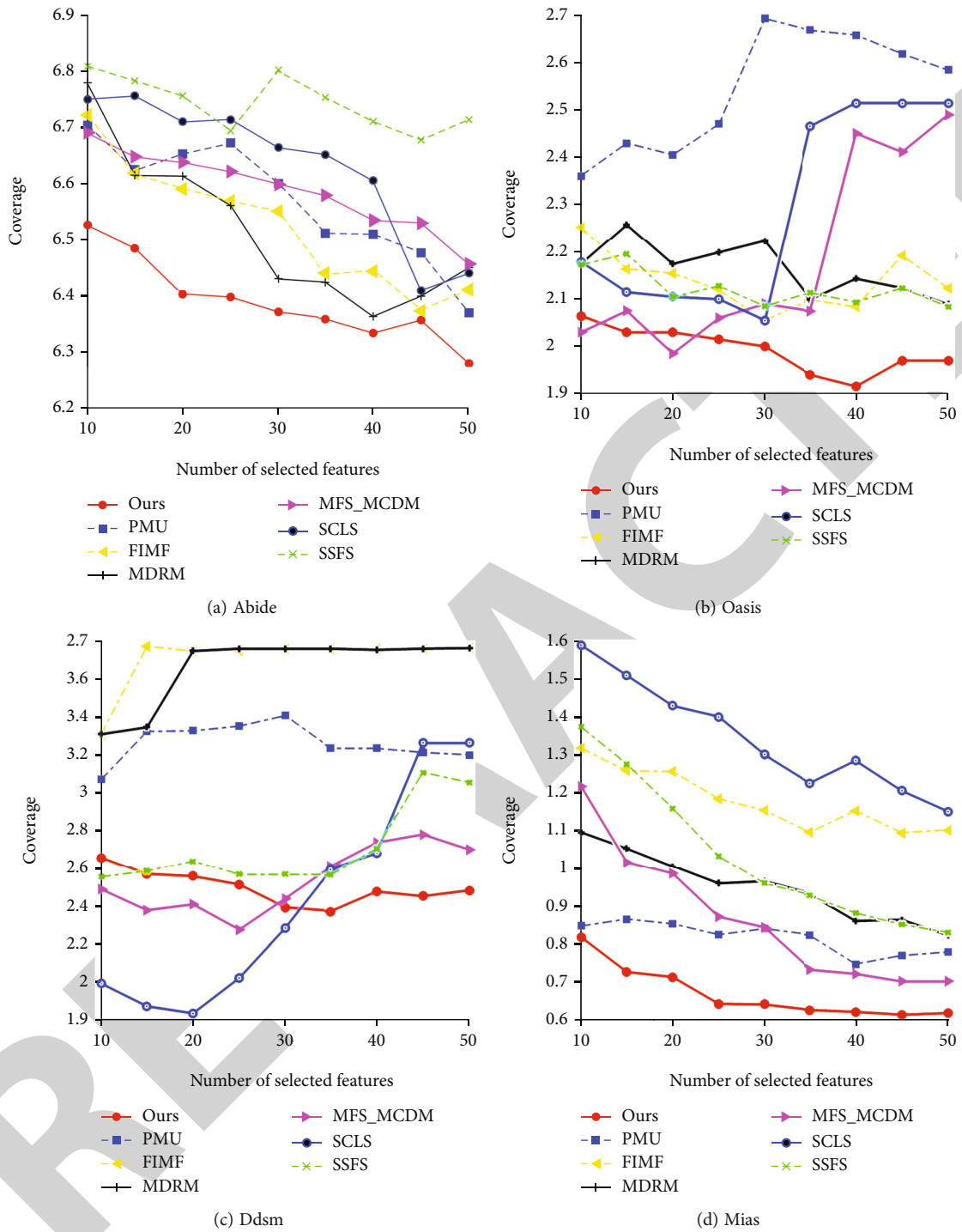


FIGURE 5: Continued.

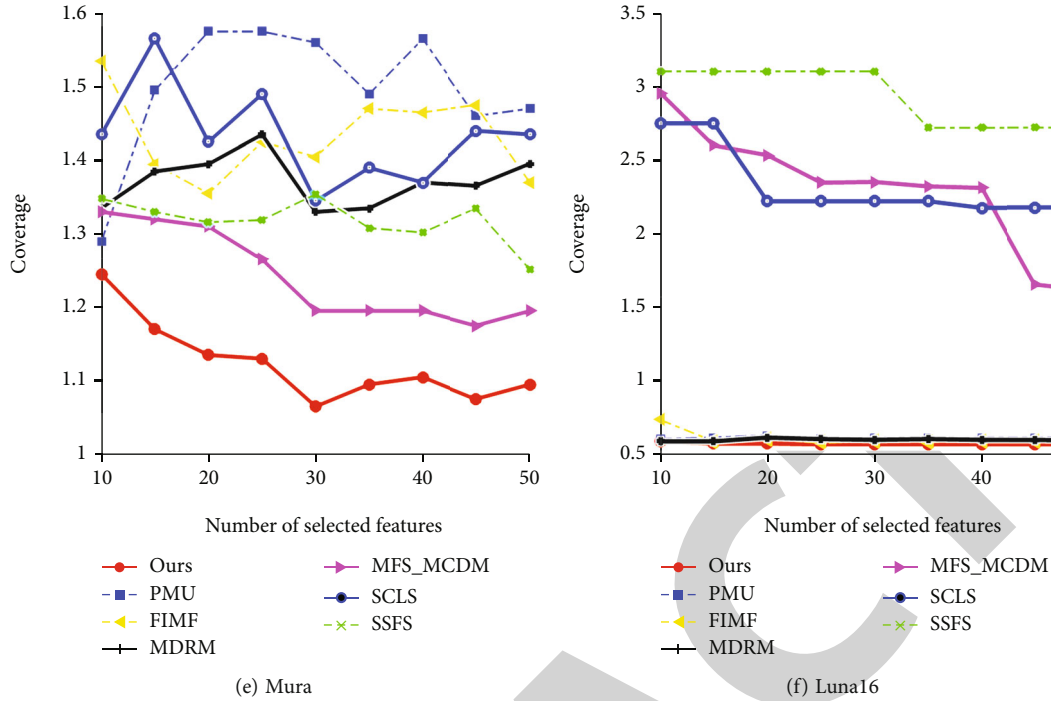


FIGURE 5: Coverage: the quality of each algorithm varies with the selected feature quantity.

is, the better the algorithm quality is, while the smaller the value of other indexes is, the better the algorithm quality is.

4.2. Analysis of Experimental Results. Tables 3–7 show the experimental comparison results of the suggested algorithm with six advanced multilabel feature selection algorithms, and the optimal results are shown in bold. The experiment was carried out under five commonly used multilabel feature selection evaluation indexes. Tables 3 and 5 show that the hamming loss and one-error index values of the suggested algorithm on each experimental dataset are optimal. In Tables 4, 6, and 7, although the index values of ranking loss, coverage, and average precision of the suggested algorithm on Ddsm datasets are slightly inferior to those of the MFS_MCDM algorithm, the index values on the other five experimental datasets are still optimal. Therefore, the overall quality of the suggested algorithm is better than that of the comparison algorithms.

In addition, in Tables 3–7, we can find that the quality of the suggested feature selection algorithm has been improved to some extent compared with that of the comparison algorithm under different data and indicators.

It can be calculated in Table 3 that, on Mias and Luna16 datasets, the quality of the feature selection algorithm suggested in this paper under the hamming loss index improves by a maximum of 35.48% and 95.33%, respectively, compared with the comparison algorithm. As can be seen in Table 4, on Mias and Mura datasets, the quality of the feature selection algorithm suggested in this paper under the ranking loss index is 48.5%–90.3% and 26.6%–98.3% higher than that of the comparison algorithm, respectively.

In addition, it can be calculated in Table 5 that on the Mura dataset, the quality of the suggested algorithm under

the one-error index is improved by 27%–87.5% compared with that of the comparison algorithm. In Table 6, the quality of the suggested algorithm under index coverage improves 43.8%–78.9% on dataset Mias. It can be calculated in Table 7 that, on Ddsm and Mura datasets, the quality of the feature selection algorithm suggested in this paper under the ranking loss index is 2.6%–26.7% and 2%–64.4% higher than that of the comparison algorithm, respectively.

In addition, to display the comparison of various algorithms in a more specific and intuitive way, the number of selected features is taken as the horizontal axis. The quality value under the corresponding index is taken as the vertical axis to show the change of the corresponding index value of the selected features in each algorithm within the range of [10, 15, 20, 25, 30, 35, 40, 45, 50].

Specifically, in Figures 2–6, we can intuitively observe the change of the optimal result of the comparison algorithm under the optimal parameter as the number of selected features increases and the quality comparison of each algorithm under the same dataset and the same indicator when the number of selected features is [10, 15, 20, 25, 30, 35, 40, 45, 50]. As can be seen in Figure 2, under the hamming loss index, the curves of the suggested algorithm on all datasets are below the curves of the comparison algorithm, thus indicating the superiority of the suggested algorithm in the hamming loss index. It can be seen in Figures 3–5 that the curve of the suggested algorithm in Abide, Oasis, Mias, and Mura datasets is obviously below the angle of the suggested algorithm. It shows the superiority of the suggested algorithm in ranking loss, one-error, and coverage index. It can be seen in Figure 6 that the curve of the suggested algorithm in Abide, Oasis, Mias, and Mura datasets is obviously above

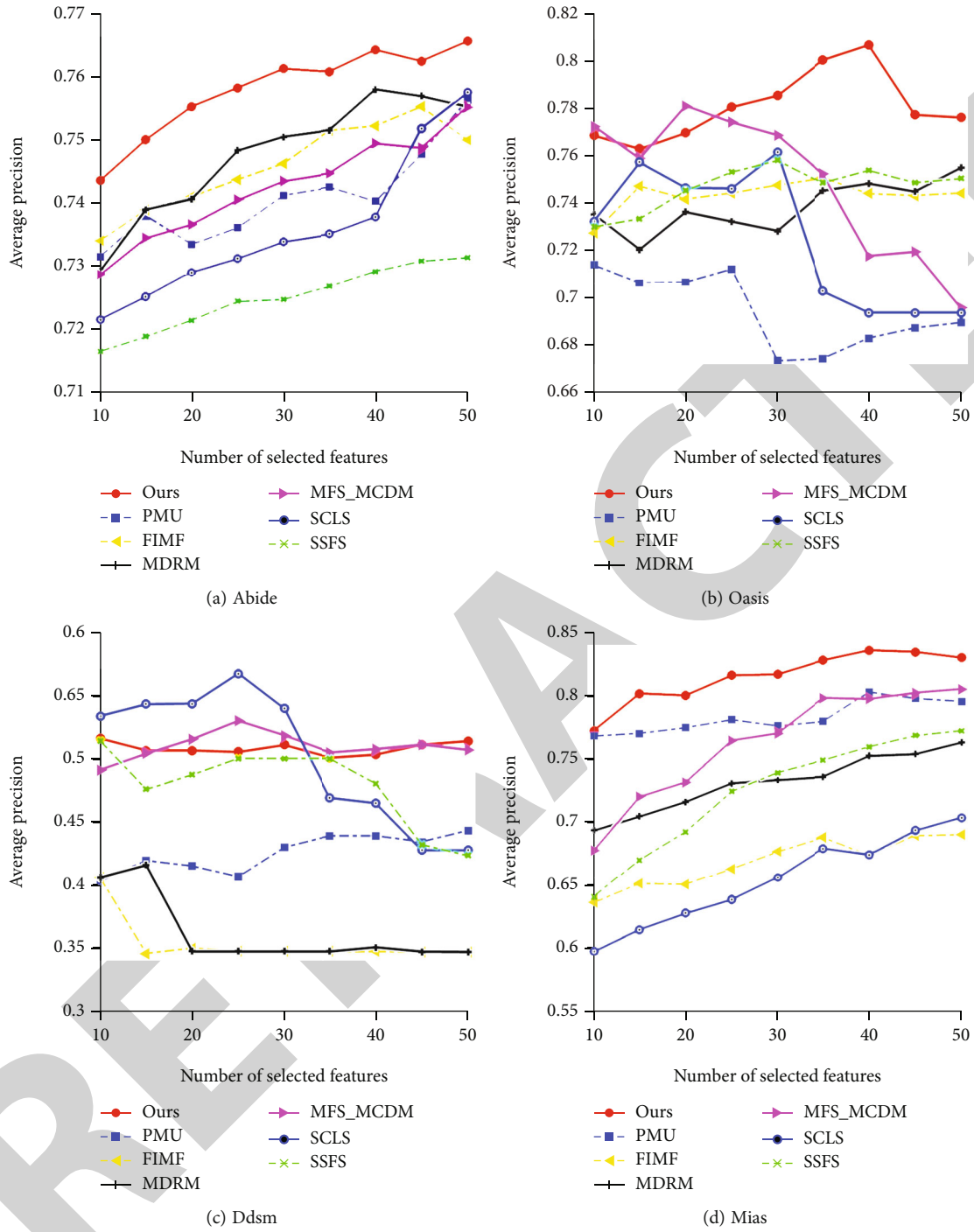


FIGURE 6: Continued.

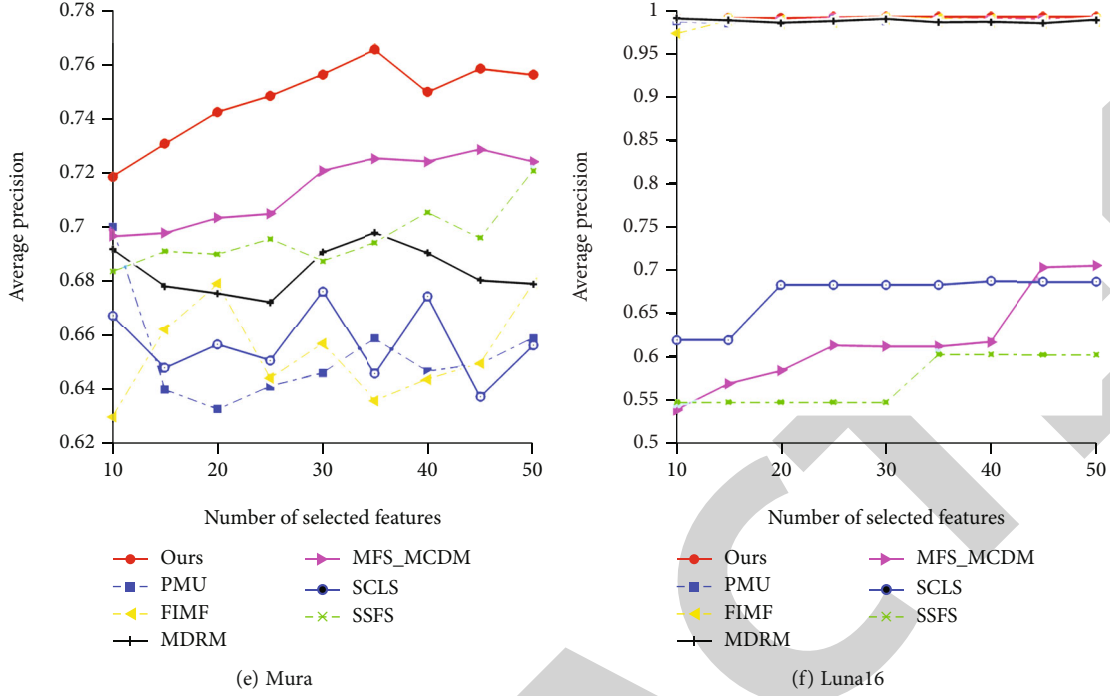


FIGURE 6: Average precision: the quality of each algorithm varies with the selected feature quantity.

the angle of the suggested algorithm, thus demonstrating the superiority of the suggested algorithm in the average precision index.

The multilabel feature selection algorithm suggested in this paper effectively deals with feature selection problems and has certain advantages over the comparison algorithm.

In addition, ranking and significant differences among experimental algorithms are shown in Figure 7 [23]. In Figure 7, the horizontal axis represents rankings, from left to right, ranking higher and higher, and a horizontal line connects algorithms with no significant difference. Specifically, it can be observed in Figure 7 that the ranking of the suggested algorithm under each index always remains the first, which indicates that the overall quality of the suggested algorithm is better than that of the comparison algorithm. Although the suggested algorithm has no significant difference with SCLS and MFS_MCDM algorithms under each index, it significantly differs from other comparison algorithms.

4.3. Performance Analysis of the Proposed Algorithm. To prove the feasibility of feature selection, a convergence experiment of the proposed algorithm was designed and carried out. In the experiment, regular term parameters $\alpha = 1$ and $\beta = 1$ were set and the algorithm iteration times were 50. The experimental results are shown in Figure 8:

In addition, we also analyze the time complexity of the suggested algorithm. As can be seen in the pseudocode in Algorithm 1, if the number of iterations is T , the time complexity of updating \mathbf{W} of the suggested algorithm is $\mathcal{O}(\mathbf{td}^2\mathbf{n})$; the time complexity of update \mathbf{D} is $\mathcal{O}(\mathbf{tdm})$; and the time complexity of updating \mathbf{H} is $\mathcal{O}(\mathbf{tdn})$. Therefore, the total time complexity of the suggested algorithm is $\mathcal{O}(\mathbf{td}^2\mathbf{n} + \mathbf{tdm} + \mathbf{tdn})$. As can be seen in Figure 8, the value of t is gener-

ally less than 10, so the time complexity of the suggested algorithm is greatly affected by the sample number n and feature number d .

Finally, we compared the running time of experimental algorithms on six experimental datasets and the results are shown in Table 8:

As shown in Table 8, although the running time of the proposed algorithm in Oasis, Mias, and Luna16 data is not as good as the MFS_MCDM algorithm, it is better than other algorithms. In addition, the running time of the proposed algorithm is also optimal on the whole.

4.4. Feasibility Analysis of the Suggested System

4.4.1. Technical Feasibility Analysis. First, under the rapid development of computer technology and science and technology, the number of computer operation talents is increasing, which can play a role of talent support for big data analysis techniques in the design of sports health management systems. Secondly, with the rapid development of computer technology in recent years, the continuous improvement of relevant hardware and software technology can lay an excellent technical guarantee for the data analysis link in the design of sports and health management systems, continuously improving the application value of unstructured data, and expand the path for users to obtain data information. Third, the rapid development of cloud storage, computing technology, data storage technology, and the IoT provide intelligent and diversified support for constructing a sports health management system.

They have targeted feasibility analysis. Sports health management systems' construction, innovation, and development are closely related to data analysis, feedback, and

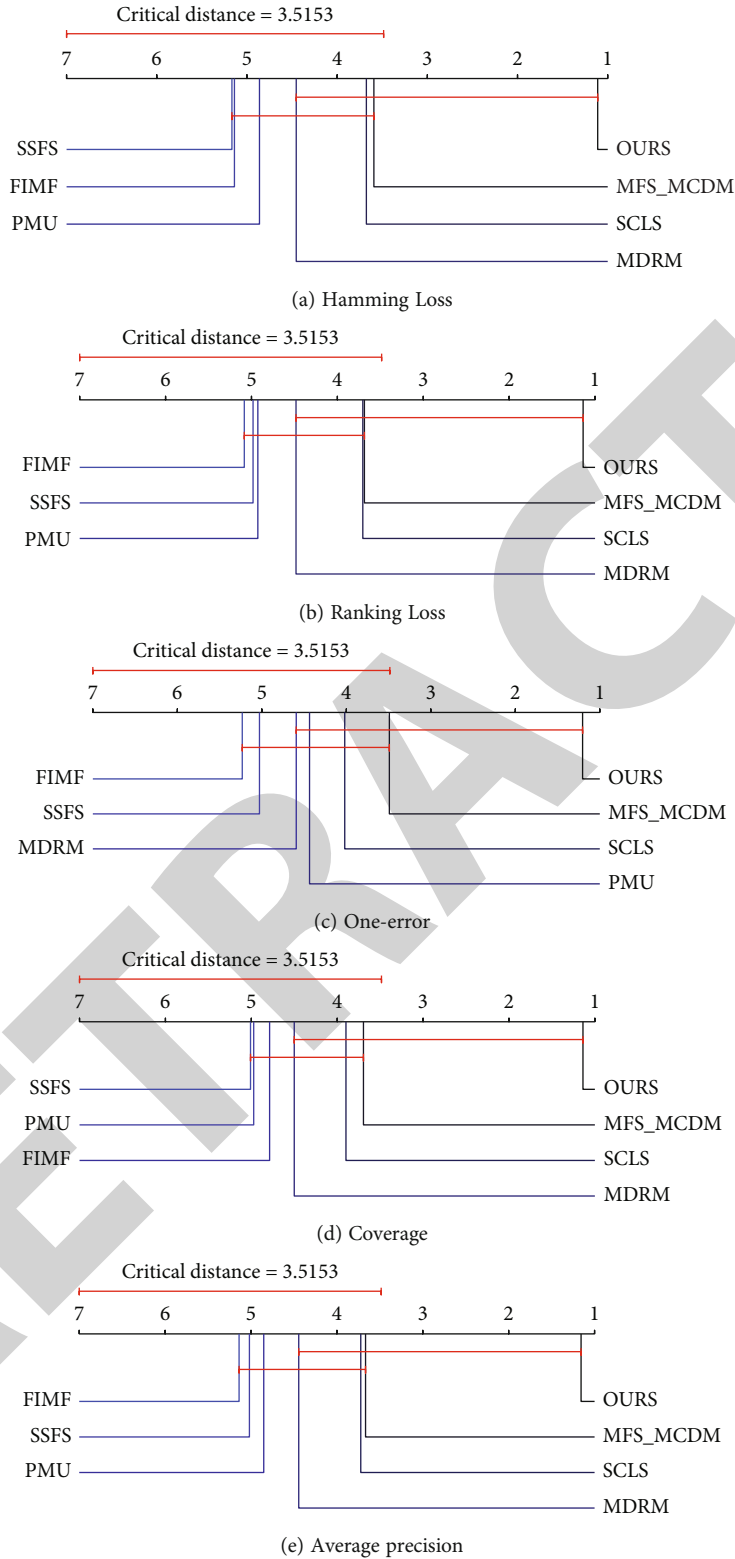


FIGURE 7: The form of the mean rank graph of the results of Bonferroni-Dunn test.

application. In the era of big data, data mining can play the role of economic evaluation. The application of big data analysis technology to analyze user health data and design fitness programs can comprehensively analyze the physical

health indicators of each user and collect and master the fitness situation of each user, as well as the fitness experience of relevant experts, to provide users with scientific and reasonable fitness programs and optimize the application of

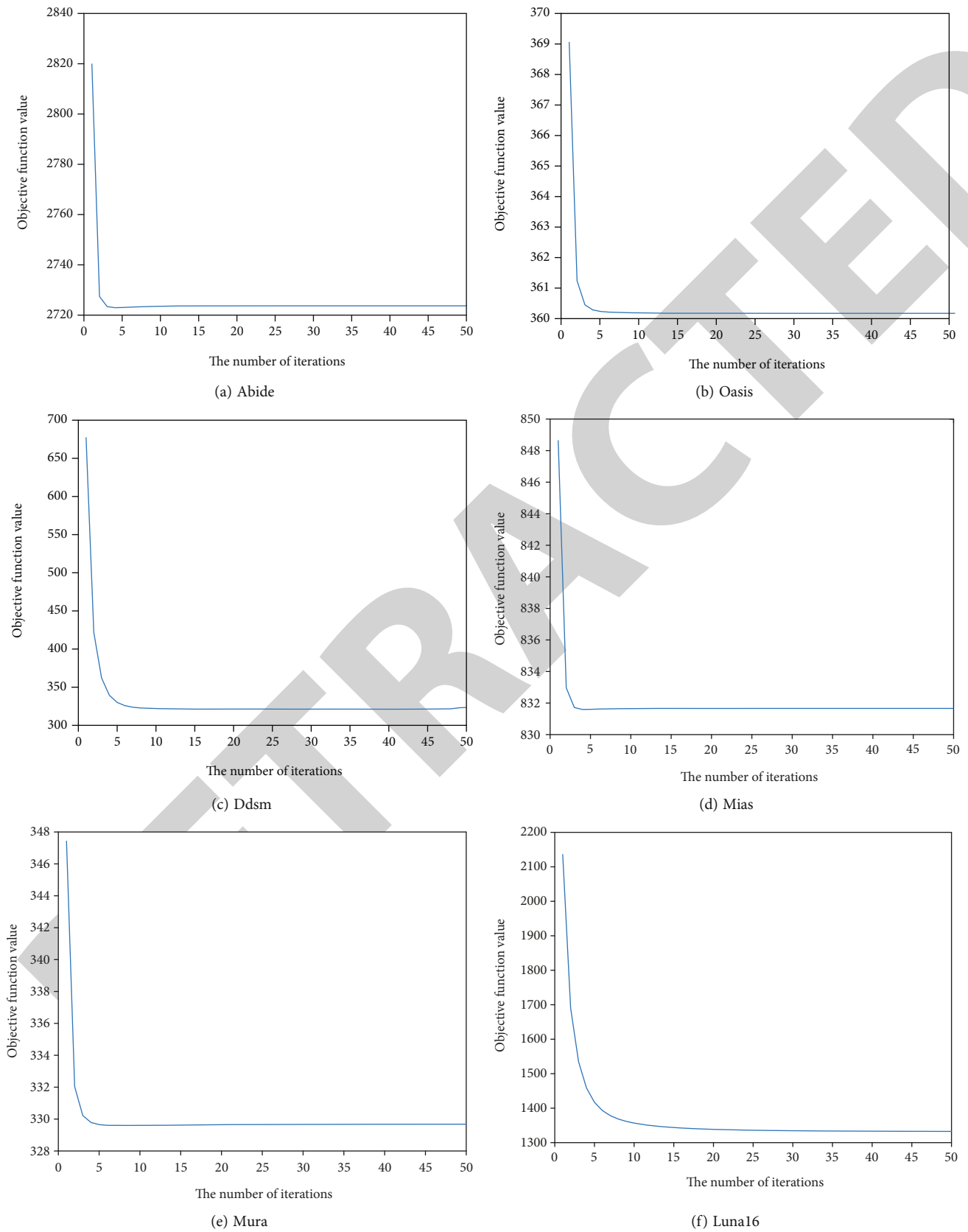


FIGURE 8: Convergence of the suggested algorithm on experimental data.

TABLE 8: Running time (sec) of the multilabel feature selection methods.

Algorithm	MFS_MCDM	SSFS	SCLS	MDMR	PMU	FIMF	Ours
Abide	0.0081	4.0996	0.5040	34.7042	131.9868	4.2689	0.0076
Oasis	0.0021	0.2829	0.0820	2.9552	9.1605	0.2407	0.1820
Ddsm	0.0212	0.6920	0.3622	35.5024	67.2589	5.6493	0.0020
Mias	0.0413	7.6726	0.9843	40.1615	161.0445	1.8438	0.1431
Mura	0.0292	0.9856	0.3069	12.4732	49.8773	0.5106	0.0193
Luna16	0.0010	0.0290	0.0201	0.2026	0.8025	0.0443	0.0204

various fitness methods. Data mining technology can integrate user health information, expert experience, medical index information, and the IoT information, play the role of data retrieval, application, and interaction, pay full attention to users' physical health, and provide targeted efficient scientific and reasonable fitness programs.

5. Conclusion

Through a lot of research and experiments, a new type of intelligent sports management system is proposed in this paper. The system was built by combining big data analysis with IoT technology. In the system, the relevant information is collected and pushed through the IoT technology, ensuring the reliability of collected information and the timeliness of pushed information. Information analysis is achieved through the multilabel feature selection algorithm proposed in this paper, which can better analyze and extract the main factors affecting user health and provide a solid basis for designing targeted fitness programs. In addition, we not only compared the proposed feature selection algorithm with SSFS, SCLS, and other algorithms but also carried out the feasibility analysis of the proposed intelligent sports health management system. The experimental results and system feasibility analysis show that the proposed intelligent sports health management system is feasible and superior. We will combine more big data analysis techniques to make the suggested system more targeted and effective 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 they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] C. C. Ding, M. Zhao, J. Lin, and J. Jiao, "Multi-objective iterative optimization algorithm based optimal wavelet filter selection for multi-fault diagnosis of rolling element bearings," *ISA Transactions*, vol. 82, pp. 199–215, 2019.
- [2] M. Labani, P. Moradi, F. Ahmadizar, and M. Jalili, "A novel multivariate filter method for feature selection in text classification problems," *Engineering Applications of Artificial Intelligence*, vol. 70, pp. 25–37, 2018.
- [3] C. Yao, Y. F. Liu, B. Jiang, J. Han, and J. Han, "LLE score: a new filter-based unsupervised feature selection method based on nonlinear manifold embedding and its application to image recognition," *IEEE Transactions on Image Processing*, vol. 26, no. 11, pp. 5257–5269, 2017.
- [4] J. Gonzalez, J. Ortega, M. Damas, P. Martín-Smith, and J. Q. Gan, "A new multi-objective wrapper method for feature selection - accuracy and stability analysis for BCI," *Neurocomputing*, vol. 333, pp. 407–418, 2019.
- [5] J. Swati, H. Hongmei, and J. Karl, "Information gain directed genetic algorithm wrapper feature selection for credit rating," *Applied Soft Computing*, vol. 69, pp. 541–553, 2018.
- [6] S. Maldonado and J. Lopez, "Dealing with high-dimensional class-imbalanced datasets: embedded feature selection for SVM classification," *Applied Soft Computing*, vol. 67, pp. 94–105, 2018.
- [7] Y. C. Kong and T. W. Yu, "A graph-embedded deep feedforward network for disease outcome classification and feature selection using gene expression data," *Bioinformatics*, vol. 34, no. 21, pp. 3727–3737, 2018.
- [8] Y. Zhang, Y. C. Ma, and X. F. Yang, "Multi-label feature selection based on logistic regression and manifold learning," *Applied Intelligence*, vol. 52, no. 8, pp. 9256–9273, 2022.
- [9] W. F. Gao, Y. H. Li, and L. Hu, "Multi-label feature selection with constrained latent structure shared term," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–10, 2021.
- [10] J. C. Hu, Y. H. Li, G. C. Xu, and W. Gao, "Dynamic subspace dual-graph regularized multi-label feature selection," *Neurocomputing*, vol. 467, pp. 184–196, 2022.
- [11] Y. Zhang and Y. C. Ma, "Non-negative multi-label feature selection with dynamic graph constraints," *Knowledge-Based Systems*, vol. 298, article 107294, 2022.
- [12] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MFS-MCDM: multi-label feature selection using multi-criteria decision making," *Knowledge-Based Systems*, vol. 206, article 106365, 2020.
- [13] J. Lee and D. W. Kim, "SCLS: multi-label feature selection based on scalable criterion for large label set," *Pattern Recognition*, vol. 66, pp. 342–352, 2017.
- [14] Y. Lin, Q. Hu, J. Liu, and J. Duan, "Multi-label feature selection based on max-dependency and min-redundancy," *Neurocomputing*, vol. 168, no. 30, pp. 92–103, 2015.
- [15] J. Lee and D. W. Kim, "Feature selection for multi-label classification using multivariate mutual information," *Pattern Recognition Letters*, vol. 34, no. 3, pp. 349–357, 2013.

Retraction

Retracted: Abnormal-Aware Multiperson Evaluation System with Improved Fuzzy Weighting

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Ni, "Abnormal-Aware Multiperson Evaluation System with Improved Fuzzy Weighting," *Journal of Function Spaces*, vol. 2022, Article ID 4899831, 8 pages, 2022.

Research Article

Abnormal-Aware Multiperson Evaluation System with Improved Fuzzy Weighting

Shutong Ni 

School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

Correspondence should be addressed to Shutong Ni; 20193088@aufe.edu.cn

Received 8 May 2022; Accepted 30 May 2022; Published 4 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Shutong Ni. 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.

There exists a phenomenon that subjectivity highly lies in the daily evaluation process. Our research primarily concentrates on a multiperson evaluation system with anomaly detection to minimize the possible inaccuracy that subjective assessment brings. We choose the two-stage screening method, which consists of rough screening and score-weighted Kendall- τ distance to winnow out abnormal data, coupled with hypothesis testing to narrow global discrepancy. Then we use fuzzy synthetic evaluation method (FSE) to determine the significance of scores given by reviewers as well as their reliability, culminating in a more impartial weight for each reviewer in the final conclusion. The results demonstrate a clear and comprehensive ranking instead of unilateral scores, and we get to have an efficiency in filtering out abnormal data as well as a reasonably objective weight determination mechanism. We can sense that through our study, people will have a chance of modifying a multiperson evaluation system to attain both equity and a relatively superior competitive atmosphere. A preprint has previously been published (Ni, 2022).

1. Introduction

The evaluation system has long been an indispensable part of measuring the performance of particular behaviour. For years, subjective evaluation and objective assessment have been rather separated in their respective fields. However, with the booming improvement in science and technology, these two indicators are somehow gradually intertwined and have the objective one taken the lead. Even so, subjective evaluation cannot be erased for good, on the account that it has its unique characteristics indeed, which can be generalized as minute scope, fair adaptability, low cost, and high randomness. When it comes to examinations or appraisals, expertise revision of contributions, personnel recruitment, project bidding, judgments on equipment's function or merchandise's quality, and even government's policymaking, subjective assessment operates in every tiny aspect of the society, paving the way for its unremitting upswing.

Practically, the two evaluation strategies have their leanings. To minimize the repercussion of their defects, under a particular circumstance, there are scholars wedded to incorporating the two assessment methods together, in the hope

that the results can be much fairer as well as more reliable [1–3]. Nevertheless, in many scenarios, objective evaluation data is hard to obtain, and many a strength of subjective evaluation make it a more practical means. A simple way to increase the credibility of the evaluation is to summarize the information of multiple reviewers, which will lead to the inconsistency of results so that a significant number of researchers address themselves into how to integrate various information to make the ultimate review authentic as much as possible. To avert a mixture of standards, experts should conduct an evaluation with respect to various indices in different situations. Xing [4] proposes correlation analysis [5] to measure the reliability of reviewers, screening out the discrepant values and, at the same time, streamlining the assessment indices that possess a strong correlation with each other. Regarding that different reviewers have different standards over objects, Xing [4] chooses to apply fuzzy analysis hierarchy process (FAHP) [6] to the problem, in the hope of a comparable and definite weight factor for each evaluation index. That is to say, by forming a fuzzy judgment matrix that consists of experts' appraisals, we can determine the weight factors in the assessment system.

Nevertheless, most works fail to consider the confidence of different reviewers, and they just take average to obtain the score of a specific assessment index for each individual. Moreover, those proposed assessment indices will also be affected both subjectively and professionally, and on a certain condition, we cannot even put forth a scientific index, for instance, when teachers in schools rate their students, the only factor worth referring to is the final score. Thus, our research will prioritize the information-limited problems and attempt to refine the disadvantages of subjective evaluation as below:

To start with, the evaluation standard differs from person to person, so there is no absolute right or wrong, and every authority has his/her own precept and preference. In other words, for a certain wide range of scoring standards, diverse experts have enormous differences in the understanding of the evaluation standards and the grasp of the assessment scales, which simultaneously give rise to a conspicuous contrast. The second lies in the psychological impact of each expert during the evaluation process. When scoring, the experts will inevitably be influenced more or less by the scores of other students he has given. That is to say, subsequent scores will be subject to all of the previous scoring results, which conduce to the essential variation.

According to the imperfections that exist in the subjective evaluation, our study intends to improve the multiperson subjective evaluation method through the lens of mathematical modeling, which is widely applied in various kinds of engineering and simulation problems [7–10]. We endeavour to provide resolutions in the design of assessment procedure as well as assessment approaches, making evaluation results and objective facts coordinate as much as possible. In this case, we may help stamp out the bias and constraints of individual evaluation and demonstrate impartiality as well as authority, so as to shape a superior competitive atmosphere.

In our research process, we use two-step screening as a quick start to examine the anomalous data. Given that the standard Q-test method [11] and $3\text{-}\sigma$ principle [12] fail to work well when the samples are inadequate, we can utilize these methods for a rough selection and then apply Kendall- τ distance to examine the data winnowed out for the sake of advancing the secondary screening procedure. In light of the drawbacks that original Kendall- τ distance can hardly fully contemplate the differences among values of scores, we rework the idea as score-weighted Kendall- τ distance and regard it as an objective function, screening out the abnormal data which will result in a minor decrease in the objective. Moreover, still in the data preprocessing stage, we propose to mitigate the discrepancy of scores given by the same reviewer between two classes through hypothesis testing. Later, considering the fuzzy relationships among reviewers' judging criteria, we interpret the outcomes from different experts as different judging indices of objects evaluated, so as to use the fuzzy synthetic evaluation model [13–16]. When weighing those experts' reviews, we not only take the weight stemmed from the fuzzy synthetic evaluation into consideration but also calculate another type of weight derived from the scale of anomalous data excluded, which

can tell the reliability of the evaluation. Coupling the two weights with each other, we take the average as the final weight for the index, thereby measuring the accuracy of evaluation more efficiently [17].

In the following sections, we first present the formulation of the entire problem. And then, we narrate the formation process of our evaluation method, including the two-stage screening method, hypothesis test, and fuzzy synthetic evaluation method. Subsequently, we conduct experiments and show the effectiveness of our methods. Furthermore, we ultimately draw a conclusion of the questions and elucidate the notion of our research and modification for a more equitable evaluation system.

2. Problem Formulation

In this section, we introduce standard problem formulation that we aim to tackle via our proposed model. Moreover, we also specify certain assumptions which contribute to the completeness of our definition.

Let g_{ij}^k denote the grade given by reviewer i to student j in class k , where $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n_k$, and $k = 1, 2, \dots, C$. Let π_i denote the ranking of all students, whose papers are reviewed by reviewer i , in descending order of their grades.

We mainly focus on two representative assessment situations. The first is that we are given one class with n students whose papers need to be graded by m reviewers. The second situation is that we are given k classes, where $k \geq 2$, and each of the m reviewers must grade all the classes. In both cases, the task is to estimate the actual grades of all the students and determine their final rankings as fair and objective as possible, without knowing about the evaluation principles or preferences of each reviewer. It is worth mentioning the following assumptions that help avoid certain intricate controversies:

- (1) Each reviewer grades papers under the same external conditions
- (2) All papers are kept secret before reviewing
- (3) Reviewers are not allowed to discuss with each other
- (4) Students' rankings are only determined by their final synthesized scores

3. The Proposed Evaluation System

3.1. Anomaly Detection. Since the only difference between the two situations is the number of classes, we decide to tackle this at the end of the analysis. For the screening of the abnormal data, the most common method is indubitably the trimmed mean, which strikes out a certain proportion of the highest and lowest scores and averages the remaining ones. However, in our research, we actually have only a small amount of reviewers for the evaluation task, where the trimmed method cannot operate to its full potential and has low robustness [18]. On the one hand, under the premise that students have only a few scores, removing high

scores or low scores can reduce extreme data to a certain extent, but it also causes a significant loss of data. In this case, for some superficially extreme scores, we decide to winnow them out instead of directly discarding them. This effectively considers that despite the extremity of the lateral comparison among scores given by different reviewers, this kind of extremity can be rather valuable in the whole rankings given by a specific reviewer. Taking this factor into consideration and then designing a rational objective function, we can reach out to a relatively optimal two-step screening method. On the other hand, each reviewer may generally rate students from different classes high or low due to various evaluation standards, and there remains a certain contrast among scoring intervals. Therefore, the scores given by the reviewers do not comply with the normal distribution of their true average score, so the horizontal comparison can be meaningless. As a consequence of that, if we substitute students' scores with their rankings, this new indicator can also make the two-step screening method well-performed.

3.1.1. Rough Screening. There are plenty of typical means to filter outliers, such as $3\text{-}\sigma$ principle, quantile method, and Q-test. However, in terms of our research, we notice that there are not enough samples, because each student have gained scores only from few reviewers, and at the same time, the degree of anomaly fails to reach the standards of the methods mentioned above. For the second situation, after we transform their original scores into rankings as discusses above, it will face the same trouble.

Aware of this problem, we alter our perspective from designing an efficient one-round screening method to a two-stage method. In the first step, we roughly calculate the average and variance of students' scores and then subsume data into a set \mathcal{S}_1 if its deviation from the average is greater than α times of its variance, that is,

$$\mathcal{S}_1 = \left\{ (j, i, g_{ij}) : |g_{ij} - \text{avg}_j| \geq \alpha \cdot \text{std}_j \right\}, \quad (1)$$

where avg_j and std_j are the average and standard deviation of student j 's scores, respectively. We call \mathcal{S}_1 the anomaly set after proceeding step i , and α is a hyperparameter required to be fine-tuned. The second step is much more pivotal and will be specified in the following section.

3.1.2. Second Screening via Score-Weighted Kendall- τ Distance. To conduct the second step of winnowing out abnormal data, we would like to introduce the definition of Kendall- τ distance [19] to you for a quick start.

Definition 1. We define the Kendall- τ distance between τ_1 and τ_2 as below:

$$\delta_K(\tau_1, \tau_2) = \sum_{u >_{\tau_1} v} I[\{v >_{\tau_2} u\}], \quad (2)$$

where $\{\cdot\}$ denotes certain events, $I[\cdot]$ denotes the indicative function, and $x >_{\tau} y$ denotes that student x is ranked higher than student y in rank τ .

The ranking of students can be directly obtained through their scores. According to Definition 1, we can sense that if two reviewers' reviews on a particular student differ a great deal, then the Kendall- τ distance will be relevantly larger.

Having a deeper insight into the problem, the given definition of Kendall- τ distance practically remains some drawbacks because it simply contains the order of two students but overlooks the specific difference in scores. However, in our research, the problem encountered has siccar grades statistics, so that we can make a modification and obtain a score-weighted Kendall- τ distance as below:

$$\delta_{SK}(\pi_k, \pi_l) = \sum_{u >_{\pi_k} v} \frac{(g_{ku} - g_{kv}) + (g_{lv} - g_{lu})}{2} \cdot I[\{v >_{\pi_l} u\}] \cdot I[\{u, v \in \pi_k\} \cap \{u, v \in \pi_l\}], \quad (3)$$

where π_k and π_l are rankings given by reviewers k and l , respectively, and $u \in \pi_k$ denotes that the score of student u from reviewer k has not been filtered out. We hope that the following objective function will become as small as possible after we winnow out the target data.

$$L(\pi_1, \dots, \pi_m) = \sum_{1 \leq k < l \leq m} \delta_{SK}(\pi_k, \pi_l). \quad (4)$$

If we, respectively, consider how much the objective function L will decline after removing a subset of the abnormal set \mathcal{S}_1 , it will become an exponential time complexity algorithm, which is pretty impractical. Hence, we apply a greedy method that only needs to figure out how much the objective function will decline when one abnormal score $g_{ij} \in \mathcal{S}_1$ is deleted. Those who cause significant decline are more likely to be eliminated. Here, we do not necessarily use formula (4) for calculation every single time but merely compute the decline value for each g_{ij} as below:

$$\begin{aligned} \Delta_{ij} &= \sum_{k \neq i} \left(\delta_{SK}(\pi_k, \pi_i) - \delta_{SK}(\pi_k \setminus \{g_{kj}\}, \pi_i \setminus \{g_{ij}\}) \right) \\ &= \sum_{k \neq i} \left(\sum_{j >_{\pi_k} v} \frac{(g_{kj} - g_{kv}) + (g_{iv} - g_{ij})}{2} \cdot I[\{v >_{\pi_i} j\}] \right. \\ &\quad \left. + \sum_{j >_{\pi_k} v} \frac{(g_{kj} - g_{kv}) + (g_{ij} - g_{iv})}{2} \cdot I[\{j >_{\pi_i} v\}] \right) \quad (5) \\ &= \sum_{k \neq i} \sum_v \frac{|g_{kj} - g_{kv}| + |g_{iv} - g_{ij}|}{2} \\ &\quad \cdot I[\{j >_{\pi_i} v, v >_{\pi_k} j\} \cup \{j <_{\pi_i} v, v <_{\pi_k} j\}]. \end{aligned}$$

After sorting $\Delta = \{\Delta_{ij} : g_{ij} \in \mathcal{S}_1\}$ in descending order, we select the largest $d\Delta_{ij}$ and find their corresponding $g_{ij} \in \mathcal{S}_1$ to form another abnormal set \mathcal{S}_2 , where the size of d can be heuristically tuned. Note that if an anomaly g_{ij} has already been blanked out, theoretically, the corresponding

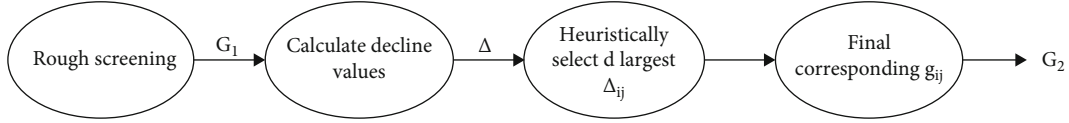


FIGURE 1: Architecture of two-stage screening method.

decline values of the other anomalies should be recalculated with respect to a smaller scores set. To put it more clearly, if a $g_{i_0j_0}$ is moved away, terms related to reviewer i_0 , or say π_{i_0} , in equation (3) will become zero due to the existence of the second indicative function $I[\{u, v \in \pi_k\} \cap \{u, v \in \pi_l\}]$. However, if we denote the set recalculated decline values as $\Delta' = \{\Delta'_{ij} : g_{ij} \in \mathcal{G}_1 \setminus \{g_{i_0j_0}\}\}$, we can easily verify that if $\Delta_{i_1j_1} \leq \Delta_{i_2j_2}$, then $\Delta'_{i_1j_1} \leq \Delta'_{i_2j_2}$. Therefore, getting rid of the anomalies greedily merely with respect to Δ is reasonable. We present a concise architecture in Figure 1 to summarize the two-stage screening.

3.2. Fuzzy Synthetic Evaluation. Based on the problem analysis in Part 2, after the screening of abnormal data, we need to synthesize all the information to determine the final score of each student. Inspired by the methods in [20, 21], we cleverly adapt the fuzzy synthetic evaluation method to our less informatics problem setting, treating each reviewer as an evaluation index. We have the observation matrix $G = (g_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$, where m is the number of reviewers and n is the number of students. Since the scores given by each reviewer can be understood as a benefit indicator, we can establish a fuzzy benefit matrix $B = (b_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$, where

$$b_{ij} = \frac{g_{ij} - \min_j g_{ij}}{\max_j g_{ij} - \min_j g_{ij}}. \quad (6)$$

It should be noted that, in accordance with Section 4.1, some g_{ij} has been screened out. We do not fill in the blanks for those missing values but choose to ignore them, which means that if g_{ij} is missing, b_{ij} is also missing in matrix B .

Coming up then, we need to establish the weight w_i , $1 \leq i \leq m$ of each reviewer. Firstly, the coefficient of variation is adopted to fully consider the influence of the size of evaluation intervals on the degree of differentiation of students. The coefficient of variation corresponding to each reviewer is calculated according to formula (7), aiming at fully considering the unknown influence of the size of evaluation intervals on ranking:

$$v_i = \frac{\sigma_i}{\bar{g}_i}, i = 1, \dots, m, \quad (7)$$

where

$$\bar{g}_i = \frac{1}{n} \sum_{j=1}^n g_{ij}, \quad (8)$$

$$\sigma_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (g_{ij} - \bar{g}_i)^2}, i = 1, 2, \dots, n. \quad (9)$$

Note that when calculating σ_i and \bar{g}_i , the sample size is taken as the number of scores given by the evaluation reviewer i , excluding the screening ones. Then, v_i is normalized to obtain the first component of the weight:

$$w_i^{(1)} = \frac{v_i}{\sum_{i=1}^m v_i}, i = 1, \dots, m. \quad (10)$$

Secondly, we consider the credibility of each reviewer and use n_i , the number of reviewer i 's screened scores, to obtain another part of the weight:

$$w_i^{(2)} = \frac{n - n_i}{\sum_{i=1}^m (n - n_i)}, i = 1, \dots, m. \quad (11)$$

Finally, the weight of each reviewer is determined by the following formula:

$$w_i = \frac{w_i^{(1)} + w_i^{(2)}}{2}, i = 1, \dots, m. \quad (12)$$

Then, $F_j = \sum_{i=1}^m w_i \cdot b_{ij}$ is calculated for student j , and their ranking can be obtained after sorting $\{F_j\}$.

Actually, the ultimate goal of our calculations is to determine the ranking of students. Now that the ranking has been obtained, if we want to output the final score, we only need to select a reference value, for example, $\min_j g_{1j}$, and further plus $F_j (\max_j g_{1j} - \min_j g_{1j})$. Then, the final score f_j of student j is

$$f_j = F_j (\max_j g_{1j} - \min_j g_{1j}) + \min_j g_{1j}. \quad (13)$$

We also provide a clear flow chart of how to utilize FSE in our problem as shown in Figure 2.

3.3. Normal Hypothesis Tests. For the second situation, we propose to perform the normal hypothesis test before screening abnormal data. For simplicity, we only consider the case with two classes. Since we can give feasible solution, it can be easily generalized to multiclass scenarios by fixing one target class and conducting hypothesis test between the target and other classes one by one. Specifically, we test the mean and variance of the grades given by one reviewer to two classes. We assume that the grades given by reviewer i to class k follow the normal distribution $N(\mu_i^k, \sigma_i^k)$, $i = 1, \dots, m, k = 1, 2$; we need to test the following two questions:

$$(1) H_0 : \mu_i^2 - \mu_i^1 = 0 \leftrightarrow H_1 : \mu_i^2 - \mu_i^1 \neq 0$$

$$(2) H_0 : \sigma_i^2 / \sigma_i^1 = 1 \leftrightarrow H_1 : \sigma_i^2 / \sigma_i^1 \neq 1$$

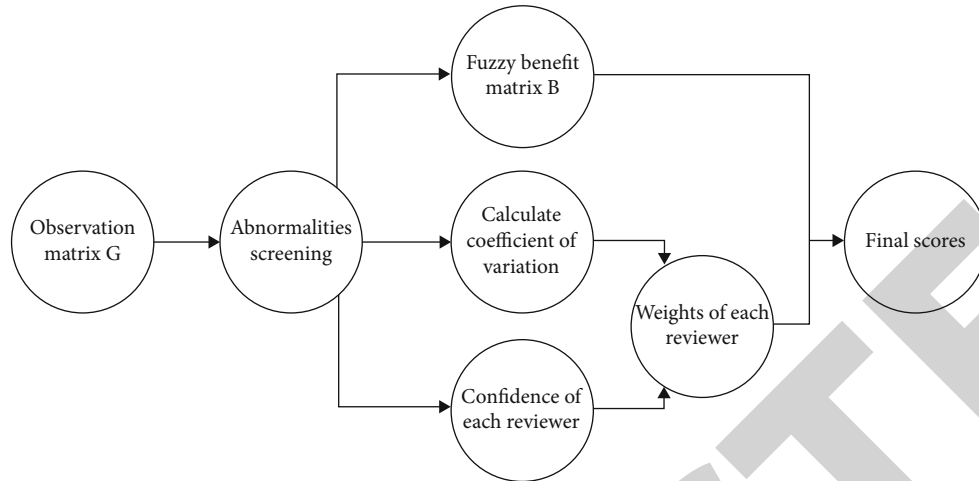


FIGURE 2: Flow chart of calculating students' final scores via fuzzy synthetic evaluation.

Set the confidence level to α . Note that μ_i^k and σ_i^k are unknown, and here, we have large samples. Therefore, we apply t -test to the first question and F -test to the second, and the rejection region of the test for question (a) and question (b) are, respectively, as below:

$$D_1 = \left\{ (g_{i1}^1, \dots, g_{im_1}^1, g_{i1}^2, \dots, g_{im_1}^2) : \left| \frac{G_i^2 - G_i^1}{S_i^w} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \right| > t_{n_1 + n_2 - 2} \left(\frac{\alpha}{2} \right) \right\},$$

$$D_2 = \left\{ (g_{i1}^1, \dots, g_{im_1}^1, g_{i1}^2, \dots, g_{im_1}^2) : \frac{(S_i^2)^2}{(S_i^1)^2} > F_{n_2 - 1, n_1 - 1} \left(1 - \frac{\alpha}{2} \right) \text{ or } \frac{(S_i^2)^2}{(S_i^1)^2} > F_{n_2 - 1, n_1 - 1} \left(\frac{\alpha}{2} \right) \right\},$$

(14)

where

$$G_i^k = \frac{1}{n_k} \sum_{j=1}^{n_k} g_{ij}^k,$$

$$(S_i^w)^2 = \frac{1}{n_1 + n_2 - 2} [(n_1 - 1)(S_1)^2 + (n_2 - 1)(S_2)^2],$$

$$(S_i^k)^2 = \frac{1}{n_k - 1} \sum_{j=1}^{n_k} (g_{ij}^k - G_i^k)^2, \quad i = 1, \dots, 5, k = 1, 2.$$

(15)

All the statistics mentioned above can be obtained from the existing data. Based on the results of the two tests as well as the practical meaning of confidence rate, we scale the scores differently:

- (1) If both of the test results are rejection, we apply the following transformation to the scores of class 2 given by the reviewer i :

$$g_{ij}^2 = (1 - \alpha) \left[\frac{S_i^1}{S_i^2} (g_{ij}^2 - G_i^2) + G_i^1 \right] + \alpha g_{ij}^2 \quad (16)$$

- (2) If the result of question (a) is rejection and that of question (b) is acceptance, we apply the following transformation to the scores of class 2 given by the reviewer i :

$$g_{ij}^2 = (1 - \alpha) (g_{ij}^2 - G_i^2 + G_i^1) + \alpha g_{ij}^2 \quad (17)$$

- (3) If the result of question (b) is rejection and that of question (a) is acceptance, we apply the following transformation to the scores of class 2 given by the reviewer i :

$$g_{ij}^2 = (1 - \alpha) \left[\frac{S_i^1}{S_i^2} (g_{ij}^2 - G_i^2) + G_i^1 \right] + \alpha g_{ij}^2 \quad (18)$$

- (4) If both of the test results are acceptance, we keep them unchanged

After rescaling, we can merge the two classes' grades into one table and then conduct abnormal data screening and fuzzy synthetic evaluation.

4. Experiments

In this section, we conduct two experiments that testify to the effectiveness of our methods. We will specifically describe the dataset we use and present the results in detail.

4.1. *Settings.* For each aforementioned situation in Section 2, we collect two corresponding datasets for evaluating the

TABLE 1: Statistics for the two datasets.

Statistics	Dataset I	Dataset II
Reviewer number (R)	3	5
Class number (C)	1	2
Student number (S)	107	61,50

effect of our proposed method. Table 1 summarizes some critical statistics of the datasets.

Notably, all the scores are rated following the percentage system, ranging from 0 to 100, preventing the uncertainties introduced from other level ranking systems. Apart from that, we do not have any knowledge of the reviewers' criteria for judging.

4.2. Results

4.2.1. Experiment I. In the first scenario, the intervals of the grades given by the three reviewers are 65-95, 62-99, and 60-98, respectively, which are almost identical. Therefore, there is no need to use the student's ranking instead of grades to filter out abnormal data.

We first apply the method proposed in Section 3.1.1 to screen out 33 anomalous scores. After that, we calculate the quantity for each anomalous score based on equation (5) and sort the results in decreasing order as below: $\Delta = \{(91, 3, 2667.0), (57, 3, 1811.5), (6, 3, 1503.5), (53, 3, 1447.5), (7, 3, 1444.5), (62, 3, 1429.5), (60, 3, 1384.5), (94, 1, 1361.0), (8, 2, 1121.5), (68, 3, 1120.0), (74, 3, 1104.0), (82, 2, 1017.0), (90, 2, 952.0), (102, 1, 871.0), (71, 3, 831.0), (48, 3, 816.0), (69, 3, 764.5), (27, 2, 757.5), (20, 3, 731.0), (19, 1, 726.0), (25, 3, 715.0), (47, 3, 696.0), (86, 1, 658.5), (33, 3, 648.0), (73, 3, 645.5), (89, 2, 642.5), (58, 3, 625.0), (51, 3, 624.0), (84, 3, 571.5), (17, 2, 548.5), (2, 1, 452.0), (78, 1, 334.0), (97, 2, 128.0)\}$. " (j, i, x) " denotes $\Delta_{ij} = x$.

We set the confidence to 0.80; that is, select the first $33 \times 0.80\% \approx 26$ elements in collection $\{\Delta_{ij}\}$ and take their corresponding g_{ij} as the final anomalous scores: $G_2 = \{(91, 3, 98), (57, 3, 65), (6, 3, 60), (53, 3, 68), (7, 3, 67), (62, 3, 67), (60, 3, 63), (94, 1, 95), (8, 2, 85), (68, 3, 70), (74, 3, 63), (82, 2, 75), (90, 2, 85), (102, 1, 85), (71, 3, 74), (48, 3, 70), (69, 3, 71), (27, 2, 69), (20, 3, 72), (19, 1, 88), (25, 3, 62), (47, 3, 74), (86, 1, 90), (33, 3, 61), (73, 3, 73), (89, 2, 95)\}$. " (j, i, x) " denotes $g_{ij} = x$.

The next step is to proceed with the fuzzy synthetic evaluation. The observation matrix can be directly obtained from the dataset, and simple calculation leads to the fuzzy benefit matrix. When these are all done, we can calculate the two types of weights:

$$\begin{aligned} w^{(1)} &= (0.325675, 0.377674, 0.296651), \\ w^{(2)} &= (0.349153, 0.345763, 0.305085), \\ w &= (0.337414, 0.361718, 0.300868). \end{aligned} \quad (19)$$

Finally, we calculate the final score using formula (13), and the quicksort algorithm [22] can be harnessed to

accelerate ranking. Because the weights are all decimals, we assume the final score should be kept to two decimal places. Students with the same score are regarded as the same ranking (for the specific final score table, please refer to Supplementary 1).

4.2.2. Experiment II. There are apparent differences between the evaluation intervals of the five teachers in the second dataset. As mentioned above, we propose to use the students' initial ranking as the score to screen out outliers and then conduct fuzzy synthetic evaluation and other operations. However, the greatest problem here is to mitigate the deviation of evaluation intervals. Thus, we are supposed to first implement the method in Section 3.3. After rescaling, we can fairly compare the grades of different classes; thus, fuzzy synthetic evaluation can be conducted to get final scores and rankings. For the sake of emphasis, we do not present the results of the screening operation. The final scores are presented in Supplementary 2.

5. Discussions

In this section, we reflect on the strengths and weaknesses of our methods in detail.

- (1) Based on the experiment results in Section 4.2.1, we can see that the two-stage method shows great performance in screening out the abnormal scores. The scores we selected are abnormal at first glance, matching the criterion of human's intuitive judgement. Moreover, their corresponding decline values are rigorously verified to be large, which goes beyond human intuition and is more trustworthy
- (2) Based on the experiment results in Section 4.2.2, we can see that the hypothesis test is suitable for multi-class situations. The distribution of rating is similar to the normal distribution in the real world. The renormalization maintains the scores of different classes in similar level, which is revealed through the rational distributed top students in two classes
- (3) We can design concise and efficient algorithms according to the two flow charts we presented in Sections 3 and 4, which demonstrate the feasibility of applying our methods on larger datasets
- (4) There are some hyperparameters existing in our methods. We mainly tuned them via heuristic methods, but it contains certain biases. It is better to design a mechanism that can automatically fine-tune the parameters
- (5) When performing the screening of abnormal data, slightly different approaches are used in the two scenarios, respectively, but these methods cannot wholly override all possible abnormalities. Supposing that multiple criterion are used to eliminate the outliers in data for a specific problem and some data appear to be an exception in all methods, we can

more confidently determine the data as abnormal ones

- (6) The fuzzy synthetic evaluation (FSE) method is a comprehensive evaluation method based on the problems of fuzziness and uncertainty in the evaluation criteria, evaluation factors, and the problem of difficulty in quantifying qualitative indicators. It can express a fuzzy object with a precise number so that the evaluation of fuzzy events is scientific and reasonable. Nevertheless, the application of the FSE method usually brings about relatively high subjectivity in the determination of indicators, fuzzy relation matrix, weight, etc. Besides, there is no clear or systematic method for determining the membership function, which conspires to a specific difference in results

6. Conclusions

The evaluation problem we concentrate on in our study is based on the fact that for schooling and appraisals, there exist a certain range of situations where no uniform standard is contained. In those cases, assessment results appear to be entirely subjective and divergent. It turns out that the evaluation problem has an inextricable connection with people's daily lives, and this is why our research is intended to center on this question and make an expansion. Our research proposes a modified score-weighted Kendall- τ distance as the judging criterion, adopts FSE and normal hypothesis test to be the principle investigating methods, and uses Python and MATLAB as auxiliary tools for implementation and testing. Under the auspices of fundamental scientific materials, we ultimately get to winnow out anomalies and then synthesize different factors for a comprehensive evaluation, culminating in a relatively equitable judging system. We believe that our work can give inspiration to improve the evaluation system under certain situations lacking in objective criteria. Besides, the two-stage screening method can be further extended to multistage version by properly examining other important characteristics of the reviews, which is left as a future work.

Data Availability

All the data used in this article are available at <https://github.com/Ciao-Yvette/Multi-person-Evaluation-System>.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

Upon the completion of this article, I would like to take this great opportunity to express my gratitude to Professor Guiyuan Yang for offering this interesting topic. At the same time, I would also like to sincerely acknowledge Professor Zhenmu Hong, who has provided meaningful and constructive suggestions.

Supplementary Materials

Table 1: the final scores and rankings of all the students in both classes in the first situation. Table 2: the final scores and rankings of the students in class 1. Rankings are calculated among all the students in both classes. Table 3: the final scores and rankings of the students in class 2. Rankings are calculated among all the students in both classes. (*Supplementary Materials*)

References

- [1] J. L. Hossain, L. W. Reinish, R. J. Heslegrave et al., "Subjective and objective evaluation of sleep and performance in daytime versus nighttime sleep in extended-hours shift-workers at an underground mine," *Journal of Occupational and Environmental Medicine*, vol. 46, no. 3, pp. 212–226, 2004.
- [2] C. A. Gumussoy, A. Pekpazar, M. Esengun, A. E. Bayraktaroglu, and G. Ince, "Usability evaluation of TV interfaces: subjective evaluation vs. objective evaluation," *International Journal of Human-Computer Interaction*, vol. 38, no. 7, pp. 661–679, 2021.
- [3] D. A. Crolla, D. C. Chen, J. P. Whitehead, and C. J. Alstead, "Vehicle handling assessment using a combined subjective-objective approach," *SAE Transactions*, vol. 107, pp. 386–395, 1998, <http://www.jstor.org/stable/44740965>.
- [4] R. Xing, *Research on the Methods of Subjective Evaluation for Handling and Stability of a Car* [PhD thesis], Jilin University, 2010.
- [5] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *Journal of Chiropractic Medicine*, vol. 15, no. 2, pp. 155–163, 2016.
- [6] E. E. Haji, A. Azmani, and M. E. Harzli, "Using the FAHP method in the educational and vocational guidance," *International Journal of Modern Education & Computer Science*, vol. 10, no. 12, 2018.
- [7] J. B. Liu and X. F. Pan, "Minimizing Kirchhoff index among graphs with a given vertex bipartiteness," *Applied Mathematics and Computation*, vol. 291, pp. 84–88, 2016.
- [8] J. B. Liu, J. Zhao, J. Min, and J. D. Cao, "The Hosoya index of graphs formed by a fractal graph," *Fractals*, vol. 27, no. 8, p. 1950135, 2019.
- [9] J. B. Liu, J. Zhao, H. He, and Z. Shao, "Valency-based topological descriptors and structural property of the generalized Sierpiński networks," *Journal of Statistical Physics*, vol. 177, no. 6, pp. 1131–1147, 2019.
- [10] J. B. Liu, Y. Bao, W. T. Zheng, and S. Hayat, "Network coherence analysis on a family of nested weighted n-polygon networks," *Fractals*, vol. 29, no. 8, p. 276, 2021.
- [11] R. B. Dean and W. J. Dixon, "Simplified statistics for small numbers of observations," *Analytical Chemistry*, vol. 23, no. 4, pp. 636–638, 1951.
- [12] F. Pukelsheim, "The three sigma rule," *The American Statistician*, vol. 48, no. 2, pp. 88–91, 1994.
- [13] Z. Zou, Y. Yi, and J. Sun, "Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment," *Journal of Environmental Sciences*, vol. 18, no. 5, pp. 1020–1023, 2006.
- [14] N. Chang, H. Chen, and S. Ning, "Identification of river water quality using the fuzzy synthetic evaluation approach," *Journal*

Retraction

Retracted: Numerical Simulation of Particle Deposition in Arterial Bifurcation via Lattice Boltzmann Method

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Yi, "Numerical Simulation of Particle Deposition in Arterial Bifurcation via Lattice Boltzmann Method," *Journal of Function Spaces*, vol. 2022, Article ID 3873484, 6 pages, 2022.

Research Article

Numerical Simulation of Particle Deposition in Arterial Bifurcation via Lattice Boltzmann Method

Houhui Yi 

School of Intelligent Manufacturing, Weifang University of Science and Technology, Weifang 262700, China

Correspondence should be addressed to Houhui Yi; yihouhui@wfust.edu.cn

Received 19 May 2022; Revised 13 June 2022; Accepted 15 June 2022; Published 4 July 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Houhui Yi. 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.

Arterial bifurcation plays a key role in cardiovascular system, so studying the characteristic of the blood flows near bifurcated arteries is of great importance in hemodynamics. The lattice Boltzmann (LB) model is used to observe the behavior of the particle deposition near the bifurcated artery. The mechanical quantity, including particle trajectories, velocities, and angular velocities, is studied numerically by LB simulations. The particle is prone to stasis as it is close to the wall of bifurcated vessel for small flow Reynolds number. Larger branch angle leads to higher possibility of particle stagnation. The numerical results are consistent with the clinical observation. The study provides a basis for understanding the mechanism of hemodynamics near bifurcations and will provide a research basis for clinical diagnosis and treatment of patients with atherosclerosis.

1. Introduction

Arterial branching is a major feature in the cardiovascular system. The most common branching pattern is that one stream of blood divides into two independent streams. As a stream-dividing unit, arterial bifurcation plays a key role in cardiovascular system. Atherosclerosis refers to the deposition of lipids (cholesterol, cholesterol esters, and phospholipids) in the intima of the artery, accompanied by the proliferation of smooth muscle cells and fiber components, and gradually develops into localized plaques. Clinical and autopsy studies show that arterial disease such as atherosclerosis has high selectivity of the lesion in complex flow areas such as bifurcation, confluence, and bend in the coronary artery, abdominal aorta, femoral artery, and carotid artery. Investigating the characteristic of the blood flows near bifurcated arteries is of great importance in hemodynamics and may provide a better understanding of the relation between arterial disease and structure of arterial bifurcations. Due to the complexity and limitations of blood vessels and the human body, numerical simulation is considered to be an effective method to measure blood circulations.

As a branch of fluid mechanics, Computational Fluid Dynamic (CFD) is the use of numerical methods to solve hydrodynamic problems. It adopts numerical methods to solve the governing equations of fluid dynamics and predict the laws of fluid motion. Especially in complex situations, measurement is often difficult or even impossible, but CFD can easily provide detailed information about flow fields. In recent years, it has been widely used in chemical, metallurgy, architecture, environment, and other related fields.

Since the proposal of the lattice Boltzmann method (LBM) [1], it has been reckoned as an effective method in CFD. The LBM has the advantages of simple algorithm, easy processing of complex boundary, and suitable for parallel computing. Recently, the LBM has been extensively applied in theoretical research and engineering fields and has been successfully used to implement blood flow simulations. The research contents mainly include transient blood flow in arteries with an artificial heart valve [2], blood flow in cerebral aneurysms [3], multicomponent blood flows [4], leukocyte rolling in virtual blood vessel [5], vesicle shape changes [6], and blood membrane dynamics [7]. The lattice Boltzmann biviscosity model was also further proposed to study blood flow problems [8]. Recently, the LBM was developed

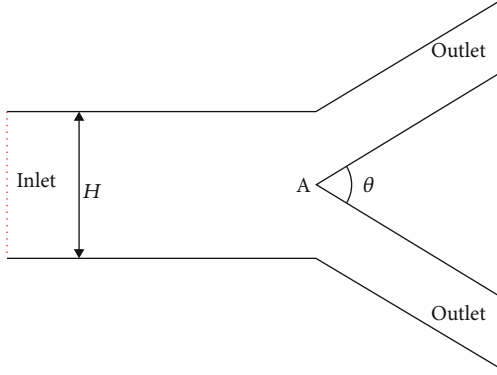


FIGURE 1: Schematic diagram of an arterial bifurcation in two dimensions.

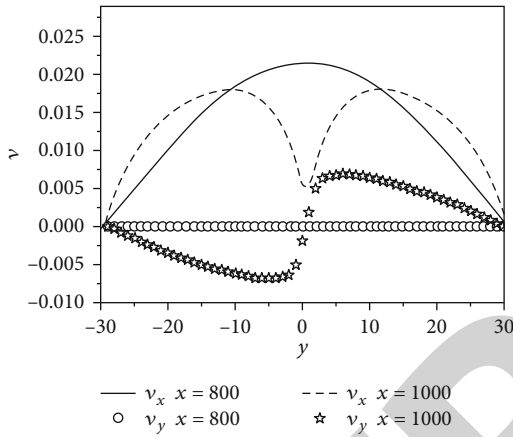


FIGURE 2: The velocities at position $x = 800$ (near straight vessel) and $x = 1000$ (near bifurcation vessel), respectively, in the case of without particle in the artery.

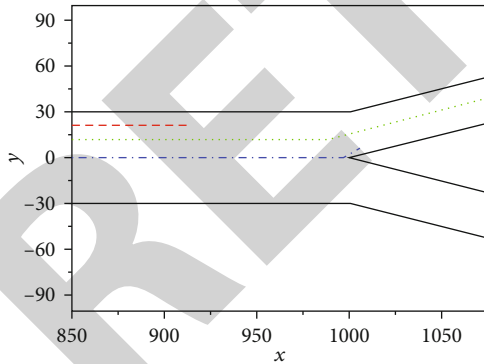


FIGURE 3: The trajectories of the particle settled at different positions for $Re = 38.7$. The angles between two branches is $\theta = 33.4^\circ$.

to study blood flow in arteries with aneurysm [9], advection-diffusion of chemicals and applications to blood flow [10], and deformable blockages in an elastic vessel [11].

The characteristics of blood flow near branches have been studied to understand the occurrence and development of diseases. The distribution of flow dynamic factors in two-

dimensional symmetric bifurcation [12] and in an optimum bifurcation geometry [13] is studied numerically. The separation of red blood cells at microvascular bifurcations [14] and the effect of stenosis growth on blood flow at the bifurcation [15] were also studied by using the lattice Boltzmann method.

Most studies concentrate on the dynamic factors, such as flow velocity, shear stress, and pressure and shear rate, by assuming the blood is Newtonian fluid. But actually, the blood flow is composed of plasma, erythrocytes, leukocytes, platelets, and lipids (cholesterol, cholesterol esters, and phospholipids). The transportation of blood cells and lipids in blood vessels performs a very important duty in cardiovascular system. Here, the particles are used to substitute for blood cells or lipids. This paper focuses on the motion of the particles in the fluid flow near the bifurcated artery. The dynamic properties of the particle near symmetric arterial bifurcations for different positions, Reynolds numbers, and branch angles are studied numerically. The study provides a basis for understanding the mechanism of hemodynamics near bifurcations of the cardiovascular system and the deposition of lipids in the intima of the artery.

2. The Lattice Boltzmann Model

This paper uses a D2Q9 (two-dimensional square lattice with nine velocities) lattice Boltzmann model. $f_i(\mathbf{x}, t)$ is a nonnegative real number which describes the distribution function of fluid at site \mathbf{x} at time t . The nine possible moving directions are $\mathbf{e}_0 = (0, 0)$; $\mathbf{e}_i = (\cos((\pi(i-1))/2), \sin((\pi(i-1))/2))$ for $i = 1, 2, 3, 4$; and $\mathbf{e}_i = \sqrt{2}(\cos((\pi(2i-1))/4), \sin((\pi(2i-1))/4))$, for $i = 5, 6, 7, 8$. The distribution functions satisfy the following equation according to the lattice Boltzmann model [1]:

$$f_i(\mathbf{x} + \mathbf{e}_i, t + 1) + f_i(\mathbf{x}, t) = -\frac{1}{\tau} [f_i - f_i^{(eq)}]. \quad (1)$$

The fluid density and macroscopic velocity are denoted by ρ and \mathbf{u} separately, which obey the following equations:

$$\begin{aligned} \rho &= \sum_i f_i = \sum_i f_i^{(eq)}, \\ \rho \mathbf{u} &= \sum_i f_i \mathbf{e}_i = \sum_i f_i^{(eq)} \mathbf{e}_i. \end{aligned} \quad (2)$$

Using Chapman-Enskog multiscale expansion, a suitable choice of the local equilibrium distribution function is

$$f_i^{eq} = \alpha_i \rho \left[1 + 3\mathbf{e}_i \cdot \mathbf{u} + \frac{9}{2} (\mathbf{e}_i \cdot \mathbf{u})^2 - \frac{3}{2} u^2 \right], \quad (3)$$

where $\alpha_0 = 4/9, \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1/9$, and $\alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = 1/36$. The pressure is defined as $p = (1/3)\rho$, and the kinematic viscosity is denoted by $\nu = (1/6)(2\tau - 1)$, respectively [1].

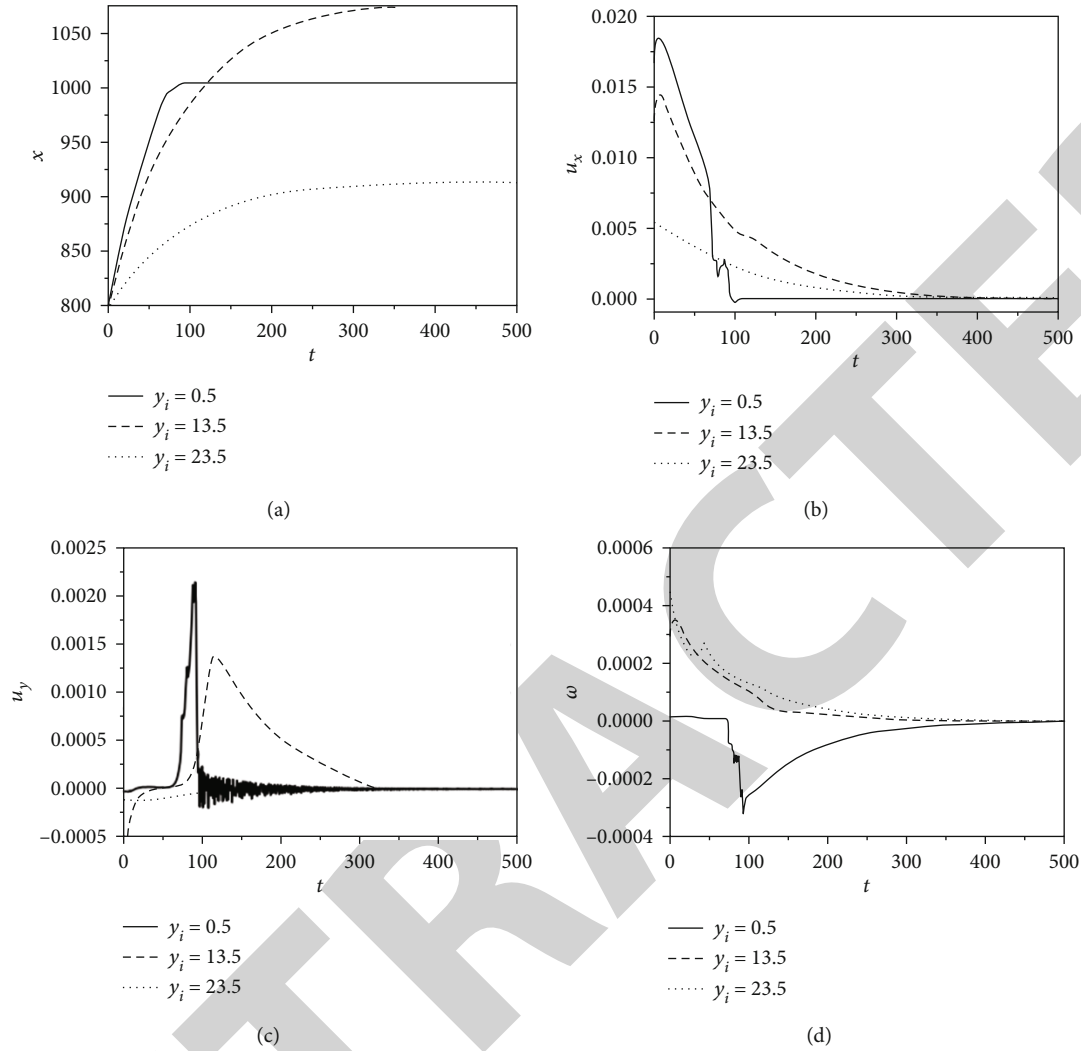


FIGURE 4: The particle x -directional position (a), x -directional velocity u_x (b), y -directional velocity u_y (c), and angular velocity ω (d) at different time steps t for the particle settled at $x_i = 800$, $y_i = 0.5$; $x_i = 800$, $y_i = 13.5$; and $x_i = 800$, $y_i = 23.5$, respectively. The angles between two branches is $\theta = 33.4^\circ$, and the flow Reynolds number is 38.7.

3. Results and Discussion

In this study, the two-dimensional configuration of the symmetric bifurcated artery is shown in Figure 1, which consists of a main vessel with diameter H of 60 lattice unit and two branches. The length of the main vessel is set to be 1000 lattice unit. The angles between two branches are denoted by θ . Here, the plasma is considered a Newtonian fluid and the particles are used to replace blood cells or lipids. In the study, the mass density of the plasma is 1 and the radius of the particle is 4.5 in lattice unit. $\tau = 0.55$ in each simulation. At the inlet of the main vessel, the velocity profile is supposed to be a parabolic distribution along the axial direction with the maximum center velocity v_0 . The flow at outlets of two branches is considered to be fully developed. At the beginning, the distribution functions at all the fluid nodes are supposed to be zero velocity equilibrium distribution functions except for those at inlet of the main vessel. The curved boundary condition depicted in references [16, 17]

is applied for the boundaries of the particle. The hydrodynamic force and moment on the particle boundary are calculated based on the method of stress-integration [18, 19].

A half-step “leapfrog” scheme is used to update the translational and rotational motion of particles at each Newtonian dynamic time step [20].

$$\begin{aligned} \mathbf{V}\left(t + \frac{1}{2}\delta t\right) &= \mathbf{V}\left(t - \frac{1}{2}\delta t\right) + \frac{\delta t \mathbf{F}(t)}{M}, \\ \mathbf{R}(t + \delta t) &= \mathbf{R}(t) + \delta t \mathbf{V}\left(t - \frac{1}{2}\delta t\right) + \frac{\delta t^2 \mathbf{F}(t)}{M}, \end{aligned} \quad (4)$$

where \mathbf{V} and \mathbf{R} are the velocity and displacement of particle centroid movement, respectively, and M are the mass of particle. At 10000 time steps, the flow field at the bifurcation of blood vessels reaches stability, at which time the particle is released. The flow Reynolds number at the inlet of the parent vessel is defined as $Hv_0/2\nu$.

3.1. The Velocity Field Near the Bifurcation. Figure 2 displays the velocity difference between the straight vessel and bifurcated vessel in the artery. The velocity v_x at $x = 800$ (near straight vessel) is a parabola while the velocity v_y is negligible compared with v_x . The existence of the bifurcation greatly hinders the x -directional flow v_x and leads to the sharp increase of outward velocity v_y near the bifurcation. The velocity is continuously distributed in the blood vessel without sudden change or eddy current, but a low velocity region is formed at the bifurcation and junction of the blood vessel. Clinical experiments also find the sudden increase of wall shear stress, and wall pressure happens on the wall near this position. Platelet particles and blood cell particles gather at this position, coupled with the synergistic effect of coagulase. The platelet and blood cell particles will grow continuously, so this position becomes the most suitable thrombus forming position.

3.2. Effect of Particle Initial Position on Its Deposition. If the blood vessel is straight, the fluid at these sites can be assumed as Poiseuille flow. When the particles flow into a straight pipe, the particles will stabilize in a position with a fixed distance from the center of the pipe after a certain distance of flow. This feature of motion is known as the Segré-Silberberg effect [21]. The phenomenon shows that the particles are separated by the driving force of the fluid along the mainstream direction and the lateral force perpendicular to the mainstream. This lateral force is the main cause of agglomeration movement of particles. The change of blood flow field near the bifurcation results in the change of the force on the particle, which leads to the particle's moving state. In order to compare the effect of vascular bifurcation for different particle positions, the particle is released at $y_i = 0.5, y_i = 13.5$, and $y_i = 23.5$ for $x_i = 800$, respectively. In the simulations, $Re = 38.7$ and $\theta = 33.4^\circ$. Figure 3 (dot line) shows the trajectory of the particle which is released at $y_i = 13.5$. The particle settled at this place can move forward and pass through the whole vessel eventually. Unlike the behavior in straight tube, the particles released at $y_i = 0.5$ and $y_i = 23.5$ have large difference in the final positions. But they all stop in the end. The particle stagnates at the inner wall of the branch for $y_i = 0.5$ while the particle stops near the main vessel wall for $y_i = 23.5$. The stagnation of the particles can make the adhesion of particle to the vessel. This adhesive obstruction is one of the reasons of atherosclerosis or thrombus. It is observed experimentally that the atherosclerosis or thrombus has a higher possibility at the location near bifurcations than straight vessel [22, 23].

The particle x -directional position, u_x , u_y , and ω are also shown in Figure 4, respectively, in order to describe the process of particle changing with the time in detail. The particle released at $y_i = 0.5$ (center of the main vessel) moves the fastest and stops the earliest while the particle settled at $y_i = 23.5$ (boundary of the main vessel) is just the opposite. But they all stagnate at the junction between the straight vessel and the bifurcations. The particle released at $y_i = 13.5$ reaches at the junction at about 100

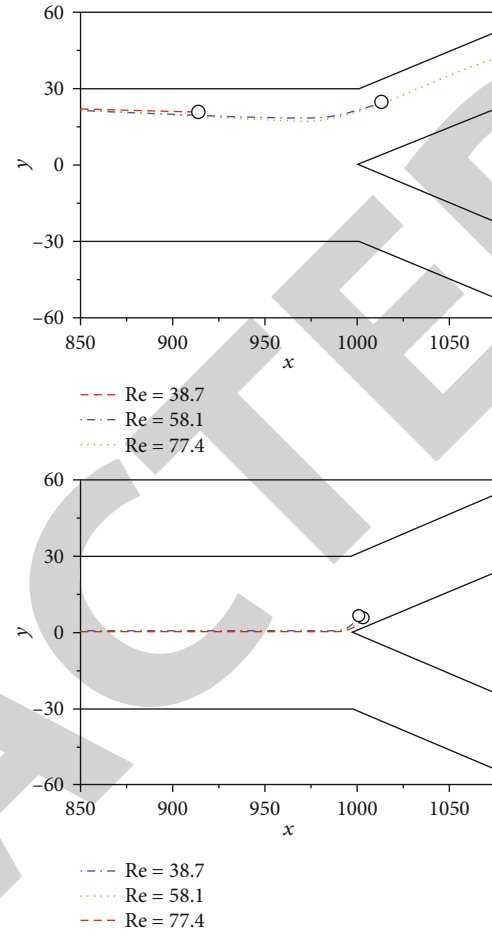


FIGURE 5: The trajectories of the particle for different flow Reynolds numbers for the particle are settled at $x_i = 800, y_i = 0.5$ and $x_i = 800, y_i = 23.5$ separately. The angles between two branches is $\theta = 33.4^\circ$.

time steps. The bifurcation decreases the motion of the particle in the x -direction while it increases the particle motion in the y -direction. Although the particle velocity in the branch is much smaller than that of in the main vessel, the particle can move forward and pass through the whole vessel successfully.

3.3. Effect of Flow Reynolds Number on Particle Deposition. Figure 5 displays the behavior of particle near symmetric bifurcation for different Reynolds numbers. The circles denote the particle stagnation position. Smaller Reynolds number causes the particle stagnates in the main vessel while larger Reynolds number leads to the particle stop in the branch if the particle is settled near the upper boundary. The particle passes through the whole vessel for Reynolds number is even higher, for example, of $Re = 77.4$. The increasing of fluid velocity can reduce particle deposition near the boundary of the artery. The particle released near the center of the main vessel deposits near the angle of the bifurcation when the Reynolds number is even higher. This result explains why this position is the most suitable thrombus forming position.

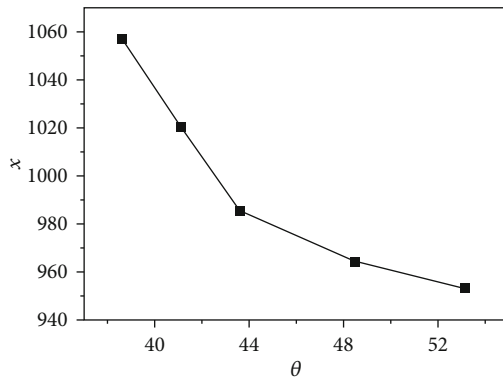


FIGURE 6: The particle final x -axis positions for different bifurcated angles θ for the particle is released at $x_i = 800, y_i = 13.5$ at $Re = 38.7$.

3.4. Effect of Bifurcation Angle on Particle Deposition. Figure 6 displays the behavior of particle near symmetric bifurcation for different branch angles θ . The particle is released at $x_i = 800, y_i = 13.5$ for flow Reynolds number which is 38.7. It is deduced from Figure 6 that the particle finally settles down at the junction of between the main vessel and the branches if θ is about 42° . Larger branch angle θ results in the particle stagnating in the main vessel. The particle can pass through the whole vessel if θ is smaller than 36° . The particle finally stops in the branches if θ is between 36° and 46° . Smaller flow Reynolds number and larger branch angle bring about higher possibility of particle stagnation near the bifurcations, which contributes to intimal thickening and atherosclerosis development. The artery with small flow Reynolds number and large branch angle is prone to develop atherosclerosis, which is consistent with the clinical observation [22, 23].

4. Conclusions

In summary, we used a lattice Boltzmann method to study the properties of particle moving near the arterial branches. The following conclusions are obtained by the simulation results.

- (1) In the straight vessel, the fluid at these sites can be assumed as Poiseuille flow. The particles finally stabilize in a position with a fixed distance from the center of the pipe. Compared with the straight vessel, the change of blood flow field near the bifurcation results in the change of the force on the particle, which leads to the change of particle's moving state.
- (2) The bifurcation has the function of hindering the particle motion. The particle is prone to stagnate near the wall of bifurcated vessels. Larger branch angle and smaller flow Reynolds number lead to larger possibility of the particle residence.
- (3) The adhesive obstruction of the blood vessel has large reason of stasis of particles. Our study will con-

tribute to the understanding of hemodynamic mechanisms around bifurcation and will contribute to the understanding of some cardiovascular diseases.

However, the current simulation is only a preliminary study of two-dimensional flow near arterial branches. Due to the complexity of blood flows and arterial structure, more studies combined with clinical observation are needed, and we are paying attention to this area.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Scientific Research Foundation of Weifang University of Science and Technology under Grant No. KJRC2022002.

References

- [1] Y. H. Qian, D. D'Humières, and P. Lallemand, "Lattice BGK models for Navier-Stokes equation," *Europhysics Letters*, vol. 17, no. 6, pp. 479–484, 1992.
- [2] M. Krafczyk, M. Schulz, E. Rank, and M. Cerrolaza, "Analysis of 3D transient blood flow passing through an artificial aortic valve by Lattice-Boltzmann methods," *Journal of Biomechanics*, vol. 31, no. 5, pp. 453–462, 1998.
- [3] J. Bernsdorf and D. Wang, "Non-Newtonian blood flow simulation in cerebral aneurysms," *Computers and Mathematics with Applications*, vol. 58, no. 5, pp. 1024–1029, 2009.
- [4] M. M. Dupin, I. Halliday, and C. M. Care, "Multi-component lattice Boltzmann equation for mesoscale blood flow," *Journal of Physics A*, vol. 36, no. 31, pp. 8517–8534, 2003.
- [5] C. H. Sun, C. Migliorini, and L. L. Munn, "Red blood cells initiate leukocyte rolling in postcapillary expansions: a lattice Boltzmann analysis," *Biophysical Journal*, vol. 85, no. 1, pp. 208–222, 2003.
- [6] H. B. Li, H. H. Yi, X. W. Shan, and H. P. Fang, "Shape changes and motion of a vesicle in a fluid using a lattice Boltzmann model," *Europhysics Letters*, vol. 81, no. 5, p. 54002, 2008.
- [7] M. M. Dupin, I. Halliday, C. M. Care, and L. L. Munn, "Lattice Boltzmann modelling of blood cell dynamics," *International Journal of Computational Fluid Dynamics*, vol. 22, no. 7, pp. 481–492, 2008.
- [8] Y. H. Liu, "A lattice Boltzmann model for blood flows," *Applied Mathematical Modelling*, vol. 36, no. 7, pp. 2890–2899, 2012.
- [9] H. H. Afrouzi, M. Ahmadian, M. Hosseini, H. Arasteh, D. Toghraie, and S. Rostami, "Simulation of blood flow in arteries with aneurysm: lattice Boltzmann approach (LBM)," *Computer Methods and Programs in Biomedicine*, vol. 187, p. 105312, 2020.

Retraction

Retracted: Mathematical Models for Analysis the Impact of Institutional Quality and Multilateral Financial Institutions on the Success of Green PPP Projects

Journal of Function Spaces

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 2023

Copyright © 2023 Journal of Function Spaces. 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. Li, R. Chen, and M. Ma, "Mathematical Models for Analysis the Impact of Institutional Quality and Multilateral Financial Institutions on the Success of Green PPP Projects," *Journal of Function Spaces*, vol. 2022, Article ID 3825211, 12 pages, 2022.

Research Article

Mathematical Models for Analysis the Impact of Institutional Quality and Multilateral Financial Institutions on the Success of Green PPP Projects

Wu-Yan Li ¹, Ruo-Yu Chen ¹, and Ming Ma²

¹School of Finance, Anhui University of Finance and Economics, Bengbu 233030, China

²Anhui Supervision Bureau, China Securities Regulatory Commission, Hefei 230088, China

Correspondence should be addressed to Ruo-Yu Chen; ruoyuchen@aufe.edu.cn

Received 16 May 2022; Accepted 14 June 2022; Published 30 June 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Wu-Yan 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.

Green PPP projects are of great significance to green infrastructure construction and green economic development. This study analyzes 552 Green PPP projects related to sewage treatment and clean energy in African countries during 2005-2019 to determine how institution quality and multilateral financial institutions influence Green PPP projects' success. We find that (1) successful (failed) Green PPP projects experience increase (decrease) new Green PPP projects' success rate. (2) Higher private capital risk-taking degree increases the likelihood of failure. (3) Institution quality affects project success through a mixed effect on private capital's risk-taking degree: *rule of law*, *government effectiveness*, *control of corruption*, and *regulatory quality* positively promote private capital's risk taking; *voice* has an opposite effect. (4) Multilateral financial institutions' participation in Green PPP projects promotes higher private capital risk taking.

1. Introduction

Green infrastructure construction can break through the limitations of traditional ecological protection and help realize the coordinated and sustainable development of ecology, society, and economy. Traditionally, most green infrastructure construction capital comes from the public, especially the governments. However, after the global financial crisis, developed economies' governments face the pressure of fiscal deficit and deleverage, and developing economies' financial systems cannot meet the capital demands of green infrastructure construction. Moreover, existing international financial organizations provide insufficient support to fill the gap [1]. Determining ways to introduce new investment subjects and innovating financing channels in green infrastructure construction is an urgent problem to solve the capital shortage. In recent years, some economies have been exploring ways to introduce private capital into green infrastructure construction, of which the green public-private partnership (Green PPP) model is considered the best. Green PPP means that government capital and social capital provide green

technology facilities and services to the public in a cooperative manner to meet social development and public demand. Currently, Green PPP is considered the most effective way to address capital shortage for green infrastructure construction. In developed countries, there is a significant growth in private capital participation in green infrastructure construction, and Green PPP infrastructure construction projects are emerging in developing countries [2, 3].

African countries are developing rapidly and face great demand for green infrastructure construction. IMF's African department noted that PPP is an important model to solve the problem of Africa's great infrastructure construction demand. Companies in various countries are actively "going global," and private capital is increasingly willing to participate in African Green PPP infrastructure construction projects. However, there are huge differences among African countries, and there are many investment risks. Private capital investment in Green PPP projects in Africa is still in the exploratory stage. Thus, it is necessary to study existing Green PPP projects in African countries and explore their success factors. This is of great significance to global private

capital's participation and success in African Green PPP projects. Although multilateral financial institutions such as the Asian Infrastructure Investment Bank (AIIB) were established to promote infrastructure construction in emerging market countries, some countries have reservations. Despite several studies on the AIIB's necessity, there is lack of studies demonstrating the role of multilateral financial institutions in promoting green infrastructure construction using empirical research methods.

Our study focuses on three issues: (1) Why do Green PPP projects have different results (success or failure) in different countries? (2) Which factors affect Green PPP projects' success or failure? (3) Which factors affect private capital's participation degree in Green PPP projects? Answering these questions is essential for global private capital willing to participate in African countries' Green PPP projects. Our study examines 552 Green PPP projects in African countries for 2005-2019 using the Probit and Cross-section regression models to answer these questions. The main conclusions are as follows: (1) Green PPP project experience affects new Green PPP projects' results. Countries with successful Green PPP project experience increase in the new Green PPP projects' success rate, while countries with failed Green PPP project have experience the opposite effect. (2) The more risks private capital takes in Green PPP projects, the higher is the likelihood of project failure. (3) Institution quality can influence the success of a Green PPP project by affecting private capital's choice of risk-taking degree in the project, and this latter is a mixed effect: *rule of law, government effectiveness, control of corruption, and regulatory quality* promote private capital's risk taking in Green PPP projects; *voice* has an opposite effect. (4) Multilateral financial institutions' involvement in Green PPP projects' financing channels promote private capital's higher risk taking.

This study contributes academically by conducting further empirical research on Green PPP issues. It identifies how institution quality and multilateral financial institutions affect Green PPP projects' internal risk structure and results (success or failure). Although prior literature studied similar problems, there were shortcomings. For example, qualitative research was often adopted to discuss institution quality's impact on PPP project success, but the channel affecting this success is a black box. This study's marginal contribution lies in quantifying institution quality through empirical research and internalizing it into the private capital's choice of a Green PPP project's internal risk structure. The logic is that institution quality affects private capital's risk-taking degree in a Green PPP project, which is reflected by the project's internal risk structure—an important factor determining Green PPP projects' success or failure. Green PPP projects' internal risk structure reflects the degree of risk transfer from public to private capital and is based on private capital's choice of risk-taking degree in the project. As a policy implication, this study recommends the problems global private capital should focus on when participating in African green infrastructure construction, and it demonstrates the importance of multilateral financial institutions, such as AIIB, in promoting and guaranteeing

global private capital participation in African green infrastructure construction.

The rest of the study is organized as follows. Section 2 reviews the literature and develops research hypotheses. Section 3 presents data sources and empirical research methods. Section 4 discusses the regression results. Section 5 concludes with some suggestions.

2. Literature Review and Research Hypotheses

2.1. Green PPP and Its Internal Risk Structure Classification.

The internal risk structure is the core of Green PPP research; different internal risk structures of Green PPP projects reflect different degrees of private capital participation. In the broad sense, Green PPP refers to all types of cooperative relationships the government establishes with private capital in the provision of green technology facilities and services. In the narrow sense, Green PPP is a contractual arrangement involving cooperation between government and private capital with risk and profit sharing [4–7]. Thus, Green PPP projects' process operation involves the local government and private capital teaming up to set up a project company called the special purpose vehicle, and they cooperate in the whole project implementation process by forming a contractual community to optimize risk and profit sharing. Lopes and Caetano [5] summarized Green PPP projects' characteristics: task bundling, risk transfer, and long-term contracting. Given Green PPP projects' long-term traits, Martimort and Straub [8] proposed that their contract arrangement should feature moral hazard, limited liability, and nonreversible contract constraints. In Green PPP projects, different contract arrangements between government and private capital represent private capital's degree of risk taking in the project, thus indicating the different internal risk structures of the Green PPP projects. Private capital can choose different contract arrangements according to the Green PPP project location's institution quality, and so Green PPP projects have different internal risk structures.

According to the World Bank's Private Participation in Infrastructure (PPI) Database, PPP projects' contract arrangement as well as internal risk structure can each be divided into 4 types and 12 subtypes. The 4 types are management contract and lease, brownfield, greenfield, and divestiture. Private capital takes on increased risks in each of these PPP project types. The 12 subtypes include management contract; lease contract; rehabilitate, operate, and transfer (ROT); rehabilitate, lease, or rent and transfer (RLT); build, rehabilitate, and transfer (BROT); built, lease, and transfer (BLT); build, operate, and transfer (BOT); build, own, and operate (BOO); merchant; rental; full; and partial.

2.2. Factors Influencing the Result (Success or Failure) of a Green PPP Project.

A Green PPP project's result (success or failure) reflects its risks. Some of its risks are endogenous and vary from project to project. Private capital can control them to some extent [9, 10]. Some risks of Green PPP projects are exogenous, which are related to the project location's economic environment and political system [11–14].

Existing literature found that the following factors generally affect Green PPP projects' results.

2.2.1. Successful Green PPP Operation Experience in Green PPP Project Location. Successful Green PPP operation experience is usually an open signal of a good cooperation between the government and private capital, which is crucial to new Green PPP projects' success [15–17]. Some studies suggested that countries with successful Green PPP experience have a higher probability of success when launching new Green PPP projects, whereas countries with failed Green PPP project experience may face the same obstacles and greater risk of failure in new Green PPP projects [18].

2.2.2. Internal Characteristics of a Green PPP Project. The internal risk structure, total amount of project investment, and cycle of the project may all affect Green PPP projects' results [4, 5, 18]. Specifically, the probability of failure of green PPP projects increases as the degree of risk sharing in the private sector increases, the total project investment increases, and the project cycle lengthens.

2.2.3. Institution Quality of the Green PPP Project Location. A country's institution quality—comprising the rule of law, government efficiency, corruption, control of private capital, etc., as well as regional and historical factors such as religion and culture—is not easy to quantify. Some studies used the World Bank's WGI and International Country Risk Guide's (ICRG) Country Risk Index—the most influential international databases for measuring institution quality—to empirically study the relationship between the PPP location's institution quality and PPP projects' success rate [13, 19, 20]. However, none analyzed the relationship's mechanism. Further research suggests that a Green PPP project's internal risk structure is related to its location's institution quality. Its internal risk structure reflects private capital's participation degree in the project, which is positively correlated with the project location's institution quality [21, 22].

2.2.4. Financial Channels for a Green PPP Project. Most infrastructure construction projects need external financing. Estache [23] argued that commercial banks always play a self-serving role in PPP projects, and so PPP projects' success rate cannot improve if commercial banks provide the financing. However, Galilea and Medda [18] empirically discovered that Green PPP projects are more likely to succeed if financed by multilateral financial institutions such as the World Bank or Asian Development Bank. These tend to be nonprofit, and their participation in Green PPP projects anchors the government and private capital into a better contract arrangement and execution.

2.2.5. Macroeconomic Environment of the Green PPP Project Location. A country's macroeconomic conditions and economic openness are closely related to the infrastructure construction's efficiency. Private capital tends to invest in countries with stable macroeconomic environments. Development of a financial market and construction of credit culture in the PPP project location are also essential [16].

2.3. Research Hypotheses. Many Green PPP projects fail due to lack of private capital's evaluation of such projects' long-term uncertainty, making their actual risks significantly higher than expected, resulting in default or renegotiation [24]. When making Green PPP investment decisions, private capital prioritizes governments with experience. If a government has successful (failed/no) Green PPP project experience, a new Green PPP project with this government is more likely to succeed (face risks). Therefore, we conclude Hypothesis 1.

Hypothesis 1. For African countries, new Green PPP projects' success rate is positively correlated with experience of successful Green PPP projects.

A Green PPP project's internal risk structure reflects the degree of risk transfer from the government to private capital, and private capital chooses its risk-sharing degree according to project information and own characteristics. To achieve higher profit, private capital must find a balance between taking more risks and making more profits: it may have a subjective desire to increase risk taking, thus increasing Green PPP projects' failure rate.

Hypothesis 2. Private capital's choice of risk-sharing degree is negatively correlated with a Green PPP project's success rate.

A good political and judicial system reduces the transaction cost and political risk of private capital's participation in Green PPP projects. When choosing the degree of risk sharing in a Green PPP project, private capital judges whether the local government has an adequately credible commitment to ensure Green PPP projects' progress, there are clear measures and schedules to implement the Green PPP contract, there are reasonable incentive pricing mechanisms, and so on. According to the WGI database, institution quality includes six specific indicators—control of corruption, voice and accountability, political stability/no violence, rule of law, government effectiveness, and regulatory quality—and the higher is their value, the better can the country's institution quality be considered and quantified. Some literature empirically tested these indicators' positive effects on private capital's risk-sharing degree in PPP projects [18–20]. However, our study argues that, among these indicators, “voice and accountability” has certain particularity; as its judgment criteria is subjective, the databases judge “voice and accountability” based on developed countries' principle of the political system. The abovementioned empirical literature was also based on developed countries' PPP projects, while our study focuses on Green PPP projects in African countries; therefore, the judgment criteria for “voice and accountability” is different. More importantly, there has been controversy about the impact of democratic decentralization and government centralization on the supply of public goods. The concentration of government resources and power can, to a certain extent, increase the government's mobilization and action power, reduce uncertain risks and transaction costs of Green PPP projects, and

create greater incentives for private capital, so as to promote private capital's participation in Green PPP projects to a greater extent.

Institution quality influences Green PPP projects' internal risk structure by influencing private capital's degree of participation, thus influencing the result (success or failure) of Green PPP projects. When selecting Green PPP projects' risk structure, rational private capital selects the most suitable risk-sharing degree according to the institution quality of the Green PPP project location; that is, the institution quality is internalized in a Green PPP project's internal risk structure. Accordingly, we hypothesize the following:

Hypothesis 3. Institution quality has a mixed effect on a Green PPP project's success. The five institution quality indicators of control of corruption, political stability/no violence, rule of law, government effectiveness, and regulatory quality positively affect private capital's risk-sharing degree in a Green PPP project, while the indicator of voice and accountability has the opposite effect.

Generally, private capital is at a disadvantage when cooperating with the government. It is unable to strictly supervise the government, making the government prone to abuse its power and increase Green PPP projects' risks. Multilateral financial institutions' participation in Green PPP projects is conducive to forming a multilateral cooperation mechanism, implementing international standards, and effectively supervising the government. For example, commercial financial institutions can provide financing for a Green PPP project, but cannot effectively supervise the government, but multilateral financial institutions are non-governmental organizations, which usually pool sovereign wealth from different countries and negotiate with governments on behalf of donor countries to reach a Green PPP project contract arrangement. Benefitting from their underlying sovereign power, multilateral financial institutions can urge governments to ensure Green PPP projects' smooth operation to reduce private capital's risk in participating in the Green PPP project. Therefore, we propose the following.

Hypothesis 4. Multilateral financial institutions' participation increases private capital's risk-sharing degree in a Green PPP project.

3. Data and Research Methods

3.1. Green PPP Project Data. The Green PPP project data used in our study were taken from the World Bank's PPI database, an authoritative database of private capital's infrastructure construction involvement in various countries. It covers the major sectors of energy, transportation, communication, and sewage. Our study focuses on Green PPP projects in African countries, and so we selected all 2005-2019 Green PPP project data in sewage and clean energy related sectors for 53 African countries from the PPI database as our empirical research object (Table 1).

The Green PPI database mainly provides the following information about Green PPP projects.

3.1.1. Financial Closure Year. This is the year of a Green PPP project's beginning. This is when the government and private capital formed a legally binding contract arrangement in which both share project risks. It is also the completion date of the Green PPP project-financing plan.

3.1.2. Project Status. This refers to the Green PPP project's current status. The PPI database defined the "project status" based on whether the project operates normally within the Green PPP contract term, and it assigned six status types to all Green PPP projects: operational, construction, concluded, cancelled, distressed, and merged. Among them, a "merged" Green PPP project is one that has been renamed or replaced; that is, its observed data are invalid. Therefore, Green PPP projects with "merged" status are excluded from our empirical analysis. Most Green PPP projects' status is "operational" or "construction," meaning they are not really finished; they will continue to be affected by different factors in the following period and may eventually succeed or fail miserably. Given the high uncertainty of ongoing Green PPP projects, our definition of their "success" or "failure" should be limited to completed projects. Specifically, the definition of a Green PPP project's "success" or "failure" in our study is as follows: if a Green PPP project's project status is "concluded," that project is a "success," and the value of dependent variable status is 1 in the empirical analysis; if the project status is "distressed" or "cancelled," the Green PPP project is a "failure," and status equals 0. The summary of project statuses of the Green PPP projects in our study is presented in Table 2.

3.1.3. Type/Subtype of PPI. This refers to contract arrangement the type/subtype of each Green PPP project. Different contract arrangements represent different Green PPP projects' internal risk structures and different degrees of risk taking by private capital. In increasing order of private capital's risk-taking degree, the PPI database divides the PPIs into four types, that is, management contract and lease, concession, greenfield projects, and divestiture. These Green PPP project types were divided into 12 subtypes, also based on private capital's risk-taking degree. According to the type/subtype of PPI, we constructed the dependent regression variables (Type_level: 1-4 and SubType_level: 1-12). Specific construction results are as follows: management contract and lease, Type_level = 1; concession, Type_level = 2; greenfield projects, Type_level = 3; and divestiture, Type_level = 4. The 12 subtypes were assigned the following values: management contract, SubType_level = 1; full, SubType_level = 2; the values in the middle increase in order. Summary of Green PPP projects' type/subtype of PPI in our empirical test sample is presented in Tables 3 and 4.

3.1.4. Sponsors. Sponsors are private capital investors who are the real risk takers in the Green PPP project. In our study, we constructed dependent regression variables (DummySponsors = 0 or 1) to measure whether Green PPP projects involve one or more private capital investors. DummySponsors equals 0 if a PPP project has only one private capital investor and 1 for more than one investor.

TABLE 1: Countries and income groups of sample Green PPP projects.

Country	Income level	Country	Income level	Country	Income level
Algeria	UM	Gabon	UM	Rwanda	Low
Angola	UM	Gambia	Low	Senegal	LM
Benin	Low	Ghana	LM	Seychelles	UM
Botswana	UM	Guinea	Low	Sie. Leone	Low
Burkina Faso	Low	Guinea-Bissau	Low	Somalia	Low
Burundi	Low	Kenya	Low	S. Africa	UM
Cameroun	LM	Lesotho	LM	S. Sudan	LM
Cape Verde	LM	Liberia	Low	Sudan	LM
C. African	Low	Madagascar	Low	Swaziland	LM
Chad	Low	Malawi	Low	STP	LM
Comoros	Low	Mali	Low	Tanzania	Low
Dem. Rep. Congo	Low	Mauritania	Low	Togo	Low
Rep. Congo	LM	Mauritius	UM	Tunisia	UM
Ivory Coast	LM	Morocco	LM	Uganda	Low
Djibouti	LM	Mozambique	Low	Zambia	LM
Egypt	LM	Namibia	UM	Zimbabwe	Low
Eritrea	Low	Niger	Low	—	—
Ethiopia	Low	Nigeria	LM	—	—

Source: Private Participation in Infrastructure Dataset, World Bank. Notes: Green PPP = green public-private partnership; LM = lower-middle-income country; UM = upper-middle-income country.

TABLE 2: The summary of the project status of the sampled Green PPP projects.

	Ongoing	Completed	Cancelled	Crisis	Total
Low income	173	18	13	8	212
LM income	162	13	7	8	190
UM income	138	9	2	1	150
Total	473	40	22	17	552

Source: Private Participation in Infrastructure Dataset, World Bank. Notes: Green PPP = green public-private partnership; LM = lower-middle-income country; UM = upper-middle-income country.

3.1.5. Multilateral Support. This refers to whether a Green PPP project involves multilateral financial institutions, and we expressed this using the Multilender (=0 or 1)-dependent regression variable. Specifically, Multilender equals 1 if a Green PPP project involves multilateral financial institutions and 0 otherwise.

3.1.6. Investment and Period. The PPI database provides information on Green PPP projects' investment amount and project period, which are two important indicators to control different Green PPP projects' risks. A higher investment amount or longer project period means the Green PPP project will face more uncertainties. Therefore, we constructed two control variables: investment and period.

3.2. Other Relevant Data

3.2.1. Green PPP Project Experience. According to the PPI database, our study used the number of successful Green PPP projects in a location country before launching a new Green PPP project as the measurement index of successful

project experience (SuccessPPI). Similarly, we used the number of failed projects as a measure of experience with failed projects (NoSuccessPPI).

3.2.2. Institution Quality. The quantification of institution quality is often subjective and difficult. Our study used the World Bank WGI database, which includes six specific indicators to measure institution quality (Table 5). Among these, (a) control of corruption (Anti_Corruption) reflects a government's ability to monitor, prevent, and control corruption. The variable Anti_Corruption ranges from -2.5 to 2.5, with a higher value indicating the government's ability to control corruption. (b) Voice and accountability (Voice) reflects a country's degree of democracy. It mainly includes two aspects: whether people have the right to elect the government and whether the public media are not restricted in the process of news reporting. Voice ranges from -2.5 to 2.5, with a higher value indicating the people's right to speak and the public's influence on social and economic affairs. (c) Political stability/no violence (Stability) reflects a government's possibility of being disturbed/overturned by a violation of the constitution or violence. Stability ranges from -2.5 to 2.5, with a higher value indicating the degree of government's stability. (d) Rule of law (Law) reflects the degree of compliance with the laws of social subjects, quality of contract execution, property rights protection, police and courts, and degree of violence and crime. Law ranges from -2.5 to 2.5, with a higher value indicating the degree of the rule of law. (e) Government effectiveness (Effectiveness) reflects the quality of public service provided by the government, including the quality of policy setting and implementation and credibility of government commitments. Effectiveness ranges from -2.5 to 2.5, with a higher value indicating the

TABLE 3: Type/subtype of sampled Green PPP projects and their variable value.

<i>Type_level</i>	Contracts			Brownfield				
Variable assignment	1			2				
<i>Subtype_level</i>	Management contract		Lease contract		ROT		RLT BROT	
Variable assignment	1		2		3		4 5	
<i>Type_level</i>	Greenfield			Divestiture				
Variable assignment	3			4				
<i>Subtype_level</i>	BLT	BOT	BOO	Merchant	Rental	Partial	Full	
Variable assignment	6	7	8	9	10	11	12	

Source: Private Participation in Infrastructure Dataset, World Bank. Notes: Green PPP = green public-private partnership; ROT = rehabilitate, operate, and transfer; RLT = rehabilitate, lease, or rent and transfer; BROT = build, rehabilitate, and transfer; BLT = built, lease, and transfer; BOT = build, operate, and transfer; BOO = build, own, and operate.

TABLE 4: Summary of the type/subtype of PPI of the sampled Green PPP projects.

Type_level	Subtype_level	Contracts	Low income	LM income	UM income	Total
1	1	Management contract	10	12	10	32
	2	Lease contract	3	6	0	9
	3	ROT	11	13	4	28
2	4	RLT	5	20	0	25
	5	BROT	4	13	6	23
	6	BLT	0	0	1	1
3	7	BOT	18	32	16	66
	8	BOO	36	32	79	147
	9	Merchant	92	47	21	160
4	10	Rental	15	2	7	24
	11	Partial	14	14	6	34
	12	Full	1	2	0	3
Total			209	193	150	552

Source: Private Participation in Infrastructure Dataset, World Bank. Notes: PPI = private participation in infrastructure; Green PPP = green public-private partnership; LM = lower-middle-income country; UM = upper-middle-income country; ROT = rehabilitate, operate, and transfer; RLT = rehabilitate, lease, or rent and transfer; BROT = build, rehabilitate, and transfer; BLT = built, lease, and transfer; BOT = build, operate, and transfer; BOO = build, own, and operate.

TABLE 5: The summary of the Institution Quality.

Variable	Index	Range	Definition
<i>Anti_Corruption</i>	Control of corruption	[-2.5,2.5]	The ability of a government to prevent and control corruption
<i>Voice</i>	Voice and accountability	[-2.5,2.5]	The level of democratization in a country, including whether civilians have the right to vote
<i>Stability</i>	Political stability no violence	[-2.5,2.5]	The regime's stability
<i>Law</i>	Rule of law	[-2.5,2.5]	Institutionalized level in one country
<i>Effectiveness</i>	Government effectiveness	[-2.5,2.5]	Quality of public service and credibility of government
<i>Regulation</i>	Regulatory quality	[-2.5,2.5]	Whether the government is willing to implement policies to improve the private sector

Source: Worldwide Governance Indicators database, World Bank.

government's effectiveness. (f) Regulatory quality (Regulation) reflects the government's ability to set and implement policies that benefit private capital development. Regulation ranges from -2.5 to 2.5, with a higher value indicating the government's positive role in promoting private capital development.

The World Bank WGI database gives specific values to these six variables in countries where the Green PPP projects in our study are located. Considering the multicollinearity problem, we calculated these six variables' correlation coefficients and found that Anti_Corruption had high positive correlation with Law, Effectiveness, and Regulation.

Therefore, only Anti_Corruption, Voice, and Stability were retained in the following regression analysis.

3.2.3. *Control Variables.* The control variables include the following.

(1) *National Income Level (Income).* The level of a country's income usually affects investors' enthusiasm for investment. We divided the sample countries into low income (*Low_Income*), lower middle income (*Lower_Middle_Income*), and upper middle income (*Upper_Middle_Income*) according to the PPI classification criteria for national income levels and assigned them values 1, 2, and 3, respectively, in the first empirical regression analysis. In the second empirical regression analysis, taking low-income as the benchmark, we introduced dummy variables *i_Lower_Middle_Income* and *i_Upper_Middle_Income*.

(2) *GDP Growth Rate (Ln_L_GDP_Growth).* GDP growth rate in the year of a Green PPP project's start can affect the demand for green infrastructure construction and enthusiasm of private capital investment. We construct variable *Ln_L_GDP_Growth* and use GDP growth rate data for a Green PPP project's financial closure year.

(3) *Capital Opening Index (L_Open).* The degree of capital openness in the Green PPP project locations can affect a Green PPP project's results (success or failure) and private capital's degree of participation. Therefore, we introduce capital opening related variables into the empirical regression. We constructed the capital opening index, calculated as $Open_index_t = FDI_t / GDP_t$. We construct control variable *L_Open*, whose value is one lag phase value of the capital opening index.

(4) *Temporal Trend (i_Year).* To eliminate the temporal trend's influence, we introduce the time control variable *i_Year* into the empirical regression, which takes the specific value of 1 when the year is 2005 and 15 when the year is 2019; the rest may be deduced by analogy.

3.3. Research Methods

3.3.1. *Regression Model 1: Factors Influencing Green PPP Projects' Success or Failure.* The first regression model empirically analyzes factors influencing Green PPP projects' success or failure of a (Status). The dependent variable is Green PPP projects' success or failure, which is a binary variable. Probit model is mostly used for binary variables. Its advantage lies in its simple form and easy processing. It is the most basic and effective statistical model. In many cases, there is no linear relationship between independent variables and dependent variables, so the model can achieve the original linear hypothesis by adjusting the fitting appropriately. Following Percoco [21], we adopt the Probit model in the first regression:

$$P(Y = 1|X) = \frac{e^{X'\beta}}{1 + e^{X'\beta}}, \quad (1)$$

where *X* is the independent variable and β is the regression coefficient. In this regression model, Green PPP project experience (SuccessPPI or NoSuccessPPI), type/subtype of PPI (Type_level and SubType_level), sponsors (DummySponsors), investment (Investment), institution quality (Corruption, Voice, and Stability), and other control variables (e.g., Income and L_Open) are independent variables. The dependent variable is project status (Status). If the project status of a Green PPP project is "concluded," Status takes the value of 1 and 0 if the project status is "distressed" or "cancelled." Table 6 describes the variables statistics in regression model 1.

3.3.2. *Regression Model 2: Factors Influencing Green PPP Projects' Internal Risk Structures.* Regression model 1 reveals which factors will contribute to the success of a Green PPP project. Among them, we focus on the influence of internal risk structures. Then, the next question is, which factors affect a Green PPP project's internal risk structure? Which factors affect private capital's risk-taking degree in a Green PPP project? To answer these questions, we adopted cross-section regression model 2 to analyze factors influencing a Green PPP project's internal risk structure. The advantage of cross-section model is that it can effectively solve the problem of insufficient sample size, estimate the impact of some factors that are difficult to measure on the explained variables, and help correctly understand the relationship between economic variables.

$$\begin{aligned} RiskSharing = & \alpha + \sum_{i=1}^m \eta_i \cdot QI_t + \gamma \cdot Multilender \\ & + \sum_{j=1}^n \sigma_j \cdot Prop_j + \sum_{l=1}^q \theta_l \cdot Control_l + \xi, \end{aligned} \quad (2)$$

where *RiskSharing* is the dependent variable representing private capital's risk-taking degree in a Green PPP project. Private capital's different risk-taking degrees in Green PPP projects are given by different contract arrangements of Green PPP projects, represented by variables Type_level and SubType_level. In regression model 2, the independent variables include institution quality (Corruption, Voice, and Stability), multilateral support (Multilender), *Prop_j* (DummySponsors, Investment, and Period), *Control_l* (*L_Open*, *i_Low_Income*, *i_Lower_Middle_Income*, *i_Upper_Middle_Income*, *Ln_L_GDP_Growth*, and *i_Year*). Regression model 2 only uses successful projects as research samples. Table 7 describes the variables' statistics in regression model 2.

4. Regression Result Analysis

4.1. *Which Factors Determine the Success or Failure of a Green PPP Project?* Table 8 presents results of regression model 1. In column (1), we empirically tested the effect of Green PPP project experience (SuccessPPI or NoSuccessPPI), subtype of PPI (SubType_level), sponsors (DummySponsors), investment (Investment), institution quality (Corruption, Voice, and Stability), and other control variables (*L_Open* and *i_Income*) on new Green PPP projects' success. We found

TABLE 6: Summary statistics of variables in regression model 1.

Variable	Observation	Mean	Std. dev	Minimum	Maximum
<i>Status</i>	79	0.242	0.335	0.000	1.000
<i>SuccessPPI</i>	79	1.612	1.881	0.000	7.000
<i>NoSuccessPPI</i>	79	2.012	2.115	0.000	5.000
<i>DummySponsors</i>	79	0.312	0.437	0.000	1.000
<i>Investment</i>	79	104.666	397.082	0.000	2,990.883
<i>Type_level</i>	79	1.926	0.894	1.000	4.000
<i>Subtype_level</i>	79	5.320	3.978	1.000	12.000
<i>Voice</i>	79	-0.109	0.302	-1.972	0.402
<i>Anti_Corruption</i>	79	-0.692	0.298	-1.084	0.320
<i>Stability</i>	79	-0.908	0.569	-2.263	0.414
<i>L_Open</i>	79	11.131	11.755	1.228	84.198
<i>Income</i>	79	1.658	0.732	1.000	3.000

TABLE 7: Summary statistics of variables in regression model 2.

Variable	Observation	Mean	Std. dev	Minimum	Maximum
<i>Type_level</i>	513	1.426	0.513	1.000	4.000
<i>Subtype_level</i>	513	5.088	1.818	1.000	12.000
<i>Voice</i>	513	-0.524	0.767	-1.875	0.463
<i>Anti_Corruption</i>	513	-0.689	0.254	-1.465	0.547
<i>Stability</i>	513	-0.973	0.407	-2.440	1.209
<i>Multilender</i>	513	0.021	0.167	0.000	1.000
<i>DummySponsors</i>	513	0.226	0.462	0.000	1.000
<i>Investment</i>	513	649.559	1,108.240	0.000	4,093.656
<i>Period</i>	513	25.470	12.483	1.000	99.000
<i>L_Open</i>	513	15.198	13.013	0.762	109.427
<i>I_Year</i>	513	4.304	3.182	1.000	12.000
<i>i_Lower_Middle_Income</i>	513	0.402	0.383	0.000	1.000
<i>i_Upper_Middle_Income</i>	513	0.447	0.394	0.000	1.000
<i>Ln_L_GDP_Growth</i>	513	2.984	1.480	-1.500	5.500

that subtype of PPI (*SubType_level*) has a statistically negative effect on Green PPP project success at the 5% significance level, meaning higher risk-taking degree of private capital in a Green PPP project increases the likelihood of project failure. Institution quality (*Corruption*, *Voice*, and *Stability*) and other variables (*SuccessPPI*, *NoSuccessPPI*, *DummySponsors*, *L_Open*, and *i_Income*) do not have a statistically significant impact on a Green PPP project's success.

In column (2), we remove these control variables and conduct regression considering a correlation between the control variables and institution quality. We found that subtype of PPI (*SubType_level*) has a statistically negative effect on new Green PPP projects' success at the 1% significance level, Green PPP project experience (*SuccessPPI*) has a statistically positive effect at the 5% significance level, and Green PPP project experience (*NoSuccessPPI*) has a statistically negative effect at the 1% significance level. Investment (*Investment*) has a statistically negative effect at the 5% significance level, and institution quality (*Corruption*, *Voice*, and *Stability*) does not have a statistically significant impact.

In column (3), we remove institution quality (*Corruption*, *Voice*, and *Stability*) from the regression. We found that Green PPP project experience (*SuccessPPI*) has a statistically positive effect on a new Green PPP project's success at the 5% significance level, Green PPP project experience (*NoSuccessPPI*) has a statistically negative effect at the 1% significance level, investment (*Investment*) has a statistically negative effect at the 1% significance level, and subtype of PPI (*SubType_level*) has a statistically negative effect at the 10% significance level.

Through the above three regression results, we can find that Hypothesis 1 and Hypothesis 2 are both accepted. And the control variables also play a corresponding role in the regression model. To further study institution quality's influence on a Green PPP project's success, we introduced interaction variables between the institution quality and subtype of PPI variables (*Voice*Subtype_level*, *Corruption*Subtype_level*, and *Stability*Subtype_level*). In column (4) regression, we added institution quality (*Corruption*, *Voice*, and *Stability*), subtype of PPI (*SubType_level*), and their

TABLE 8: Results of regression model 1: factors determining the success or failure of a Green PPP project.

Variable	(1)	(2)	(3)	(4)	(5)
<i>SuccessPPI</i>	0.03061 (0.0304)	0.00562** (0.00218)	0.00424** (0.00172)	0.000305 (0.00164)	0.00914** (0.00406)
<i>NoSuccessPPI</i>	-0.7378 (0.553)	-0.1824*** (0.0579)	-0.1108*** (0.0411)	-0.0775** (0.0341)	-0.3111** (0.1251)
<i>DummySponsors</i>	-0.865 (0.759)	-0.789 (0.507)	-0.767 (0.538)		
<i>Investment</i>	-12.49 (8.825)	-3.298** (1.442)	-3.058*** (0.758)		-4.699* (2.426)
<i>Subtype_level</i>	-0.467** (0.204)	-0.211*** (0.0573)	-0.147* (0.797)	-0.306** (0.118)	-0.218** (0.113)
<i>Voice</i>	-7.438 (5.358)	-0.708 (0.509)		-0.663 (0.623)	-1.359 (0.923)
<i>Anti_Corruption</i>	7.089 (4.561)	0.616 (0.711)		0.735 (1.041)	0.845 (1.447)
<i>Stability</i>	5.656 (4.109)	0.224 (0.673)		-0.0716 (0.664)	1.141 (0.817)
<i>Voice * Subtype_level</i>				0.0491 (0.0889)	-0.252 (0.195)
<i>Corruption * Subtype_level</i>				-0.211 (0.274)	-0.0410 (0.253)
<i>Stability * Subtype_level</i>				0.0678 (0.152)	0.152 (0.153)
<i>L_Open</i>	-0.198 (0.158)		-0.0203 (0.0159)		
<i>i_Income</i>	-5.725 (4.796)		-0.522 (0.445)		
<i>Constant</i>	20.22 (14.63)	2.290 (0.609)	2.819 (1.665)	1.262 (0.772)	1.039 (1.527)
Observation	79	79	79	79	79
R²	0.746	0.598	0.638	0.495	0.597

Notes: ***, **, and * indicate the significance levels of 1%, 5%, and 10%, respectively.

interaction variables (*Voice*Subtype_level*, *Corruption*Subtype_level*, *Stability*Subtype_level*), and we removed other control variables from the regression. We found that subtype of PPI (*SubType_level*) has a statistically negative effect on a Green PPP project's success at the 5% significance level, while institution quality (*Corruption*, *Voice*, and *Stability*) and the interaction variables (*Voice*Subtype_level*, *Corruption*Subtype_level*, and *Stability*Subtype_level*) do not have a statistically significant impact. Such regression results show that institution quality does not directly influence Green PPP project success, and the influences of institution quality and Green PPP projects' internal risk structure are not affected by each other.

In column (5) regression, we added control variables with a significant influence in the previous regression. Thus, institution quality (*Corruption*, *Voice*, and *Stability*) and the interaction variables (*Voice*Subtype_level*, *Corruption*Subtype_level*, and *Stability*Subtype_level*) do not have a statistically significant impact on Green PPP project success.

The regression results show that, although scholars deem that institution quality affects Green PPP project's success, in our empirical test samples, institution quality's influence on a Green PPP project's success is not expressed.

4.2. Which Factors Determine the Internal Risk Structure of a Green PPP Project? Regression model 2 tests factors that influence private capital's risk-taking degree in a Green PPP project, namely, factors influencing Green PPP projects' internal risk structure (*Type_level* and *SubType_level*). The full sample contains 513 Green PPP projects with project status of concluded, operational, and construction. We conducted regression in two subindustries: clean energy and sewage. The result of regression model 1 is presented in Table 9.

From the full sample regression results, the degree of democracy in a country (*Voice*) has a statistically negative effect on private capital's risk-taking degree in a Green PPP project (*Subtype_level*) at the 1% significance level.

TABLE 9: Results of regression model 2: factors determining the internal risk structure of a Green PPP project.

Variable	Full sample	Energy	Sewage
<i>Voice</i>	-0.586*** (0.146)	-0.0227 (0.161)	-0.182 (0.310)
<i>Anti_Corruption</i>	0.867** (0.397)	-0.0479 (0.403)	-0.182 (1.173)
<i>Stability</i>	-0.168 (0.183)	0.204 (0.177)	-0.173 (0.428)
<i>Multilender</i>	0.520** (0.246)	0.407** (0.186)	0.420 (0.573)
<i>DummySponsors</i>	0.126 (0.107)	0.221** (0.162)	0.119 (0.141)
<i>Investment</i>	0.0362*** (0.0112)	0.00107 (0.0153)	-0.0127 (0.162)
<i>Period</i>	0.0318*** (0.00488)	0.00361 (0.00576)	0.0601*** (0.0161)
<i>L_Open</i>	0.0129* (0.00711)	0.0136*** (0.00490)	-0.0155 (0.0178)
<i>I_Year</i>	-0.0103 (0.0166)	-0.0307** (0.0102)	-0.0762** (0.0303)
<i>I_Lower_Middle_Income</i>	-1.365*** (0.305)	-1.011*** (0.203)	1.054* (0.692)
<i>I_Upper_Middle_Income</i>	-1.849*** (0.494)	-0.778* (0.499)	0.840 (0.899)
<i>Ln_L_GDP_Growth</i>	-0.582*** (0.163)	-0.609*** (0.148)	0.296 (0.366)
<i>Constant</i>	5.203*** (1.226)	6.344*** (1.205)	-1.344 (1.668)
Observation	513	312	201
R ²	0.147	0.191	0.299

Notes: ***, **, and * indicate the significance levels of 1%, 5%, and 10%, respectively.

Thus, the more democratic is a country, the less willing is private capital to take risks in its Green PPP projects. The government's ability to control corruption (*Anti_Corruption*) has a statistically positive effect on private capital's risk-taking degree in a Green PPP project (*Subtype_level*) at the 5% significance level. Thus, the Hypothesis 3 is accepted and a country can increase private capital's willingness to participate in Green PPP projects by improving its ability to control corruption. Since *Anti_Corruption* is highly positively correlated with Law, Effectiveness, and Regulation, the above result also means that if a country has high degree to compliance with laws, high-quality public services, and high ability to set private capital development policies, private capital would be willing to take more risks in its Green PPP projects. Multilateral financial institutions involved in Green PPP projects (*Multilender*) have a statistically positive effect on private capital's risk-taking degree in a Green PPP project (*Subtype_level*) at the 5% significance

level. The regression results show that the Hypothesis 4 is accepted and it implies that the private capital tends to participate in the Green PPP project which involves multilateral financial institution. A Green PPP project's investment and period of (*Investment* and *Period*) have a statistically positive effect on private capital's risk-taking degree in a Green PPP project (*Subtype_level*) at the 1% significance level. The control variables (*L_Open*, *i_Low_Income*, *i_Lower_Middle_Income*, *i_Upper_Middle_Income*, *Ln_L_GDP_Growth*, and *i_Year*) all have a statistically positive effect on private capital's risk-taking degree in a Green PPP project (*Subtype_level*).

There are differences in the regression results of the two subindustries (clean energy and sewage). Multilateral financial institutions' involvement (*Multilender*) and private capital investors (*DummySponsors*) positively influences private capital participation in the clean energy sector's Green PPP projects, while the influence on sewage sector is relatively weak. The main reason is that Green PPP project investment in the clean energy sector is large, and participation of multilateral financial institutions and multiple private capital investors can increase risk. On the contrary, the Green PPP project's period (*Period*) positively influences private capital participation in Green PPP projects in the sewage sector, while the influence on the clean energy sector is relatively weak. This is because Green PPP projects' period in the sewage sector is long; the risk-sharing mechanism is particularly sensitive to Green PPP projects' period.

4.3. Robustness Test. The multicolumn regression in Tables 8 and 9 has the effect of a robustness test. To further verify the models' robustness, we replaced *Subtype_level* with *Type_level* in regression models 1 and 2 and found no changes in the conclusion. We also used Corruption, Democratic Accountability, and Government Stability from the ICRG database to replace the *Anti_Corruption*, *Voice*, and *Stability* in both regression models: the regression result did not change substantially.

5. Conclusion

Our study used the dataset of 552 Green PPP projects to analyze the effects of institution quality and multilateral financial institutions on Green PPP projects' results (success or failure). We focused on African countries that have huge Green PPP projects demands for green infrastructure construction. We show that the institution quality of a Green PPP project location leads private capital to take more risks in the Green PPP project and indirectly influences the project's results. Our study is the first to connect the institution quality with the result of Green PPP projects through its internal risk structure.

Our analysis used a Probit regression model to estimate factors influencing a Green PPP project's success or failure. We found that a country with successful Green PPP project experience is associated with an increase in success rate when launching a new Green PPP project, and private capital's risk-taking degree in a Green PPP project negatively influences Green PPP project success. However, institution

quality has no statistically significant impact on Green PPP project success. To explore the role of institution quality, we introduced interaction variables between the institution quality and subtype of PPI variables into the regression model. We found that although scholars deem that institution quality influence the success of a Green PPP project, this influence is not seen directly in our empirical test samples.

Then, we used a cross-section regression model to estimate factors affecting Green PPP projects' internal risk structure. The regression results showed that institution quality has a mixed effect on Green PPP projects' internal risk structure. Voice has a negative impact on private capital's risk taking in a Green PPP project. This implies that concentration of government resources and power can, to a certain extent, increase the government's mobilization and action power, reduce uncertain risks and transaction costs of Green PPP projects, and create greater incentives for private capital. This will promote private capital's participation in Green PPP projects to a greater extent in African countries.

The regression results also found that multilateral financial institutions' participation can affect the internal risk structure of Green PPP projects. This implies that multilateral financial institutions can urge governments to ensure Green PPP projects progress smoothly and reduce the risk in private capital's participation in Green PPP projects.

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 in part by the Anhui Natural Science Fund Youth Project under Grant 2008085QG344.

References

- [1] I. Demirag, I. Khadaroo, P. Stapleton, and C. Stevenson, "Risks and the financing of PPP: perspectives from the financiers," *The British Accounting Review*, vol. 43, no. 4, pp. 294–310, 2011.
- [2] E. Iossa and D. Martimort, "Risk allocation and the costs and benefits of public-private partnerships," *The Rand Journal of Economics*, vol. 43, no. 3, pp. 442–474, 2012.
- [3] Y. Sun, Y. Sun, and X. Li, "Constructing a green financial innovation system with the PPP environmental protection industry fund," *International Journal of Technology Management*, vol. 85, no. 2/3/4, pp. 319–332, 2021.
- [4] L. Bing, A. Akintoye, P. J. Edwards, and C. Hardcastle, "The allocation of risk in PPP/PFI construction projects in the UK," *International Journal of Project Management*, vol. 23, no. 1, pp. 25–35, 2005.
- [5] A. I. Lopes and T. T. Caetano, "Firm-level conditions to engage in public-private partnerships: what can we learn?," *Journal of Economics and Business*, vol. 79, pp. 82–99, 2015.
- [6] J. B. Liu, C. Wang, S. Wang, and B. Wei, "Zagreb indices and multiplicative Zagreb indices of Eulerian graphs," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 42, no. 1, pp. 67–78, 2019.
- [7] C. Cui, Y. Liu, A. Hope, and J. Wang, "Review of studies on the public-private partnerships (PPP) for infrastructure projects," *International Journal of Project Management*, vol. 36, no. 5, pp. 773–794, 2018.
- [8] D. Martimort and S. Straub, "How to design infrastructure contracts in a warming world: a critical appraisal of public-private partnerships," *International Economic Review*, vol. 57, no. 1, pp. 61–88, 2016.
- [9] F. Corielli, S. Gatti, and A. Steffanoni, "Risk shifting through nonfinancial contracts: effects on loan spreads and capital structure of project finance deals," *Journal of Money, Credit and Banking*, vol. 42, no. 7, pp. 1295–1320, 2010.
- [10] Y. Choi, "Intermediary propositions for green growth with sustainable governance," *Sustainability*, vol. 7, no. 11, pp. 14785–14801, 2015.
- [11] J. B. Liu, X. F. Pan, L. Yu, and D. Li, "Complete characterization of bicyclic graphs with minimal Kirchhoff index," *Discrete Applied Mathematics*, vol. 200, pp. 95–107, 2016.
- [12] M. De Jong, R. Mu, D. Stead, Y. Ma, and B. Xi, "Introducing public-private partnerships for metropolitan subways in China: what is the evidence?," *Journal of Transport Geography*, vol. 18, no. 2, pp. 301–313, 2010.
- [13] C. Sharma, "Determinants of PPP in infrastructure in developing economies," *Transforming government: people, process and policy*, vol. 6, no. 2, pp. 149–166, 2012.
- [14] Z. Cheng, H. Wang, W. Xiong, D. Zhu, and L. Cheng, "Public-private partnership as a driver of sustainable development: toward a conceptual framework of sustainability-oriented PPP," *Environment, Development and Sustainability*, vol. 23, no. 1, pp. 1043–1063, 2021.
- [15] J. B. Liu, J. Zhao, H. He, and Z. Shao, "Valency-based topological descriptors and structural property of the generalized Sierpiński networks," *Journal of Statistical Physics*, vol. 177, no. 6, pp. 1131–1147, 2019.
- [16] M. Hammami, J. Ruhashyankiko, and E. B. Yehoue, "Determinants of public-private partnerships in infrastructure," *IMF Working Papers*, vol. 99, 2006.
- [17] E. Iossa and D. Martimort, "The simple microeconomics of public-private partnerships," *Journal of Public Economic Theory*, vol. 17, no. 1, pp. 4–48, 2015.
- [18] P. Galilea and F. Medda, "Does the political and economic context influence the success of a transport project? An analysis of transport public-private partnerships," *Research in Transportation Economics*, vol. 30, no. 1, pp. 102–109, 2010.
- [19] S. G. Banerjee, J. M. Oetzel, and R. Ranganathan, "Private provision of infrastructure in emerging markets: do institutions matter?," *Development Policy Review*, vol. 24, no. 2, pp. 175–202, 2006.
- [20] P. M. Panayides, F. Parola, and J. S. L. Lam, "The effect of institutional factors on public-private partnership success in ports," *Transportation Research Part A: Policy and Practice*, vol. 71, pp. 110–127, 2015.
- [21] M. Percoco, "Quality of institutions and private participation in transport infrastructure investment: evidence from

Retraction

Retracted: Optimization of UAV Cooperative Path Planning Mathematical Model Based on Personalized Multigroup Sparrow Search Algorithm in Complex Environment

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Yan, W. Liu, P. Yang, F. Wu, D. Zhu, and G. Chen, "Optimization of UAV Cooperative Path Planning Mathematical Model Based on Personalized Multigroup Sparrow Search Algorithm in Complex Environment," *Journal of Function Spaces*, vol. 2022, Article ID 2521737, 25 pages, 2022.

Research Article

Optimization of UAV Cooperative Path Planning Mathematical Model Based on Personalized Multigroup Sparrow Search Algorithm in Complex Environment

Shaoqiang Yan ¹, Weidong Liu,¹ Ping Yang ¹, Fengxuan Wu ¹, Donglin Zhu,²
and Gang Chen³

¹*Xi'an Research Institute of High Technology, Xi'an, Shaanxi 710025, China*

²*School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China*

³*College of Physics and Information Engineering, Fuzhou University, Fuzhou, Fujian 350108, China*

Correspondence should be addressed to Ping Yang; yyp_ing@163.com

Received 7 May 2022; Revised 6 June 2022; Accepted 8 June 2022; Published 29 June 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Shaoqiang 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.

Sparrow search algorithm has the problem of redundancy of convergence speed due to its fast convergence speed, and it is easy to fall into local optimum in multimodal environment. To solve the above problem, this paper presents a personalized multipopulation sparrow search algorithm (MPSSA). By introducing multiple population mechanisms to reduce the probability of falling into the local optimum due to single-population search, by using a personalized subpopulation strategy to improve the personalized differences of subpopulations and balance the exploratory ability of algorithm development, then by using weighted center-of-gravity communication strategy to improve the quality of communication between populations, and finally by using dimension by dimension dynamic reverse learning to improve the accuracy of search. The superiority of MPSSA is validated by comparing the benchmark function and CEC2017. Finally, the algorithm solves the problem of poor quality due to the dimension increase of the UAV cooperative track. MPSSA helps the UAV to quickly plan a better and stable track group to ensure the UAV to complete the cooperative task safely and stably.

1. Introduction

With the development of the times, science and technology is increasingly showing the characteristics of cross-cutting and penetration, the complexity of various engineering problems is increasing, and the calculation dimension is increasing, so the efficiency of optimization has become a major difficulty in soft computing [1–5]. As the core means to solve optimization problems, algorithm is the focus of many scholars' attention and research. For common optimization problems, there are often discontinuity, nonlinearity, nonunique variables, and constraints, and modeling is not easy. Traditional optimization algorithms such as exhaustion, integer programming, constraint planning, graph theory, etc. need to traverse the entire space to solve such problems, resulting in high time and space complexity,

space-occupying and low efficiency of the algorithm. At the same time, traditional algorithms require high objective functions and constraints and can only be used in differentiable cases. Therefore, the solution obtained by the algorithm is generally only a local optimal solution, and the accuracy of the solution cannot reach the actual requirements. As a new evolutionary computing technology, the swarm intelligence algorithm overcomes the drawbacks of the traditional optimization algorithm perfectly and enters a new stage for solving optimization problems with its own self-organization and adaptive characteristics.

The swarm intelligence optimization algorithm is a random optimization algorithm (also known as probability search algorithm) constructed to simulate the group behavior of natural organisms [6–10]. Compared with most gradient-based optimization algorithms and traditional

algorithms, the intelligence of swarm intelligence optimization algorithm is mainly because the algorithm is independent of the optimization problem itself, insensitive to initial conditions, self-organizing, and adaptive. The swarm intelligence optimization algorithm is simple in overall design, requires fewer parameters, is easy to implement, and can be processed in parallel, so it has the advantages of good fault tolerance, strong robustness and stability.

With the development of the whole intelligent algorithm system, more and more classical swarm intelligence optimization algorithms appear, such as grey wolf optimization (GWO) [11], whale optimization algorithm (WOA) [12], Harris hawks optimization (HHO) [13], salp swarm algorithm (SSA) [14], artificial bee colony algorithm (ABC) [15], firefly algorithm (FA) [16], bat algorithm (BA) [17], chicken swarm optimization (CSO) [18], bird swarm algorithm (BSA) [19], and pigeon-inspired optimization (PIO) [20]. They have become brilliant stars in the study of swarm intelligence algorithms.

The sparrow search algorithm (SSA) is a new population intelligence optimization algorithm proposed by Xue et al. based on the foraging, predatory, and antipredatory behaviors of sparrow population [21]. Like other traditional swarm intelligence optimization algorithms, the sparrow search algorithm has some drawbacks such as uneven initial population distribution, insufficient convergence ability at the end of iteration, easy to fall into local optimum, and premature stagnation [22]. Compared with the research history of classical algorithms such as the particle swarm optimization (PSO) [23] and genetic algorithm (GA) [24], the research and application of this algorithm is still in the initial stage and to be developed. With the advantages of SSA, compared with other swarm intelligence algorithms, the algorithm has faster convergence speed, stronger stability, and higher search accuracy, and it has great research potential and development prospects. At the same time, the improvement of the sparrow search algorithm's own defects, whether in theoretical research, algorithm design, or engineering practice, has the necessity and value of further research [25, 26].

With the rapid development of modern warfare and the more complex operational tasks, the cooperative operation of multiple unmanned aerial vehicles has become an inevitable trend. Collaborative track planning is based on the path planning of multiple UAVs and further considers the collaborative constraints to make the path shortest. It also brings new problems, such as the difficulty of fast convergence with increasing dimension of solution, the difficulty of high-precision time coordination of tracks, and the resolution of complex spatial conflicts. As a result, the solution of this kind of more complex optimization problem has higher requirements on the performance of the algorithm. Literature [27] proposed a parallel genetic algorithm to obtain a stable collaborative planning path through parallel computing of genetic algorithm. Literature [28] proposed an improved pigeon swarm algorithm based on social class, which uses social class strategy to accelerate the convergence of path planning. Literature [29] proposed an improved sparrow search algorithm based on logarithmic spiral strat-

egy and adaptive step size strategy, which improved the optimization quality of collaborative path planning.

In this paper, a personalized multipopulation sparrow search algorithm (MPSSA) is proposed and applied to UAV cooperative track planning. The results show that MPSSA algorithm has better optimization performance than the comparison algorithm and can help UAV safely and steadily plan better cooperative path. The main contributions of this paper are as follows:

- (1) In Section 1, the development of swarm intelligence algorithm in the context of big data is briefly introduced, and the main contributions of the article are introduced
- (2) In Section 2, the paper briefly introduces the principle and mathematical model of sparrow search algorithm
- (3) In Section 3, a variety of population mechanisms are applied to SSA, and a personalized subpopulation strategy is proposed to make different subpopulations have different parameter settings, so that each subpopulation has a different exploration ability
- (4) In Section 3, a weighted center-of-gravity communication strategy is proposed to reduce the disturbance of one population, reduce the risk that one subpopulation will fall into local optimum, which will result in all subpopulations falling into local optimum, and prevent other high-quality solutions from being ignored because of the fast convergence to the optimal solution
- (5) In Section 3, dimensional dynamic reverse learning is introduced to update the scouters' position in SSA, enhance their feeding back behavior, and reduce the shortcomings of slow convergence speed and inadequate convergence accuracy caused by the reduction of factor population;
- (6) In Section 4, MPSSA is tested and compared with six other algorithms in standard test function and CE2017 [30, 31], and the performance of these algorithms is analyzed by ranking of algorithms, Wilcoxon rank sum test [32]
- (7) In Section 5, the above algorithm is applied to the UAV track planning, further verifying the applicability of the algorithm, and MPSSA is applied to the UAV cooperative track planning to help improve the quality of the UAV cooperative track planning
- (8) In Section 6, the experimental results are summarized, and the next research content is pointed out

2. Sparrow Search Algorithm

As a group bird, sparrows are active in places where humans live. They are very active, intelligent, and have a good memory. They are bold and approachable, but they are very

vigilant. During the sparrows' feeding process, individual populations have a clear division of labor and can be divided into discoverers and participants according to their suitability for the environment. Discoverers have a high degree of environmental adaptability and need to search extensively to discover food, guide individuals to obtain food, and master the search direction of the entire population. Participants were less adaptable to the environment than the discoverers, and to improve their own fitness, they followed the discoverers to obtain food. At the same time, the sparrow population is bound to encounter various threats from the external natural environment such as natural enemies during feeding. In order to improve the survival probability, the sparrow population will randomly allocate a part of the individual as scouts, keep alert to the surrounding environment, and alert the population to flee whenever a threat is found.

In the case of sparrow populations searching for lost grains in harvested fields, the discoverer is responsible for skipping across a wide range of fields to find lost grains and stopping eating when they discover them. When the participants notice that the discoverers have found food, they jump straight to the location of the grain, grab the food with the discoverer, and eat it in competition. However, due to the limited number of grains, there is no guarantee that each sparrow will be free from hunger among the participants, so sparrows farther away from the grain (i.e., less adaptable to the environment) will give up competing with the population for food and choose to fly elsewhere to feed alone. At the same time, in order to ensure the safe feeding of the population, the sparrow population will randomly arrange a certain number of sparrows for sentinel investigation in the periphery and interior of the population, so that the whole population can escape in time in response to emergencies. This is how the sparrow population feeds.

In the process of sparrows' foraging, the discoverer-participant model was used, and the two behavioral strategies were predatory behavior and anti-predatory behavior. Individuals with better locations in the sparrow population were considered discoverers, while the remaining individuals are participants, and 20% of the individuals were randomly assigned as scouts. Predatory behavior means that the discoverer is responsible for leading the population in search direction and discovering food, while the participant follows the discoverer to seize food. Antipredatory behavior means that scouts are always vigilant against environmental threats, mainly natural enemies, and sends timely dangerous signals to alert sparrow populations to move closer to safe areas in order to prevent them from being preyed.

Set the current iteration number T and the maximum iteration number M . The current position of the i th sparrow in the j th dimension is $X_{i,j}^t$.

When the discoverer did not find the threat ($R_2 < ST$), they were responsible for guiding the population to forage and conduct extensive search. When individuals in the population have found predators (natural enemies) and issued an alarm ($R_2 \geq ST$), guide the population to the

location of the safe area. The location update is described as follows:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(\frac{-i}{\alpha \cdot M}\right), & R_2 < ST, \\ X_{i,j}^t + Q \cdot L, & R_2 \geq ST, \end{cases} \quad (1)$$

where α is a random number belonging to $[0, 1]$. $R_2 \in [0, 1]$ represents the early warning value. $ST \in [0.5, 1]$ represents the security threshold of the current environment. Q is responsible for controlling the step size, which is a random number subject to normal distribution. L is a matrix of $1 \times d$, and all elements are 1, and d represents dimension.

In order to obtain food, the participants follow and supervise the discoverer to grab food ($i \leq N/2$) or look for food alone ($i > N/2$). Therefore, the location update description of the participants is as follows:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{i^2}\right), & i > \frac{n}{2}, \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L, & \text{otherwise,} \end{cases} \quad (2)$$

where X_{worst} represents the worst position of the current population and X_p is the best position currently occupied by the discoverer. A is responsible for controlling the direction $1 \times d$ matrix, the element is only 1 or -1, and $A^+ = A^T(AA^T)^{-1}$.

When aware of the danger, the sparrow population will make anti predation behavior. When $f_i \neq f_g$, it means that the current sparrow is on the edge of the population and aware of the danger and needs to move closer to the population center to reduce the risk of predation. When $f_i = f_g$, it indicates that the sparrow in the center of the population is aware of the danger and needs to escape from its current position. The location update description of the scouts is as follows:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta \cdot |X_{i,j}^t - X_{\text{best}}^t|, & f_i \neq f_g, \\ X_{i,j}^t + K \cdot \left(\frac{|X_{i,j}^t - X_{\text{worst}}^t|}{(f_i - f_w) + \varepsilon}\right), & f_i = f_g, \end{cases} \quad (3)$$

where X_{best} represents the optimal location of the current population. β is responsible for controlling the step size, which is a random number subject to standard normal distribution. K controls the direction of sparrow movement and the moving step length. It is a random number belonging to $[-1, 1]$. f_i , f_g , and f_w represent the fitness value of the i th individual and the best and worst fitness values of the current population, respectively. To prevent the denominator from being 0, ε , take a minimal positive real number.

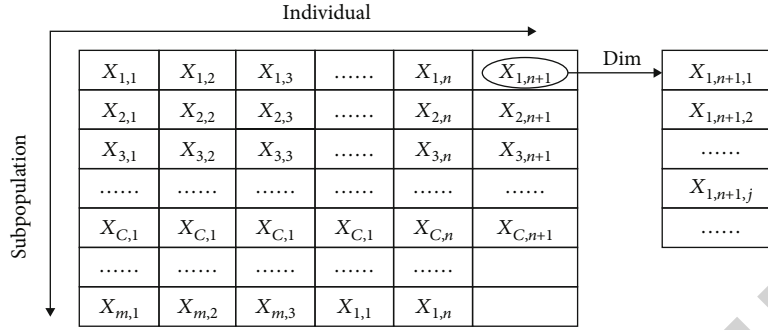


FIGURE 1: Individual distribution in multiple populations.

3. Improved Sparrow Search Algorithm

3.1. Multigroup Search Mechanism. The current SSA algorithm adopts the single-population mode to search, and all the search processes are limited to one population. Once the population falls into the local optimization and cannot escape the local optimization trap, the optimization performance of the whole algorithm will be reduced. For the SSA algorithm, the convergence speed is very fast due to its discoverer-participant model. The participants will gather to the position of the discoverer for food, and a large number of individuals will converge to the optimal position currently occupied by the discoverer (formula (2)), which will increase the risk of falling into local optimization. Moreover, for the single-population sparrow algorithm and multipopulation sparrow algorithm, the number of individuals in a single population is multiple times that of multiple populations, which will lead to the redundancy of convergence ability of the single-population sparrow algorithm and premature stagnation.

The efficiency of traditional single-population search is low. It is easy to fall into the trap of local optimization in the search space with more local optimization traps or multi-peaks, resulting in the reduction of the overall performance of the algorithm. Compared with single-population search, multipopulation search has the following advantages [33–36]:

- (1) Different parameters can be set for each population. Take the DE algorithm as an example, such as the setting of crossover probability and mutation probability, which can make each population evolve in different directions and comprehensively enhance the search ability
- (2) Each subpopulation can communicate and share. For example, the DE algorithm can set immigration operator to introduce the optimal individual into other subpopulations, so as to realize the learning from the current subpopulation to other subpopulations, which is conducive to the convergence of the algorithm
- (3) The population is not easy to fall into the local optimum. Different subpopulations search separately, and the search method is more flexible. When a pop-

ulation falls into the local optimum, it can escape the attraction of the local optimum by communicating with other subpopulations

The multipopulation search mechanism is to divide the population individual N into m subpopulations, then the individual of each subpopulation is round $(N/m) = n$, and round is a downward integer operation. If $\text{mod}(N, m) \neq 0$ and mod is the remainder operation, the remainder C is evenly divided into the first C subpopulation. In the search process, the j th dimension of the i th individual of the k th subpopulation in the t th iteration is expressed as $X_{k,i,j}^t$ at this time, $i \leq NG_k$, NG_k is the number of individuals of the k th subpopulation, and $\sum_{k=1}^m NG_k = N$, the details are shown in Figure 1.

3.2. Personalized Subpopulation Strategy. At the same time, the discoverer in the SSA algorithm is responsible for guiding the direction of the population and conducting extensive search. The participants directly jump to the current global optimal position occupied by the discoverer and conduct detailed search, so the discoverer and participant control the global search and local search of the population, respectively. At present, the general SSA algorithm sets the proportion of discoverer to individual population to be 20%, that is, the proportion of discoverer to entrant is 1:4, so SSA has strong local search ability and high search accuracy.

For multigroup SSA, we can set different discoverer proportions PD according to different populations and set the discoverer proportion of the k th population as PD_k , so that different populations have different development and exploration capabilities, which can further balance the development and exploration stages. The subpopulation with strong global search ability can more easily find the approximate location of the global optimal solution. At this time, information sharing can be carried out according to the exchange between populations, so that other subpopulations can converge, and the subpopulation with strong local search ability can be used to search the location of the global optimal solution more carefully, so as to enter the development stage and further improve the optimization accuracy.

The principle is shown in Figure 2. In (a), sparrow individuals have less proportion of discoverers and weak global search ability and can only find local optima, but the convergence speed is fast. In (b), sparrow individuals have a high

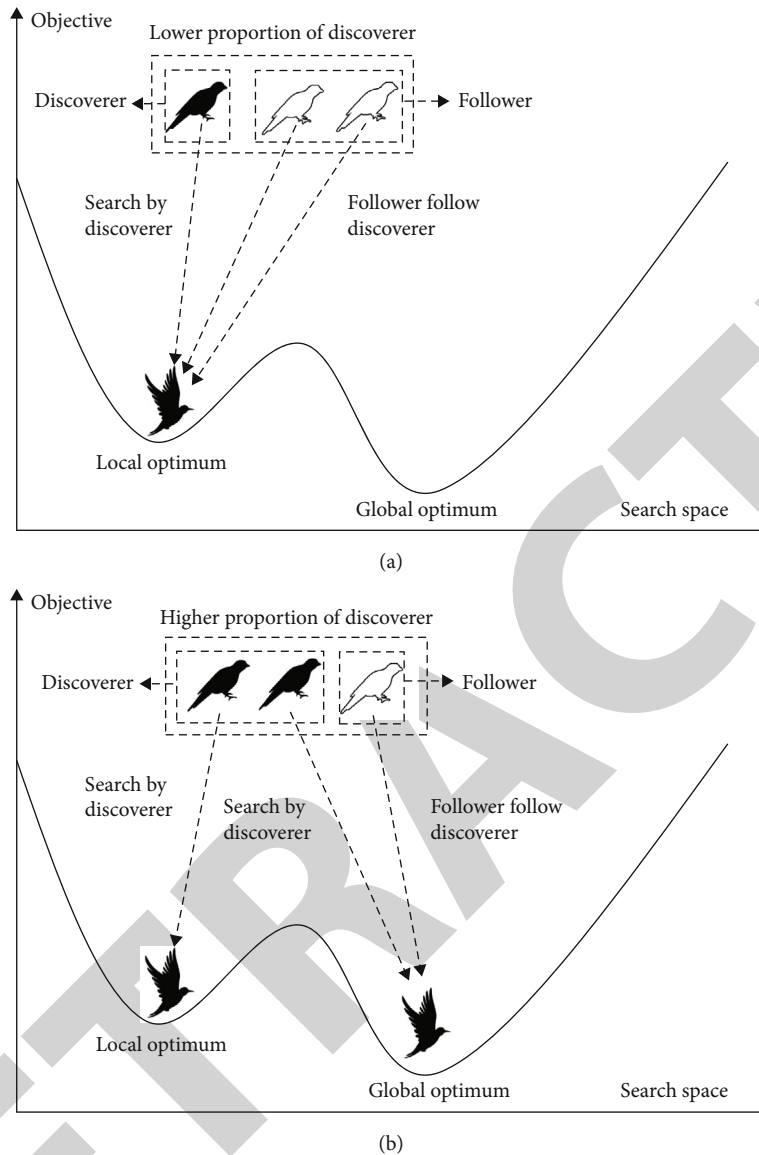


FIGURE 2: Personalized subpopulation.

proportion of discoverers and high global search ability and can find the global optimum, but the convergence speed is slow.

3.3. Weighted Center of Gravity Communication Strategy. The population is divided into several subpopulations, and each individual searches in their own subpopulation. This search method can reduce the risk of the whole population falling into local optimization. However, when the whole population size remains unchanged, the number of individuals in the subpopulation decreases and searches independently and converges to the optimal solution found by each subpopulation. Once the optimal solutions in each subpopulation are inconsistent, the solution efficiency will be reduced. Therefore, the subpopulations communicate with each other, that is, the elite individuals among the subpopulations share information, which helps the subpopulations converge to the global optimal solution.

The general population exchange mechanism is to compare the best individuals in each subpopulation [37], select the best individual X , and replace the best or worst individual in each subpopulation with this individual, so as to improve the population quality. However, once the selected optimal individual X is local optimal, all subpopulations will fall into local optimal. Another common population exchange mechanism is the random recombination strategy [38], which means that the whole population will be randomly reorganized every certain number of iterations, and each individual in the population will be divided into a new subpopulation and start a new search. This strategy greatly improves the communication between populations, but the existence probability causes a chaotic state. For example, the subpopulation individuals who have fallen into the local optimum may lead to more new subpopulations falling into the local optimum. The individuals with poor quality are divided into the same subpopulation, resulting

in the low quality of the subpopulation. Therefore, this kind of strategy has the defect of instability.

Therefore, this paper proposes a weighted center of gravity communication strategy, that is, the worst position of each subpopulation is replaced by the weighted center of gravity of the optimal position in all other subpopulations, and the weight is the proportion of the fitness of the optimal position of each subpopulation to the total fitness. It is worth noting that the subpopulation of the worst position replaced does not participate in the selection of the center of gravity. The role of this strategy is as follows:

- (1) The worst position is replaced by the weighted center of gravity of the best position of other subpopulations. First, the worst position is replaced. If the quality of the subpopulation is poor, it will help to improve the population. If the quality of the subpopulation is good, it is helpful to help the population test whether it falls into local optimization
- (2) Selecting and replacing the weighted center of gravity of the optimal position in all other subpopulations that do not include the subpopulation can reduce the interference of its own population, reduce the risk that a subpopulation falls into the local optimum and all subpopulations fall into the local optimum, and prevent the situation that the convergence speed to the optimal solution is too fast and other high-quality solutions are ignored

Taking the four subpopulations as an example, $G_1, G_2, G_3,$ and G_4 are the four populations, respectively; $W_1, W_2, W_3,$ and W_4 are the positions of the worst individuals of the four subpopulations after weighted center of gravity communication; and the black circle is the global optimum, as shown in Figure 3. In (a), because G_1 and G_4 are close to the global optimum and their weights are high, W_3 is closer to G_1 and G_4 . In (b), W_3 is located closer to G_2 because G_2 is closer to the global optimum with its higher weight. In both cases, the individual after the weighted center of gravity exchange will be closer to the global optimum, but W_2 (b) is far from the global optimum, which helps to test whether it is a local optimum.

The worst position formula for the k th subpopulation is updated as follows:

$$X_{\text{worst}}^k = \left(\frac{\sum_{j=1}^n \frac{f_g^j}{\left(\sum_{i=1}^n f_g^i\right) - f_g^k} \cdot X_{\text{best}}^j}{\left(\sum_{i=1}^n f_g^i\right) - f_g^k} \right) - \frac{f_g^j}{\left(\sum_{i=1}^n f_g^i\right) - f_g^k} \cdot X_{\text{best}}^k. \quad (4)$$

The above superscript represents the i th, j th, and k th subpopulations, which can be further simplified as

$$X_{\text{worst}}^k = \sum_{j=1}^n \frac{f_g^j}{\sum_{i=1}^n f_g^i} \cdot X_{\text{best}}^j, \quad (i, j \neq k). \quad (5)$$

3.4. Dimension by Dimension Dynamic Reverse Learning. As a result of the multiple population mechanisms dividing the

population into multiple subpopulations, the number of individuals in the subpopulation is smaller than that in the original population, which leads to a decrease in convergence rate and optimization accuracy, which is also a disadvantage of the multiple population mechanisms. The reverse learning strategy helps to accelerate the convergence speed and improve the optimization accuracy, and the solution after reverse learning can be closer to the optimal solution. General reverse learning can only find the optimal solution [39–41] in a fixed spatial search, while dimension by dimension dynamic reverse learning has better search ability than general direction learning, reduces the interference between dimensions, and can continuously converge to the optimal solution in dynamic space. At the same time, the alert has antipredatory behavior to help the population jump out of local optimum and accelerate convergence (Formula (3)). In this paper, dimension by dimension dynamic reverse learning is added to the scouters' location update to enhance the antipredatory behavior and reduce the drawbacks of slow convergence speed and inadequate convergence accuracy caused by the reduction of the number of factor population.

The principle is as follows:

$$X^* = a + b - X. \quad (6)$$

Set the original position to X and the reverse learning position to X^* . a and b are upper and lower boundaries, respectively, which is the general reverse learning strategy.

This paper expands reverse learning to each dimension and reverse learning to each dimension. The formula is expanded as follows:

$$X_{:,j}^* = a_j + b_j - X_{:,j}, \quad (7)$$

where j is the dimension, a_j is the lower boundary of the j th dimension, and b_j is the upper boundary of the j th dimension.

At the same time, the dynamic boundary is adopted in this paper:

$$a_j = \min(X_{:,j}), \quad b_j = \max(x_{:,j}), \quad (8)$$

where $\min(X_{:,j})$ is the minimum of the j th dimension in all individuals and $\max(X_{:,j})$ is the maximum of the j th dimension in all individuals.

Figure 4 is a dimension by dimension dynamic reverse learning diagram for the i th individual. The black circle is the global optimal solution, and the squares are the largest and smallest values in each dimension of the current population, that is, the upper and lower boundaries of each dimension. $(a + b)/2$ is the midpoint of each dimension boundary, and it is worth noting that in practice it is not a straight line, which is abstracted as a straight line. X is the original location and X^* is the inverse solution of X under one-dimensional dynamic reverse learning. It can be seen that the dimension by dimension dynamic reverse learning, like general reverse learning, can approach the optimal

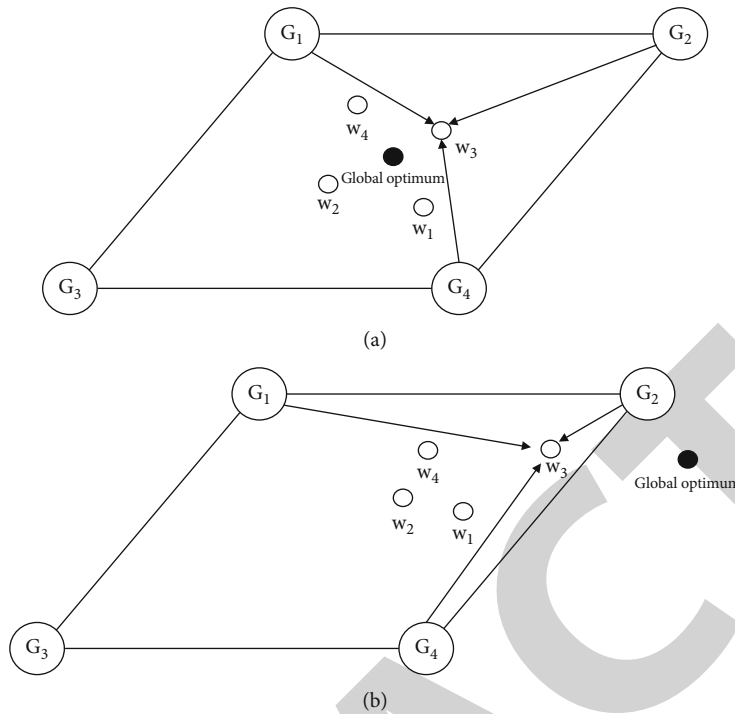


FIGURE 3: Weighted center of gravity exchange.

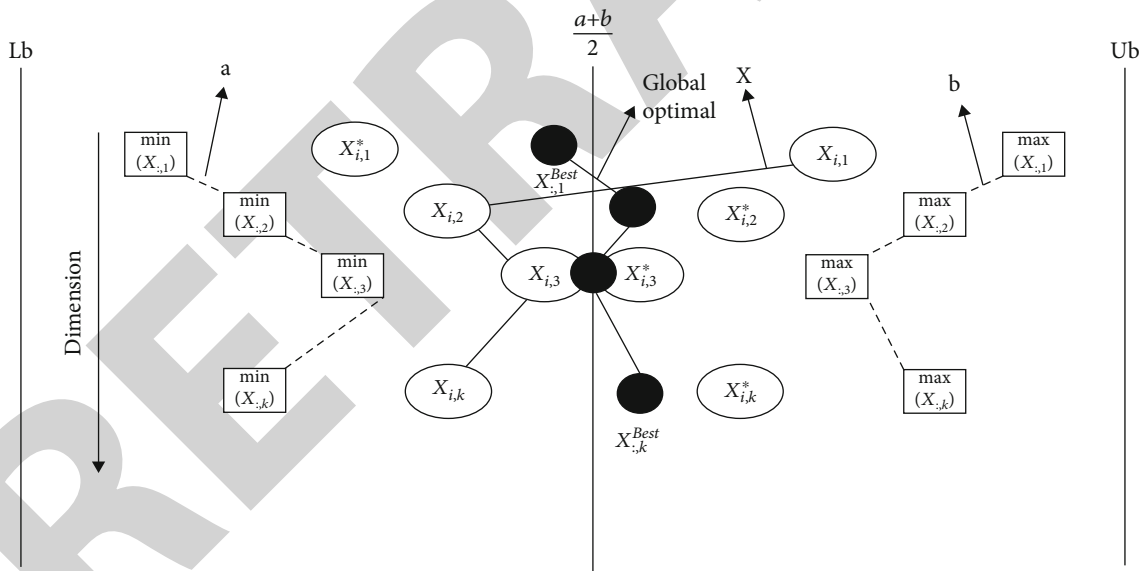


FIGURE 4: Dynamic per dimension reverse learning.

solution, accelerate the convergence process, and further eliminate the interference between dimensions so that each dimension has no influence on each other.

3.5. MPSSA Algorithm Flow. The MPSSA algorithm adopts a variety of population search mechanisms and divides the population into equal sized subpopulations as evenly as possible after initialization. We set different “discoverer-participant” ratios in different subpopulations to ensure that each algorithm has different development and exploration capabilities, increase the personalized differences among

populations, and make the search mode of populations more flexible. Then, dimension by dimension dynamic reverse learning is added when the location of the scouter is updated, which can accelerate the convergence speed of the subpopulation and have more detailed search accuracy. Finally, after all subpopulations are updated, the weighted center of gravity communication strategy is adopted to exchange among populations and transfer high-quality solutions, so as to help improve the quality of populations and jump out of local optima. The specific implementation steps are shown in Algorithm 1.

```

Set populations  $N$ 
Set the number of subpopulations  $n$ 
Set maximum number of iterations  $M$ 
Set alert value  $ST$ 
Set the proportion  $PD_k$  of discoverers in the  $k$ th subpopulation
Set proportion of scouters in the population  $SD$ 
 $t = 1$ ;
Initialize population.
Randomize the population into  $n$  equal-sized subpopulations.
While (  $t \leq M$ )
  For  $k = 1 : n$ 
    In the  $k$ th subpopulation, sort fitness values and mark the positions of the best and worst sparrows.
     $R_2 = \text{rand}(1)$ .
    For  $i = 1 : PD_k * N/n$ 
      Using formula (1) to update the location of the discoverers;
    End for
    For  $i = PD_k * N/n + 1 : N$ 
      Using formula (2) to update the location of the participants;
    End for
    For  $l = 1 : SD$ 
      Using formula (3) to update the location of the scouters;
      Update the location of the sparrows according to formula (7) and (8);
    End for
    Get the new optimal individual;
    If the new location is better than the original location, update the location;
  End for
  Update the worst position of each subpopulation according to the formula (5);
  The optimal solution of each subpopulation is compared, and the best one is selected as the optimal solution and optimal location of the whole population;
   $t = t + 1$ ;
End while
Return:  $X_{\text{best}}, f_g$ 

```

ALGORITHM 1: The framework of the MPSSA.

3.6. *Algorithm Complexity Analysis.* As shown in the above algorithm steps, MPSSA mainly performs the following operations more than SSA:

- (1) The population number n is divided into n subpopulations, and the time complexity of this task is $O(n)$
- (2) Although the location update is carried out according to n subpopulations, the time complexity of individual update is $O(N \times M \times D)$ as the original SSA because the total number remains unchanged, D is the dimension. At the same time, due to the increase of subpopulation, the sorting times before individual location update will be increased, but the same number of population remains unchanged, and the computational complexity remains unchanged
- (3) Personalized multigroup strategy only changes the initial parameters and the proportion of discoverers and participants and does not increase the complexity of the algorithm
- (4) The weighted center of gravity AC strategy updates only the worst individuals of the n subpopulations, and its algorithm complexity is $O(n \times M)$
- (5) The dimension by dimension dynamic reverse learning only operates when the scouter locations in the population are updated with a time complexity of $O((SD * N) \times M \times D)$, SD as a proportion of the population of scouters, did not add an order of magnitude to the temporal complexity of the algorithm, despite the added level of complexity

In summary, MPSSA has a computational complexity of $O(N \times D \times M)$.

4. Simulation Experiment

4.1. *The Basic Test Function.* In this paper, nine kinds of test functions were selected to carry out simulation experiments to verify the performance of MPSSA, among which F1-F3 are high-dimensional unimodal functions, F4-F6 are high-dimensional multimodal functions, F7-F9 are fixed dimensional functions, their dimensions, boundaries, theoretical minima, etc. are shown in Table 1, and the function space is shown in Figure 5. Meanwhile, PSO [23], DE [42], GWO [1], SSA [22], BSSA [25], and CSSA [26] were selected as the comparison algorithms in this paper, and the parameters of the algorithms were set as shown in Table 2.

TABLE 1: Test function [22].

	Function	Dimensions	Boundary	Min
High-dimensional single peak	$F_1(x) = \sum_{i=1}^n x_i^2$	30	[-100, 100]	0
	$F_2(x) = \sum_{j=1}^n \left(\sum_{i=1}^j x_i \right)^2$	30	[-100, 100]	0
	$F_3(x) = \sum_{i=1}^n (x_i + 0.5)^2$	30	[-100, 100]	0
High-dimensional multipeak	$F_4(x) = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i })$	30	[-500, 500]	-418.98n
	$F_5(x) = \frac{\pi}{n} \left\{ 10 \sin(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] + (y_n - 1)^2 \right\} + \sum_{i=1}^n u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{x_i + 1}{4}$ $u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m & x_i > a \\ 0 & -a < x_i < a \\ k(-x_i - a)^m & x_i < -a \end{cases}$	30	[-50, 50]	0
Fixed dimension	$F_6(x) = 0.1 \left\{ \sin^2(3\pi x_1) + \sum_{i=1}^n (x_i - 1)^2 [1 + \sin^2(3\pi x_i + 1)] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)] \right\} + \sum_{i=1}^n u(x_i, 5, 100, 4)$	30	[-50, 50]	0
	$F_7(x) = \sum_{i=1}^{11} (a_i - x_i (b_i^2 + b_1 x_2) / b_i^2 + b_1 x_3 + x_4)^2$	4	[-5, 5]	0.0003
	$F_8(x) = \sum_{i=1}^7 [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	[0, 10]	-10.4029
	$F_9(x) = \sum_{i=1}^{10} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	[0, 10]	-10.5364

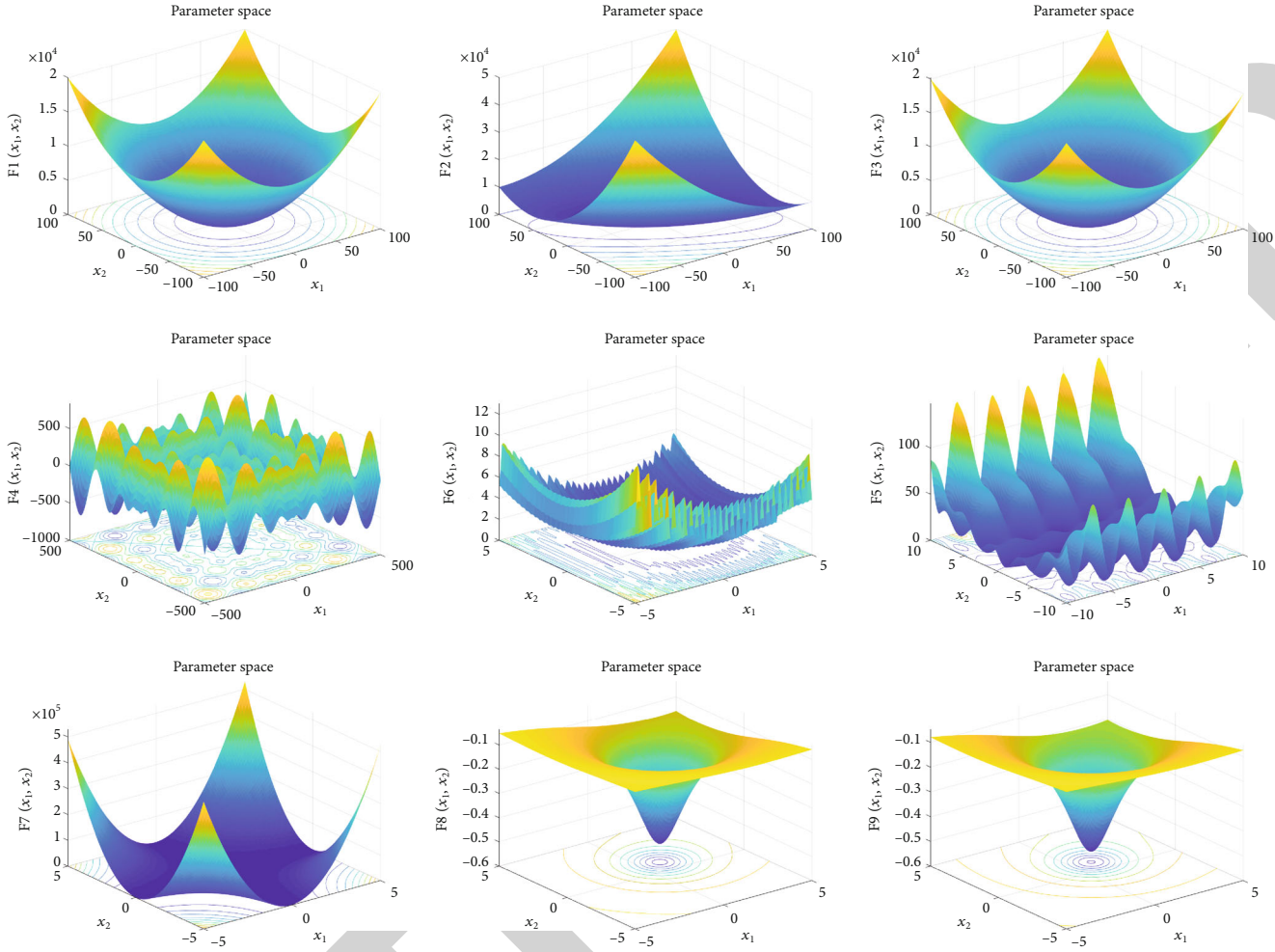


FIGURE 5: Function diagram.

TABLE 2: Algorithm parameter setting.

Algorithm	PSO	DE	GWO	SSA	BSSA	CSSA	MPSSA
Parameter	$c1 = 2$ $c2 = 2$ $W_{min} = 0.2$ $W_{max} = 0.9$	$CR=0.2$ $Fmin=0.2$ $Fmax=0.8$	$a = (2 \rightarrow 0)$	$ST = 0.8$ $PD = 0.2$ $SD = 0.2$	$ST = 0.8$ $PD = 0.2$ $SD = 0.2$	$ST = 0.8$ $PD = 0.2$ $SD = 0.2$	$n = 4$ $ST = 0.8$ $PD_k = 0.1, 0.2, 0.3, 0.4$ $SD = 0.2$

Setting the number of populations in the experiment N to 100 and the maximum number of iterations M to 500, each algorithm was run independently 30 times, and the optimal, worst, mean, standard deviation 4-term outcomes were recorded, in which the best outcome in each indicator was coarsened; then, each algorithm was ranked in different functions (depending on the average, or the standard deviation if the average is equal), and finally, the ranking results were averaged to give a total ranking, as detailed in Table 3. “A/I” means “Algorithm/Index” in Table 3. The average convergence of each algorithm in each function is shown in Figure 6.

In order to test the role of each strategy, the sparrow algorithm (SSA), the sparrow algorithm with multi group mechanism (SSA1), the sparrow algorithm with personal-

ized multigroup strategy (SSA2), the sparrow algorithm with weighted barycenter exchange mechanism based on SSA2 (Ssa3), and the sparrow algorithm with dynamic reverse learning based on SSA3 (MPSSA) are tested and simulated in $F1$, as shown in the first figure in Figure 6. The two algorithms can be used as the control group to test the necessity of the newly added mechanism or strategy. It can be seen that the performance of SSA, SSA1, SSA2, SSA3, and MPSSA shows an increasing trend, indicating that the addition of each mechanism or strategy has a certain improvement on the performance of the previous algorithm.

It can be seen from the experimental results in Table 3 that MPSSA has found the theoretical optimal solution in $F1, F2, F3, F4, F7, F8$, and $F9$, and in addition to the above functions, the optimal solution in the comparison algorithm

TABLE 3: Function experimental results.

<i>F</i>	<i>A/I</i>	PSO	DE	GWO	SSA	BSSA	CSSA	MPSSA
<i>F1</i>	Best	$5.95E-09$	$3.81E-04$	$8.30E-41$	0	0	0	0
	Worst	$1.69E-06$	$9.87E-04$	$1.19E-37$	0	0	0	0
	Ave	$1.78E-07$	$6.30E-04$	$1.16E-38$	0	0	0	0
	Std	$3.24E-07$	$1.55E-04$	$2.18E-38$	0	0	0	0
	Rank	6	7	5	1	1	1	1
<i>F2</i>	Best	$5.48E-05$	$2.02E-03$	$5.69E-23$	0	0	0	0
	Worst	$1.21E-03$	$3.73E-03$	$5.58E-22$	$7.14E-243$	0	0	0
	Ave	$4.02E-04$	$2.78E-03$	$1.92E-22$	$2.38E-244$	0	0	0
	Std	$2.96E-04$	$4.18E-04$	$1.18E-22$	0	0	0	0
	Rank	6	7	5	4	1	1	1
<i>F3</i>	Best	$3.98E-09$	$3.88E-04$	$2.97E-08$	$3.70E-32$	$1.35E-13$	$1.35E-13$	0
	Worst	$9.02E-07$	$1.20E-03$	$6.96E-01$	$5.83E-09$	$6.87E-10$	$6.87E-10$	$4.12E-16$
	Ave	$1.27E-07$	$6.73E-04$	$1.73E-01$	$3.93E-10$	$4.84E-11$	$4.84E-11$	$1.98E-16$
	Std	$1.77E-07$	$1.68E-04$	$2.03E-01$	$1.16E-09$	$1.31E-10$	$1.31E-10$	$1.50E-16$
	Rank	5	6	7	4	2	3	1
<i>F4</i>	Best	$-8.32E+03$	$-9.69E+03$	$-8.12E+03$	$-1.26E+04$	$-1.26E+04$	$-1.26E+04$	$-1.26E+04$
	Worst	$-4.33E+03$	$-8.78E+03$	$-3.53E+03$	$-7.25E+03$	$-8.41E+03$	$-8.41E+03$	$-1.07E+04$
	Ave	$-6.72E+03$	$-9.21E+03$	$-6.38E+03$	$-1.09E+04$	$-1.15E+04$	$-1.15E+04$	$-1.21E+04$
	Std	$8.75E+02$	$2.23E+02$	$8.32E+02$	$1.72E+03$	$8.14E+02$	$8.14E+02$	$5.15E+02$
	Rank	6	5	7	4	2	3	1
<i>F5</i>	Best	$1.55E-10$	$3.65E-05$	$1.20E-08$	$1.57E-32$	$2.48E-15$	$2.48E-15$	$1.57E-32$
	Worst	$1.04E-01$	$1.83E-04$	$4.67E-02$	$2.95E-10$	$7.65E-12$	$7.65E-12$	$4.71E-11$
	Ave	$3.46E-03$	$8.86E-05$	$1.57E-02$	$3.45E-11$	$1.46E-12$	$1.46E-12$	$1.57E-12$
	Std	$1.89E-02$	$3.57E-05$	$1.20E-02$	$6.69E-11$	$2.14E-12$	$2.14E-12$	$8.60E-12$
	Rank	6	5	7	4	2	3	1
<i>F6</i>	Best	$1.60E-09$	$2.66E-04$	$1.25E-07$	$2.56E-14$	$1.60E-14$	$1.60E-14$	$1.35E-32$
	Worst	$1.10E-02$	$6.02E-04$	$4.91E-01$	$5.55E-09$	$7.01E-10$	$7.01E-10$	$6.55E-17$
	Ave	$7.91E-04$	$4.01E-04$	$1.37E-01$	$9.16E-10$	$6.23E-11$	$6.23E-11$	$1.09E-17$
	Std	$2.79E-03$	$8.12E-05$	$1.29E-01$	$1.58E-09$	$1.38E-10$	$1.38E-10$	$2.48E-17$
	Rank	6	5	7	4	2	3	1
<i>F7</i>	Best	$3.61E-04$	$3.56E-04$	$3.07E-04$	$3.07E-04$	$3.07E-04$	$3.07E-04$	$3.07E-04$
	Worst	$1.07E-03$	$7.80E-04$	$2.04E-02$	$3.32E-04$	$3.07E-04$	$3.07E-04$	$3.07E-04$
	Ave	$7.78E-04$	$6.51E-04$	$2.40E-03$	$3.08E-04$	$3.07E-04$	$3.07E-04$	$3.07E-04$
	Std	$2.05E-04$	$8.89E-05$	$6.09E-03$	$4.39E-06$	$1.31E-10$	$1.31E-10$	$6.59E-11$
	Rank	6	5	7	4	2	3	1
<i>F8</i>	Best	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$
	Worst	$-5.09E+00$	$-1.04E+01$	$-5.09E+00$	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$
	Ave	$-9.17E+00$	$-1.04E+01$	$-1.00E+01$	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$	$-1.04E+01$
	Std	$2.27E+00$	$6.71E-04$	$1.35E+00$	$1.96E-05$	$9.33E-16$	$9.33E-16$	$1.87E-15$
	Rank	7	5	6	4	1	2	3
<i>F9</i>	Best	$-1.05E+01$	$-1.05E+01$	$-1.05E+01$	$-1.05E+01$	$-1.05E+01$	$-1.05E+01$	$-1.05E+01$
	Worst	$-5.13E+00$	$-1.05E+01$	$-1.05E+01$	$-5.13E+00$	$-1.05E+01$	$-1.05E+01$	$-1.05E+01$
	Ave	$-1.02E+01$	$-1.05E+01$	$-1.05E+01$	$-1.04E+01$	$-1.05E+01$	$-1.05E+01$	$-1.05E+01$
	Std	$1.37E+00$	$1.54E-07$	$6.24E-08$	$9.87E-01$	$2.29E-15$	$2.29E-15$	$5.14E-15$
	Rank	7	4	5	6	2	3	1
Total rank		6.11	5.44	6.22	3.89	1.67	2.44	1.22

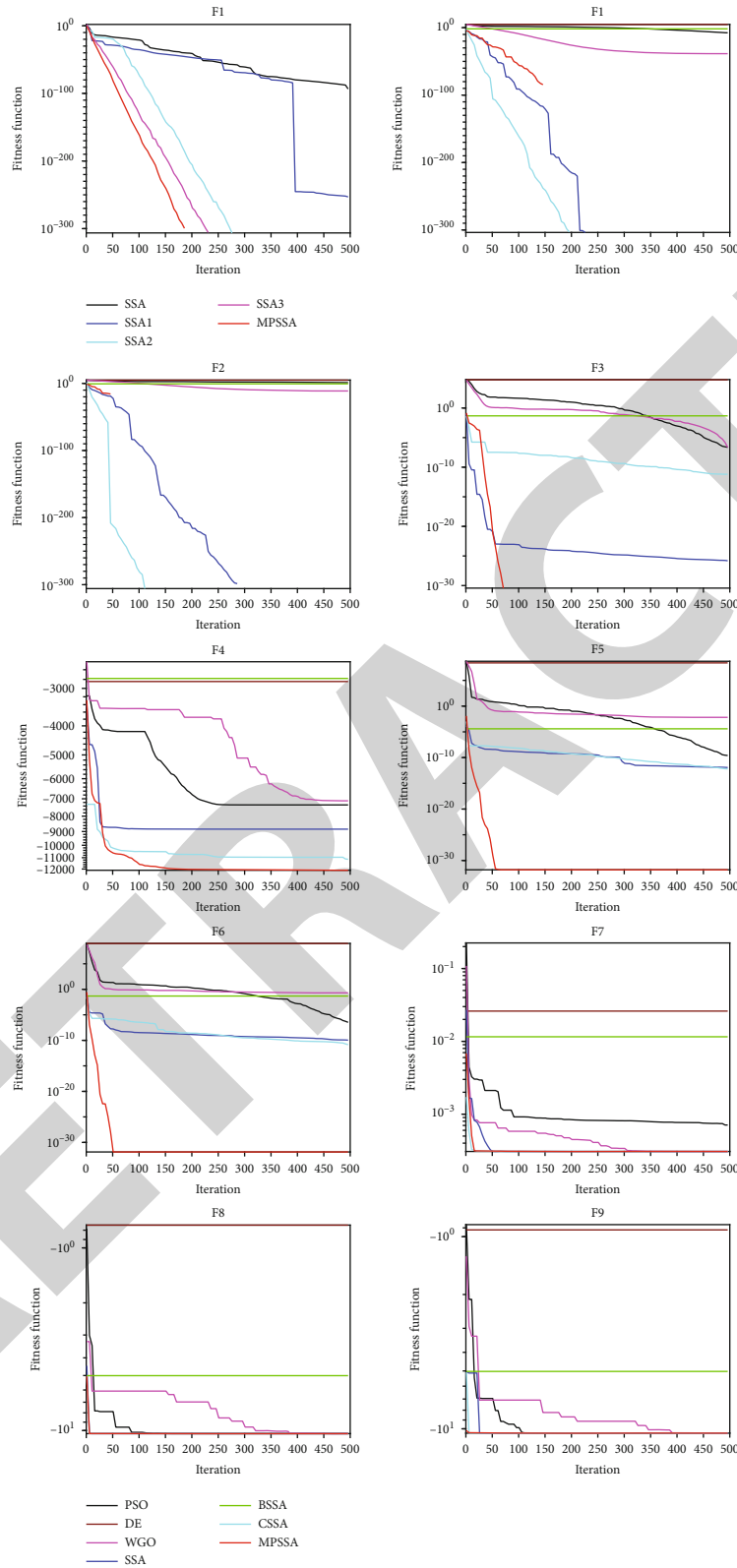


FIGURE 6: Comparison of algorithm convergence.

is found in $F5$ and $F6$, indicating that MPSSA algorithm has strong optimization performance. In terms of function types, in high-dimensional unimodal functions, MPSSA has strong search ability with other algorithms such as SSA, BSSA, and

CSSA in $F1$ and $F2$, and only MPSSA has found the theoretical optimal solution in $F3$. In the high-dimensional multimodal function, MPSSA ranks first in all indicators, and its overall performance is better than other algorithms, which

is also the advantage of multiple swarm algorithms in the case of multimodal. In the fixed dimension function, MPSSA has little difference from other SSA improved algorithms, only a small difference in standard deviation.

From the convergence of each algorithm in Figure 6, in the high-dimensional unimodal function, the convergence speed of MPSSA is not fast, but due to its superior optimization performance, it can directly find the optimal value and end the convergence. In the high-dimensional multimodal function, the advantage of MPSSA is highlighted. The convergence speed is significantly faster than all comparison algorithms, and it does not fall into local optimization due to the attraction of multimodal.

To sum up, PSO, DE, and GWO perform poorly in the standard test function, which is significantly worse than the SSA and SSA improved algorithms. Moreover, MPSSA has a strong search ability in the basic test function, ranking first, followed by BSSA, CSSA, and SSA. Especially in the case of high-dimensional and multi peak, MPSSA is outstanding.

4.2. CEC 2017. In order to test the general adaptability of the algorithm and prevent the randomness of the selected function in the above experiment, it is necessary to select a more complex test function, that is, the CEC 2017 test function used in the international algorithm competition, to test the above algorithm. Due to the defect of F2 [43], the article will not adopt it. The evaluation times are set to $1000 \cdot \text{dim}$, and other parameters are consistent with the above experiments. The experimental results after 30 times of operation are shown in Table 4.

From the results in Table 4, MPSSA finds the optimal value in all algorithms in $F1, F3, F4, F12, F15, F18, F22,$ and $F28$, and the average value is the closest to the theoretical optimal value in $F3, F4, F12, F14, F18, F22, F23,$ and $F28$. Among the 29 test functions, MPSSA ranks in the top three except $F7, F15, F20, F25,$ and $F30$. In the comprehensive ranking, MPSSA ranks first, followed by De, PSO, GWO, CSSA, BSSA, and SSA.

The values in Table 5 represent the p value of Wilcoxon rank sum test for the results of 30 operations. When $p < 0.05$, it can be considered that MPSSA is significantly different from the algorithm, “+”, “=”, and “-”, respectively, represent that MPSSA is superior, equal, and inferior to the comparison algorithm [44]. It can be seen that most algorithms are inferior to the MPSSA algorithm, which shows that MPSSA is significantly different from other algorithms, and the performance is better than other algorithms, which shows that MPSSA has better advantages than other algorithms.

5. UAV Cooperative Track Planning

Track planning is a problem of path optimization in continuous space, which is mainly divided into single track planning [45] and cooperative track planning [46].

5.1. Single UAV Track Planning Model

5.1.1. Flight Fuel Cost. In the actual combat mission, the track length can reflect the fuel consumption. Let L_i be the

track length of i th segment and N be the number of track points [47]. That is, the fuel consumption cost of UAV flight can be expressed as the track length:

$$f_{\text{path}} = \sum_{i=1}^N L_i. \quad (9)$$

5.1.2. Flight Altitude Change Cost. In order to avoid radar search and prevent collision with mountains, the UAV must adjust the flight altitude in time. A stable flight altitude can reduce the burden of the control system and save more fuel. The variance of track altitude change can describe the stability of flight altitude, which can be expressed as

$$f_{\text{height}} = \frac{\sum_{i=1}^N \left(z_i - \frac{1}{N} \sum_{i=1}^N z_i \right)^2}{N}, \quad (10)$$

where z_i is the Z coordinate of the i th track point.

5.1.3. Smoothing Cost. The smaller the deflection angle during flight, the more stable the flight state of UAV and the smoother the flight path. The smoothing cost can be expressed by the change degree of deflection angle δ , and the function is set as follows:

$$f_{\text{smooth}} = \sum_{i=1}^N |\delta_i - \delta_{i-1}|. \quad (11)$$

5.1.4. Comprehensive Threat Constraint. When passing through the enemy area, UAV will encounter the enemy's air defense system, including detection radar, air defense antiaircraft gun, ground to air missile, and other threats. The above threats are approximately regarded as a cylindrical area on the three-dimensional plane, and the detection range or attack range is taken as its radius R . The current track segment L_i is divided into five segments, M represents the comprehensive threat, a total of n_M comprehensive threats are set, k_M represents the current k th comprehensive threat, R_{kM} represents the radius of the current threat, and $d_{k,i}$ represents the distance from the current threat point to the five equally divided segments. The threat constraint principle of the current comprehensive threat point to track segment L_i is shown in Figure 7, and its threat constraint function is set as follows:

$$\text{Constraint} = \sum_{i=1}^{5N} \sum_{k_M=1}^{n_M} \max(R_{k_M} - d_{k_M,i}, 0). \quad (12)$$

Therefore, the cost function of UAV with single track is

$$\min(\text{fitness}) = \omega_1 \cdot f_{\text{path}} + \omega_2 \cdot f_{\text{height}} + \omega_3 \cdot f_{\text{smooth}} + \eta \cdot \text{Constraint}. \quad (13)$$

5.2. Multi-UAV Collaborative Planning Model. In the process of collaborative track planning, it is necessary to plan multiple candidate tracks that meet the flight constraints of

TABLE 4: CEC 2017 [30] experimental results

(a)

F	F1			F3			F4			F5		
	Best	Ave	Std	Rank	Best	Ave	Std	Rank	Best	Ave	Std	Rank
PSO	2.79E+02	4.44E+03	2.55E+03	6	3.00E+02	3.00E+02	5.62E-02	2	4.13E+02	4.72E+02	1.88E+01	3
DE	1.28E+02	9.52E+02	7.60E+02	1	6.60E+04	8.11E+04	7.30E+03	7	4.87E+02	4.89E+02	1.19E+00	6
GW0	4.89E+07	1.06E+09	8.07E+08	7	8.00E+03	2.73E+04	9.92E+03	5	4.94E+02	5.31E+02	3.32E+01	7
SSA	1.12E+02	3.17E+03	4.07E+03	3	2.25E+04	3.84E+04	1.43E+04	6	4.71E+02	4.89E+02	1.71E+01	4
BSSA	2.25E+02	3.96E+03	5.44E+03	5	3.00E+02	3.03E+02	2.70E+00	3	4.63E+02	4.89E+02	2.14E+01	5
DSSA	1.07E+02	3.54E+03	3.86E+03	4	3.00E+02	4.67E+02	1.57E+02	4	4.00E+02	4.53E+02	4.04E+01	2
MPSSA	1.02E+02	1.40E+03	1.75E+03	2	3.00E+02	3.00E+02	2.60E-02	1	4.00E+02	4.33E+02	3.58E+01	1

(b)

F	F6			F7			F8			F9		
	Best	Ave	Std	Rank	Best	Ave	Std	Rank	Best	Ave	Std	Rank
PSO	6.19E+02	6.31E+02	5.36E+00	4	7.72E+02	8.39E+02	2.98E+01	1	8.72E+02	9.05E+02	1.93E+01	3
DE	6.00E+02	6.00E+02	1.06E-13	1	8.42E+02	8.66E+02	6.98E+00	3	9.10E+02	9.31E+02	8.38E+00	5
GW0	6.01E+02	6.03E+02	1.26E+00	2	7.83E+02	8.50E+02	5.34E+01	2	8.55E+02	8.81E+02	2.13E+01	1
SSA	6.50E+02	6.64E+02	7.56E+00	7	1.21E+03	1.32E+03	4.10E+01	7	9.48E+02	1.01E+03	2.61E+01	7
BSSA	6.46E+02	6.55E+02	4.79E+00	6	1.23E+03	1.32E+03	3.03E+01	6	9.19E+02	9.91E+02	2.00E+01	6
DSSA	6.20E+02	6.35E+02	1.27E+01	5	9.50E+02	1.11E+03	8.63E+01	5	8.80E+02	9.25E+02	2.66E+01	4
MPSSA	6.00E+02	6.06E+02	3.43E+00	3	8.19E+02	8.94E+02	8.63E+01	4	8.65E+02	9.03E+02	2.70E+01	2

(c)

F	F10			F11			F12			F13		
	Best	Ave	Std	Rank	Best	Ave	Std	Rank	Best	Ave	Std	Rank
PSO	2.60E+03	4.46E+03	7.00E+02	1	1.15E+03	1.19E+03	2.27E+01	1	1.17E+04	4.89E+04	3.51E+04	2
DE	5.73E+03	6.31E+03	2.65E+02	6	1.20E+03	1.21E+03	6.61E+00	2	1.55E+06	5.86E+06	2.05E+06	6
GW0	3.22E+03	4.53E+03	1.17E+03	2	1.21E+03	1.32E+03	1.06E+02	7	9.91E+05	2.30E+07	2.51E+07	7
SSA	5.04E+03	6.38E+03	1.13E+03	7	1.15E+03	1.24E+03	3.69E+01	4	3.11E+04	1.10E+06	1.18E+06	5
BSSA	4.40E+03	5.56E+03	4.94E+02	4	1.16E+03	1.25E+03	6.11E+01	5	2.76E+04	5.92E+05	5.12E+05	4
DSSA	4.31E+03	5.60E+03	7.00E+02	5	1.18E+03	1.26E+03	6.49E+01	6	1.56E+04	1.41E+05	7.75E+04	3
MPSSA	3.70E+03	4.57E+03	5.91E+02	3	1.15E+03	1.22E+03	4.17E+01	3	6.84E+03	3.36E+04	1.16E+04	1

(d)

F	F14			F15			F16			F17		
	Best	Ave	Std	Rank	Best	Ave	Std	Rank	Best	Ave	Std	Rank
PSO	1.50E+03	6.20E+03	3.59E+03	2	1.72E+03	8.20E+03	8.88E+03	3	2.13E+03	2.63E+03	2.42E+02	4
DE	1.75E+04	6.05E+04	2.88E+04	5	5.18E+03	1.92E+04	1.18E+04	6	1.79E+03	2.15E+03	1.78E+02	1
GW0	2.69E+03	1.26E+05	2.56E+05	7	1.37E+04	5.07E+04	3.26E+04	7	1.73E+03	2.16E+03	2.86E+02	2
SSA	4.71E+03	7.26E+04	1.82E+05	6	2.02E+03	7.11E+03	5.15E+03	1	2.87E+03	3.65E+03	7.78E+02	7
BSSA	3.46E+03	1.27E+04	5.00E+03	4	1.82E+03	8.63E+03	8.59E+03	4	2.39E+03	2.90E+03	2.64E+02	6
DSSA	1.82E+03	6.59E+03	5.09E+03	3	2.00E+03	7.38E+03	9.86E+03	2	2.07E+03	2.84E+03	3.20E+02	5
MPSSA	1.76E+03	5.13E+03	2.62E+03	1	1.61E+03	8.84E+03	7.48E+03	5	2.13E+03	2.56E+03	3.29E+02	3

(e)

F	F18			F19			F20			F21		
	Best	Ave	Std	Rank	Best	Ave	Std	Rank	Best	Ave	Std	Rank
PSO	4.41E+04	1.51E+05	1.56E+05	4	2.11E+03	1.34E+04	9.97E+03	5	2.28E+03	2.60E+03	1.46E+02	5
DE	1.15E+05	3.98E+05	1.94E+05	5	1.25E+04	2.04E+04	7.83E+03	6	2.07E+03	2.17E+03	6.17E+01	1
GW0	4.47E+04	4.79E+05	6.51E+05	6	1.06E+04	6.04E+05	6.97E+05	7	2.10E+03	2.30E+03	1.33E+02	2
SSA	2.34E+04	8.34E+05	3.96E+06	7	2.22E+03	1.12E+04	8.28E+03	4	2.33E+03	2.87E+03	2.53E+02	7
BSSA	1.26E+04	8.34E+04	6.04E+04	2	2.12E+03	9.39E+03	1.27E+04	3	2.44E+03	2.69E+03	1.27E+02	6
DSSA	2.90E+04	1.06E+05	5.57E+04	3	2.07E+03	5.81E+03	4.91E+03	1	2.17E+03	2.34E+03	1.59E+02	3
MPSSA	1.13E+04	4.68E+04	3.06E+04	1	2.84E+03	7.96E+03	8.10E+03	2	2.09E+03	2.44E+03	2.12E+02	4

(f)

F	F22			F23			F24			F25		
	Best	Ave	Std	Rank	Best	Ave	Std	Rank	Best	Ave	Std	Rank
PSO	2.30E+03	3.35E+03	1.77E+03	4	2.87E+03	3.05E+03	9.78E+01	5	2.95E+03	3.15E+03	7.74E+01	5
DE	2.82E+03	3.24E+03	2.96E+02	3	2.76E+03	2.78E+03	6.60E+00	3	2.95E+03	2.98E+03	1.04E+01	3
GW0	2.35E+03	3.48E+03	1.50E+03	5	2.69E+03	2.76E+03	5.73E+01	2	2.86E+03	2.91E+03	6.06E+01	1
SSA	2.31E+03	7.28E+03	1.18E+03	7	2.96E+03	3.21E+03	1.68E+02	7	3.11E+03	3.40E+03	1.20E+02	7
BSSA	2.30E+03	5.91E+03	2.44E+03	6	2.91E+03	3.09E+03	7.70E+01	6	3.08E+03	3.26E+03	9.03E+01	6
DSSA	2.30E+03	2.51E+03	7.99E+02	2	2.80E+03	2.89E+03	5.85E+01	4	2.96E+03	3.05E+03	8.57E+01	4
MPSSA	2.30E+03	2.46E+03	8.66E+02	1	2.71E+03	2.75E+03	3.17E+01	1	2.88E+03	2.95E+03	3.55E+01	2

(g)

F	A/I	F26			F27			F28			F29						
		Best	Ave	Std	Rank	Best	Ave	Std	Rank	Best	Ave	Std	Rank				
PSO		2.80E+03	5.66E+03	1.78E+03	4	3.15E+03	3.34E+03	1.88E+02	5	3.10E+03	3.21E+03	5.50E+01	4	3.49E+03	3.86E+03	2.05E+02	4
DE		4.74E+03	4.88E+03	9.30E+01	3	3.20E+03	3.21E+03	1.79E+00	1	3.22E+03	3.23E+03	1.08E+01	5	3.53E+03	3.64E+03	6.84E+01	1
GW0		4.06E+03	4.45E+03	1.67E+02	1	3.21E+03	3.23E+03	1.41E+01	2	3.28E+03	3.34E+03	4.10E+01	7	3.39E+03	3.65E+03	8.43E+01	2
SSA		7.49E+03	8.83E+03	9.13E+02	7	3.39E+03	3.63E+03	1.83E+02	7	3.20E+03	3.23E+03	3.03E+01	6	4.08E+03	5.58E+03	8.03E+02	7
BSSA		4.01E+03	7.63E+03	9.11E+02	6	3.27E+03	3.49E+03	1.46E+02	6	3.10E+03	3.16E+03	5.63E+01	3	3.99E+03	4.48E+03	4.84E+02	6
DSSA		2.90E+03	5.75E+03	6.14E+02	5	3.22E+03	3.30E+03	8.42E+01	4	3.10E+03	3.13E+03	5.74E+01	1	3.56E+03	4.15E+03	3.09E+02	5
MPSSA		2.90E+03	4.69E+03	7.16E+02	2	3.21E+03	3.24E+03	2.04E+01	3	3.10E+03	3.14E+03	6.68E+01	2	3.49E+03	3.79E+03	2.23E+02	3

(h)

F	A/I	F30			Total rank		
		Best	Ave	Std	Rank	Std	Rank
PSO		3.38E+03	5.64E+03	2.50E+03	1		3.43
DE		1.44E+04	2.63E+04	8.94E+03	3		3.4
GW0		1.05E+06	3.28E+06	1.75E+06	7		4.13
SSA		1.25E+04	3.28E+04	1.52E+04	5		5.97
BSSA		7.21E+03	3.48E+04	1.09E+05	6		4.97
DSSA		5.29E+03	9.78E+03	3.60E+03	2		3.67
MPSSA		5.93E+03	2.76E+04	1.07E+05	4		2.43

TABLE 5: Wilcoxon rank sum test.

	PSO	DE	GWO	SSA	BSSA	CSSA
<i>F1</i>	$9.99E-07$	$8.42E-01$	$2.93E-11$	$5.29E-03$	$3.50E-02$	$1.07E-02$
<i>F3</i>	$9.59E-09$	$2.71E-11$	$2.79E-11$	$2.83E-11$	$4.70E-11$	$2.38E-11$
<i>F4</i>	$4.93E-05$	$6.44E-11$	$2.42E-11$	$6.65E-08$	$6.92E-07$	$4.54E-01$
<i>F5</i>	$1.00E-03$	$5.38E-01$	$1.52E-07$	$2.37E-11$	$2.51E-11$	$9.10E-07$
<i>F6</i>	$2.61E-11$	$4.55E-12$	$4.80E-06$	$2.62E-11$	$2.56E-11$	$2.60E-11$
<i>F7</i>	$1.76E-04$	$2.83E-01$	$2.33E-04$	$3.86E-11$	$3.46E-11$	$1.32E-09$
<i>F8</i>	$8.18E-01$	$1.20E-05$	$7.84E-03$	$2.75E-11$	$7.89E-11$	$3.97E-03$
<i>F9</i>	$3.75E-02$	$2.15E-11$	$1.93E-06$	$1.34E-10$	$3.59E-10$	$5.43E-07$
<i>F10</i>	$7.96E-01$	$2.55E-11$	$1.11E-01$	$2.61E-10$	$5.98E-09$	$3.94E-07$
<i>F11</i>	$1.97E-05$	$7.05E-01$	$1.67E-06$	$2.63E-02$	$1.99E-01$	$2.38E-01$
<i>F12</i>	$2.57E-01$	$2.69E-11$	$2.66E-11$	$1.89E-10$	$1.95E-08$	$6.33E-09$
<i>F13</i>	$1.45E-01$	$7.87E-09$	$4.50E-06$	$4.82E-01$	$3.62E-01$	$3.54E-01$
<i>F14</i>	$2.21E-01$	$2.78E-11$	$2.36E-08$	$3.26E-09$	$3.15E-08$	$3.70E-01$
<i>F15</i>	$6.84E-01$	$9.69E-05$	$1.62E-09$	$6.20E-01$	$9.23E-01$	$2.05E-01$
<i>F16</i>	$1.26E-01$	$7.32E-07$	$1.94E-05$	$5.05E-09$	$1.14E-04$	$4.36E-04$
<i>F17</i>	$8.19E-02$	$9.19E-10$	$1.22E-06$	$1.77E-10$	$7.57E-11$	$1.73E-06$
<i>F18</i>	$2.25E-07$	$3.01E-11$	$4.71E-09$	$7.33E-06$	$2.58E-02$	$2.18E-06$
<i>F19</i>	$4.01E-02$	$2.88E-08$	$9.17E-11$	$3.24E-01$	$2.50E-01$	$1.44E-01$
<i>F20</i>	$1.15E-02$	$2.47E-07$	$7.22E-03$	$4.84E-08$	$7.04E-06$	$6.10E-02$
<i>F21</i>	$6.18E-10$	$2.77E-11$	$6.09E-01$	$2.80E-11$	$2.69E-11$	$8.12E-03$
<i>F22</i>	$3.78E-03$	$4.87E-10$	$4.81E-10$	$6.36E-11$	$3.40E-09$	$4.29E-01$
<i>F23</i>	$2.72E-11$	$5.17E-06$	$7.28E-01$	$2.63E-11$	$2.74E-11$	$4.80E-11$
<i>F24</i>	$2.43E-10$	$1.61E-05$	$4.86E-05$	$2.88E-11$	$2.65E-11$	$2.51E-07$
<i>F25</i>	$6.05E-05$	$1.61E-06$	$2.81E-05$	$1.08E-01$	$5.58E-01$	$2.54E-05$
<i>F26</i>	$4.79E-03$	$2.79E-02$	$5.05E-04$	$2.79E-11$	$7.52E-10$	$5.99E-09$
<i>F27</i>	$9.94E-01$	$4.40E-09$	$5.68E-02$	$2.59E-11$	$3.29E-11$	$1.75E-05$
<i>F28</i>	$9.85E-04$	$6.54E-05$	$2.83E-11$	$2.38E-05$	$1.14E-03$	$9.73E-03$
<i>F29</i>	$1.70E-01$	$9.68E-03$	$1.25E-01$	$4.04E-11$	$1.05E-09$	$2.60E-05$
<i>F30</i>	$1.40E-05$	$1.46E-09$	$2.84E-11$	$1.00E-09$	$3.63E-07$	$1.42E-02$
+/-/-	10/0/19	4/0/25	5/0/24	4/0/25	5/0/24	9/0/21

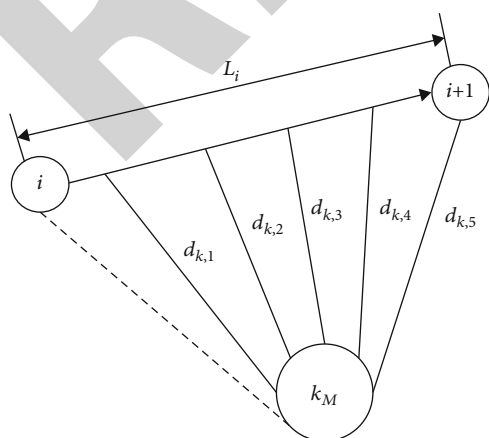


FIGURE 7: Comprehensive threat constraint.

single UAV for each UAV at the single-plane planning level in advance and then establish the cooperative constraint relationship to plan the cooperative track of multiple UAVs. Cooperative constraints mainly include spatial cooperative constraints and temporal cooperative constraints among multiple UAVs. Spatial cooperative constraint means to avoid collision between multiple UAVs while meeting the track planning of a single UAV. Time cooperative constraint means that multiple UAVs can reach the specified target point at the same time or within a certain time difference.

The specific operations of multi-UAV cooperative track planning are as follows: firstly, generate multiple candidate tracks of each UAV, then calculate the time range of reaching the target point according to the minimum and maximum flight speed and track length of each UAV, and calculate the time intersection of reaching the target area

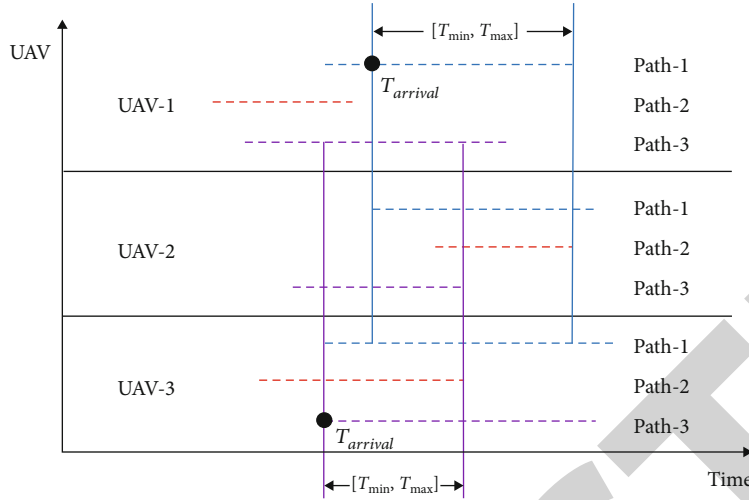


FIGURE 8: Time collaboration constraint.

between different UAVs, then determine the cooperative time range, and finally select the flight scheme corresponding to the minimum cooperative cost that meets the constraints.

5.2.1. Spatial Cooperative Constraint. When large UAVs perform combat tasks in the modern battlefield, it is necessary to specify that the distance of each UAV at the same time should be greater than the safe distance D , so as to avoid the damage caused by the collision between UAVs and ensure the UAVs to complete the combat tasks safely. Assuming that the European distance between the i th UAV and the j th UAV at a certain time is $D_{i,j}$, where $i, j \in [1, n]$ and n are the number of UAVs, the space constraints that multiple UAVs should meet at this time are as follows.

For $\forall i, j \in [1, n]$, there is $D - D_{i,j} \leq 0$. Then, set the spatial collaboration constraint function as follows:

$$\text{Constraint}_{\text{space}} = \sum_{i=1}^n \sum_{j=1}^n \max(D - D_{i,j}, 0). \quad (14)$$

5.2.2. Time Cooperative Constraint. When multiple UAVs perform tasks cooperatively, if they cannot reach the task target area at the same time, it will not only reduce the cooperative work efficiency of UAVs but also increase the probability that a single UAV will be destroyed, resulting in the reduction of task completion rate. Therefore, it is necessary to restrict the time when UAVs arrive at the target area, that is, set the cooperative time, that is, meet the intersection of the time intervals when all UAVs arrive at the target area. When the intersection is not empty, it indicates that there is a period of time to enable the UAV to reach the target point at the same time by adjusting the flight speed [28].

There are n UAVs in total, and each UAV has m alternative tracks. The j th flight path of the i th UAV is $L_{i,j}$, and the speed of each UAV is $v_i \in [v_{\min}, v_{\max}]$; then, the flight time range of the j th flight path of the i th UAV reaching the target area is $T_{i,j} = [L_{i,j}/v_{\max}, L_{i,j}/v_{\min}] = [T_{i,j}^{\min}, T_{i,j}^{\max}]$, $T_{i,j}^{\min}$, and $T_{i,j}^{\max}$ respectively, represent the shortest and longest time

of the j th flight path of the i th UAV reaching the target area, then the time range of the i th UAV is $T_i = T_{i,1} \cup T_{i,2} \cup \dots \cup T_{i,m}$; then, the requirement is $T = \bigcap_{i=1}^N T_i \neq \emptyset$; then, there is a feasible solution in the track alternative group of the current UAV.

Let T_i^{\min} and T_i^{\max} , respectively, be the shortest and longest time for a track selected from the candidate track group of the i UAV to reach the target area. When the intersection of the time intervals for all UAVs to reach the target area is not empty, there is a feasible solution, that is, the time cooperation constraints of multiple UAVs are

$$\max(T_1^{\min}, T_2^{\min}, \dots, T_m^{\min}) \leq \min(T_1^{\max}, T_2^{\max}, \dots, T_m^{\max}). \quad (15)$$

Then, set the time collaboration constraint function as

$$\text{Constraint}_{\text{time}} = \max(\max(T_1^{\min}, T_2^{\min}, \dots, T_m^{\min}) - \min(T_1^{\max}, T_2^{\max}, \dots, T_m^{\max}), 0). \quad (16)$$

Figure 8 is the schematic diagram of time cooperative constraint of three UAVs, in which each UAV has three alternative tracks. In the figure, only the same order track of each UAV is used to explain. It can be seen that the blue and purple lines have time intersection, that is, they meet the cooperative time constraint, while the red line does not meet the cooperative time constraint.

5.2.3. Cooperative Cost Function. The cooperative objective function includes two parts. One part is the sum of the cost $\sum_{i=1}^n \text{fitness}_i$ of the track objective function of a flight track selected from the alternative track group by each UAV, where fitness_i is the cost of the track objective function selected by the i th aircraft. The other part is the planned coordination time cost T_{arrival} , in order to minimize the coordination cost of UAV, the minimum value in the intersection of coordination time is taken as T_{arrival} , as shown in Figure 7. Therefore, the coordination cost function is as follows.

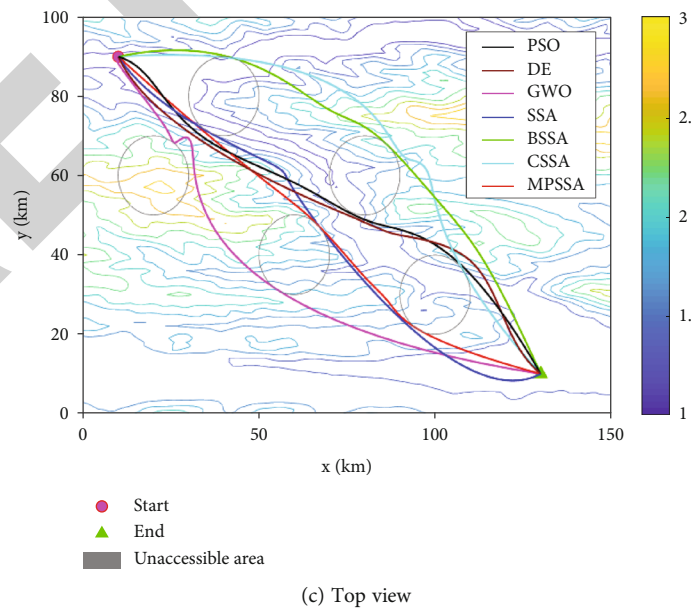
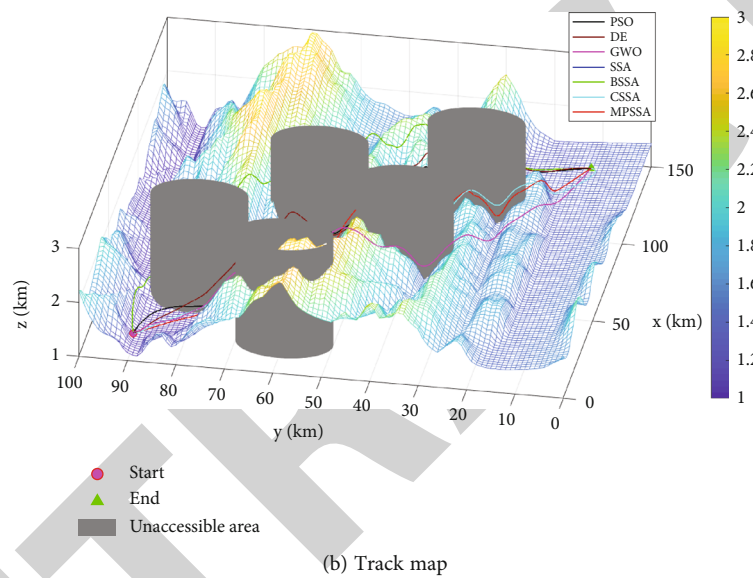
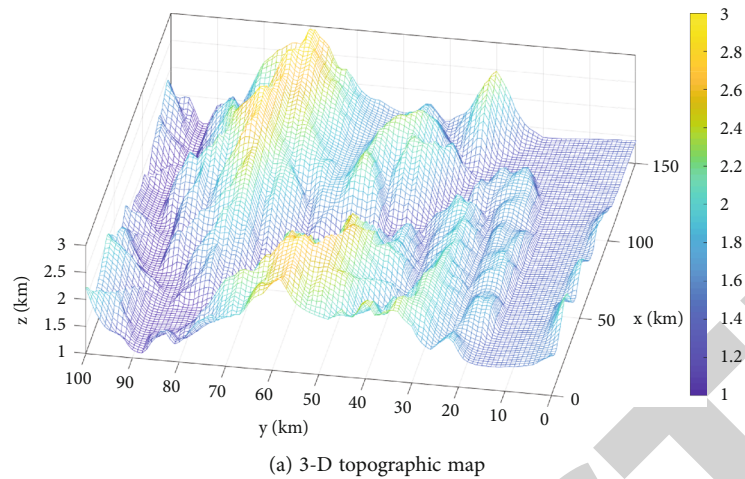
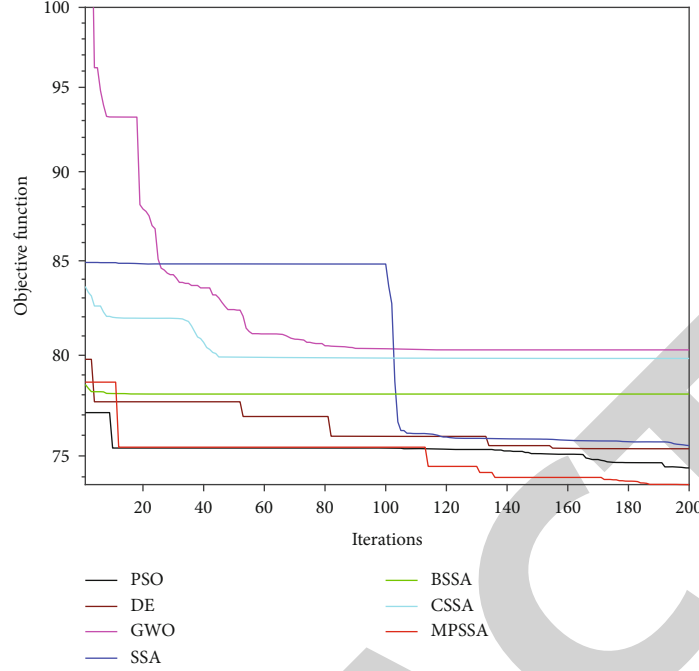


FIGURE 9: Continued.



(d) Cost function convergence

FIGURE 9: Single UAV track planning.

$$\min (F) = \sum_{i=1}^n fitness_i + T_{arrival} + \eta(Constraint_{space} + Constraint_{time}) \quad (17)$$

5.3. Experimental Setup. The map environment adopts the establishment of three-dimensional topographic map within the range of $100 \text{ km} \times 150 \text{ km} \times 3 \text{ km}$, and the track is smoothed by cubic B-spline curve.

An individual S_i in the population is defined as a path connected by multiple track points, $S_i = [s_{x1}, s_{y1}, s_{z1}, s_{x2}, \dots, s_{xn}, s_{yn}, s_{zn}]$, in which every three constitute the three-dimensional coordinates of a track point, a total of n track points, and the dimension of each individual is $3 * n$. If the starting point coordinate S is (x_s, y_s, z_s) and the ending point coordinate E is (x_E, y_E, z_E) , the three-dimensional coordinates of the track point d_i are

$$\begin{aligned} x_{di} &= x_s + s_{xi} \cdot (x_s - x_E), \\ y_{di} &= y_s + s_{yi} \cdot (y_s - y_E), \\ z_{di} &= z_{d_i}^* + s_{zi} \cdot (z_{\max} - z_{d_i}^*), \end{aligned} \quad (18)$$

where $z_{d_i}^*$ can be obtained in the elevation map according to x_{di} and y_{di} and z_{\max} refers to the upper bound of coordinate Z in the elevation map.

5.4. Experimental Simulation of Track Planning. In single UAV track planning, the coordinates of starting point and ending point are set as $(10, 90, 1.1)$ and $(130, 10, 1.1)$. There are five nonflying areas with comprehensive threats. The

coordinates of the central points are $(20, 60)$, $(40, 80)$, $(60, 40)$, $(80, 60)$, and $(100, 30)$, and the radius is 10 km , $\eta = 10^7$, $\omega_1 = 0.5$, $\omega_2 = 0.3$, and $\omega_3 = 0.2$. The algorithm parameters, iteration times, and population number are consistent with the above experiments. The track point is set to 10 and runs independently for 20 times. Figure 9(a) is the original topographic map, (b) is the optimal three-dimensional track map of each algorithm, (c) is the contour top view, and (d) is the objective function convergence map. The objective function results after the simulation experiment are shown in Figure 10.

As can be seen from Figure 9, MPSSA can cling to the ground, perfectly avoid the threat area, quickly get rid of the constraints, and the planned route is smoother. PSO, De, GWO, SSA, BSSA, and CSSA fall into local optimization after a certain number of iterations and cannot jump out. The planned line is obviously not a good line, while MPSSA converges faster and finds a better solution than other algorithms.

Figure 10 shows the results of the objective function after the above seven algorithms independently conduct track planning for 20 times. It can be seen that the four indexes of MPSSA are better than other algorithms, indicating that it has strong optimization ability and stability. Although the time consumed by the algorithm is not the shortest, it gains better performance at the cost of time.

To sum up, MPSSA can quickly avoid the constraints of threat areas and complete the track planning, and the planned route is shorter and smoother than other comparison algorithms. At the same time, the convergence speed, optimization accuracy, and robustness of the algorithm are better than other algorithms.

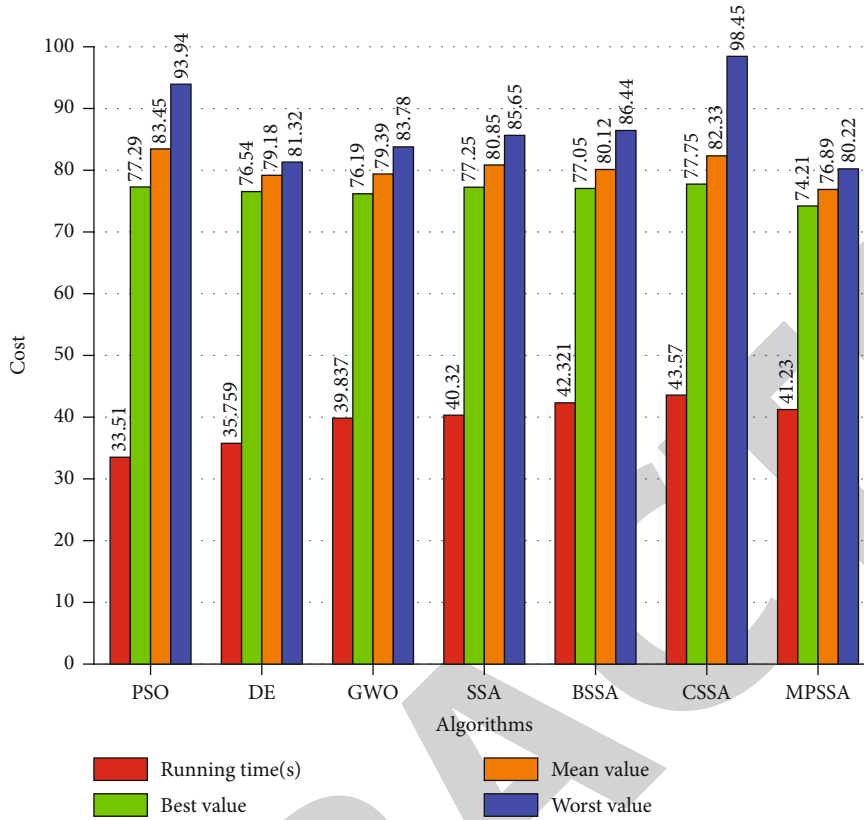


FIGURE 10: Single UAV planning results comparison.

5.5. Experimental Simulation of Cooperative Track Planning.

In the collaborative track planning, three UAVs are selected for experiments, and there are 10 alternative tracks in each alternative track group. The coordinates of starting point 1, starting point 2, starting point 3, and ending point are set as (10, 90, 1.1), (10, 70, 2.2), (10, 90, 1.2), and (130, 10, 1.1), respectively. The settings of other parameters such as comprehensive threat are the same as above, and the safety distance D is set as 0.2 km, $V_{\max} = 1$ km/s, and $V_{\min} = 0.5$ km/s. The objective function results after the simulation experiment are shown in Table 6. Figure 11(a) is the original topographic map, (b) is the three-dimensional route map of cooperative track planning, (c) is the top view of contour line, and (d) is the convergence map of cooperative objective function. Table 6 shows the results of UAV objective function of single UAV track planning. Table 7 shows the results of multi-UAV cooperative track planning cooperative objective function.

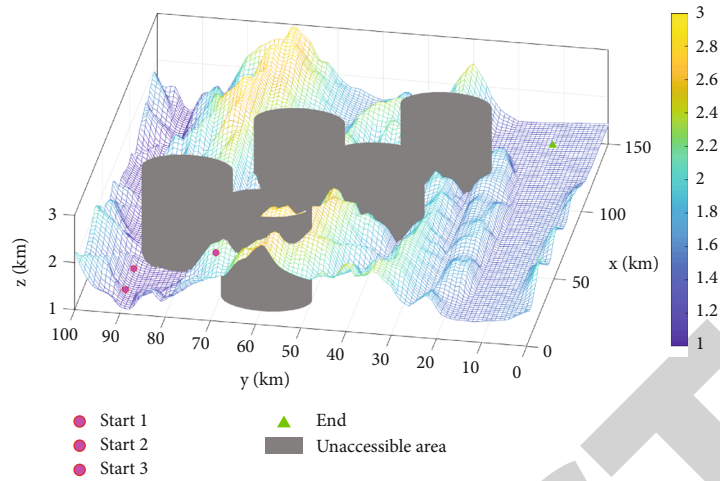
As can be seen from (b) and (c), MPSSA can still avoid the threat area and get rid of constraints in the case of collaborative planning of multiple UAVs to fly smoothly at a lower altitude relative to the terrain. Although UAV1 and UAV3 seem to coincide from the top view, it can be seen from (b) that the two UAVs are still outside the safe distance and have a time difference. It can be seen that due to the constraints of cooperative space and time, the tracks of UAVs are not always optimal. It can be seen from the graph that the optimal solution (SSD) can still converge faster and jump out of the optimal solution.

TABLE 6: Track planning for each UAV.

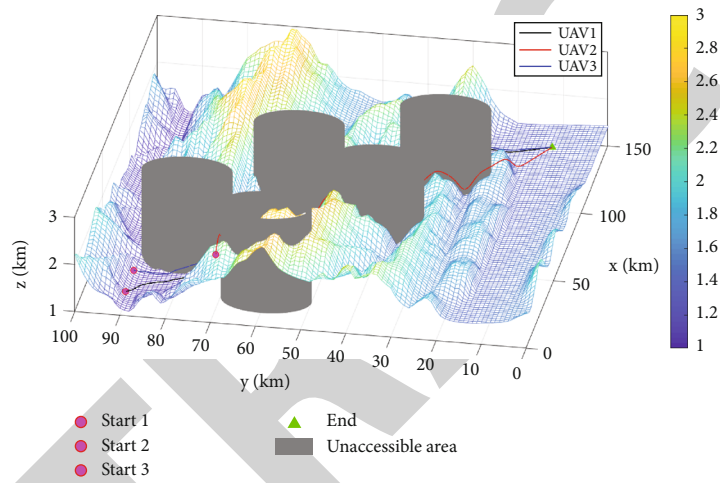
UAV/cost	Best	Worst	Ave	Std
UAV-1	74.31	77.61	75.96	1.55
UAV-2	69.43	76.19	72.81	3.10
UAV-3	69.10	72.87	70.99	1.75

It can be seen from Table 6 that the standard deviation of the cost function value of the alternative track under MPSSA is small, and the alternative track group meeting the constraints can be planned stably for each UAV at different starting points. Table 7 shows the multi-UAV collaborative planning scheme, which shows that each UAV can reach the end point within the time range of [148.05 s, 275.94 S], and the collaborative cost is 360.99. It can be seen that the track selected by each UAV is not the best in the alternative track group, indicating that the cooperative track planning scheme is not necessarily composed of the best track in the alternative track group of each UAV but selects the relatively better track that meets the coordination requirements to perform the task.

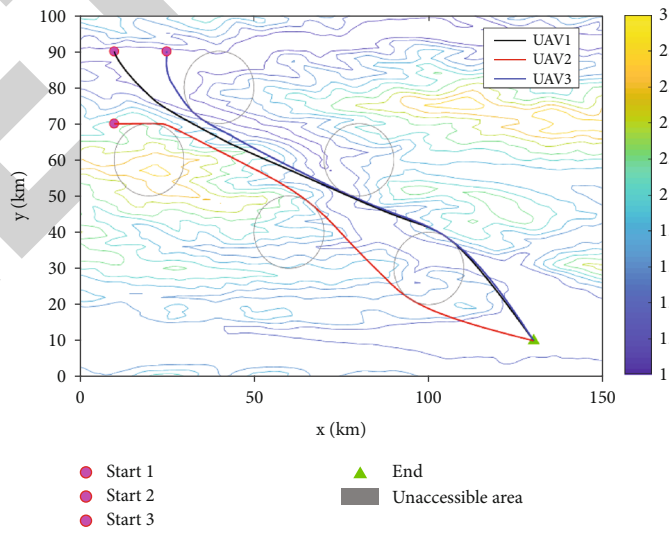
In conclusion, the MPSSA algorithm adds a variety of population search mechanisms and sets different “discoverer enrollee” ratios in different subpopulations, ensuring that each algorithm has different development and exploration capabilities. By adding dynamic dimension by dimension dynamic reverse learning, it accelerates the convergence speed of subpopulations and has more detailed search



(a) 3-D topographic map

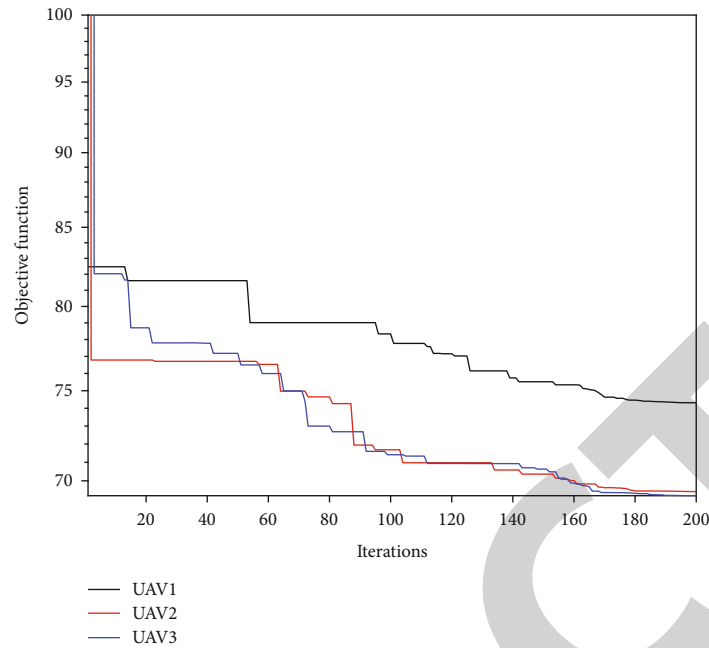


(b) Track map



(c) Top view

FIGURE 11: Continued.



(d) Cost function convergence

FIGURE 11: Collaborative track planning.

TABLE 7: Collaborative track planning results.

UAV/index	Cost	Track length (km)	Arrival time (s)	Collaborative arrival time (s)	Collaborative cost
UAV-1	74.31	148.05	[148.05, 296.10]		
UAV-2	69.43	137.99	[137.99, 275.99]	[148.05, 275.94]	360.99
UAV-3	69.21	137.97	[137.97, 275.94]		

accuracy. After all subpopulations are updated, it adopts the weighted center of gravity exchange strategy, The exchange between populations and the transfer of high-quality solutions help to improve the quality of the population and jump out of the local optimum. The above strategies make MPSSA's search mode more flexible and its comprehensive search capability enhanced, so that MPSSA can help UAVs quickly bypass threat areas and obtain better track routes under complex environments and mathematical models, and help multiple UAVs to plan better, stable, and smooth track schemes, ensure that UAVs can complete cooperative tasks safely and stably, and further verify MPSSA's good algorithm performance.

6. Conclusion

In this paper, a sparrow search algorithm based on personalized multipopulation is proposed to solve the original defects of SSA and help UAV complete complex cooperative track planning. Through the introduction of multipopulation mechanism, personalized subpopulation strategy, weighted center of gravity communication strategy, and dimension by dimension dynamic reverse learning, it can prevent falling into local optimization, enhance the personalized difference of subpopulation, improve the quality of population

communication, accelerate convergence and improve optimization accuracy, and effectively improve the performance of the original SSA algorithm. The experimental results of benchmark function and CEC2017 show that MPSSA can stably find better solutions than PSO, GWO, DE, SSA, and SSA improved algorithms, has good advantages, and can adapt to more complex optimization problems. Through the simulation experiment of UAV cooperative track, the results show that MPSSA is helpful to the static planning of UAV cooperative track, with shorter time, better planning path, and better stability.

At the same time, there is still a lot of work to be done in the future, such as checking the impact on the performance of the algorithm when more constraints are added. The algorithm is applied to more complex dynamic path planning and real-time obstacle avoidance. The algorithm is applied to multiobjective programming problem and so on.

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 financially supported by the Regional foundation of the National Natural Science Foundation of China (No. 61703411).

References

- [1] W. L. Liu, Y. J. Gong, W. N. Chen, Z. Liu, H. Wang, and J. Zhang, "Coordinated charging scheduling of electric vehicles: a mixed-variable differential evolution approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 12, pp. 5094–5109, 2020.
- [2] S. Zhou, L. Xing, X. Zheng, N. Du, L. Wang, and Q. Zhang, "A self-adaptive differential evolution algorithm for scheduling a single batch-processing machine with arbitrary job sizes and release times," *IEEE transactions on cybernetics*, vol. 51, no. 3, pp. 1430–1442, 2019.
- [3] F. Zhao, R. Ma, and L. Wang, "A self-learning discrete Jaya algorithm for multiobjective energy-efficient distributed no-idle flow-shop scheduling problem in heterogeneous factory system," *Cybernetics*, pp. 1–12, 2021.
- [4] F. Zhao, X. He, and L. Wang, "A two-stage cooperative evolutionary algorithm with problem-specific knowledge for energy-efficient scheduling of no-wait flow-shop problem," *IEEE transactions on cybernetics*, vol. 51, no. 11, pp. 5291–5303, 2021.
- [5] F. Zhao, L. Zhao, L. Wang, and H. Song, "An ensemble discrete differential evolution for the distributed blocking flowshop scheduling with minimizing makespan criterion," *Expert Systems with Applications*, vol. 160, article 113678, 2020.
- [6] W. Guo, N. Xiong, H. C. Chao, S. Hussain, and G. Chen, "Design and analysis of self-adapted task scheduling strategies in wireless sensor networks," *Sensors*, vol. 11, no. 7, pp. 6533–6554, 2011.
- [7] X. Wang, Q. Li, N. Xiong, and Y. Pan, "Ant colony optimization-based location-aware routing for wireless sensor networks," in *International Conference on Wireless Algorithms, Systems, and Applications*, pp. 109–120, Berlin, Heidelberg, 2008.
- [8] Y. Jiang, G. Tong, H. Yin, and N. Xiong, "A pedestrian detection method based on genetic algorithm for optimize XGBoost training parameters," *IEEE Access*, vol. 7, pp. 118310–118321, 2019.
- [9] R. Wan, N. Xiong, and N. T. Loc, "An energy-efficient sleep scheduling mechanism with similarity measure for wireless sensor networks," *Human-centric Computing and Information Sciences*, vol. 8, no. 1, pp. 1–22, 2018.
- [10] Y. Lu, S. Wu, Z. Fang, N. Xiong, S. Yoon, and D. S. Park, "Exploring finger vein based personal authentication for secure IoT," *Future Generation Computer Systems*, vol. 77, pp. 149–160, 2017.
- [11] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
- [12] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.
- [13] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019.
- [14] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp swarm algorithm: a bio-inspired optimizer for engineering design problems," *Advances in Engineering Software*, vol. 114, pp. 163–191, 2017.
- [15] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," *Journal of Global Optimization*, vol. 39, no. 3, pp. 459–471, 2007.
- [16] X. S. Yang, "Firefly algorithm, stochastic test functions and design optimisation," *International journal of bio-inspired computation*, vol. 2, no. 2, pp. 78–84, 2010.
- [17] X. S. Yang and A. H. Gandomi, "Bat algorithm: a novel approach for global engineering optimization," *Engineering Computations*, vol. 29, no. 5, pp. 464–483, 2012.
- [18] X. Meng, Y. Liu, X. Gao, and H. Zhang, "A new bio-inspired algorithm: chicken swarm optimization," in *International Conference in Swarm Intelligence*, pp. 86–94, Cham, 2014.
- [19] X. B. Meng, X. Z. Gao, L. Lu, Y. Liu, and H. Zhang, "A new bio-inspired optimisation algorithm: bird swarm algorithm," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 28, no. 4, pp. 673–687, 2016.
- [20] H. Duan and P. Qiao, "Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning," *International journal of intelligent computing and cybernetics*, vol. 7, no. 1, pp. 24–37, 2014.
- [21] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: sparrow search algorithm," *Systems Science & Control Engineering*, vol. 8, no. 1, pp. 22–34, 2020.
- [22] S. Yan, P. Yang, D. Zhu, W. Zheng, and F. Wu, "Improved sparrow search algorithm based on iterative local search," *Computational Intelligence and Neuroscience*, vol. 2021, 31 pages, 2021.
- [23] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *1997 IEEE International conference on systems, man, and cybernetics. Computational cybernetics and simulation*, vol. 5, Orlando, FL, USA, 1997.
- [24] D. Whitley, "A genetic algorithm tutorial," *Statistics and Computing*, vol. 4, no. 2, pp. 65–85, 1994.
- [25] X. Lv, X. Mu, and J. Zhang, "Multi-threshold image segmentation based on improved sparrow search algorithm," *Systems Engineering and Electronics*, vol. 43, no. 2, pp. 318–327, 2021.
- [26] X. Lv, X. Mu, and Z. Jun, "Chaotic sparrow search optimization algorithm," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 47, no. 8, pp. 1712–1720, 2021.
- [27] H. Shorakaei, M. Vahdani, B. Imani, and A. Gholami, "Optimal cooperative path planning of unmanned aerial vehicles by a parallel genetic algorithm," *Robotica*, vol. 34, no. 4, pp. 823–836, 2016.
- [28] D. Zhang and H. Duan, "Social-class pigeon-inspired optimization and time stamp segmentation for multi-UAV cooperative path planning," *Neurocomputing*, vol. 313, pp. 229–246, 2018.
- [29] T. Han, A. D. Tang, and H. Zhou, "Multi UAV cooperative path planning method based on Lassa algorithm," *Systems engineering and electronic technology*, vol. 44, no. 1, pp. 233–241, 2022.
- [30] G. Wu, R. Mallipeddi, and P. N. Suganthan, *Problem definitions and evaluation criteria for the CEC 2017 competition on constrained real-parameter optimization*, National University of Defense Technology, Changsha, Hunan, PR China and Kyungpook National University, Daegu, South Korea and Nanyang Technological University, Singapore, Technical Report, 2017.
- [31] N. H. Awad, M. Z. Ali, and J. J. Liang, *Problem definitions and evaluation criteria for the CEC 2017 special session and*

Retraction

Retracted: Simulation of Subject Coordination for a Smart Campus Based on Complex Network and Evolutionary Game Theory

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Guo and T. Pang, "Simulation of Subject Coordination for a Smart Campus Based on Complex Network and Evolutionary Game Theory," *Journal of Function Spaces*, vol. 2022, Article ID 4141475, 9 pages, 2022.

Research Article

Simulation of Subject Coordination for a Smart Campus Based on Complex Network and Evolutionary Game Theory

Shaoyong Guo  and Ting Pang 

Network and Information Center, Xinxiang Medical University, Henan, Xinxiang 453003, China

Correspondence should be addressed to Shaoyong Guo; gsy@xxmu.edu.cn

Received 9 April 2022; Revised 26 April 2022; Accepted 14 June 2022; Published 26 June 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Shaoyong Guo and Ting Pang. 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 construction of smart campuses can help realize the modernization of higher education. However, the subjects and mechanisms are easily ignored in the construction of smart campuses in colleges and universities. The purpose of this paper is to construct an evolutionary game model based on the complex network structure characteristics of smart campus subjects and the dynamic relationship of collaborative games. Taking the scale-free network as the carrier, the decision-making behavior among subjects is explained. Then, the importance of subject coordination and the effectiveness of the mechanisms of leadership organization, special funds, assessment rewards, and punishments are analyzed through the simulation platform, i.e., Simpy. The results show that (1) moderate subject collaboration helps to avoid the phenomenon of *getting something for nothing*, (2) appropriate leadership and organization mechanisms can promote the construction process, (3) special fund mechanisms can ensure sustainable development, and (4) assessment reward and punishment mechanisms can increase the popularity of achievements. The proposed method enriches relevant theories of smart campus research and provides a reference for decision-making in smart campus construction.

1. Introduction

With the continuous intelligence of network information technology and the more diversified needs of teachers and students, universities have vigorously advocated for wisdom education, and information construction has entered the smart campus stage. The COVID-19 epidemic has made the development of smart campuses perform a “bend overtake,” and it is becoming more urgent for smart campuses to reach a high level with high quality. A smart campus refers to the deep integration of scattered information systems and educational resources on a university campus using cloud computing, big data, artificial intelligence, the Internet of Things, mobile interconnections, and other new generations of network information technology to construct a comprehensive, perceptive, efficient, and convenient environment for teaching, scientific research, management, clothing, and living [1–3].

The construction of a smart campus is a difficult task involving a wide range of areas, with heavy workloads and

a long duration [4]. After years of research and practice, the smart campus has made a preliminary development. However, current research ignores two important issues: the subjects and the mechanism [5, 6]. Most of the work of constructing smart campuses is undertaken by the network information department. For the sake of responding to current interests, IT (information technology) enterprises may not keep up with subsequent maintenance. In addition, the relevant subjects of universities are not fully involved, resulting in a weak construction force, difficult coordination of data resources, and poor user experience. Research on smart campuses has exposed the shortcomings of *head-heavy and feet-light* [7, 8]. At the beginning, construction focuses on the excavation of the technical level and neglects the importance of the construction mechanism, which enacts the construction and implementation of smart campuses without solid guarantees. In addition, many subjects are involved in the construction of smart campuses in universities, and they have a competitive and cooperative relationship, which has the structural characteristics of a

complex network and the dynamic relationship of a collaborative game [9–11]. However, research from the perspective of a complex network evolutionary game is still lacking.

To alleviate the above constraints, this paper studied the construction of smart campuses in high-level universities based on the theory of complex networks and evolutionary games. Taking the scale-free network as the carrier, the evolutionary game model of public goods was constructed to explain decision-making behavior among subjects. Then, the importance of subject coordination and the effectiveness of the mechanisms of leadership organization, special funds, assessment rewards, and punishments are analyzed through the simulation platform, i.e., Simpy.

This work enriches the relevant theories of smart campus research and provides suggestions and references for the construction of smart campuses in universities. The contributions of this work are fourfold. (1) Moderate subject collaboration helps to prevent the phenomenon of getting something for nothing. (2) Appropriate leadership and organizational mechanisms can promote the construction process. (3) Special fund mechanisms can ensure sustainable development. (4) Assessment reward and punishment mechanisms can increase the popularity of achievements.

The structure of this paper is as follows. In Section 2, we summarize and analyze related works on the construction of smart campuses. Our review covers three main aspects: theoretical connotations, strategic proposals, and technical methods. In Section 3, methods utilizing the theory of complex networks and evolutionary games are proposed. In Section 4, we describe the process and results of the simulation. In Section 5, the importance of subjects' coordination and the effectiveness of the mechanism guarantees are discussed in detail. Finally, in Section 6, brief conclusions are provided.

2. Related Works

At present, the construction of smart campuses represents the direction of information development in universities, and it is also a trending issue among experts in related fields at home and abroad. It encompasses the following three main aspects.

Theoretical connotations: Atif et al. studied the establishment of a smart campus environment to provide a ubiquitous learning mode [12]. Wang analyzed the wisdom concept, main characteristics, and design of the smart campus [13]. In discussing the essential elements and characteristics of smart campuses, Li highlighted existing problems and gaps [14]. Nasro et al. proposed the management evaluation and positioning of the existing infrastructure of higher education institutions based on the concept of the smart campus [15].

Strategic proposals: Kang et al. expounded on the solutions of smart campuses, such as campus network construction, hardware infrastructure construction, security system construction, intelligent multimedia teaching application system construction, multimedia conference system construction, and software platform construction [16]. Chen carried out research from the aspects of organizational struc-

ture, follow-up capital investment, team construction, system perfection, and incentive policy and put forward corresponding suggestions for safeguard mechanisms [17]. Gan et al. elaborated the construction of smart campuses in foreign universities and focused on Chinese universities from seven dimensions: top-level design, construction vision, construction mode, smart teaching environment, smart teaching resources, smart campus management, and smart campus services [18]. Yan et al. pointed out that in the process of carrying out theoretical research on smart campuses, universities should always establish *people-oriented concepts* [19].

Technical methods: Mar et al. designed a smart campus system including intelligent education, parking, and classroom using information technology [20]. Ikrisi et al. focused on several vulnerabilities and vulnerable attacks affecting the data and information security of smart campuses [21]. Luo et al. designed an overall architecture model of smart campuses in a big data environment. From bottom to top, the model has an infrastructure layer, data layer, application support layer, business application layer, and terminal display layer [22]. Yu et al. built smart campus private networks based on 5G edge-cloud fusion, sinking the services concentrated in the remote cloud center from *core* to *edge* [23].

3. Method

3.1. Basic Analysis. In this study, it is assumed that the subjects involved in the construction of smart university campuses include enterprises, network information departments (NIDs), business departments, faculties, and affiliated units (AU). Among them, the task of the enterprise is to provide comprehensive design and technology development services. The network information department coordinates all aspects of relations and leads the construction process. The business department is in charge of different matters and mastering different kinds of big data resources. These subjects need to communicate and connect. This study considers the party committee office (PCO), organization department (OD), publicity department (PD1), labor union (LU), youth league committee (YLC), principal's office (PO), development planning department (DPD), personnel department (PD2), finance department (FD), science and technology department (STD), audit department (AD), asset management department (AMD), infrastructure department (ID), security department (SD1), logistics department (LD), retirement department (RD), student department (SD2), academic affairs department (AAD), enrollment and employment department (EED), library, and archives. Although there are a large number of departments and affiliated units, in most cases, they use the resources only after the completion of the smart campus, and each of them is regarded as a subject. Consequently, a total of 25 subjects were identified.

A smart university campus is a kind of public good. First, the subject plays a very important role in the process of building a smart campus. We cannot ignore any party. Only by sharing resources and collaborative construction can we

complete the task. Second, the products after completion are shared by each subject. For example, enterprises make profits and continue to provide updates and maintenance for deep-seated technology; the network information departments continue to lead the upgrading and construction of smart campus and basic technical maintenance services in the later stage; the business departments should constantly provide updated data and needs; colleges, departments, and affiliated units can enjoy the resources of a smart campus to carry out teaching, learning, scientific research, and life more conveniently. However, in the process of building smart campuses in universities, enterprises tend to pay attention to interests and neglect quality, the network information department lacks the leading force, the business department lacks a sense of ownership, and the departments and affiliated units are dependent. Therefore, there is both competition and cooperation between subjects, and there is a dynamic cooperative game relationship, which is also an evolutionary game. The game of public goods represents a typical study of cooperative behavior between multiple individuals [24]. Each individual has only two strategic choices, i.e., cooperation or noncooperation (betrayal), and there will be a gain coefficient to increase the gains of each collaborator and betrayer by a certain multiple. Therefore, cooperation is an altruistic behavior, and the betrayer can also obtain gains, so selfish individuals will betray, resulting in the phenomenon of *hitchhiking* or *getting something for nothing*. However, all individuals will not benefit if they choose not to cooperate, so there will be individual cooperation. Therefore, the game of public goods can not only improve the enthusiasm of individual cooperation but also inhibit the emergence of *hitchhiking*.

At the same time, the subject of building a smart campus in universities and the cooperative game relationship between them can be regarded as a typical complex network. Nodes represent the subject, edges represent the relationship, and weights represent the strength of the actual connection [25]. In the process of building a smart campus in universities, enterprises are only in contact with the network information department, asset management department (bidding), etc., and departments and affiliated units have contact or communication only with individual business departments and network information departments. The business department contains most of the subjects, but there are only individual connections between them. The network information department is connected with all other subjects due to the coordination of relations in all aspects, which is in line with the characteristics of *most nodes and few nodes are connected, and few nodes are connected with many nodes* of the scale-free network. A scale-free network is a network with a power-law distribution, as opposed to a random network. A large number of real-world networks are scale-free networks [23].

To sum up the experience of the construction of smart university campuses in recent years, the network information department alone is advocating for this change, but its ability and power are limited. The follow-up services of enterprises cannot be guaranteed, and the sense of ownership of relevant departments of the school is not strong.

Hence, the effect and quality of smart campus construction are not ideal, and the construction process is slow. Even if there is investment and some achievements, the popularity of smart campus construction is not high. The latter leads to a waste of resources and school accountability, and the campus cannot be upgraded and updated, so construction falls into a bad cycle. To break the cycle, on the one hand, we should ensure sufficient subjective initiative and strengthen active cooperation between subjects. On the other hand, we should establish a sound mechanism, especially the strong guarantee of leadership organizations, the continuous support of special funds, and scientific and reasonable assessment, reward, and punishment. Next, it will be verified by model and simulation.

3.2. Model Construction. Based on the analysis presented in Section 3.1, this paper uses the game of public goods to establish an evolutionary game model and takes the scale-free network as the complex network carrier to study the importance of subject coordination in the construction of smart university campuses [26] and the effectiveness of leading organizations, special funds, assessment, rewards, and punishments. Based on realistic considerations, the following assumptions are put forward for the construction of a complex network evolution game model of smart university campuses:

H1: heterogeneous complex network. It has different types of nodes, and the impact on each node is different after completion.

H2: weighted complex networks. Collaborative strength is represented by the weight of nodes.

H3: the number of nodes in the network does not change.

H4: each node becomes a game opponent only with its neighbors that have direct contact with it, and its income is only affected by the game opponent.

H5: the game strategy of each node is only cooperation and noncooperation.

H6: the input cost of each node per round of the game is set to 1.

H7: nodes exhibit bounded rationality, and there is a possibility of choosing game strategy mistakes.

3.2.1. The Complex Network. Let $G = \{N, E, W\}$ represent the main network of smart campus construction in universities, where N represents the number of nodes, E represents the number of edges, and W represents the weight. The weighted degree of node i is

$$k_i^\omega = k_i \omega_{ij}, \quad (1)$$

where k_i is the degree of node i , node j is the neighbor of node i , and $\omega_{ij} (0 < \omega_{ij} < 1)$ is the weight, which represents the cooperative strength between nodes i and j .

Section 3.1 shows that N is 25. The complex network built on the smart university campus is quantified as a relationship matrix of 25×25 , where the connection between nodes is determined according to the direct relationship

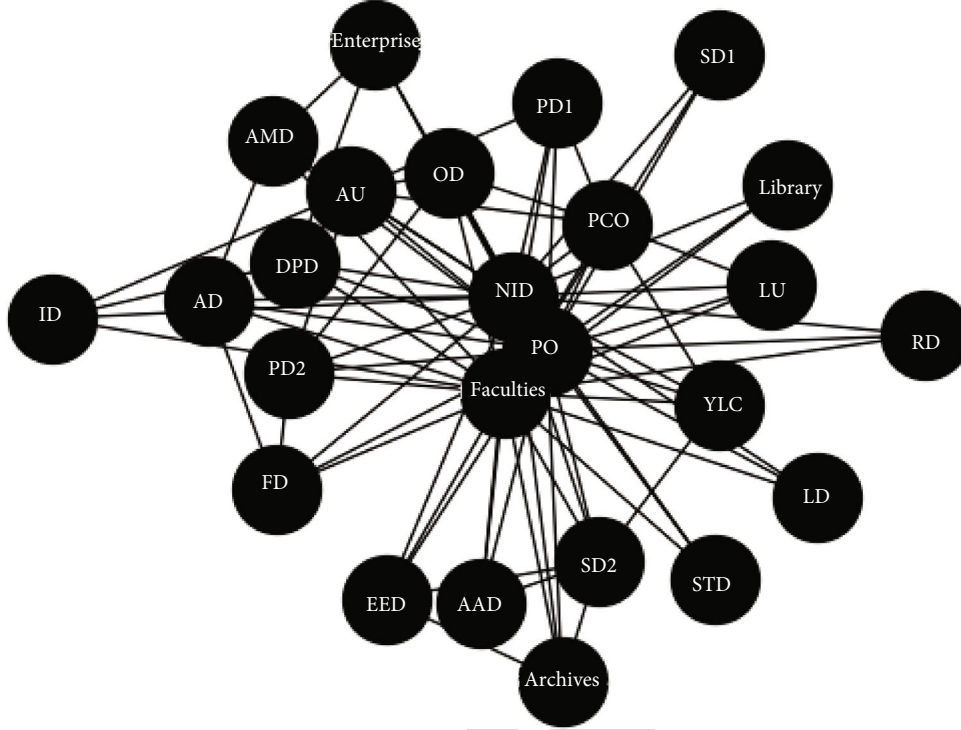


FIGURE 1: Complex network diagram of smart campus construction for universities.

between subjects. The complex network diagram is obtained by using NetworkX [27], as shown in Figure 1.

3.2.2. *Evolutionary Game.* Let A denote the adoption of a cooperative strategy, and let the number of nodes be n_A ; then, the proportion of the nodes is as follows.

$$x_A = \frac{n_A}{N}. \quad (2)$$

Let B denote the adoption of a noncooperative strategy, and the number of nodes is n_B ; then, the proportion of nodes is as follows.

$$x_B = \frac{n_B}{N}, \quad (3)$$

where $x_A + x_B = 1$. The symmetric game gains of node $i(1, 2, \dots, i)$ and neighbor node $j(1, 2, \dots, j)$ are shown in Table 1.

The letters a , b , c , and d represent the benefit obtained by the first individual in the strategy pair (A, A) , (A, B) , (B, A) , and (B, B) , respectively. In the evolutionary game, each individual plays with its neighbors like the previous two games, and the sum of the income obtained is called the total income of the individual.

In the first case, the average gains of node i selecting A at time step t are

$$f_{Ai} = \sum_{m=0}^{k_i} C_{k_i}^m x_A^m (1-x_A)^{k_i-m} \frac{[ma + (k_i - m)b](1 + \gamma)(\alpha t^2 + \beta t)}{k_i^\omega}, \quad (4)$$

TABLE 1: The matrix of gains for the symmetric game.

Node i	Node j	
	A	B
A	a	b
B	c	d

where γ indicates the guarantee strength of the leadership organization mechanism, α indicates the guarantee strength of the special fund mechanism, and β indicates the guarantee strength of the evaluation, reward, and punishment mechanism.

In the second case, the average gains of node i selecting B at time step t are

$$f_{Bi} = \sum_{m=0}^{k_i} C_{k_i}^m x_B^m (1-x_B)^{k_i-m} \frac{[mc + (k_i - m)d](1 + \gamma)(\alpha t^2 + \beta t)}{k_i^\omega}. \quad (5)$$

We adopt the synchronous learning game strategy, that is, each node carries out strategy learning at each time step [28]. At the same time, the game strategy of pairing comparison is used to update the rules; that is, node i randomly selects a neighbor node j to compare the gains, and there is probability p_{ij} of updating the strategy of node j .

$$p_{ij} = \frac{1}{1 + \exp \left[-\left(f_j - f_i \right) / \epsilon \right]}, \quad (6)$$

where ε is a noise term that characterizes the decision-making error and bounded rationality of the subject [29]. When $\varepsilon \rightarrow 0$, the subject tends to make irrational decisions and decides to update to the neighbor strategy. When $\varepsilon \rightarrow \infty$, the subject cannot make rational decisions and randomly learn neighbor decisions. The meaning of probability p_{ij} is that when the income of node i is less than that of node j , the decision of node j is accepted with a great probability. In contrast, the decision of node i is accepted with a small probability.

4. Simulation and Results

4.1. Simulation Flowchart and Parameter Settings. (1) After the start of simulation, first, build a scale-free network, set the initial parameters, and randomly assign strategies to each node. Then, each node plays a game with its neighbors to calculate the income of each node. Next, the strategy of each node was updated synchronously, and the game turned back. Finally, the simulation ends when the predetermined time T is over. The process is shown in Figure 2

(2) The Simpy [30] platform was used for simulation analysis. The subject is simulated by process, and the game is simulated by event. The parameter settings are shown in Table 2

4.2. Simulation Results

4.2.1. The Importance of Subjects' Coordination to the Construction of Smart University Campuses. We assumed that the initial number of partners is 10, $\gamma = 0.5$, $\alpha = 0.5$, and $\beta = 0.5$. We used different coordinative intensity indices ω , that is, $\omega = 0.2$, $\omega = 0.4$, $\omega = 0.6$, and $\omega = 0.8$, to observe the changing trend in the number of partners. The simulation results are shown in Figure 3.

As seen from the comparison in Figure 3, when $\omega = 0.2$, the number of partners first increased and then decreased and finally was less than the initial number of partners. When $\omega = 0.4$, although the number of partners also decreased to a certain extent, it still exceeded the initial number of partners in the end. When $\omega = 0.6$, the number of partners increased steadily, and finally, it realized cooperation among all partners. When $\omega = 0.8$, although the number of partners also showed an upward trend, it finally stabilized at approximately 22.

4.2.2. The Effectiveness of the Mechanism Guarantees the Construction of Smart University Campuses. We assumed that the initial number of partners is 10. When $\omega = 0.6$, $\alpha = 0.5$, and $\beta = 0.5$, we used different values of γ for the guarantee strength of the leadership organization mechanism. We observed the changing trend of the average gains of the network information department (which plays a key role in connecting the preceding and the following) for $\gamma = 0.3$, $\gamma = 0.6$, and $\gamma = 0.9$, respectively, and the simulation results are shown in Figure 4(a). When $\omega = 0.6$, $\gamma = 0.9$, and $\beta = 0.5$, we used different values of α for the special fund mechanism guarantee strength. We observed the changing trend in the average gains of the network information depart-

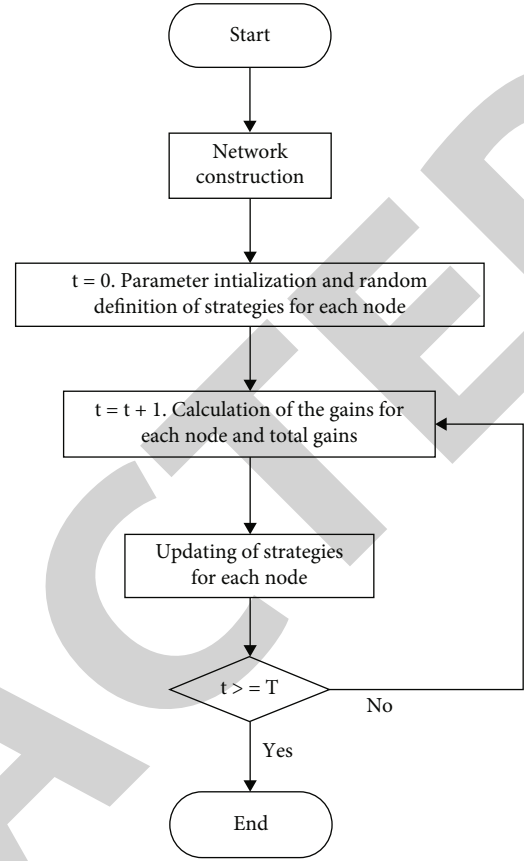


FIGURE 2: Flowchart of simulation execution.

TABLE 2: Parameter settings.

a	b	c	d	ε	T
5	7	7	3	0.5	50

ment for $\alpha = 0.3$, $\alpha = 0.6$, and $\alpha = 0.9$, and the simulation results are shown in Figure 4(b). When $\omega = 0.6$, $\gamma = 0.9$, and $\alpha = 0.9$, we used different values of β for the guarantee strength of the assessment, reward, and punishment mechanism. We observed the change trend in the number of partners in $\beta = 0.2$, $\beta = 0.4$, $\beta = 0.6$, and $\beta = 0.8$, and the simulation results are shown in Figure 4(c).

As seen from the comparison in Figure 4(a), when the value of γ is increasing, the average gains of the network information department are growing steadily, and the convergence speed is accelerating. As seen from the comparison in Figure 4(b), when $\alpha = 0.3$, the average income of the network information department first decreased and then increased; when $\alpha = 0.6$, the average gains increased steadily. When $\alpha = 0.9$, the average gains first increased and then decreased. As seen from the comparison in Figure 4(c), when $\beta = 0.2$, the number of partners decreased. When $\beta = 0.4$, the number of partners first increased and then decreased, finally exceeded the initial number of partners and stabilized at approximately 15. When $\beta = 0.6$, the number of partners continued to rise and finally realized

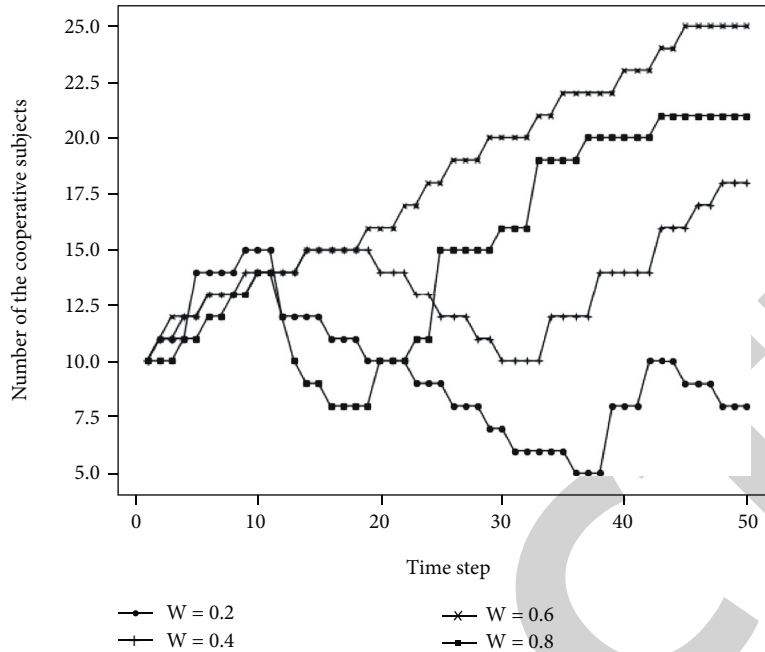


FIGURE 3: The impact of cooperation on the number of partners.

whole network cooperation. When $\beta = 0.8$, the number of partners first increased, then decreased, then increased, and finally stabilized at 23.

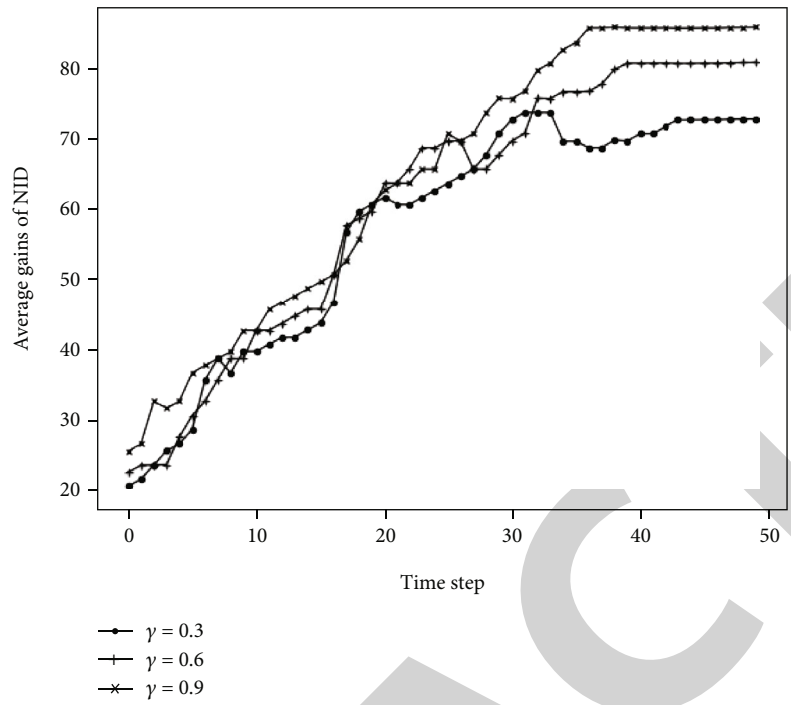
5. Discussion

A certain collaborative intensity can promote cooperation between subjects. When everyone is actively involved in the construction of smart campuses, it will help to prevent the phenomenon of “getting something for nothing” to improve the satisfaction of all staff as much as possible, form a good interactive cycle, and promote the construction progress of smart campuses. However, when the intensity is too high, the frequent interaction between the subjects is prone to contradictions and disputes. At the same time as coordinating between subjects, trust and cooperation need to be established through multiple games to finally achieve win-win results.

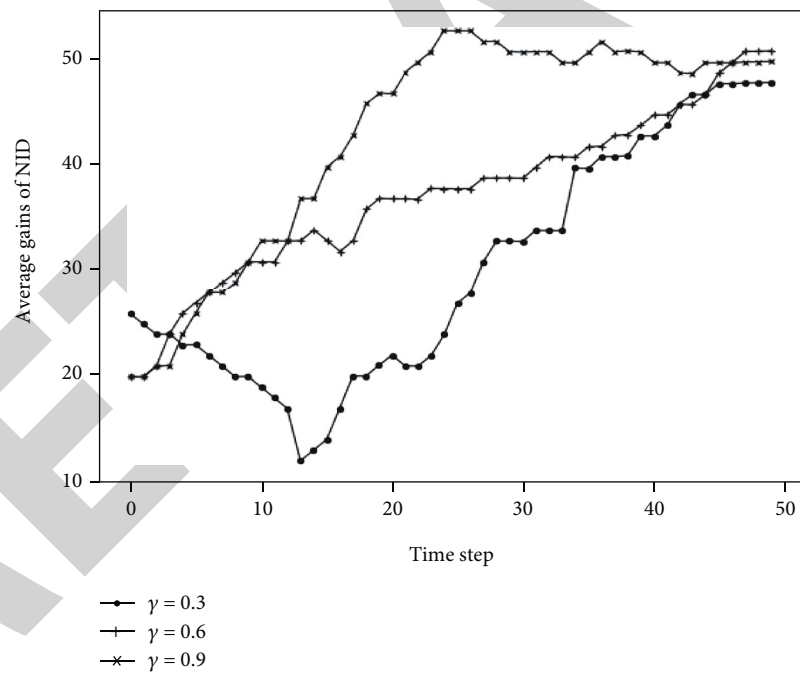
An appropriate leadership organization mechanism can play a key leading role in the whole construction process of smart campuses, strengthen supervision, and accelerate the construction process of smart university campuses. An appropriate special fund mechanism can clarify the source of funds and ensure the stability and sustainability of smart campus construction. However, when the investment is too large, it can easily lead to the decentralization of control power and the dissatisfaction of other departments at a later stage. An appropriate assessment, reward, and punishment mechanism can mobilize subjects' enthusiasm to participate and willingness to use the achievements after construction and can encourage them to constantly put forward suggestions for improvement. However, too many strengths can easily lead to rebellious psychology and are not conducive to development.

In general, the appropriate subject coordination and organization mechanism, fund mechanism, and reward and punishment mechanism all directly affect the income function of each node subject in the smart campus construction network, thus affecting its behavior choice for smart campus construction. Suggestions for promoting the voluntary cooperation of all subjects in the construction of the smart campus are as follows. First, we should cultivate and enhance the awareness of all subjects in the construction of smart campuses. A convenient smart campus is a public property resource and requires the participation of all parties in construction. Second, from the institutional level, it strengthens the leadership of the information construction leading group, promotes the implementation of the information target assessment system, and curbs the emergence of “free riding” behavior. In addition, we should comply with future development trends and the general requirements of teachers and students, rely on the extensive participation of teachers and students, and jointly strengthen the promotion of smart campuses.

As a typical representative of universities in underdeveloped areas, Xinxiang Medical University is below average in terms of informatization level in Henan Province, and the development level of smart campuses is still relatively backward. However, with the renaming of universities and the increasingly urgent requirements of educational modernization, smart campuses are of great significance to the development of Xinxiang Medical University. At present, Xinxiang Medical University focuses on promoting informatization at the leadership level, starting with the construction of a smart campus and with the help of information technology, to realize the *curve overtaking* of educational modernization. Therefore, to carry out the construction of smart campuses more efficiently and effectively, we should appropriately



(a)



(b)

FIGURE 4: Continued.

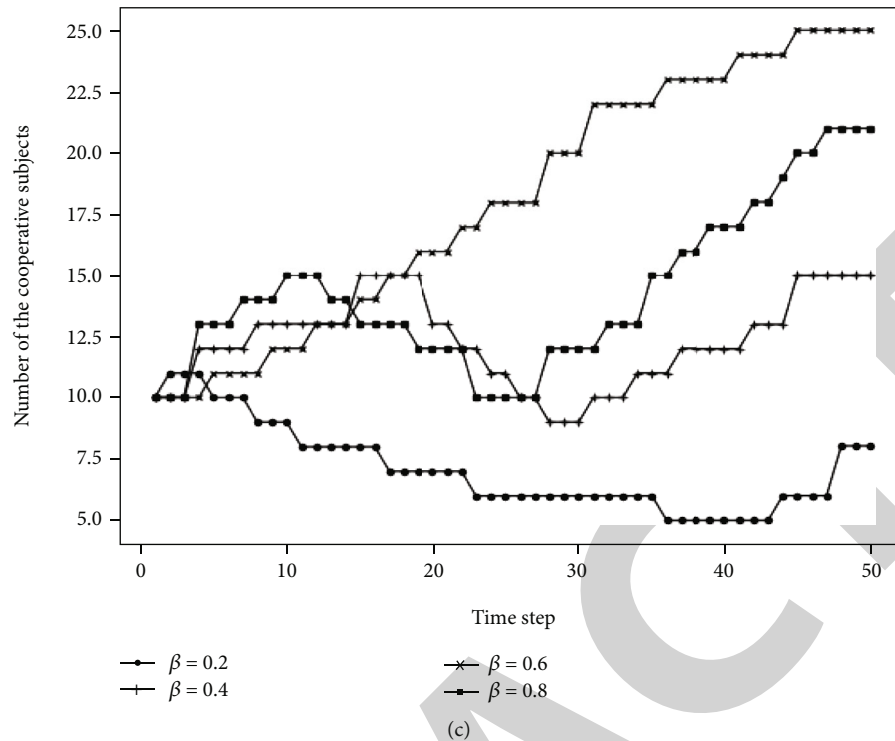


FIGURE 4: The impact of the mechanism on the construction of smart university campuses: (a) leadership organizations; (b) special funds; (c) assessment rewards and punishments.

promote the coordination of the subjects of smart campus construction while taking comprehensive planning as the foothold, including information, capital, and technology, and strengthen cooperation. In addition, we strive to make use of an appropriate organizational mechanism, funding mechanism, and reward and punishment mechanism to coordinate the relationship between construction subjects and give better play to the functions of each subject.

6. Conclusion

Taking the construction of smart campuses in universities as the research object, this paper qualitatively analyzes the subjects of construction and their relationship and quantitatively constructs an evolutionary game model based on a complex network. Twenty-five subjects are selected as nodes. According to the relationship between them and taking the synergy intensity index as the weight, a weighted scale-free network is constructed. At the same time, a cooperative evolution game model is established. Through simulation, the impact of collaboration between subjects, leadership organization, special funds, assessment, reward, and punishment on the construction of a smart university campus is deeply investigated. The results show that (1) appropriate subject coordination is very important, which is helpful to prevent the phenomenon of *getting something for nothing*, to establish trust and cooperative relations through multiple games, and finally to achieve win-win results. (2) An appropriate leadership and organization mechanism can strengthen supervision and promote the construction process of smart university campuses. An appropriate special fund mecha-

nism can clarify the source of funds and ensure sustainable development. Appropriate mechanisms of assessment, reward, and punishment can mobilize subjects' enthusiasm to participate and improve the popularity of construction achievements. The analysis and results of this paper have guiding significance for universities in promoting the construction of a complete, comprehensive, and high-quality smart campus. However, this paper is a preliminary study based on the abstraction of some factors. The influence of the changes in network structure and strategy renewal on the overall decision-making behavior of smart campuses needs to be further studied.

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 known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] S. Suxuan, Y. Xianmin, and S. Ziqiang, "Research on the construction of a new generation of smart campus infrastructure platform in universities under the background of educational informatization 2.0," *Modern Educational Technology*, vol. 29, no. 8, pp. 18–24, 2019.

Retraction

Retracted: Enterprise Management Decision and Financial Management Based on Dynamic Cost Volume Profit Model

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Guo, "Enterprise Management Decision and Financial Management Based on Dynamic Cost Volume Profit Model," *Journal of Function Spaces*, vol. 2022, Article ID 9016060, 8 pages, 2022.

Research Article

Enterprise Management Decision and Financial Management Based on Dynamic Cost Volume Profit Model

Yanlin Guo 

Department of Big Data and Accounting, Henan Institute of Economics and Trade, Zhengzhou 450000, China

Correspondence should be addressed to Yanlin Guo; guoyanlin@henetc.edu.cn

Received 25 April 2022; Revised 24 May 2022; Accepted 25 May 2022; Published 22 June 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Yanlin 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.

Under the background of today's times, the internal and external environment of enterprises is complex and changeable. It is of great significance for the survival and development of enterprises to continuously reduce the business risks and improve the economic benefits of enterprises by adopting scientific decision-making methods. CVP (cost volume profit) analysis is a model established by sorting out variables related to business decisions in enterprise production through mathematical modeling. Based on the interpretation of cost behavior by multiple cost drivers, this paper establishes a dynamic CVP model of traditional CVP analysis, breaks through the limitations of original assumptions, and improves the application value of CVP analysis in practice. The model is applied to the business decision-making and financial management of a case enterprise, and the products with different characteristics of the enterprise are modeled and analyzed, so as to provide targeted improvement suggestions or strategies for the enterprise's product production decision-making. The research shows that the dynamic CVP model can be used to formulate the financial management strategy to optimize the profits of enterprises, and it can achieve stable profits and development in the market competition.

1. Introduction

With the vigorous development of China's socialist market economy, the competition among enterprises is becoming fiercer and fiercer, and the competitive pressure is increasing. In order to improve the management level of the company's management decision and expand the company's share in the market, scientific management methods must be used to reasonably control the cost of the enterprise and maximize the production profit. The secondary industry, which is dominated by manufacturing, is still an important part of China's real economy. However, the manufacturing enterprises' sales profit rate is declining year by year because of the relatively large production cost of their products in sales revenue and the increasing expenses of raw materials and labor wages, so the development trend of the industry is not optimistic [1]. In real economic life, the number of buyers and sellers is often very limited. Under the limited number, the market cannot be completely competitive. In

order to establish CVP (cost volume profit) analysis theory, we must make some basic assumptions about the above-mentioned complex relationships, so as to strictly limit the scope of CVP analysis. If these basic assumptions are not met, we can conduct CVP extended analysis.

CVP analysis is an important management tool for enterprises to forecast, make decisions, and control costs, and it plays an important role in enterprise management. Fuksa et al. [2] hold that the prerequisite for the implementation of CVP model is to simplify the application environment and meet the requirements of convenient calculation and analysis. Agrawal and Yadav [3] have found that CVP analysis can help the workshop understand the production and operation situation, find out various related reasons that have a significant impact on the profits of enterprises, and make improvements according to the degree of impact, so as to control the operating costs of enterprises [4]. Riehle et al. [5] point out that CVP analysis is an analytical tool suitable for short-term operation and management. The

main reason why this model requires the assumption of “short-term operation” is that if the products produced by an enterprise belong to a variety with a long cycle, their habits such as fixed cost and variable cost may not remain unchanged. Propose a cost calculation method based on activity. At present, enterprises in many western countries are using activity-based costing, and it is also a method to calculate the product cost of enterprises in line with modern requirements. Xu et al. [6] think that CVP analysis can help hotel management to make correct business decisions [7] and use CVP to analyze the main indicators of hotel industry model to find and solve the problems [8]. Beugelsdijk and Jindra [9] suggest that enterprises should learn to use CVP analysis flexibly and rely on the continuous research and summary of their own conditions to find out the most appropriate method to provide effective information for the decision-making and development of enterprises. Van-dekerkhof et al. [10] embed the time factor into CVP analysis and propose to adjust the product structure to cope with the rising cost by identifying and eliminating the factors that are insensitive to product sales.

CVP analysis takes cost behavior analysis as the premise and variable cost method as the basis to calculate profit. Its basic formula reveals the relationship among price, cost, business volume, and profit [11]. Therefore, it is necessary to analyze and summarize the company’s operating conditions and operating results through financial analysis tools. Scientific financial analysis is the “compass” in the process of enterprise management. Enterprises need to apply financial analysis to the decision-making of production, operation, and management, to help the production point out the direction, check for leaks and fill vacancies, reduce waste, and improve production efficiency [12]. CVP analysis can also help units scientifically control the financial risks in the development process of units, effectively control the influence of various risk factors on the overall business development of units, and ensure the healthy development of units. On the basis of CVP-related knowledge and theory, combined with the actual operating situation of W company, this paper analyzes the correlation among cost, business volume, and profit and uses the basic CVP model to analyze the investment and financing of the new factory building of W company, so as to guide the future development direction of W company.

Based on the explanation of cost behavior by multiple cost drivers, this paper creatively establishes the dynamic CVP model of traditional CVP analysis. It breaks through the limitations of the original assumptions and improves the application value of CV P analysis in practice. The model is applied to the business decision-making and financial management of the case enterprise to model and analyze the products with different characteristics of the enterprise, so as to provide targeted improvement suggestions or strategies for the product production decision-making of the enterprise. This paper shows that the dynamic CVP model can be used to formulate financial management strategy, optimize enterprise profits, and realize stable profits and development in market competition. When CVP is used in business decision-making and financial management of enterprises, the introduction of net present value method

can improve the timeliness and accuracy of analysis and is more practical.

2. Research Method

2.1. CVP Analysis Theory. CVP analysis is the abbreviation of cost, business volume, and profit analysis. By analyzing the relationship among cost, business volume, and profit, it finds out the law of the relationship among them and determines the functional relationship, thus providing a basis for effective business decision-making and target control. Its principle is widely used in enterprise forecasting, decision-making, planning, and control.

The research of CVP analysis is based on a premise and three basic assumptions, namely, the linear relationship hypothesis, the production-marketing balance hypothesis, and the variety structure stability hypothesis.

Assumption of correlation range and linear relationship: CVP analysis starts with the division of cost habits and takes this as the first step. Under the condition that the fixed cost remains relatively stable, the variable cost is in direct proportion to the business volume. However, through the research and investigation of a large number of enterprise examples and the summary of experience, it can be known that the actual situation is filled with a large number of nonlinear relationships among total cost, product revenue, and business volume [13, 14].

Assumption of product structure stability: under the assumption that the original product structure remains unchanged, the sales fluctuation brought by market changes will inevitably lead to a big difference between profits and expected profits. Through this assumption, the managers of the company can focus on the impact of price, cost, and business volume on the company’s profits [15].

Assumption of balance between production and sales: assume that all products produced in the current period are sold in the current period, and combine the output and sales volume into one.

Based on the above basic assumptions, the basic model of CVP analysis can be expressed as

$$\Pi = S - TV - F = P \times Q - V \times Q - F, \quad (1)$$

where Π is earnings before interest and tax, S is sales revenue, TV is variable cost, F is fixed cost, P is unit price, Q is production and sales volume, and V is unit variable cost.

The model is a basic profit and loss equation, which clearly expresses the quantitative relationship among cost, quantity, and profit, and can be transformed in form to predict unknown variables through known variables.

Breakeven point is a special state in which the input cost and output income of an enterprise are balanced. Calculate the output boundary point in this balanced state. Figure 1 shows the breakeven relationship diagram under the linear cost-output function [16].

As can be clearly seen from the above Figure 1, the straight line $R = px$ is the sales revenue line, while $C = C_f + C_v x$ represents the total cost line, where C_f represents

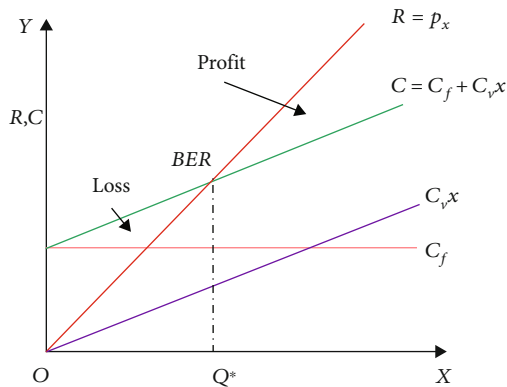


FIGURE 1: Schematic diagram of linear breakeven relationship.

the fixed cost and C_v represents the variable cost. The intersection point of the two line segments Q^* breakeven point BEP.

At the left end of the breakeven point, that is, the part of $x < Q^*$, the total cost line of the enterprise is above the sales revenue line, and the enterprise is at a loss at this time. When $x > Q^*$, the total cost line of the enterprise is below the sales revenue line, and the product revenue is greater than the cost, and then, the enterprise is profitable.

2.2. Application of Dynamic CVP Model in Enterprise Management Decision. Applying CVP analysis to specific enterprises can fully help enterprises to judge the safety situation of production and business activities, so as to control it in the most favorable situation of enterprises, master the changes of various factors, find out the problems, and solve them in time, so as to maximize the profits of enterprises [17, 18]. Through the analysis of profit target, enterprises can analyze the influence of various factors on the realization of profit target according to their own actual situation, so as to make correct decisions and adjustments to realize profit target.

CVP model reveals the relationship among cost, production, and sales volume and book accrued profit, which can help enterprises determine production and sales volume reasonably. However, sometimes, decision-makers are more concerned about the impact of production and sales on cash flow or working capital than the corresponding profits [19].

Based on the above theoretical basis of cash flow analysis and the definition of cash and noncash items, the cash flow model is obtained by transforming CVP model. The cash flow model highlights the significance of cash flow and is a modification of CVP model. The cash breakeven point formula under the cash flow model is as follows:

$$Q_C^* = \frac{F_C}{P_C - V_C}, \quad (2)$$

where Q_C^* is the cash breakeven point, F_C is a fixed cash cost, P_C is the unit cash selling price, and V_C is the unit cash variable cost.

Under the linear hypothesis, the general multivariety CVP analysis model considering various constraints is as follows:

Objective function:

$$\pi = \text{Max} \sum_{i=1}^n (p_i - b_i)x_i - \sum_{i=1}^n a_i. \quad (3)$$

Constraints:

$$\begin{cases} \sum_{i=1}^n r_{ij}x_i \leq c_j, & j = 1, 2, \dots, m, \\ x_i \geq 0, & i = 1, 2, \dots, n, \end{cases} \quad (4)$$

where π is profit and P_i is the selling price of the i th product, b_i is the unit variable cost of product i , x_i is the output of the i th product, a_i is the fixed cost of the i product, r_{ij} the j th resource used to produce the i th product, and j stands for all kinds of available resources.

This CVP analysis considers the influence of selling price, variable cost, and various constraints on decision-making but does not consider the influence of fixed cost on decision-making and does not consider the difference between avoidable cost and avoidable cost in fixed cost.

Through the above analysis, the author believes that for the multivariety CVP analysis with avoidable fixed costs, a new model must be established, which Kaplan calls the fixed cost problem. A new simple integer variable Z_i can be defined before creation.

The improved model is as follows:

Objective function:

$$\pi = \text{Max} \sum_{i=1}^n (p_i - b_i)x_i - \sum_{i=1}^n a_{i1} - \sum_{i=1}^n a_{i2}Z_i. \quad (5)$$

Constraints:

$$\begin{cases} \sum_{i=1}^n x_i \leq C_j, & j = 1, 2, \dots, m, \\ & i = 1, 2, n, \\ x_i \leq Z_i U, & x_i \geq 0, \\ & Z_i = 0, 1. \end{cases} \quad (6)$$

The U of the constraint condition is an arbitrary but very large number, and its value is greater than the maximum output of any product; that is, an upper limit is set on the output and sales of any product, and the output and sales of any product cannot be greater than the U value. The value of U can be determined as a numerical value according to the actual situation of the enterprise.

The purpose of CVP sensitivity analysis is to analyze the limits of the changes of related factors when profit target changes from profit to loss and the sensitivity of the changes of related factors to profit, so that enterprises can determine

the focus of management according to sensitive information and make timely adjustments to decisions.

However, because the two CVP analyses are based on different cost behaviors, the content and results of sensitivity analysis will be different. Sensitivity analysis of sales-related factors provides enterprises with the sensitivity of profits to sales. Different sensitive information will lead enterprises to different management directions, and wrong sensitive information may even lead to mistakes in enterprise decision-making.

Let the profit $\prod_i = 0$ determine the critical value of each related factor of product i .

$$Q_i = \frac{F_i + \sum_{j=1}^m d_j y_{ij}}{P_i - V_i}. \quad (7)$$

Similarly, the maximum allowable activity cost of any activity of the product can be deduced:

$$b_r y_{ir} = (P_i - V_i)Q_i - \sum_{j=1}^{r-1} b_j y_{ij} - \sum_{j=r+1}^m b_j y_{ij} - F_i, \quad 1 \leq r \leq m. \quad (8)$$

In the dynamic CVP analysis model, the relevant factors have different influences on profits, and the sensitivity coefficient can also be used as an index, and the percentage change of the sensitivity coefficient profit factor is the percentage change. By analyzing the sensitivity coefficient of each factor to judge the sensitivity of profit to the change of factors, it is helpful for managers to distinguish primary and secondary, take appropriate measures, and adjust and realize profit target.

2.3. Practice of CVP Analysis in Financial Management. In the process of implementing long-term investment, the unit can use CVP analysis method to accurately grasp the capital turnover of the unit, control the risk of capital use, and lay a foundation for the unit to obtain better operating benefits. Based on this, CVP analysis can provide support for the long-term scientific decision-making of the company. In the process of unit development, the operating profit will be affected by many factors, among which the cost has a great influence. And scientifically control the use of cash flow according to the risk control situation, provide support for strengthening the overall control effect of the unit and optimizing the fund allocation effect, and promote the development of the overall operating efficiency within the unit.

We should continue to produce this product when the company has sufficient resources. If the production is stopped, the fixed cost will not be reduced, but there will be no sales revenue, and the company's loss will increase instead. Because the marginal contribution of each product is different, if the sales structure of the product changes, it will lead to the change of expected profit and actual profit, which will affect CVP analysis. Therefore, when CVP analysis is conducted, it is necessary to assume that the variety structure is stable, which is beneficial for enterprise man-

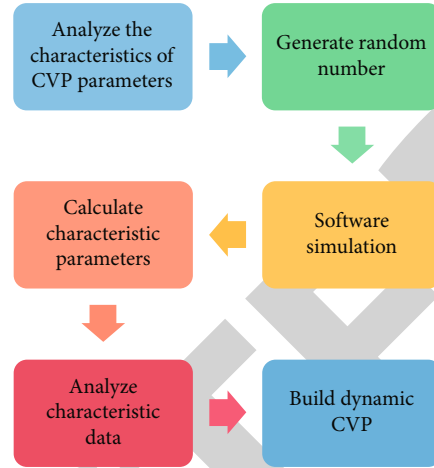


FIGURE 2: Frame diagram of dynamic CVP model.

agers to analyze the impact of price, cost, and business volume on operating profit.

Whether it is a linear model or a nonlinear model, it is assumed that the relevant variables are deterministic variables. However, in real decision-making, many factors in the future are not deterministic, but random or uncertain. For example, the influence of price market factors will produce random fluctuations. The unit cost also has randomness. The simulation process is shown in Figure 2.

In this paper, Monte Carlo fitting method, breakeven analysis, and graphic method will be used to help managers make decisions. By generating discrete random numbers, CVP graphs under uncertain environmental conditions will be given dynamically.

The complexity and variability of the economy in reality make it difficult to predict the sales unit price, unit variable cost, fixed cost, and sales volume with certain values, and all factors are uncertain. The main elements of the economic system are people with limited rationality and irrationality. Therefore, strictly speaking, the economic system is composed of all complex activities of people and the complex relationship between people and the environment. The biggest feature of the economic system lies in people's participation. Any economic activity and economic behavior are carried out under the guidance of man, which depends on man's limited rational and irrational abilities, such as thought, will, preference, knowledge, psychology, and values. Because people's bounded rationality and irrationality are extremely complex and nonlinear, the economic system is bound to show complexity.

It is assumed that the selling unit price, unit variable cost, and fixed cost are certain values, and the sales volume is a random variable.

If the sales volume Q_i of product i obeys the normal distribution $Q \sim N(\mu, \sigma^2)$ with expected value μ and variance σ^2 , according to the expected profit,

$$E\left(\prod_i\right) = (P_i - V_i) \times \mu - \sum_{j=1}^m b_j y_{ij} - F_i. \quad (9)$$

TABLE 1: Marginal contribution rate of products in workshop.

Product category	Unit price (yuan)	Sales	Sales revenue (yuan)	Sales proportion (%)	Unit variable cost (yuan)	Marginal contribution rate (%)	Fixed cost (yuan)
A	101	213660	21,579,660	81.326	90	10.66	455280
B	40	110240	4,409,600	18.601	38	11.87	

TABLE 2: Breakeven point and margin of safety.

Product number	Sales proportion (%)	Breakeven point sales (yuan)	Capital preservation amount	Margin of safety
A	81.326	3403316	33151	-12868
B	18.601	832269	18697	-7201

In which $\mu = \sum_{j=1}^n Q_{il} Y_i$ is the sales volume, n is the number of results of sales volume, Q_{il} is the sales volume of l result of product i , and Y_l is the probability of l result.

If the decision-maker is risk-averse, he will choose the scheme with higher profit probability or lowest loss probability; if it is risk-neutral, he will choose the scheme with the highest expected profit; if it is risk-taking, he will choose the scheme with higher risk but higher profit.

If the production capacity of each enterprise permits, balanced output should be included in the plan as the target output of the enterprise. At this time, the profit of the enterprise is the balanced profit. But in reality, the enterprise may not take the balanced profit as the profit target for some reason.

Assume that the i th enterprise takes L_0 as the profit target. Under the static game of complete information, other enterprises do not know its choice and still produce according to the balanced output Q_j^* ($j \neq i$). At this time, the profit target model is as follows:

$$\begin{cases} L_0 = Q_i P \left(\sum_{j \neq i} Q_j^* + Q_i \right) - C_i(Q_i), \\ L_k(Q_k^*) = Q_k P \left(\sum_{j \neq i} Q_j^* + Q_i \right) - C_k(Q_k), \quad k \neq i. \end{cases} \quad (10)$$

3. Result Analysis

3.1. Breakeven Analysis. This part takes H enterprise as the research object. The dynamic visualization of CVP analysis chart is realized by using tools such as Excel, which provides a scientific basis for enterprise management decision-making. Firstly, through a large number of literature reading and field case investigation, the collected data are sorted out by using statistical knowledge. Secondly, the CVP analysis model of different products in line with the actual situation of the enterprise is constructed.

When dividing the cost of H enterprise, it is found that the production of the two factories in H enterprise has no influence on each other, so the breakeven point of the two factories is calculated separately during the breakeven analy-

sis. Now, the analysis is based on the actual production and sales situation of the workshop of H enterprise in a certain month. Make a breakeven analysis of the products in the workshop production line (see Table 1 for the marginal contribution rate, production, and sales of various products in the workshop).

According to Table 1, calculate the weighted average marginal contribution rate of various products in the workshop: $81.326\% \times 10.66\% + 18.601\% \times 11.87\% = 10.877\%$.

So the workshop breakeven point sales are $455280 \div 10.877\% = 4,185,712.972$ (yuan).

According to the sales proportion of the two products, the guaranteed amount and safety margin of each product in the workshop are calculated, respectively (see Table 2).

According to the margin of safety in Table 2, the sales volume of the two products in the workshop did not reach the breakeven point, which did not bring profit to the enterprise. However, the marginal contribution rate of the two products in the workshop was positive, so the enterprise should increase its production and sales to make the workshop turn losses into profits.

Dynamic CVP analysis is more detailed in the analysis of fixed costs and pays attention to the process. To some extent, all fixed costs can find cost drivers, which can be the costs that change with their respective cost drivers. In the actual calculation process, the cost-benefit principle is taken into account, and all cost drivers will not be traced, but only those that the company considers to be key and feasible can be considered.

As shown in Figure 3, the plane is the combination plane of capital preservation workload, and any point on the plane represents the capital preservation combination of three workloads. This plane is also the maximum limit of the combination of three kinds of workload, and the combination of three kinds of workload should not exceed this plane; otherwise, the enterprise may lose money due to the unreasonable allocation of enterprise resources. When the combination of the guaranteed sales volume and the guaranteed operation volume of product A reaches the guaranteed state at the same time, product A will truly achieve breakeven.

Similarly, we can analyze the combination of guaranteed sales volume and guaranteed operation volume of product B, which is omitted here. From the analysis, it can be seen that

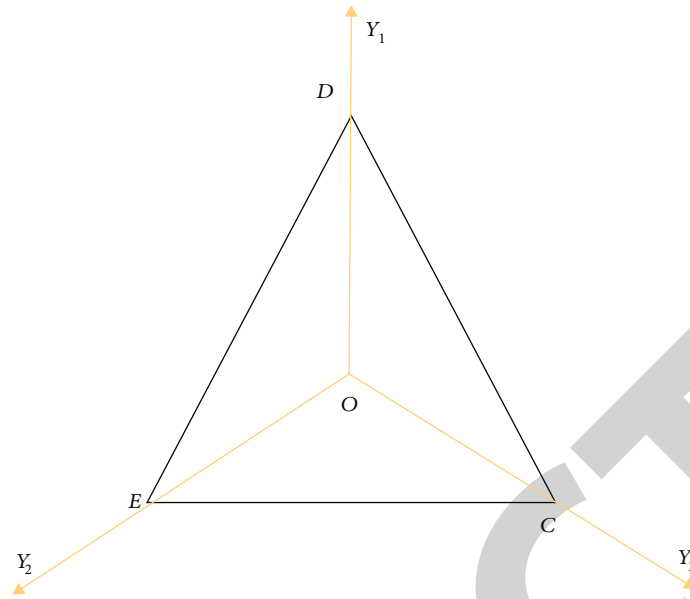


FIGURE 3: Product A breakeven workload combination.

the actual sales volume of product A has not reached the guaranteed sales volume, and the actual operation consumption has exceeded the guaranteed operation consumption. Product A is actually eroding the profits of enterprises.

Analyze the CVP decision indicators of the following categories, as shown in Figure 4.

It can be seen from the analysis of the combination of guaranteed sales volume and guaranteed operation volume of products A and B that the H weighted average product safety margin of various products of the enterprise is at a normal level. In essence, the margin of safety calculated by the traditional and dynamic CVP analysis models, which indicates the overall production and operation safety of the company, should be the same theoretically. That is to say, CVP analysis is not to change the result of the enterprise, but to trace and analyze the process leading to this result, so as to achieve the purpose of improving the cost and profit.

In the actual calculation of this case, the results of two CVP analysis and calculation are 62.95% and 64.41%, respectively, with certain errors. Because the fixed cost is roughly allocated and a large number of decimal places are reserved in the calculation process, it is normal for the two calculation results to have slight errors.

Dynamic CVP analysis is more conducive to the analysis of each product when the variety structure of an enterprise is unstable. Dynamic CVP analysis can establish the CVP analysis model of each product, analyze and judge each product in detail, avoid the existence of a single product eroding the profits of enterprises, and help enterprises to master all the information of different products and clearly reveal the profitability of various products.

3.2. *Dynamic CVP Model Analysis.* This paper selects C product data produced by H enterprise as an example. In this section, the changes of cost and unit price are set as random values for simulation analysis and model construction.

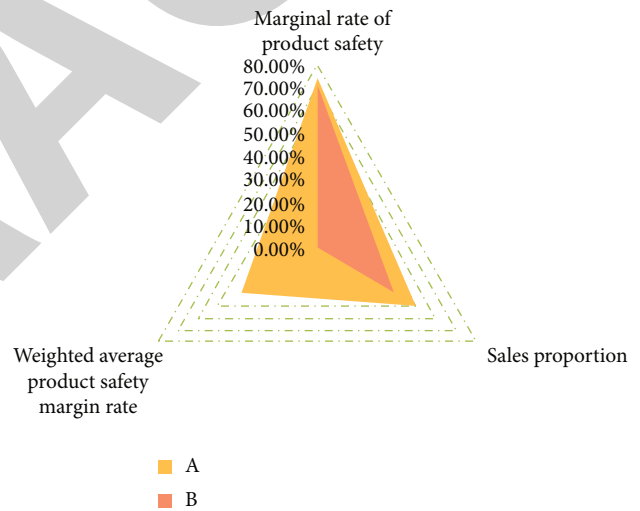


FIGURE 4: H weighted average product safety margin rate of various products of enterprises.

The fluctuation of cost is expressed by the mean value of unit variable cost and its standard deviation, and the probability of different unit prices reflects the randomness of prices. When filling in the calculation area, enter the formula of each factor into it, and calculate the sales revenue, variable cost, total cost, profit, and other values. Dynamic CVP analysis breakeven dynamic calculation results are shown in Figure 5.

This figure shows that when the random unit price is 56.39 yuan, the sales volume exceeding 43,206 can satisfy the enterprise's turning losses into profits on this product. Of course, when sales volume, fixed cost, and other variable factors are linked to controls, or unit cost is randomly assigned to other data, it can change with the data. This group of models can directly show the profit and loss of each

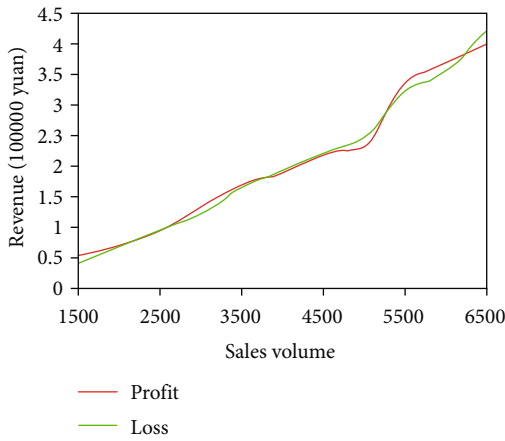


FIGURE 5: Dynamic CVP analysis of breakeven dynamics.

TABLE 3: Safety margin simulation statistics.

Margin of safety	Frequency of occurrence	Probability (%)	Cumulative probability (%)
Very safe	1367	42.01	42.01
Safe	633	20.36	67.21
Relatively safe	390	12.17	80.09
Careful	241	8.87	89.66
Danger	386	12.15	100

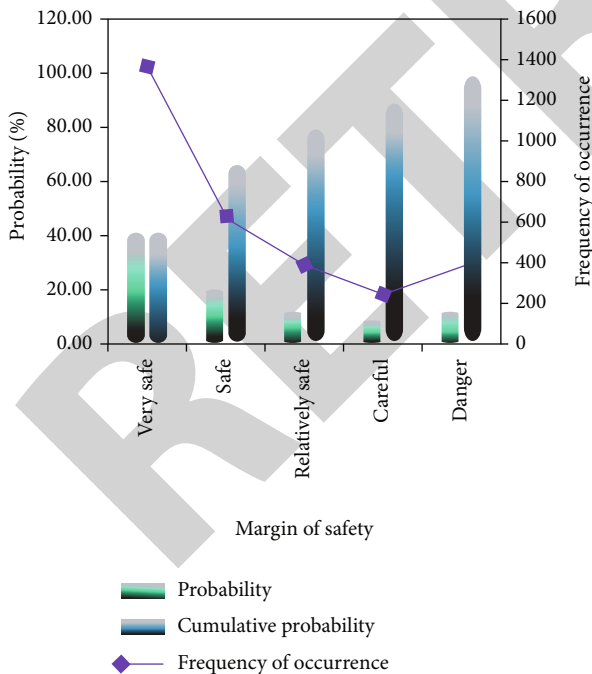


FIGURE 6: Statistical chart of safety margin simulation.

sales volume by means of controls, which is convenient for decision-makers to budget or organize production.

The purpose of dynamic concept introduction is to quantify the existing risks as much as possible and to form

statistical data on the impact of the risks on the cost and unit price and present them in a probabilistic way, so as to facilitate enterprises to choose production schemes and analyze the feasibility of related activities.

Using the above function, the simulated breakeven point is brought into the safety margin formula, and the business risks of enterprises with different unit prices under a certain sales volume can be obtained. The more tests, the higher the reliability. Because of the limited conditions, this paper simulates 3500 times in total and then counts the simulated results by using the function statistics and counts the times of each safety result. The results are shown in Table 3 and Figure 6.

As can be seen from the figure, the probability that the margin rate of safety of product C is above “safer” is 80.09%, which shows that the product has good profitability. Although the cost and price are easily affected by the outside world, its development prospect cannot be underestimated; if it can successfully open up the market, product C may become the main profit force of the case company.

When an enterprise’s operation needs to be maintained and developed, its lowest point is that the stipulated price is mainly based on the enterprise’s related cost guarantee. Under the current market conditions, on the basis of product price and cost determination, the enterprise is urged to make relevant price limits. Therefore, enterprises should calculate CVP on the premise of multicost factors. Dividing costs into short-term variable costs, long-term variable costs, and fixed costs can meet the business results of enterprises and at the same time effectively increase the accuracy of enterprise management and control. The analysis results are more strategic.

4. Conclusion

To sum up, by using CVP analysis, we can make a comprehensive and in-depth analysis of the current production and operation situation of enterprises, find out the aspects in production and operation that need to be improved urgently, and provide ideas and methods for improvement. Dynamic CVP model analysis expands the application scope of CVP analysis, not only considering the influence of variable cost, fixed cost, production, and sales volume and unit price on cost and profit but also increasing the analysis of workload. According to the functional characteristics of the product cost line, different CVP dynamic models are constructed for it in turn, which is more in line with the actual situation of the product and improves the reliability of data analysis. In the real environment, enterprises need to rely on economies of scale to win if they want to survive and develop. Therefore, when CVP is used to make business decisions and financial management for enterprises, the introduction of net present value method can increase the timeliness and accuracy of analysis, which is more practical. However, the research still has some limitations. In real economic life, the number of buyers and sellers is often very limited. In the case of limited quantity, the market cannot compete completely. In order to establish CVP (cost volume profit) analysis theory, we must make some strict restrictions

Retraction

Retracted: The Experimental Validation of a Digital Filter Mathematical Modelling for Nuclear Fuel Vibration Monitoring System

Journal of Function Spaces

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

Copyright © 2023 Journal of Function Spaces. 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. Yong, H. Shiyong, H. MengXuan et al., "The Experimental Validation of a Digital Filter Mathematical Modelling for Nuclear Fuel Vibration Monitoring System," *Journal of Function Spaces*, vol. 2022, Article ID 9741419, 11 pages, 2022.

Research Article

The Experimental Validation of a Digital Filter Mathematical Modelling for Nuclear Fuel Vibration Monitoring System

Han Yong ¹, Huang Shiyong ², Han MengXuan ³, Zhao Tao ¹, Wen Guang ¹,
Gu Mingfei ⁴, and Zhang Qi ^{1,5}

¹Chengdu Vocational & Technical College of Industry, Chengdu 610218, China

²Chengdu China Nuclear Haichuan Nuclear Technology Co., Ltd., Chengdu 610041, China

³University of Electronic Science and Technology of China, Chengdu 610041, China

⁴Science and Technology on Reactor System Design Technology Laboratory, Nuclear Power Institute of China, Chengdu 610213, China

⁵School of Intelligent Manufacturing, Panzhihua College, Panzhihua 617000, China

Correspondence should be addressed to Zhang Qi; pzhuzq@stu.scu.edu.cn

Received 4 April 2022; Revised 10 May 2022; Accepted 16 May 2022; Published 4 June 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Han Yong 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.

Aiming at vehicles will produce shocks and vibrations for various reasons in the process of transportation and handling, any shock and vibration that exceeds the regulations would cause damage to the nuclear fuel being transported. In order to monitor and analyze the shock and vibration data scientifically, a monitoring system based on MEMS acceleration sensor technology is developed to measure shocks and vibrations. The measured result involves some noise caused by the sensitivity of the sensor and the electrical noise, and a digital filter mathematical modelling is introduced to reduce noise and improve the measured result accuracy. In order to verify the effectiveness of the device designed in this paper, the device and the high-precision seismic acceleration sensor AC73 are, respectively, used for field tests; the results show that the test precision of the designed device can meet the requirements, and also, the designed shock and vibration monitoring system can record and store the three-axis acceleration, temperature, humidity, pressure, time, and other parameters that exceed the set threshold during nuclear fuel transportation in real time and complete the fuel assembly transportation status assessment and promptly remind the transportation personnel to take safety precautions and treatment measures. Therefore, it can be widely used in the monitoring, recording, and postevent data analysis of the transportation process of important products such as nuclear fuel.

1. Introduction

Nuclear fuel is expensive and easy to be impacted and vibrated during transportation, which will affect its use. Therefore, the internal or external damage monitoring of nuclear fuel caused by various shocks and vibrations during nuclear fuel transportation should be fully considered. In order to ensure the safety of nuclear fuel transportation, special vehicles and professional drivers are generally used for speed limited transportation, supplemented by mechanical accelerometers as the indication of exceeding the acceleration limit. The mechanical accelerometer, as the auxiliary monitoring of shocks and vibrations, can indicate whether the acceleration exceeds the limit, but it cannot give infor-

mation about the specific value and occurrence time of the acceleration suffered by the fuel assembly and will not give an early warning. Due to the complex transportation conditions, professional drivers need to often check the road conditions before deciding on subsequent operations; mechanical accelerometers cannot provide monitoring data records as postanalysis, which is lack of scientificity. Therefore, it is necessary to use advanced sensor technology and monitoring technology to design a safe, efficient, and scientific nuclear fuel transportation monitoring system. Senodia Technologies has independent intellectual property rights of commercial MEMS inertial sensor production technology, and MEMSIC Inc. (USA) developed products that integrate MEMS and signal processing circuits into a single chip. Liu

Huijuan of Ocean University of China and others carried out the design of vibration monitoring system based on MEMS [1]. Zhang Jianguang of China Academy of Radiation Protection and others carried out the experience summary of test verification of radioactive goods transport containers [2, 3]. Xu Lei of the Army Engineering University and others carried out research on the design of vibration acceleration measurement system based on MEMS acceleration sensor [4]. By far, the acceleration monitoring technology of ordinary commercial transportation has gradually integrated into the domestic general logistics transportation, mechanical and electronic manufacturing, automobile manufacturing, and other fields and has achieved certain application results. However, there are relatively fewer reports on the acceleration monitoring of fuel assembly transportation in nuclear power plant.

According to the requirements of fuel assembly transportation acceleration and fuel assembly special transportation container and the special requirements of fuel assembly transportation container, a low-power digital fuel assembly transportation container acceleration monitoring system (hereinafter referred to as transportation acceleration monitoring system) based on MEMS acceleration sensor and single-chip microcomputer technology is designed, as it has the function of real-time automatic recording and storage of acceleration information in the whole transportation process, and it can timely help the transportation personnel take safety precautions and facilitate subsequent analysis and summary and meet the requirements of nuclear fuel safety transportation analysis [5–9].

2. System Design

2.1. Technical Requirement of Shock and Vibration Acceleration Recorder

- (1) Acceleration (triaxial): range: $0 \sim \pm 10 \text{ g} / \pm 20 \text{ g} / \pm 40 \text{ g}$, settable; nonlinearity: $\pm 0.1\%$ FSR ($\pm 10 \text{ g}$) and $\pm 1.3\%$ FSR ($\pm 40 \text{ g}$)
- (2) It has RS232 serial computer interface and alarm indication output interface
- (3) The power is not more than 90 MW, the battery power supply mode, and the continuous working time is more than 15 days
- (4) Overall weight: no more than 1.5 kg; size: no more than $110 * 100 * 80 \text{ mm}$

2.2. System Working Process

- (1) Before the transportation of nuclear fuel and other valuables, take out the shock and vibration acceleration recorder, activate it through PC monitoring application software, and install and fix it on the transported equipment for transportation process monitoring
- (2) In the process of transportation, in case of shocks and vibrations, the shock and vibration acceleration

recorder measures its value through the integrated temperature, humidity, pressure, and triaxial acceleration sensors, triggers recording and alarm when it exceeds the set value, reminds the transportation personnel to pay attention, and stores the temperature, humidity, pressure, triaxial acceleration, and corresponding time when it exceeds the threshold

- (3) After transportation, remove the shock and vibration acceleration recorder from the transported equipment, and then, use PC monitoring application software to read the recorded data for relevant analysis, so as to mine the value of the data, improve the transportation quality, and reduce losses

Its schematic diagram of system working process is shown in Figure 1.

2.3. The Shock and Vibration Acceleration Recorder Design

2.3.1. Selection and Calculation of Key Technologies

(1) *MEMS Triaxial Acceleration Sensor*. The MEMS triaxial acceleration sensor selects the programmable maximum shock and vibration measurement range of $\pm 20 \text{ g}$ and $\pm 40 \text{ g}$.

(2) *Battery Capacity*. In the actual process of nuclear fuel transportation, fuel assemblies are often transported through the combination of special railway line and road transportation. According to the layout of China's nuclear fuel supply and use enterprises, the single transportation distance of nuclear fuel is up to 2000 or 3000 kilometers, and the transportation time lasts about 7 days. The working time of 15 days needs to be considered. The calculation formula of battery capacity is as follows:

$$Q = \frac{P \times t}{k \times U}, \quad (1)$$

where Q is the required lithium battery capacity (Ah), P is the load power (W), t is the daily electricity hours, U is the standard lithium battery voltage, and k is the discharge control coefficient (0.75~0.8).

The main chip power consumption is shown in Table 1.

According to Table 1, the total power consumption is 22.91 mA, calculated as 25 mA, and the voltage is calculated according to the working voltage of 3.7 V; then,

$$Q_1 = \frac{25 * 24 * 15}{0.75 * 3.7} = 3244 \text{ mAh}. \quad (2)$$

Considering that the circuit needs voltage transformation, and the comprehensive efficiency is taken as 50%, the following can be obtained:

$$Q = \frac{Q_1}{0.5} = 6488 \text{ mAh}. \quad (3)$$

Therefore, two 18650 type rechargeable lithium batteries with a nominal voltage of 3.7 V and capacity of 3400 mAh

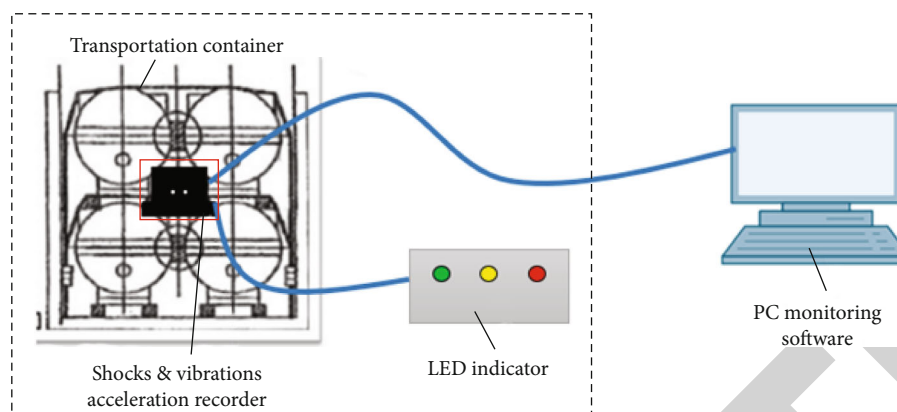


FIGURE 1: Schematic diagram of system working process.

TABLE 1: Main chip power consumption.

Name	Power consumption (mA)	Remark
MCU (ARM)	12.8	16 MHz operation
Triaxial acceleration sensor	0.15	
Storage	0.7	It is estimated that 10% of the time will be spent reading and writing data
Pressure sensor	4	
Temperature and humidity sensor	0.5	
LDO power supply and power reference	0.46	
Serial port protection chip	1	
Optical coupling relay	3.3	The outputs are mutually exclusive, with only one output at a time

can ensure the normal operation of the shock and vibration acceleration recorder for 15 days.

2.3.2. Mechanical Structure Design. In order to ensure the strength and measurement accuracy of the product, the shock and vibration acceleration recorder is designed with a compact structure, which is composed of an upper cover plate and an integrated base. The integrated base is divided into three parts according to the structure and function; from bottom to top are circuit board installation area, baffle plate, and battery area. The sensor measuring board and main board are fixed in the circuit board installation area, the upper side of the main board is equipped with a baffle plate which plays the role of support and electromagnetic interference shielding, two large capacity rechargeable lithium batteries are placed on the baffle plate, and an upper cover plate is placed on the upper part of the battery. The upper cover plate and the integrated base are compacted with 4 fastening screws through the sealant strip to ensure the reliable fixation of the battery and provide the IP65 protection function of the instrument. The integrated base is provided with a connecting hole, which can be fixed to the transportation parts with screws during operation. The overall dimension of the shock and vibration acceleration recorder is about 105 mm (length) \times 86 mm (width) \times 65 mm (high). The effect diagram is shown in Figure 2.

2.3.3. Electronic Circuit Design. The electronic circuit is the core part of the shock and vibration acceleration recorder, which undertakes the role for measurement, recording, and storage of triaxial shock and vibration acceleration, temperature, humidity, and atmospheric pressure. Its principle block diagram is shown in Figure 3.

In order to ensure the accuracy of triaxial shock and vibration measurement and reduce mutual interference, the measurement motherboard and sensor board are arranged separately. The measurement motherboard completes the measurement of temperature, humidity, atmospheric pressure, battery power, and other parameters, and the measurement motherboard is connected with the outside through the alarm interface and data interface. The sensor board adopts the three-axis integrated design to complete the three-axis shock and vibration measurement. The measurement motherboard and the sensor board are connected through a special interface with a flexible line. The PCB effect diagram of the shock and vibration acceleration recorder is shown in Figure 4.

2.3.4. Program Design

(1) Program Function Design. The shock and vibration acceleration recorder reads the data of the triaxial acceleration sensor in real time. Compared with the set alarm threshold,

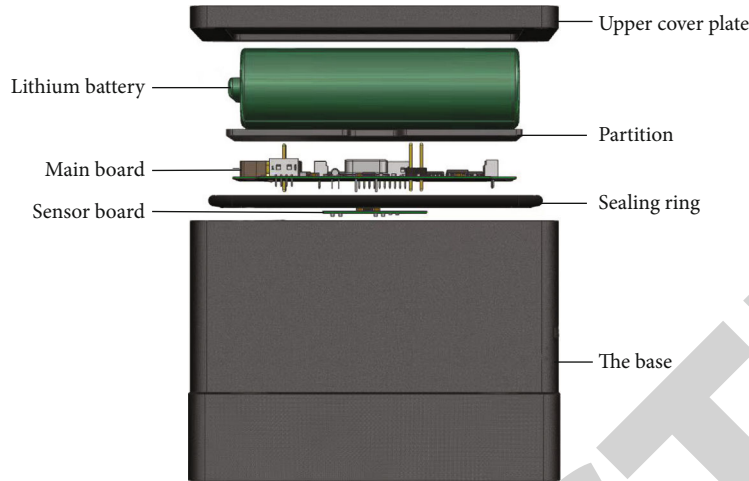


FIGURE 2: Mechanical appearance effect diagram.

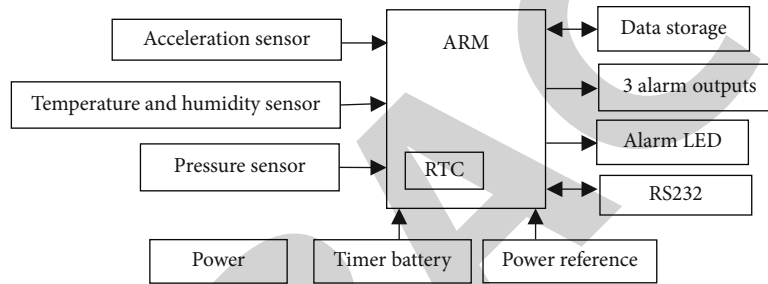


FIGURE 3: The principle block diagram of the shock and vibration acceleration recorder.

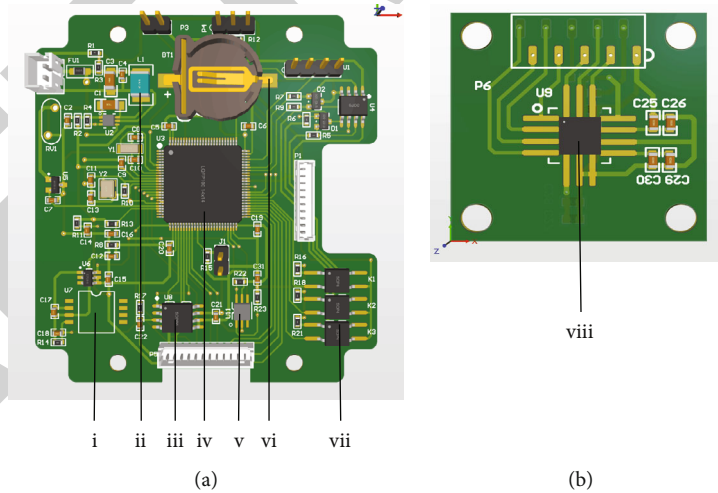


FIGURE 4: The PCB effect diagram of the shock and vibration acceleration recorder. (a) Motherboard with (A) pressure measurement, (B) LDO power, (C) storage, (D) ARM, (E) humidity and temperature measurement, (F) timer battery, and (G) alarm output. (b) Sensor board with (A) acceleration MEMS sensor.

it is greater than the set alarm threshold and drives the alarm circuit to work. At the same time, relevant parameters (current time, triaxial acceleration, temperature, humidity, and pressure) are stored and respond to computer commands and other operations at an appropriate time. The software block diagram is shown in Figure 5.

(2) Filtering method and algorithm.

- (a) The built-in low-pass antialiasing filter of the three-axis acceleration sensor realizes the hardware filtering of the analogue signal to ensure that the output signal eliminates the noise and interference caused

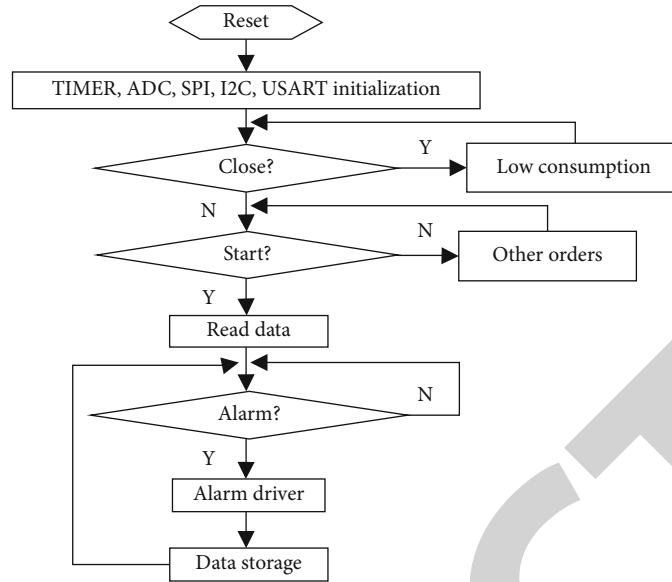


FIGURE 5: Program block diagram of the shock and vibration acceleration recorder.

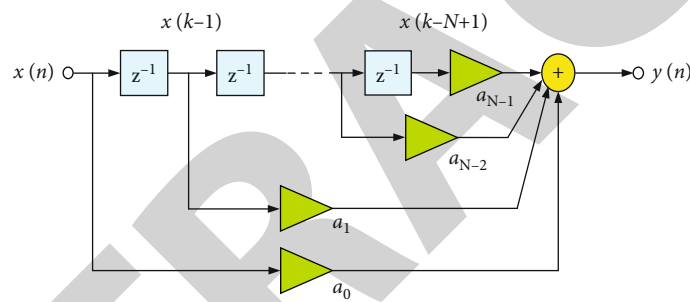


FIGURE 6: Implementation principle of FIR.



FIGURE 7: Function and performance test picture. (a) PC monitoring. (b) Vibration test console. (c) Vibration test platform. (d) Shock and vibration acceleration recorder.

by the sensor itself. After A/D conversion, ARM is used to program the low-pass and high-pass filters of MEMS sensor to realize hardware digital filtering to ensure that the output signal is clean and is a digital signal within the required frequency range

- (b) FIR filtering algorithm further removes unwanted interference signals. The algorithm to realize the filter is usually realized by FIR or IIR. The main feature of FIR filter is that there is no feedback loop and no instability problem. While the amplitude characteristic is set arbitrarily, the accurate linear phase can be guaranteed, and the software and hardware are easy to realize. The filter transition process has a finite interval. Compared with IIR filter, the disadvantage is higher order and larger delay than IIR filter with the same performance. The advantage of IIR filter is that it can achieve better filtering effect at the same order, but the disadvantage of IIR filter design method is that it cannot control the phase characteristics of the filter and the filter may be unstable. Therefore, FIR filter is generally used as adaptive filter. FIR filter is a LTI (linear time invariant) digital

TABLE 2: Repeatability test data of acceleration recorder (z -axis, unit: g).

Value (g)									
0.84	-0.92	0.82	-0.88	0.85	-0.89	0.84	-0.89	0.8	-0.87
0.81	-0.9	0.88	-0.92	0.81	-0.89	0.81	-0.92	0.85	-0.95
0.82	-0.9	0.81	-0.91	0.85	-0.95	0.84	-0.92	0.8	-0.88
0.84	-0.92	0.82	-0.9	0.81	-0.89	0.82	-0.9	0.85	-0.91
0.85	-0.89	0.85	-0.92	0.85	-0.9	0.81	-0.91	0.82	-0.91
0.85	-0.92	0.81	-0.89	0.84	-0.91	0.8	-0.89	0.82	-0.9
0.82	-0.91	0.82	-0.89	0.82	-0.91	0.82	-0.91	0.82	-0.89
0.82	-0.91	0.82	-0.9	0.81	-0.91	0.85	-0.92	0.82	-0.91
0.85	-0.89	0.81	-0.91	0.82	-0.91	0.85	-0.9	0.85	-0.9

filter with constant coefficients. Its implementation principle is shown in Figure 6.

The relationship between FIR output with length N and input time series $X(n)$ is given in the form of finite convolution sum. The specific form is as follows:

$$y(n) = \sum_{n=0}^{N-1} (a(n) \times x(k-n)), \quad (4)$$

$$a(n) = \frac{6}{N(N+1)} \times \left(1 - \frac{2n}{N-1}\right) n = 0, 1, \dots, N-1.$$

The shock and vibration acceleration recorder uses ARM as the control core and is a mobile product, which needs to balance performance and power consumption. Therefore, it is very reasonable to select FIR within the tenth order for filtering to better remove the interference signal

- (c) The piecewise linear coefficient compensation algorithm is used to calibrate the collected and processed signal to ensure that the linearity and error of the output signal in the whole measurement range are within the allowable range of the national standard. After the filtering of the above software and hardware, the signal of the impact vibration acceleration recorder has been relatively clean. In order to ensure the linearity and accuracy of the signal, we designed the linear coefficient compensation method to calibrate it. Use $y = ax + b$ to correct the acceleration values of three axes, respectively, namely,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} d1 \\ d2 \\ d3 \end{pmatrix}. \quad (5)$$

The specific method is as follows: use the data management software to write the coefficients a , b , c , $d1$, $d2$, and $d3$ into the impact vibration acceleration recorder. The recorder uses the above formula to correct and remove the

DC offset to further ensure the linearity and accuracy of acceleration measurement

- (d) The least square method is used to fit the temperature compensation algorithm. The shock and vibration acceleration recorder uses the temperature sensor to detect the working environment temperature of the product. ARM uses the measured working environment temperature and uses the least square method to fit the temperature compensation algorithm according to the temperature curve of the acceleration sensor for correction, which can ensure that the product has good linearity and accuracy in different working temperature environments

(3) *Communication Protocol*. The computer as the host adopts USB interface. As a slave, the shock and vibration acceleration recorder adopts TTL level duplex serial port communication mode. Data communication format: 115200, N , 8,1. Verification method: use cumulative sum to take the lower byte. Cumulative sum data: function code +data length+data. The command format is as follows: start bit (0x00), function code 1 byte, data length 2 bytes, data n byte, and check 1 byte.

3. Experiment Verification and Data Analysis

3.1. *Experiment Methods*. Fix the shock and vibration acceleration recorder and Ultrashock/AC-73 product on the vibration test platform and connect it to PC with communication cable; open and execute the PC monitoring application software, and start the test. The composition of the test system is shown in Figure 7.

3.2. *Repeatability Test of Acceleration Measurement*. The vibration test platform outputs a stable acceleration excitation source, and a total of 90 data are collected at the same point (as shown in Table 2). The range of positive acceleration in z -axis is 0.8~0.85, the mean value is 0.828, and the maximum error of the mean value is 0.028. The range of negative acceleration in z -axis is -0.87~-0.91, the mean value is -0.905, and the maximum error of the mean value is 0.035. It can be seen that the resolution of the device is 0.01 and the repetition accuracy is less than 0.05 g.

3.3. *Verification and Calibration Test of National Metering Station*. The shock and vibration acceleration recorder was sent to China's national metrology station for verification and calibration. Its three-axis of XYZ was tested, 2G, 4G, and 6G were tested, respectively, and each parameter was tested for 10 times. The detection parameters should meet the requirements of amplitude error less than $\pm 5\%$ specified in JJG 676-2019 verification regulation of vibration meter. The detection errors are within the allowable error range of 5%, meet the technical requirements, and pass the verification of national metrology station. According to JJG676-2019 standard test procedure, CEPREI fixes the shock and vibration acceleration recorder on the standard test platform and tests the three-axis of XYZ of the shock and vibration

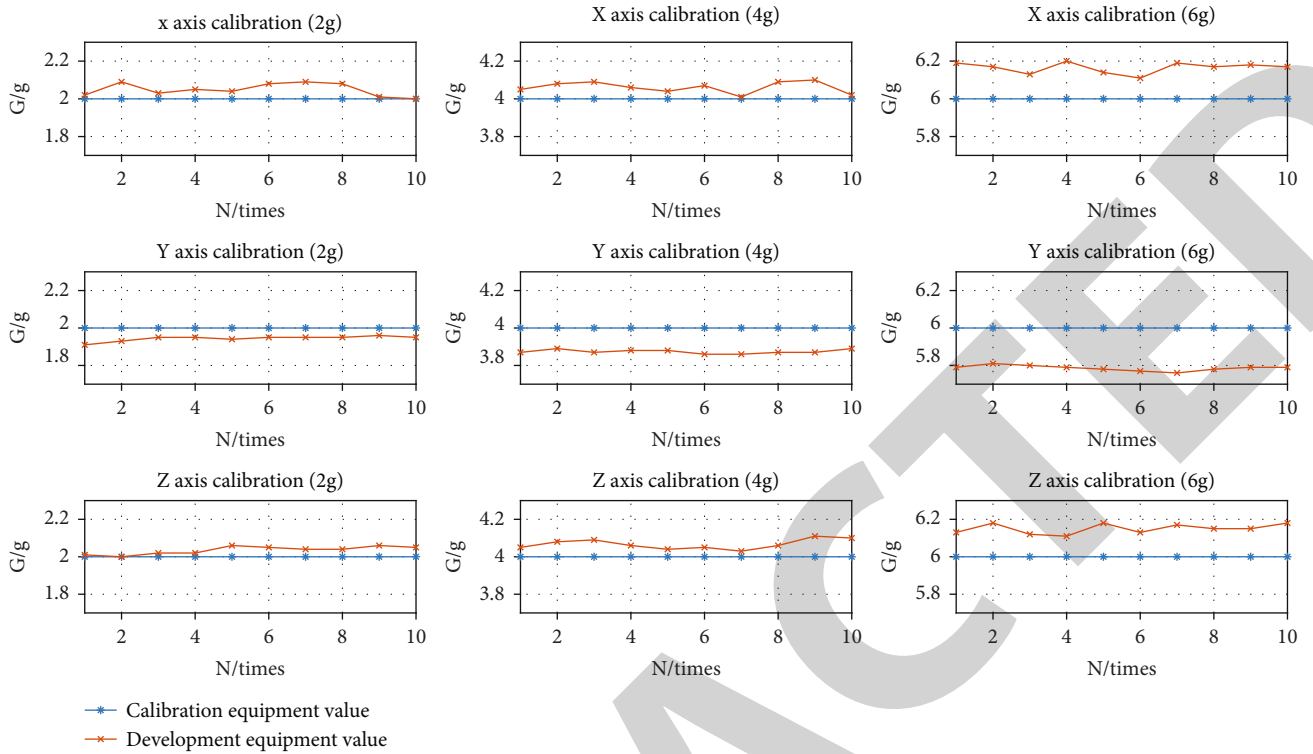


FIGURE 8: Axis calibration.

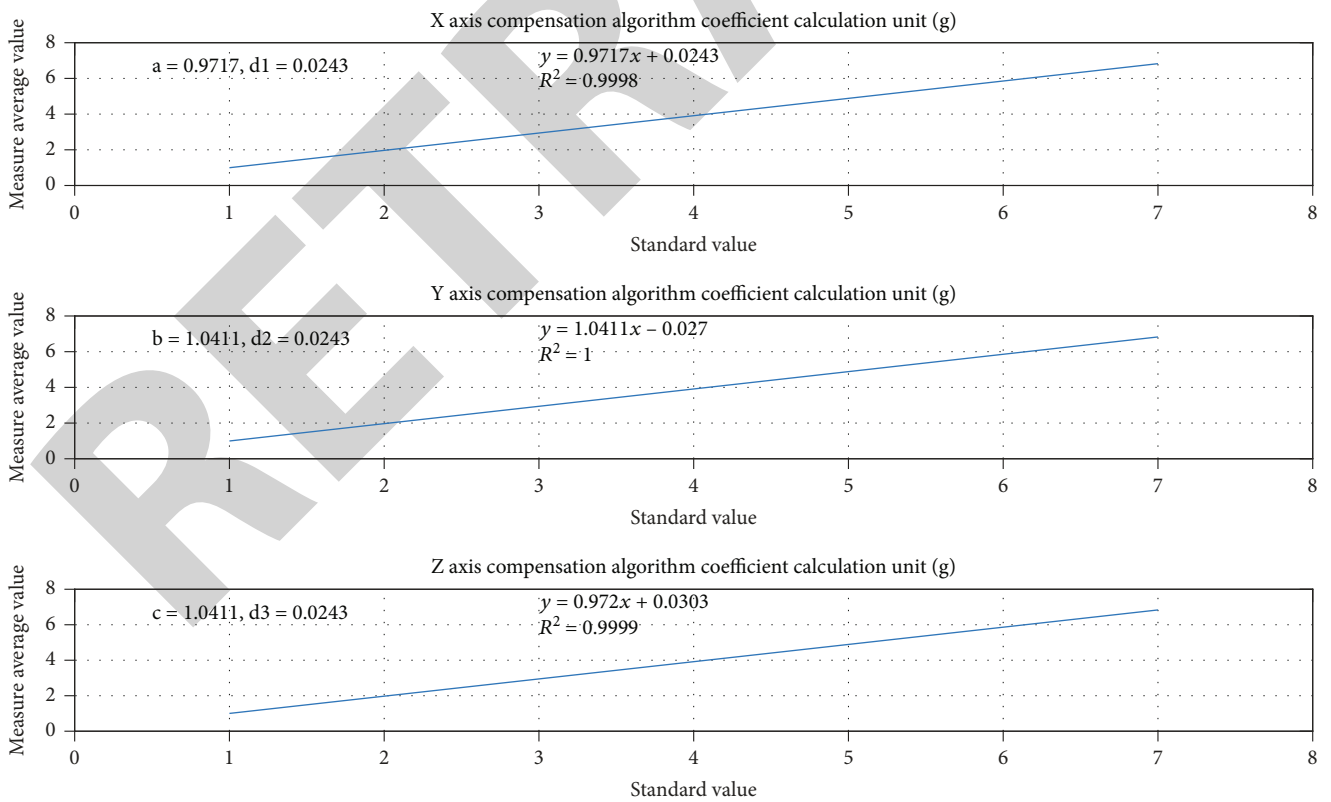


FIGURE 9: Compensation algorithm coefficient calculation of the shock and vibration acceleration recorder (top: x-axis calibration equation, middle: y-axis calibration equation, and bottom: z-axis calibration equation).

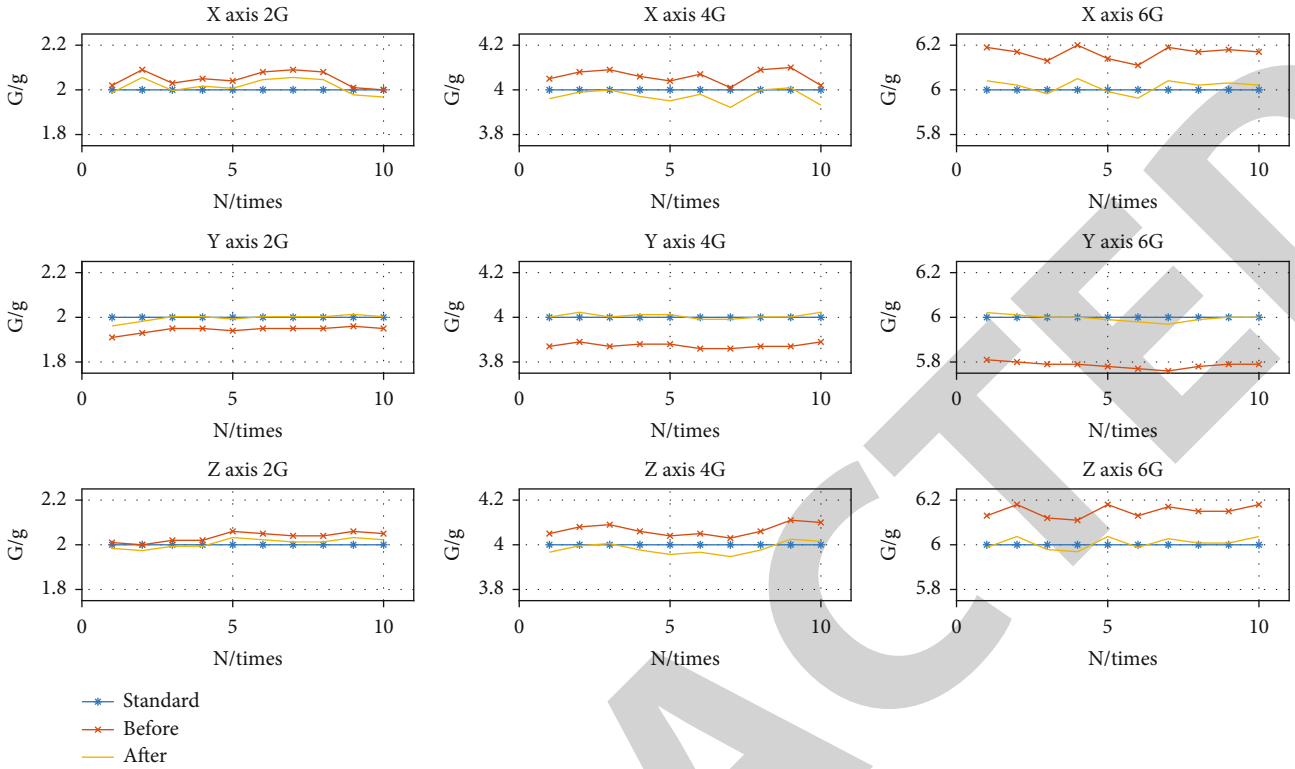


FIGURE 10: Comparison of measured and standard values of triaxial acceleration before and after linear coefficient compensation (blue: standard, red: before, and orange: after).

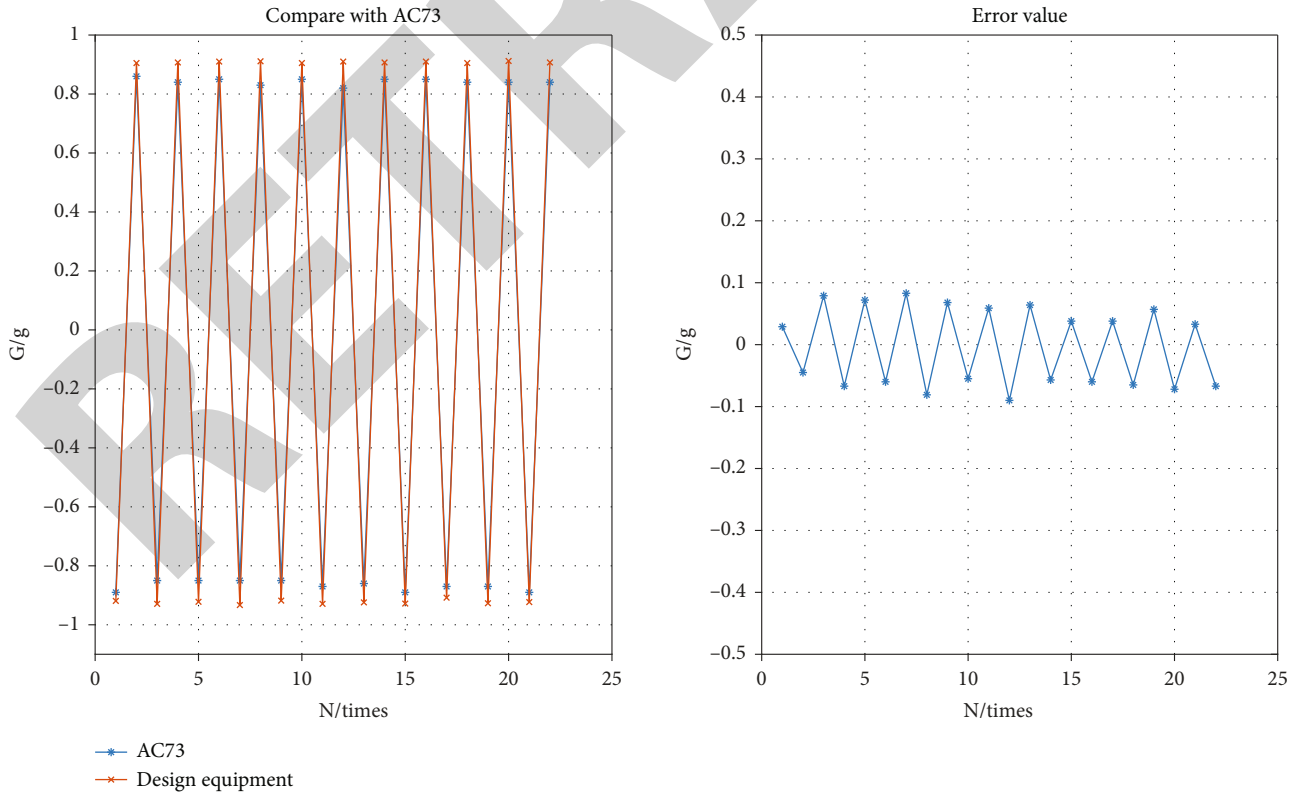


FIGURE 11: Measurement comparison test of the shock and vibration acceleration recorder and AC73 seismic acceleration sensor (vibration platform test data).

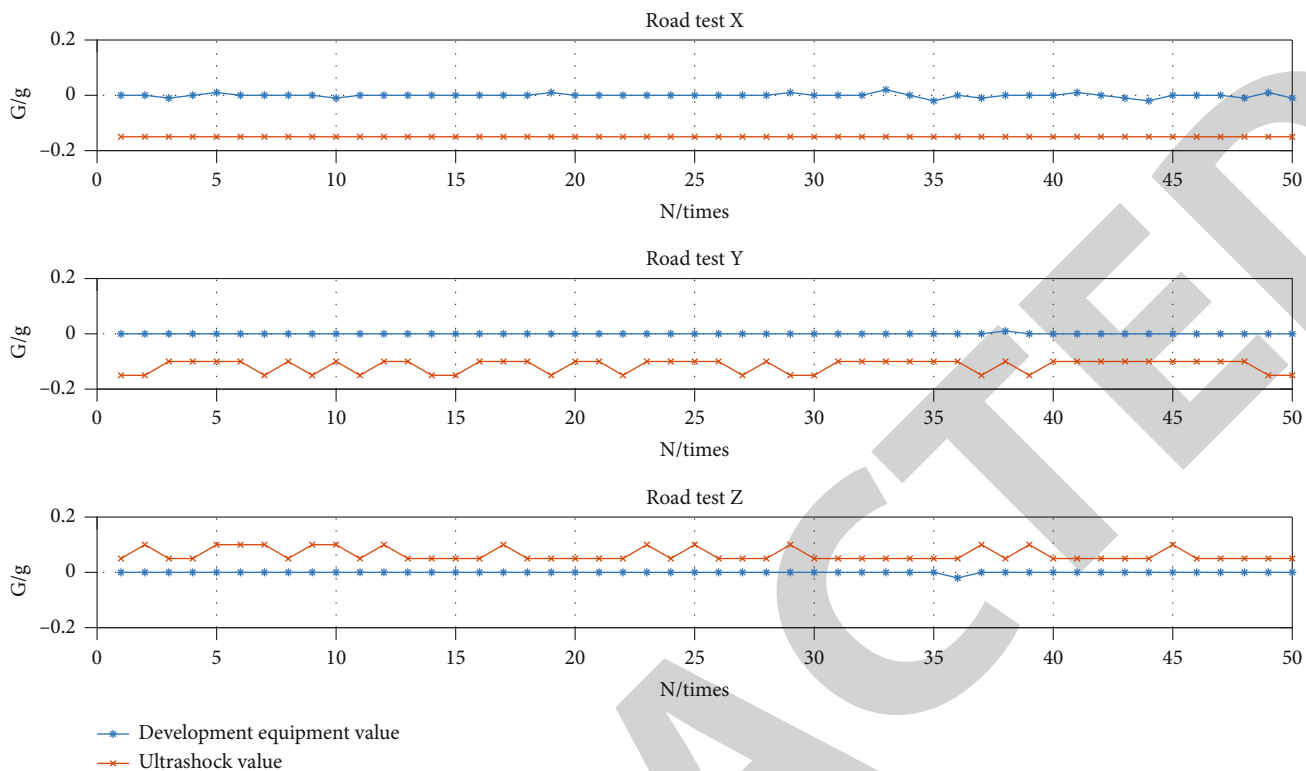


FIGURE 12: Road test comparison between the shock and vibration acceleration recorder and the similar foreign product acceleration measurement (first stage: beginning; second stage: work; and third stage: after transportation).

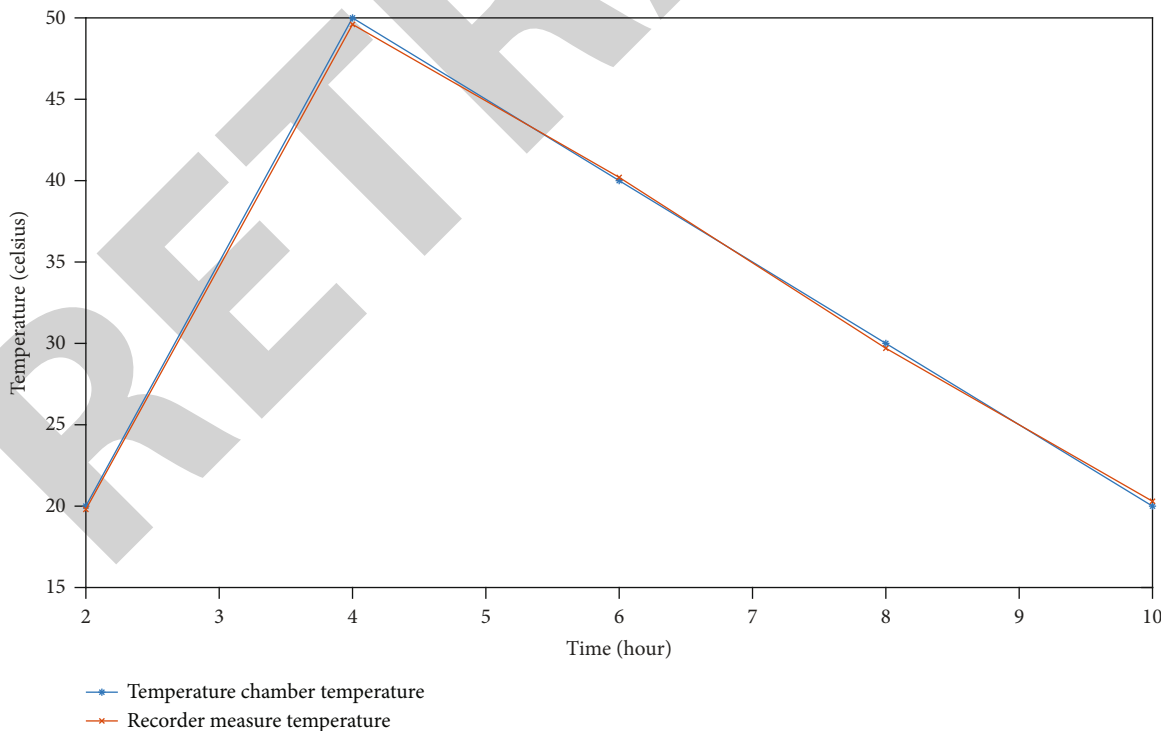


FIGURE 13: Temperature curve of environmental test.

acceleration recorder, 2G, 4G, and 6G were tested, respectively, and each parameter was tested for 10 times. The three-axis detection data is shown in Figure 8.

Taking the third-party calibration data as the standard, the linearity is calculated by Excel table, and the values of coefficient in Equation (5) are obtained, as shown in Figure 9.

Then, the linear coefficient is substituted into Excel table for simulation calculation. The results show that the accuracy of the shock and vibration acceleration recorder has been greatly improved. Except for individual discrete points, 95% numerical accuracy returns from 1%~5% to 0%~1%. Next, it is considered to implant the coefficient into the shock and vibration acceleration recorder through the computer monitoring software and send the product with the implanted coefficient to the third-party authority for recalibration to verify its effect. The comparison of measured acceleration before compensation, calculated acceleration after compensation, and standard value is shown in Figure 10.

3.4. Precision Test with High-Precision Seismic Acceleration Sensor AC73 as Reference. AC73 has higher performance than shock and vibration acceleration recorder and is very suitable for performance experiment verification as acceleration test benchmark. From the continuous vibration test in the z -axis direction, it can be seen that although the test is affected by the vibration platform, the z -axis test data of the two products have a certain deviation, the z -axis measurement values of the two products are relatively close, and the average error of the shock and vibration acceleration recorder relative to AC73 is 6%. The fluctuation error of the shock and vibration acceleration recorder is 4.7%. The fluctuation error of AC73 itself is 1.9%. Considering the different installation positions between the two and the influence of the vibration platform, after deducting the influence of the fluctuation error of AC73 itself, the average error between the two is less than 5% specified in JJG 676-2019, verification regulation of vibration meter. The sampling frequency of shock vibration acceleration recorder (512 Hz) and AC73 (1 kHz) is different, resulting in a certain time shift in the frequency reduction waveform. The actual time shift value is very small and can be ignored. After intercepting and graphing the maximum value, it can be seen that the two frequencies are consistent (see Figure 11).

3.5. Road Test with Ultrashock Logger. The shock and vibration acceleration recorder is placed on the vehicle together with similar foreign product for road test. Start the recorder before transportation and put them on the vehicle after getting on the vehicle (not rigidly fixed). After more than 1 hour of transportation (the transportation process time is 8:30~9:32), get off the vehicle, take out the two recorders, and read the data through their management software, as shown in Figure 12.

It can be seen from the comparison of measurement data of x -, y -, and z -axes:

- (1) Function: both can start the recorder smoothly through the monitoring application software, correctly record and store data during the road test, and correctly read the data again through the monitoring application software after transportation
- (2) Performance: before transportation, the data difference between the two is obvious because they are manually placed on the vehicle. During transportation, because the two recorders are not rigidly fixed on the vehicle (they are placed in different positions), the recorded data of the two recorders are different to some extent. Generally speaking, the data of x -axis and y -axis are close, and there are some differences in z -axis (in most cases, the absolute value of the shock and vibration acceleration recorder is slightly smaller than that of the similar foreign product at the same measuring point), but the recording trend of the two is basically the same (the high and low values are basically the same). After transportation, the two are removed from the vehicle due to manual placement, and the recorded data are obviously different

3.6. Environmental Test. The environment test step is described as follows: first, get the value from the shock and vibration acceleration recorder to make sure the recorder can work in normal. And keep the recorder work in the whole process. Second, put the recorder into the temperature test chamber, and manipulate the temperature test chamber to raise the temperature. Third, when the temperature reaches the highest point, make sure the recorder work in normal. Fourth, decrease the temperature. The environmental test temperature curve is in Figure 13.

After the test, it can be seen that the shock and vibration acceleration recorder can normally measure acceleration, temperature, humidity, and pressure in the whole process.

3.7. Electromagnetic Compatibility Test. According to GB/Z17799.6-2017, GB/T17626.2-2018, and GB/T17626.4-2008, the immunity bench is used for shell immunity test (considering 6kV contact or 8kV air discharge) and signal port immunity test (considering 1kV port contact). It can be seen from the EMC test results that the design of shielding with metal shell plays a good role in electromagnetic shielding and has withstood the test in the immunity test of shell and signal port.

4. Conclusions

For the purpose of engineering reliability and stability, considering the economic practicability, the shock and vibration monitoring system of nuclear fuel transportation is designed by using MEMS sensor acceleration measurement method and system monitoring application software, which can better meet the requirements of nuclear fuel transportation process monitoring.

After the monitoring system is started, the shock and vibration acceleration recorder can automatically complete the functions of measurement, recording, out of limit

Retraction

Retracted: Macro-Micro Study on Mechanical Properties of Frozen Fine Sandstone Based on DEM Mathematical Model

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. 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, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret 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 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. Ma, M. You, S. Peng, B. Zhang, and Y. Lin, "Macro-Micro Study on Mechanical Properties of Frozen Fine Sandstone Based on DEM Mathematical Model," *Journal of Function Spaces*, vol. 2022, Article ID 7176665, 10 pages, 2022.

Research Article

Macro-Micro Study on Mechanical Properties of Frozen Fine Sandstone Based on DEM Mathematical Model

Maoyan Ma ¹, Min You,¹ Shuguang Peng,¹ Biao Zhang,² and Yuan Lin³

¹Anhui Province Key Laboratory of Building Structure and Underground Engineering, Anhui Jianzhu University, Hefei 230601, China

²Jining Planning and Design Institute, Jining 272100, China

³School of Environment and Energy Engineering, Anhui Jianzhu University, Hefei 230601, China

Correspondence should be addressed to Maoyan Ma; maoyanma2019@163.com

Received 4 April 2022; Accepted 27 April 2022; Published 16 May 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Maoyan 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.

The study of freezing rock mechanical properties is getting more and more urgent because of coal mine construction in western China. Particle discrete element method (DEM) can describe discontinuous medium problem mathematically. In order to reveal the mechanical failure mechanism of frozen fine sandstone, the uniaxial compressive strength test of frozen fine sandstone was carried out, and then, DEM was used to simulate the uniaxial test of frozen fine sandstone. Furthermore, the nuclear magnetic resonance (NMR) technology was used to obtain pore distribution of the freezing sandstone. Finally, the results of NMR test and discrete element simulation were combined to reveal the microscopic mechanism of mechanical change in freezing fine sandstone. The DEM results show that the strength of frozen fine sandstone increases with the decrease of temperature. With the decrease of temperatures, strain softening occurs in frozen sandstone, which indicates that the discrete element simulation results are in good agreement with the uniaxial test results. Therefore, DEM can be used to simulate the mechanical behavior of frozen fine sandstone. At the same time, the DEM results also indicate that the formation and development of the shear band are the precursor of the failure of the sample. Furthermore, the NMR test confirms that temperature has a great impact on the pore distribution of sandstone. With the decrease of temperature, the pore ice content increases greatly, which induces a great decrease in NMR porosity and a vast decrease in the proportion of large and medium pores in all pores. Meanwhile, with the growth of the cohesion induced by increasing ice content, the uniaxial compressive strength increases macroscopically.

1. Introduction

With the development of coal mining in western China, the number of mine shafts constructed by freezing method in these areas is increasing. Shafts need to pass through deep bedrocks such as Cretaceous and Jurassic strata, which are characterized by poor rock cementation and easy to be softened in water. Different from the freezing sinking in deep alluvium in the eastern region, the previous design parameters and construction experience cannot scientifically guide the mine construction in the western region [1–5]. Furthermore, the research on the mechanical behavior of laboratory frozen rock needs to be equipped with a temperature control system, and the technical requirements are higher than those of conventional geotechnical tests. Therefore, numerical

simulation is often used as an effective and necessary supplement to the frozen rock test technology. Frozen rock is a discrete particle aggregation composed of multiphase media, and particle discrete element method (DEM) is a method for studying the characteristics of granular materials [6, 7]. Therefore, DEM can be used to simulate the mechanical properties of laboratory frozen rock [8–13].

At present, many researchers have made a series of achievements in the study of mechanical behavior of laboratory rock and soil by using discrete element method. Belheine et al. [14] simulated triaxial tests of sand under drainage conditions by using three-dimensional spherical discrete element model with rolling stiffness. Kozicki et al. [15] studied the influence of initial porosity and particle shape by simulating triaxial tests of sand under drainage

conditions using three-dimensional DEM. De Bono et al. [16] conducted discrete element simulation on the triaxial test of crushed cemented sand. Hu et al. [17] established a damage and fracture model of rock aging deformation based on discrete element method and used this model to conduct numerical simulation research on the creep instability and failure process of rock under different stress levels. Zhou et al. [18] used the two-dimensional discrete element method to numerically simulate the laboratory plane strain test of frozen sand and used the model of bond occurring within a small range of contact points to consider the role of ice in frozen soil. Yin et al. [19] used the parallel bond model to simulate the triaxial test of frozen clay under different temperatures and confining pressures. Yuan et al. [20] used the parallel bond model to describe the interaction between soil particles and between ice and soil particles. Using nuclear magnetic resonance technology (NMR), the pore structure characteristics of rock and soil can be visually displayed, which provides a powerful means for the study of micro mechanism of rock and soil [21, 22].

However, the application of discrete element in frozen rock research still needs to solve many problems, especially the selection of model which is extremely important. In this paper, DEM was used to simulate the laboratory uniaxial test of frozen fine sandstone. The model that the bond occurs in the finite range between the contact particles was used to consider the bonding effect of ice in frozen sandstone. This model can simulate the bonding process of particles after loading and is suitable for describing the constitutive characteristics of interlayer materials in the finite range between particles. In view of the limitations of numerical simulation, NMR was used to obtain the pore distribution inside the rock during freezing. Combined with the results of DEM numerical simulation, the microscopic mechanism of the frozen fine sandstone was revealed.

2. Material and Methods

2.1. Samples. The rock samples were taken from the Jurassic strata in western China and processed into standard samples of $\Phi 50 \text{ mm} \times 100 \text{ mm}$ (see Figure 1) for uniaxial compression tests. Samples were screened by HC-F800 acoustic wave tester to ensure that the test rock samples have no original cracks. At the same time, the sandstone was processed into $\Phi 25 \text{ mm} \times 60 \text{ mm}$ for NMR tests after being saturated.

2.2. Methods

2.2.1. Uniaxial Compression Strength Test. The test was carried out according to standard test procedure for physical and mechanical properties of artificial freezing rock and soil (MT/T593-1996) and geotechnical engineering test method and criterion (GB/T50123-2009). The sample was immersed in a closed water cylinder for pumping treatment, followed by continuous immersion for 24 h. The saturated rock samples were put into the low-temperature tank with preset temperature, and the freezing rate was set to $1^\circ\text{C}/\text{h}$. After freezing for 48 h, the samples were put on rubber sleeves and placed on the MTS compression test machine for uniax-

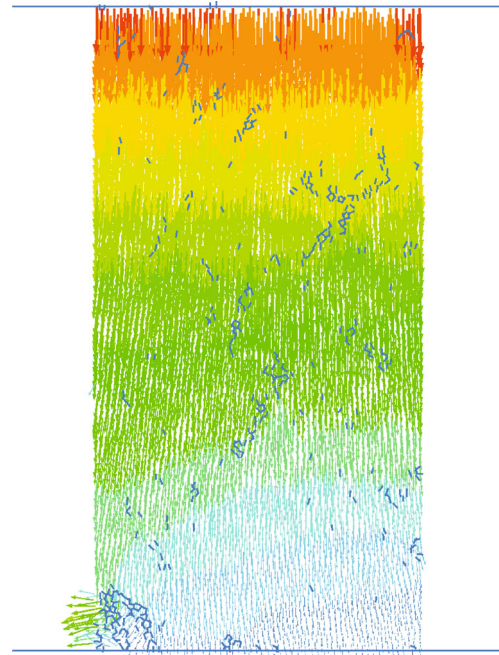


FIGURE 1: The standard sample.

ial compressive strength test. DX-40 low-temperature numerical control test chamber was used in the test. The temperature control range was -40 to 0 degrees Celsius with the temperature control accuracy of $\pm 2^\circ\text{C}$, and the temperature was automatically controlled. In order to meet the actual needs of freezing shaft construction in western China, uniaxial compression test was carried out at 20°C , -2°C , -5°C , -8°C , -11°C , and -15°C .

2.2.2. NMR Test. MesoMR23-060H-I low-temperature nuclear magnetic resonance micro structure analysis and imaging system was used (Figure 2) in this test. The processed fine sandstone samples were maintained at 1 MPa in the vacuum pressure saturation equipment for 24 h. The samples were taken out and put into the nuclear magnetic resonance equipment to measure the porosity of the rock sample at five temperature levels of 20°C , -2°C , -5°C , -8°C , -11°C , and -15°C .

3. Basic Theory of DEM

The discrete element model can discrete the simulation object into discontinuous particles and define the transmission of force or force moment between particles, which can reflect the discontinuity and anisotropy of rock materials. PFC is a discrete element software using particle aggregates to characterize materials. Based on Newton's second law and the relationship between force and displacement, it can simulate the movement of circular particles and their interaction and can also simulate the block structure problem by connecting two or more particles with their directly adjacent particles to form arbitrary shape combinations. The particle contact model in this paper adopts the parallel bond contact model. The parallel bond model can be used



FIGURE 2: The equipment for NMR test.

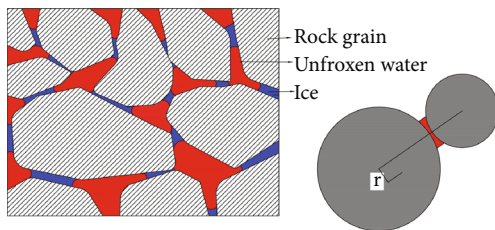


FIGURE 3: Diagram of frozen sandstone structure and parallel bond model.

to characterize the particle materials with large bonding force between particles, and the frozen rock has strong cohesive force due to the action of ice. The bonding structure diagram is shown in Figure 3. The contact model can transfer force and force moment between particles, limit relative sliding and rotation between particles, and give normal stiffness and tangential stiffness between particles. It can generate bond in a certain range of particle contact point to simulate the cementation between particles and provide normal tensile strength and shear strength of particles. The macroscopic mechanical behavior of frozen soil can be reproduced by adjusting the stiffness parameter, bond strength parameter and friction coefficient in the model.

In the parallel bond model, the bond stress follows the relationship between force and displacement. The force-displacement relationship of parallel bond is obtained by the normal and tangential stiffness, tensile and shear strength, bond radius factor, and other parameters. The force and force moment acting on the parallel bond can be expressed as F_i and M_i composing of components in normal and tangential directions and can be expressed as

$$\bar{F}_i = \bar{F}_i^n + \bar{F}_i^s, \quad (1)$$

$$\bar{M}_i = \bar{M}_i^n + \bar{M}_i^s, \quad (2)$$

where \bar{F}_i^n and \bar{M}_i^n are normal force and moment vectors, respectively, and \bar{F}_i^s and \bar{M}_i^s are tangential force and moment vectors, respectively.



FIGURE 4: The generated sample.

Once the parallel bond is formed, F_i and M_i will be initialized to zero. The forces and force moments resulting from subsequent relative displacement increments and relative rotation increments will be superimposed on the current values. The force and moment generated by relative displacement increments and relative rotation increments can be expressed as

$$\begin{cases} \Delta \bar{F}_i^n = -k^n A \Delta u_i^n, \\ \Delta \bar{F}_i^s = -k^s A \Delta u_i^s, \\ \Delta \bar{M}_i = -k^n I \Delta \theta_i. \end{cases} \quad (3)$$

Among them, $\Delta u_i = v_i \Delta t$, $\Delta \theta_i = w_i \Delta t$, $A = 2\bar{R}$, and $I = (2/3)\bar{R}^3$, where $\Delta \bar{F}_i^n$ and $\Delta \bar{F}_i^s$ are the increment of normal and tangential bonding force respectively; $\Delta \bar{M}_i$ is force

TABLE 1: Parameters of sandstone particle flow model.

Temperature (°C)	Contact stiffness ($\times 10^6$ kPa)		Bond strength ($\times 10^6$ kPa)		Coefficient of friction	Pb_emod
	Normal	Tangent	Pb_coh	Pb_ten		
20	10e8	10e8	35e6	25e6	0.50	15e8
-2	11e8	11e8	42e6	31e6	0.43	18e8
-5	12e8	12e8	45e6	33e6	0.41	20e8
-8	14e8	14e8	47e6	35e6	0.39	24e8
-11	16e8	16e8	50e6	37e6	0.35	26e8
-15	17e8	17e8	52e6	39e6	0.31	27e8

moment increment; k^n and k^s are normal and tangential stiffness, respectively; A is bonding area; Δu_i^n and Δu_i^s are displacement increments in normal and tangential directions, respectively; $\Delta \theta_i$ is relative rotation angle of contact particles; I is the inertia moment of the bonding surface on the neutral axis; v_i is the relative velocity of contact particles; w_i is relative angular velocity of contact particles; \bar{R} is contact bond radius; and Δt is time steps. The maximum tensile stress and shear stress acting on the parallel bond are

$$\bar{\sigma}_i^{\max} = \frac{\bar{F}_i^n}{A} + \frac{|\bar{M}_i^n| \bar{R}}{I}, \quad (4)$$

$$\bar{\tau}_i^{\max} = \frac{|\bar{F}_i^s|}{A}. \quad (5)$$

When the maximum tensile stress acting on the bond is larger than the ultimate tensile strength of the bond itself, the bond breaks and produces tensile cracks. Similarly, when the maximum shear stress acting on the bond exceeds the ultimate shear strength of the bond itself, the bond breaks and produces shear cracks.

4. Discrete Element Simulation of Frozen Sandstone

4.1. Simulation Process. Firstly, the wall with the same size as the laboratory test sample was defined to load the sample. The model size was 100 mm in height and 50 mm in diameter, and the particles were filled inside the wall. Then, the suspended particles outside the model were eliminated, and the particles were adjusted to achieve uniform model (Figure 4). The contact type between particles was set as parallel bonding, and the model parameters were assigned. Finally, the model was axially loaded by the relative movement of the upper and lower walls, and axial strain and axial stress were recorded.

In the process of simulation, in order to match the characteristics of laboratory sand test curve, a series of numerical simulation tests are needed. Before the numerical test results are basically consistent with the actual physical model test results, the input parameters of PFC model are adjusted repeatedly [23, 24].

For the parallel bond model used in this simulation, the mesoparameters such as the contact stiffness, bond strength, and friction coefficient of the model are adjusted to ensure

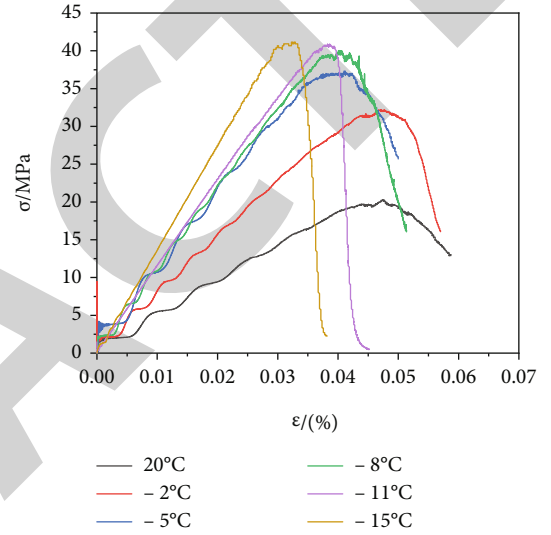


FIGURE 5: DEM calculating results under different soil temperature.

that the simulated stress-strain curves can conform to the stress-strain curves in the experiment. Due to the dependence of rock on temperature, the model parameters obtained at different freezing temperatures are different. Table 1 shows the basic parameters of the model.

4.2. Discrete Element Analysis of Stress-Strain Relationship. Discrete element simulation of frozen rock uniaxial compressive test was carried out at 20°C, -2°C, -5°C, -8°C, -11°C, and -15°C. The results have been shown in Figure 5. It can be seen from Figure 5 that the peak strength of frozen sandstone increases with the decrease of temperature. It increases from 20.3 MPa at 20°C to 41.6 MPa at -15°C. Table 1 shows that the microscopic parameters including the contact stiffness between particles and the bond strength between particles also increase accordingly; that is, the microscopic parameters have a strong temperature dependence, which is significantly different from the sandstone at normal temperature. With the decrease of temperature, water will gradually transform into ice, and the cementation of ice improves the cohesion of the sandstone. At the same time, the micro pores decrease accordingly, so that the strength of frozen sandstone is greatly improved, and with the decrease of temperatures, the failure type of the

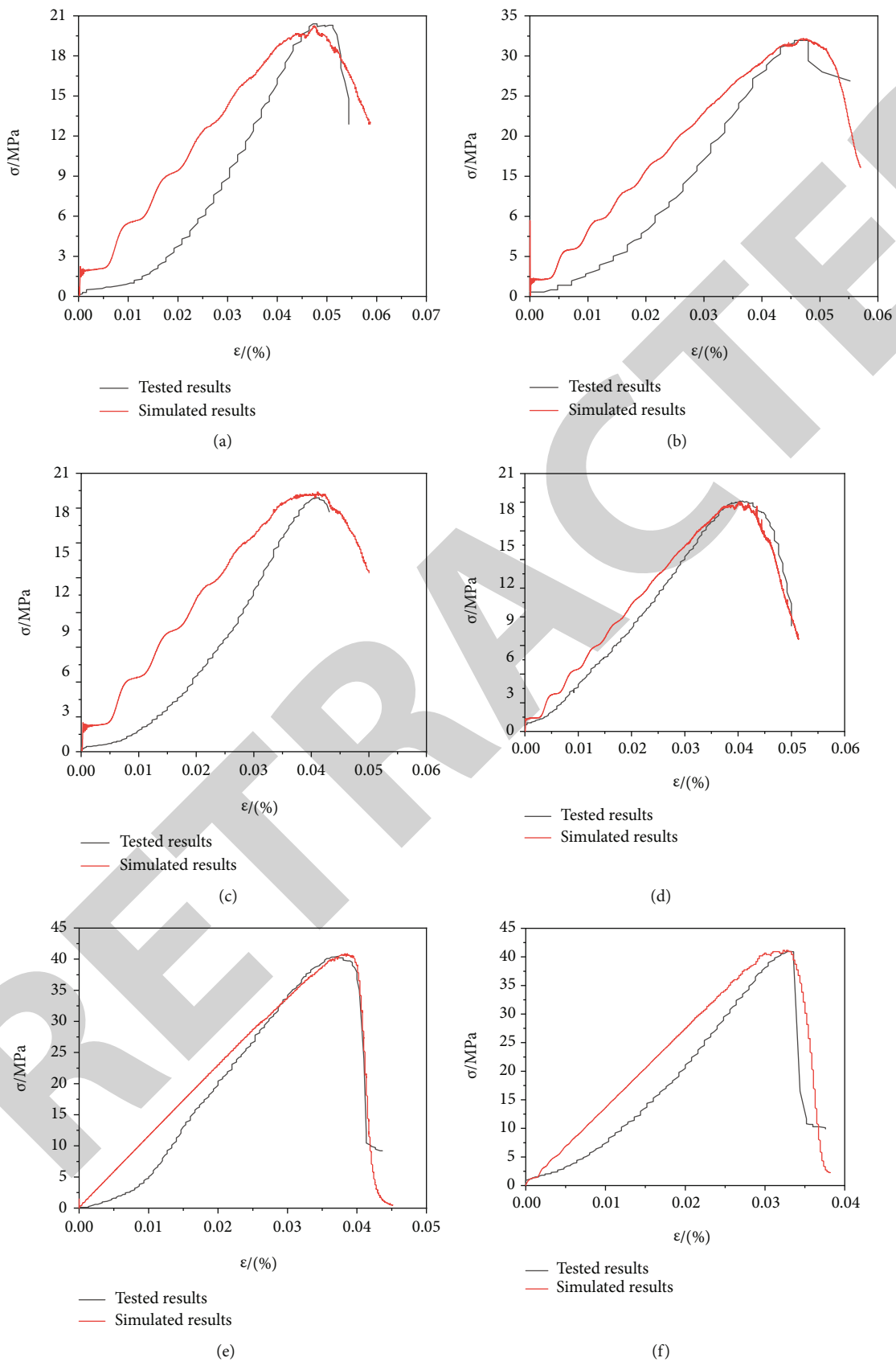


FIGURE 6: Comparison of discrete element numerical test and laboratory test: (a) 20°C; (b) -2°C; (c) -5°C; (d) -8°C; (e) -11°C; (f) -15°C.

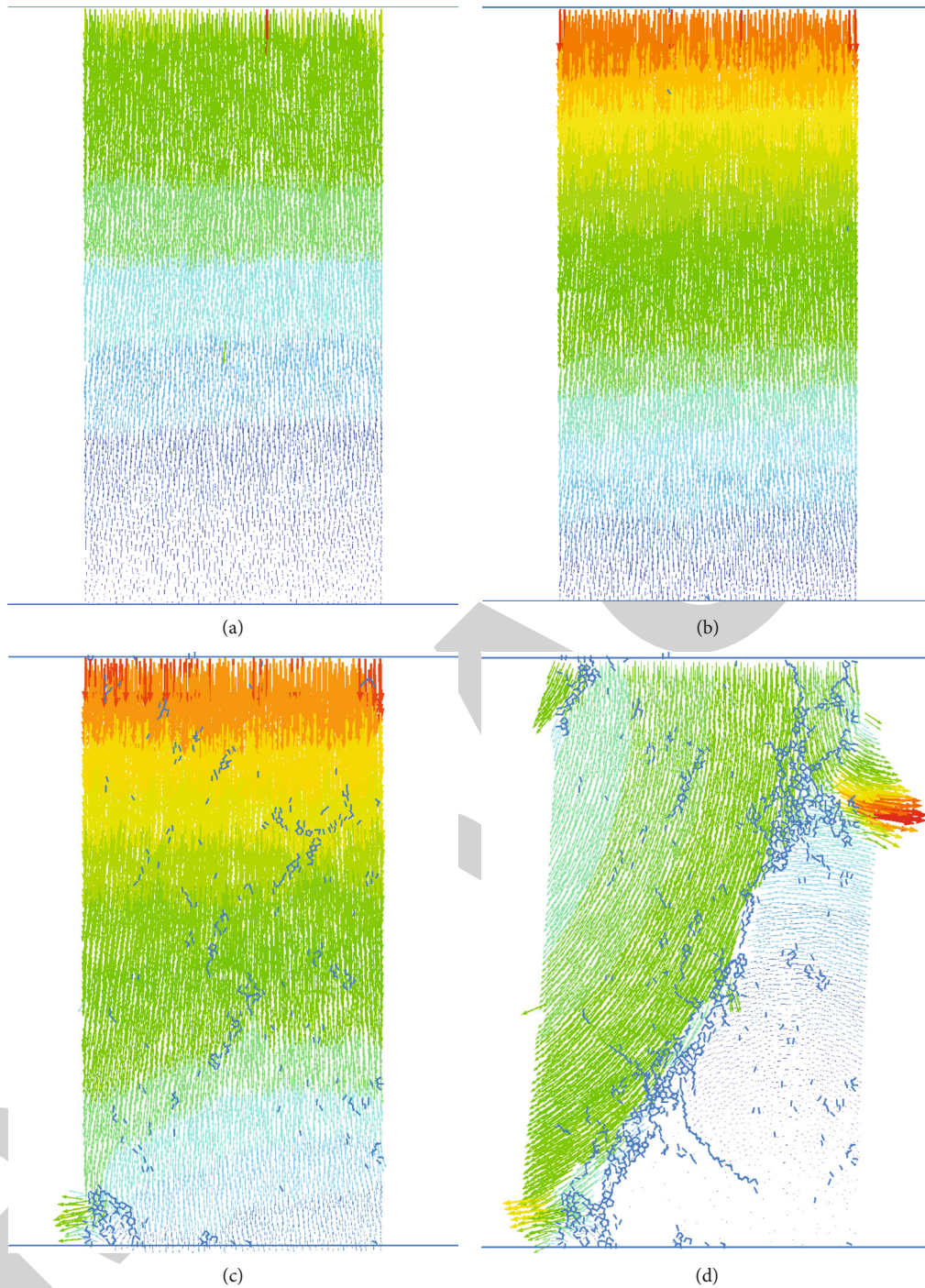


FIGURE 7: Displacement field distribution under different strains: (a) $\varepsilon = 1\%$; (b) $\varepsilon = 2.3\%$; (c) $\varepsilon = 4.8\%$; (d) $\varepsilon = 5.6\%$.

sandstone changes from plasticity at high temperatures to brittleness at lower temperatures. In addition, as shown in Figure 5, with the decrease of temperature, the strain corresponding to the peak strength decreases from 4.8% at -5°C to 3.2% at -15°C . It shows that with the decrease of temperature, strain softening occurs in frozen sandstone.

Figure 6 shows that the simulation results are consistent with the laboratory uniaxial test results, which indicates that the discrete element method can be used to simulate the mechanical behavior of frozen sandstone.

4.3. DEM Simulation of Shear Band Formation and Development. At the temperature of -5°C , the displacement field under different strains is obtained by DEM, which has been shown as Figure 7. The criterion for judging the beginning of shear band formation is that the stress-strain curve is at the peak point [25]. In this study, the strain corresponding to the peak of the stress-strain curve is 4.8%; therefore, 4.8% is taken as the failure strain. The formation and development of the displacement are corresponding to the change trend of the stress-strain curve. The curve at the temperature

TABLE 2: NMR porosity of saturated fine sandstone at different temperatures.

Temperature (°C)	20	-2	-5	-8	-11	-15
NMR porosity (%)	9.1	3.8	1.5	1.2	1.1	1.1

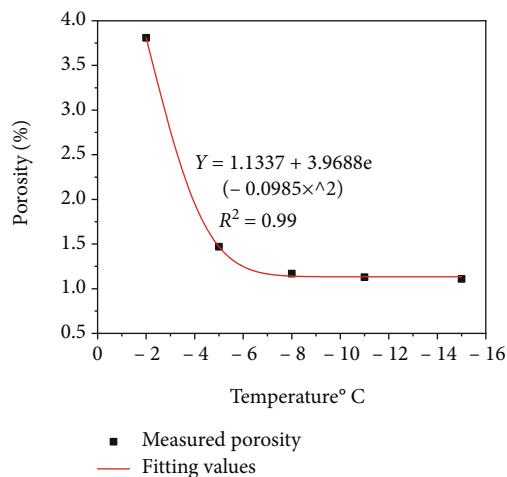


FIGURE 8: Relationship between NMR porosity and temperature.

of -5°C in Figure 5 corresponds to the change process of displacement field in Figure 7. The displacement field of strain 1% and 2.3% in Figure 7 corresponds to the rising curve of -5°C in Figure 5. When the displacement field begins to appear, obvious uneven distribution, damage, or micro cracks begin to appear (see Figure 7(b)). When the strain is 4.8%, corresponding to the peak of the curve, the micro cracks continue to increase, and the shear band begins to form (see Figure 7(c)). The strain value of 5.6% corresponds to the curve after the peak. With the increase of strain, the micro cracks gradually increase and develop into shear bands (see Figure 7(d)). At the same time, the uniform distribution of particle displacement is obvious, and the volume of the sample increases continuously. Finally, a stable shear band from left to right is formed. The corresponding macroscopic deformation occurs after reaching the peak. The displacement field distribution at other temperatures also has a similar corresponding relationship with the macroscopic stress-strain curve, indicating that the formation and development of shear bands are the precursors of specimen failure.

5. Mesopore Structure Changes of Sandstone during Freezing Process

5.1. NMR Porosity Variation of Saturated Fine Sandstone. The freezing porosity data of saturated fine sandstone at 20°C , -2°C , -5°C , -8°C , -11°C , and -15°C are shown in Table 2.

It can be seen from Figure 8 that temperature has a great influence on the porosity of fine sandstone at negative temperatures, and the porosity decreases exponentially when the temperature decreases. According to Table 2, from 20°C to -15°C , the freezing porosity decreases from 9.14% to 1.1%.

When the temperature decreases from 20°C to -8°C , the porosity decreases sharply and tends to be stable after -8°C .

5.2. Pore Throat Distribution. The micro pore throat structure is the most important factor affecting the macroscopic freeze-thaw physical properties of rock, and the pore size distribution plays an important role in the freezing progress of fine sandstone. The percentage of each pore throat distribution group in the total volume of rock under different temperature conditions of saturated sandstone samples has been shown in Table 3.

It can be seen from Table 3 that the pore throat distribution is different at different temperatures. With the decrease of temperature, the large and medium pore size ($>10\ \mu\text{m}$ and $0.1\text{--}10\ \mu\text{m}$) decreases sharply, while the small pores ($<0.1\ \mu\text{m}$) increase. The change trend of pore size slows down after -8°C . At the early freezing stage of the temperature range of 0°C to 5°C , that is, at the rapid freezing stage, the water involved in freezing is mainly free water in large pores, and ice crystals are formed. When the temperature continues to decrease, the water in medium pores freezes, and the ice crystals continue to increase. When the temperature drops below -8°C , the unfrozen water in the large and medium pore throats continues to freeze into ice, while the bound water content in the small pore throat is still high because its freezing point is low, and the freezing rate is slow. Therefore, the proportion of unfrozen water in small pores is high at lower temperature. The content of each pore size group does not always increase or decrease but fluctuates slightly. The reason for this phenomenon is related to the pore throat distribution of the fine sandstone at room temperature and the conversion rate of fine sandstone from different large pore size groups to adjacent small pore size groups from room temperature to -15°C .

5.3. The Relationship between Frozen Sandstone Strength and Micro Structure. Figure 9 shows the relationship between uniaxial compressive strength and small pore size content in sandstone during freezing. It can be seen from the figure that the proportion of small pores increases with the decrease of porosity, indicating that with the decrease of temperature, ice crystals grow mainly in large and medium pores, and the proportion of large and medium pores decreases, while the proportion of small pores increases. Due to the increase of ice content, rock cohesion and frost heaving force increase, which results in an increase in rock deformation resistance. All these lead to the improvement of uniaxial compressive strength of frozen fine sandstone.

Figure 10 shows the relationship between porosity and uniaxial compressive strength during the freezing process of sandstone. It can be seen from the figure that the porosity of sandstone decreases with the decrease of temperature, while the measured and simulated values of uniaxial compressive strength increase. With the continuous freezing of unfrozen water, the pores are occupied by ice crystals, filling the cracks and pores inside the sample, and the porosity decreases. At the same time, the bonding effect of cracks inside the fine sandstone is improved, and the strength of saturated sandstone is greatly increased. In addition, with

TABLE 3: Table of change in proportion of pore size groups at different temperatures (%).

Pore size groups (μm)	Temperature ($^{\circ}\text{C}$)					
	20	-2	-5	-8	-11	-15
<0.1	10.3	85.6	87.4	86.8	85.4	89.1
0.1-10	26.5	5.8	5.6	7.6	8.9	7.0
>10	63.2	8.6	7.1	5.7	5.8	3.9

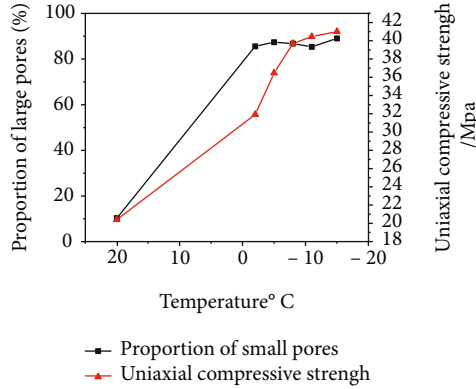


FIGURE 9: Relationship between small pores and uniaxial compressive strength.

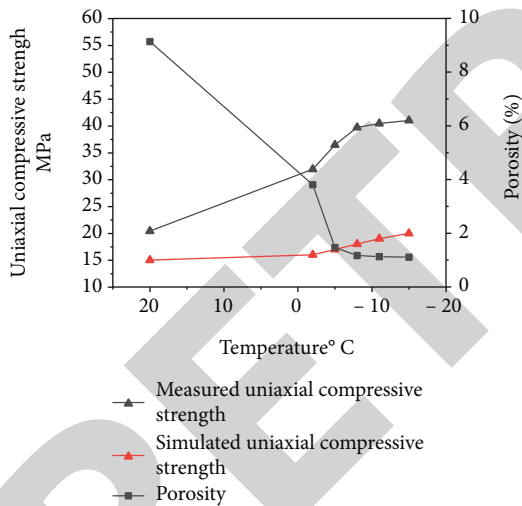


FIGURE 10: Relationship between porosity and uniaxial compressive strength.

the decrease of temperature, the volume of rock particles is compressed, and the arrangement between particles becomes closer, which also makes the strength slightly improved.

When the temperature drops from 20°C to -5°C , the freezing of the sample is mainly the rapid freezing of water in large pores, which shows a sharp decrease in porosity and also a sharp increase in uniaxial compressive strength. However, at low temperatures, there is still water in pores, and in the range of -5°C to -11°C , the unfrozen water in the surrounding micro pores continues to freeze and migrates to the ice crystals in the macro cracks due to the chemical potential. With the temperature decreases continu-

ously, ice in the pores continues to increase, and the measured porosity continues to decrease. Macroscopically, the uniaxial compressive strength continues to increase but the increase rate decreases, because with the continuous decrease of temperature, the amount of unfrozen water in pores is getting less. The content of pore ice in the sample still increases, but the increase is small. At the same time, the changes of porosity and uniaxial compressive strength tend to be stable. Previous studies show that the freezing strength of sandstone has a nonlinear relationship with temperature [26, 27], and the research results of this paper are consistent with them.

6. Conclusions

Based on the uniaxial test of frozen sandstone, the mechanical properties of fine sandstone in the freezing process are simulated by using the DEM where the bonding occurs within a limited range of contact particles. Meanwhile, the NMR technique was used to explore the pore distribution inside the frozen fine sandstone. Combined with the DEM results, the microscopic mechanism of the macroscopic mechanical change of fine sandstone was revealed. The main results are as follows:

- (1) The discrete element simulation results are in good agreement with the laboratory test results, which can be used to simulate the mechanical behavior of frozen fine sandstone. The DEM simulation results show that the strength of frozen sandstone increases with the decrease of temperature. With the decrease of temperature, the failure type of fine sandstone changes from plasticity to brittleness. When the temperature is lower than -5°C , brittle failure happens. With the decrease of temperature, strain softening occurs in frozen sandstone
- (2) DEM simulation shows that with the increase of strain, the micro cracks gradually increase and develop into shear bands. At the same time, the uneven distribution of particle displacement is obvious, and the volume of the sample increases continuously. Finally, a stable shear band from the top left to the bottom right is formed. The corresponding macroscopic deformation occurs after reaching the peak. It shows that the formation and development of shear band is the precursor of failure
- (3) The freezing strength of sandstone has a nonlinear relationship with temperature, and the change of porosity and pore throat distribution can explain the reason. With the decrease of temperature, the large and medium pore size ($>10\mu\text{m}$ and $0.1-10\mu\text{m}$) decreases sharply, and the porosity of sandstone decreases. When the pores and cracks are occupied by ice crystals gradually, the bonding effect of pores and cracks inside the fine sandstone is improved. The cohesion of fine sandstone and frost heaving force increase, which result in an increase in deformation resistance. This is the cause of the

improvement of uniaxial compressive strength of frozen fine sandstone

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 is supported by the Natural Science Foundation of Anhui Province (Grant Nos. 2008085ME165, 2108085ME185) and the Education Department of Anhui Province (KJ2020A0463).

References

- [1] Z. S. Yao, H. Cheng, and C. X. Rong, "Shaft structural design and optimization of deep freezing bedrock shaft in west area," *Journal of China Coal Society*, vol. 35, no. 5, pp. 760–764, 2010.
- [2] B. Wang, "Discussion on problems of mine construction in western Chin," *Mine Construction Technology*, vol. 35, no. 1, pp. 49–52, 2014.
- [3] G. S. Yang, J. M. Xi, H. J. Li, and L. Cheng, "Experimental study of rock mechanical properties under triaxial compressive and frozen conditions," *Chinese Journal of Rock Mechanics and Engineering*, vol. 29, no. 3, pp. 459–464, 2010.
- [4] Q. M. Wang, H. Wang, J. Yang, X. Y. Zhang, and X. L. Dong, "Hydrogeochemical characteristics of main water filled aquifers and source indicators of mine water in typical Jurassic mine area of western China," *Journal of Engineering Geology*, vol. 29, no. 4, pp. 1084–1093, 2021.
- [5] M. Y. Ma, Y. S. Huang, G. Y. Cao, J. Lin, and S. Xu, "Study on mechanical behavior of Jurassic frozen sandstone in western China based on NMR porosity," *Journal of Chemistry*, vol. 2020, 2020.
- [6] P. A. Cundall, *The measurement and analysis of acceleration in rock slopes*, Imperial College of Science and Technology, London, England, 1971.
- [7] P. A. Cundall and O. D. L. Strack, "A discrete numerical model for granular assemblies," *Geotechnique*, vol. 29, no. 1, pp. 47–65, 1979.
- [8] M. J. Jiang, "New paradigm for modern soil mechanics: geomechanics from micro to macro," *Chinese Journal of Geotechnical Engineering*, vol. 41, no. 2, pp. 195–254, 2019.
- [9] F. X. Zhou and Y. M. Lai, "Simulation of mechanical behavior for frozen sand clay by discrete element method," *Rock and Soil Mechanics*, vol. 31, no. 12, pp. 4016–4020, 2010.
- [10] N. Yin, S. Y. Li, W. S. Pei, M. Zhang, and Y. Dong, "Microscopic deformation mechanisms of triaxial test of frozen clay analyzed by discrete element method," *Journal of Glaciology and Geocryology*, vol. 38, no. 1, pp. 178–185, 2016.
- [11] C. Liu, T. C. Le, B. Shi, and Y. Zhu, "Discussion on three major problems of engineering application of the particle discrete element method," *Chinese Journal of Rock Mechanics and Engineering*, vol. 39, no. 6, pp. 1142–1152, 2020.
- [12] B. Zhou, H. B. Wang, W. F. Zhao, J. W. Li, and B. C. Zheng, "Analysis of relationship between particle mesoscopic and macroscopic mechanical parameters of cohesive materials," *Rock and Soil Mechanics*, vol. 33, no. 10, pp. 3171–3178, 2012.
- [13] C. Y. Kwok and M. D. Bolton, "DEM simulations of soil creep due to particle crushing," *Géotechnique*, vol. 63, no. 16, pp. 1365–1376, 2013.
- [14] N. Belheine, J. P. Plassiard, F. V. Donzé, F. Darve, and A. Seridi, "Numerical simulation of drained triaxial test using 3D discrete element modeling," *Computers and Geotechnics*, vol. 36, no. 1-2, pp. 320–331, 2009.
- [15] J. Kozicki, J. Teichman, and H. B. Mühlhaus, "Discrete simulations of a triaxial compression test for sand by DEM," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 38, no. 18, pp. 1923–1952, 2014.
- [16] J. P. De Bono, G. R. McDowell, and D. Wanatowski, "DEM of triaxial tests on crushable cemented sand," *Granular Matter*, vol. 16, no. 4, pp. 563–572, 2014.
- [17] G. H. Hu, T. Xu, C. F. Chen, and X. K. Yang, "A microscopic study of creep and fracturing of brittle rocks based on discrete element method," *Engineering Mechanics*, vol. 35, no. 9, pp. 26–36, 2018.
- [18] J. Zhou, Y. W. Chi, Y. Chi, and J. P. Xu, "Simulation of biaxial test on sand by particle flow code," *Chinese Journal of Geotechnical Engineering*, vol. 22, no. 6, pp. 701–704, 2000.
- [19] M. Yin, S. Y. Li, Y. H. Shi, Z. Z. Sun, and Z. H. Yin, "Discrete element analysis of cemented behavior of frozen loess under different confining pressures," *Journal of Glaciology and Geocryology*, vol. 39, no. 4, pp. 858–867, 2017.
- [20] W. Yuan, X. L. Yao, and W. L. Wang, "Study on triaxial mechanical behaviors of frozen sand based on discrete element method," *Journal of Glaciology and Geocryology*, vol. 41, no. 6, pp. 1388–1396, 2019.
- [21] T. Wang, Q. Sun, H. Jia, J. Ren, and T. Luo, "Linking the mechanical properties of frozen sandstone to phase composition of pore water measured by LF-NMR at subzero temperatures," *Bulletin of Engineering Geology and the Environment*, vol. 80, no. 6, pp. 4501–4513, 2021.
- [22] H. L. Jia, F. Zi, G. S. Yang et al., "Influence of pore water (ice) content on the strength and deformability of frozen argillaceous siltstone," *Rock Mechanics and Rock Engineering*, vol. 53, no. 2, pp. 967–974, 2020.
- [23] J. F. Lu, C. W. Zhang, and P. Jian, "Meso-structure parameters of discrete element method of sand pebble surrounding rock particles in different dense degrees," in *Proceedings of the 7th International Conference on Discrete Element Methods*, X. Li, Y. Feng, and G. Mustoe, Eds., Springer, Singapore, 2016.
- [24] T. V. Tieng, N. T. T. Hang, and N. X. Khanh, "Compressive behavior of concrete: experimental study and numerical simulation using discrete element method," in *Computational Intelligence Methods for Green Technology and Sustainable Development. GTSD 2020. Advances in Intelligent Systems and Computing*, Y. P. Huang, W. J. Wang, H. A. Quoc, L. H. Giang, and N. L. Hung, Eds., Springer, Cham, 2021.
- [25] H. R. Xiong, K. L. Yuan, M. J. Wen, A. N. Yu, and J. M. Xu, "Influence of pore structure on the moisture transport property of external thermal insulation composite system as studied by NMR," *Construction and Building Materials*, vol. 228, no. 20, pp. 1–9, 2019.