Wireless Communications and Mobile Computing

Applications of Semantic Mobile Computing in the Internet of Things and Web of Things

Lead Guest Editor: Chi-Hua Chen
Guest Editors: Lingjuan Lyu, Ting Bi, and Feng-Jang Hwang

Applications of Semantic Mobile Computing in the Internet of Things and Web of Things

Wireless Communications and Mobile Computing

Applications of Semantic Mobile Computing in the Internet of Things and Web of Things

Lead Guest Editor: Chi-Hua Chen
Guest Editors: Lingjuan Lyu, Ting Bi, and Feng-Jang Hwang

Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in “Wireless Communications and Mobile Computing.” 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

Zhipeng Cai, USA

Associate Editors

Ke Guan, China
Jaime Lloret, Spain
Maode Ma, Singapore

Academic Editors

Muhammad Inam Abbasi, Malaysia
Ghufran Ahmed, Pakistan
Hamza Mohammed Ridha Al-Khafaji, Iraq
Abdullah Alamoodi, Malaysia
Marica Amadeo, Italy
Sandhya Aneja, USA
Mohd Dilshad Ansari, India
Eva Antonino-Daviu, Spain
Mehmet Emin Aydin, United Kingdom
Parameshachari B. D., India
Kalapraveen Bagadi, India
Ashish Bagwari, India
Dr. Abdul Basit, Pakistan
Alessandro Bazzi, Italy
Zdenek Becvar, Czech Republic
Nabil Benamar, Morocco
Olivier Berder, France
Petros S. Bithas, Greece
Dario Bruneo, Italy
Jun Cai, Canada
Xuesong Cai, Denmark
Gerardo Canfora, Italy
Rolando Carrasco, United Kingdom
Vicente Casares-Giner, Spain
Brijesh Chaurasia, India
Lin Chen, France
Xianfu Chen, Finland
Hui Cheng, United Kingdom
Hsin-Hung Cho, Taiwan
Ernestina Cianca, Italy
Marta Cimitile, Italy
Riccardo Colella, Italy
Mario Collotta, Italy
Massimo Condoluci, Sweden
Antonino Crivello, Italy
Antonio De Domenico, France
Floriano De Rango, Italy
Antonio De la Oliva, Spain
Margot Deruyck, Belgium
Liang Dong, USA
Praveen Kumar Donta, Austria
Zhuojun Duan, USA
Mohammed El-Hajjar, United Kingdom
Oscar Esparza, Spain
Maria Fazio, Italy
Mauro Femminella, Italy
Manuel Fernandez-Veiga, Spain
Gianluigi Ferrari, Italy
Luca Foschini, Italy
Alexandros G. Fragkiadakis, Greece
Ivan Ganchev, Bulgaria
Óscar García, Spain
Manuel García Sánchez, Spain
L. J. García Villalba, Spain
Miguel Garcia-Pineda, Spain
Piedad Garrido, Spain
Michele Girolami, Italy
Mariusz Glabowski, Poland
Carles Gomez, Spain
Antonio Guerrieri, Italy
Barbara Guidi, Italy
Rami Hamdi, Qatar
Tao Han, USA
Sherief Hashima, Egypt
Mahmoud Hassaballah, Egypt
Yejun He, China
Yixin He, China
Andrej Hrovat, Slovenia
Chunqiang Hu, China
Xuexian Hu, China
Zhenghua Huang, China
Xiaohong Jiang, Japan
Vicente Julian, Spain
Rajesh Kaluri, India
Dimitrios Katsaros, Greece
Muhammad Asghar Khan, Pakistan
Rahim Khan, Pakistan
Ahmed Khattab, Egypt
Hasan Ali Khattak, Pakistan
Mario Kolberg, United Kingdom
Meet Kumari, India
Wen-Cheng Lai, Taiwan

Jose M. Lanza-Gutierrez, Spain
Pavlos I. Lazaridis, United Kingdom
Kim-Hung Le, Vietnam
Tuan Anh Le, United Kingdom
Xianfu Lei, China
Jianfeng Li, China
Xiangxue Li, China
Yaguang Lin, China
Zhi Lin, China
Liu Liu, China
Mingqian Liu, China
Zhi Liu, Japan
Miguel López-Benítez, United Kingdom
Chuanwen Luo, China
Lu Lv, China
Basem M. ElHalawany, Egypt
Imadeldin Mahgoub, USA
Rajesh Manoharan, India
Davide Mattera, Italy
Michael McGuire, Canada
Weizhi Meng, Denmark
Klaus Moessner, United Kingdom
Simone Morosi, Italy
Amrit Mukherjee, Czech Republic
Shahid Mumtaz, Portugal
Giovanni Nardini, Italy
Tuan M. Nguyen, Vietnam
Petros Nicopolitidis, Greece
Rajendran Parthiban, Malaysia
Giovanni Pau, Italy
Matteo Petracca, Italy
Marco Picone, Italy
Daniele Pinchera, Italy
Giuseppe Piro, Italy
Javier Prieto, Spain
Umair Rafique, Finland
Maheswar Rajagopal, India
Sujan Rajbhandari, United Kingdom
Rajib Rana, Australia
Luca Reggiani, Italy
Daniel G. Reina, Spain
Bo Rong, Canada
Mangal Sain, Republic of Korea
Praneet Saurabh, India
Hans Schotten, Germany
Patrick Seeling, USA
Muhammad Shafiq, China
Zaffar Ahmed Shaikh, Pakistan
Vishal Sharma, United Kingdom
Kaize Shi, Australia
Chakchai So-In, Thailand
Enrique Stevens-Navarro, Mexico
Sangeetha Subbaraj, India
Tien-Wen Sung, Taiwan
Suhua Tang, Japan
Pan Tang, China
Pierre-Martin Tardif, Canada
Sreenath Reddy Thummaluru, India
Tran Trung Duy, Vietnam
Fan-Hsun Tseng, Taiwan
S Velliangiri, India
Quoc-Tuan Vien, United Kingdom
Enrico M. Vitucci, Italy
Shaohua Wan, China
Dawei Wang, China
Huaqun Wang, China
Pengfei Wang, China
Dapeng Wu, China
Huaming Wu, China
Ding Xu, China
YAN YAO, China
Jie Yang, USA
Long Yang, China
Qiang Ye, Canada
Changyan Yi, China
Ya-Ju Yu, Taiwan
Marat V. Yuldashev, Finland
Sherali Zeadally, USA
Hong-Hai Zhang, USA
Jiliang Zhang, China
Lei Zhang, Spain
Wence Zhang, China
Yushu Zhang, China
Kechen Zheng, China
Fuhui Zhou, USA
Meiling Zhu, United Kingdom
Zhengyu Zhu, China

Contents

Retracted: Application Analysis of New Internet Multimedia Technology in Optimizing the Ideological and Political Education System of College Students
Wireless Communications and Mobile Computing
Retraction (1 page), Article ID 9849843, Volume 2023 (2023)

Predicting Online Consumer Transaction from Big Data: Influential Factors and Strategic Planning
Chiang-Yu Cheng, Ming-Ying Lu, and Han-Ping Tsen
Research Article (9 pages), Article ID 8834713, Volume 2021 (2021)

Research on Business English Translation Architecture Based on Artificial Intelligence Speech Recognition and Edge Computing
Yunwei Xu
Research Article (10 pages), Article ID 5518868, Volume 2021 (2021)

A Blockchain-Based Vehicle Condition Recording System for Second-Hand Vehicle Market
You-Ting Jiang and Hung-Min Sun
Research Article (10 pages), Article ID 6623251, Volume 2021 (2021)

Medical Knowledge Graph in Chinese Using Deep Semantic Mobile Computation Based on IoT and WoT
Wanheng Liu, Ling Yin, Cong Wang, Fulin Liu, and Zhiyu Ni
Research Article (13 pages), Article ID 5590754, Volume 2021 (2021)

A Security Wireless Monitoring and Automatic Protection System for CCEL
Qiao-yan Sun, Xue-jiao Liu, Yu-mei Sun, Mei-chun Wang, Xiang Han, and Xiang-guang Chen
Research Article (14 pages), Article ID 6652246, Volume 2021 (2021)

A Novel Subcarrier-Level Spectrum Sensing Method by Utilizing Fine-Grained Channel State Information in Wireless Networks
Zheng Lu, Handong Wang, Hongyu Sun, Chin-Ling Chen, and Zhenjiang Tan
Research Article (12 pages), Article ID 6639894, Volume 2021 (2021)

Construct a Teaching System Combining Image Linguistics and Multimedia Technology
Yan Yuejue, Sun Xinze, Li Bingyue, and Wang Xiaoyuan
Research Article (11 pages), Article ID 6699010, Volume 2021 (2021)

Quasi-Affine Transformation Evolutionary with Double Excellent Guidance
Tien-Wen Sung, Baohua Zhao, and Xin Zhang
Research Article (15 pages), Article ID 5591543, Volume 2021 (2021)

Computational Resource Allocation Strategy in a Public Blockchain Supported by Edge Computing
Sujie Shao, Weichao Gong, Shaoyong Guo, and Xuesong Qiu
Research Article (8 pages), Article ID 8851121, Volume 2021 (2021)

Appling an Improved Method Based on ARIMA Model to Predict the Short-Term Electricity Consumption Transmitted by the Internet of Things (IoT)
Ni Guo, Wei Chen, Manli Wang, Zijian Tian, and Haoyue Jin
Research Article (11 pages), Article ID 6610273, Volume 2021 (2021)

A New Method for Reconstructing Data on a Single Failure Node in the Distributed Storage System Based on the MSR Code
Miao Ye, Ruoyu Wei, Wei Guo, Qiuxiang Jiang, Hongbing Qiu, and Yong Wang
Research Article (14 pages), Article ID 5574255, Volume 2021 (2021)

A Study on the Influencing Factors of Mobile Digital Music Selection in Subhealthy People Based on Perceived Value Theory
Ming Jiang, Liming Lin, Weiguo Zhang, Weiqing Zhuang, Minghui Jin, and Yi Lin
Research Article (10 pages), Article ID 5547044, Volume 2021 (2021)

Research on Music Multiterminal Audio Authentication Based on Wireless Network
Dongmei Li
Research Article (10 pages), Article ID 5520009, Volume 2021 (2021)

TSN: Performance Creative Choreography Based on Twin Sensor Network
Yufeng Wu, Longfei Zhang, Gangyi Ding, Dapeng Yan, and Fuquan Zhang
Research Article (12 pages), Article ID 5532754, Volume 2021 (2021)

Automatic Classification Algorithm for Multisearch Data Association Rules in Wireless Networks
Cailing Li and Wenjun Li
Research Article (9 pages), Article ID 5591387, Volume 2021 (2021)

Multiobjective Sensor Ontology Matching Technique with User Preference Metrics
Hai Zhu, Xingsi Xue, Chengcai Jiang, and He Ren
Research Article (9 pages), Article ID 5594553, Volume 2021 (2021)

The Design of University Physical Education Management Framework Based on Edge Computing and Data Analysis
Wenze Sun and Yuchi Gao
Research Article (8 pages), Article ID 5537471, Volume 2021 (2021)

Research on Data Optimization of CT System for Patients with Severe Pneumonia Based on Filtered Back Projection and Internet of Things
Shaojie Qin and Wentao Wang
Research Article (7 pages), Article ID 5560652, Volume 2021 (2021)

[Retracted] Application Analysis of New Internet Multimedia Technology in Optimizing the Ideological and Political Education System of College Students
Haibin Yu
Research Article (12 pages), Article ID 5557343, Volume 2021 (2021)

Contents

Analysis of e-Commerce Characteristics Based on Edge Algorithm and COX Model
Jinzhong Zuo
Research Article (13 pages), Article ID 6628068, Volume 2021 (2021)

A Multilevel Single Stage Network for Face Detection
Kanghua Hui, Jin Wang, Huaiqing He, and W. H. Ip
Research Article (10 pages), Article ID 5582132, Volume 2021 (2021)

Depth and Width Changeable Network-Based Deep Kernel Learning-Based Hyperspectral Sensor Data Analysis
Jing Liu, Tingting Wang, and Yulong Qiao
Research Article (8 pages), Article ID 8842396, Volume 2021 (2021)

An E-Commerce Economic Dynamic Data Evaluation Model Based on Multiuser Demand Constraints
Chenyuan Wang
Research Article (9 pages), Article ID 6645961, Volume 2021 (2021)

Research on Supply Chain Financial Risk Assessment Based on Blockchain and Fuzzy Neural Networks
Yijie Wang
Research Article (8 pages), Article ID 5565980, Volume 2021 (2021)

Energy Efficiency Optimization of Cognitive UAV-Assisted Edge Communication for Semantic Internet of Things
Yilong Gu, Yangchao Huang, Hang Hu, Weiting Gao, and Yu Pan
Research Article (12 pages), Article ID 6682340, Volume 2021 (2021)

Design of English Diagnostic Practice Sentence Repetition Recognition System Based on Matching Tree and Edge Computing
Tanping Xi
Research Article (8 pages), Article ID 6651145, Volume 2021 (2021)

Integrating Sensor Ontologies with Global and Local Alignment Extractions
Xingsi Xue, Xiaojing Wu, Chao Jiang, Guojun Mao, and Hai Zhu
Research Article (10 pages), Article ID 6625184, Volume 2021 (2021)

Face Image Publication Based on Differential Privacy
Chao Liu, Jing Yang, Weinan Zhao, Yining Zhang, Jingyou Li, and Chunmiao Mu
Research Article (20 pages), Article ID 6680701, Volume 2021 (2021)

Analysis and Optimization of Asymmetric Wireless Power Transfer in Concrete
Zijian Tian, Dairong Liu, Wei Chen, Wenqing Wang, Shan Pan, and Ni Guo
Research Article (12 pages), Article ID 8841210, Volume 2020 (2020)

Sensor Ontology Metamatching with Heterogeneity Measures
Xingsi Xue, Jiawei Lu, Chengcai Jiang, and Yikun Huang
Research Article (10 pages), Article ID 6666228, Volume 2020 (2020)

Blockchain-Based Internet of Vehicles Privacy Protection System
Tianhong Su, Sujie Shao, Shaoyong Guo, and Min Lei
Research Article (10 pages), Article ID 8870438, Volume 2020 (2020)

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9849843, 1 page
https://doi.org/10.1155/2023/9849843

Retraction

Retracted: Application Analysis of New Internet Multimedia Technology in Optimizing the Ideological and Political Education System of College Students

Wireless Communications and Mobile Computing

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

Copyright © 2023 Wireless Communications and Mobile Computing. This is an open access article 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, "Application Analysis of New Internet Multimedia Technology in Optimizing the Ideological and Political Education System of College Students," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5557343, 12 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 8834713, 9 pages
https://doi.org/10.1155/2021/8834713

Research Article

Predicting Online Consumer Transaction from Big Data: Influential Factors and Strategic Planning

Chiang-Yu Cheng[1], **Ming-Ying Lu**[1] **and Han-Ping Tsen**[2]

[1]*School of big Data Management, Soochow University, Taiwan*
[2]*Institute of Information Management, National Yang Ming Chiao Tung University, Hsinchu, Taiwan*

Correspondence should be addressed to Han-Ping Tsen; hanping311111@gmail.com

Received 30 September 2020; Accepted 24 August 2021; Published 25 September 2021

Academic Editor: Feng-Jang Hwang

Copyright © 2021 Chiang-Yu Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Online transaction has recently benefited from coronavirus; however, the sales of e-commerce in some areas are substantially on the decline. The current study proposes a theoretically constructed and empirically viable way for predicting the relevant factors that may detract or foster e-commerce success. We apply web analytics (one of the big data techniques) to simultaneously, generalizably, and objectively measure the influential factors of e-commerce success. The findings indicate that (1) pageviews is an important key for consumers to make transactions. (2) Bounce rate of the website should not be a member factor of e-commerce success. (3) Adhesion strategy and repeatability strategy can be used to induce consumer online transaction. Several theoretical contributions and practical implications are also provided.

1. Introduction

The new type of coronavirus has invaded the world and has reduced people's willingness to go outside; however, this harmful impact fits quite well with the characteristic of e-commerce, because people do not need to actually patronize the storefronts but can still make their purchases at home. Even though online transaction may be benefited from the outbreak of such virus pandemic, the sales of e-commerce in some areas are substantially on the decline. According to the survey reported by eMarketer [1] that global retail e-commerce sales decelerated to a 16.5% growth rate in 2020 (down from 20.2% in 2019), this tendency is particularly obvious in North America and Asia-Pacific. Due to the fact that e-commerce sales in the US and China are accounted for the majority of overall global sales metrics, it is necessary to figure out the factors that can be used to boost online transactions for both countries. Previous studies have adopted some ways to explore the influential factors of e-commerce success. For example, Wang et al. [2] utilized structural equation modeling to validate the proposed e-commerce success model. Kshetri [3] applied case study to clarify the barriers of e-commerce in the developing country. Liu et al. [4] conducted an expert interview to measure the performance of e-commerce. Although these methods are applicable to understand the influential factors of e-commerce success, there exist several disadvantages that should be further considered. *First*, in most cases, structural equation modeling is subject to validate the relationship between independent and dependent variables [5], but it is rarely used to verify the relationship between independent variables. In fact, the success of e-commerce is definitely determined by multiple independent variables. Thus, structural equation modeling in this scenario may not be suitable. *Second*, case study is one of qualitative research methods which is useful for the theory building, especially when existing theories are inapplicable [6]. Many researchers have paid attention to the relevant theories in the field of e-commerce, and many related theories have also been published, even though some e-commerce proprietors are still at a financial loss. In addition, the results of case study method are only applicable to specific subjects but not for the overall e-commerce industry in general. Therefore, when the research goal is to propose a generalizable method for the whole

industry, the applicability of case study is in doubt. *Finally*, expert interview can help researchers obtain valuable references especially for those within the company. For instance, an interviewee points out that the main reason for the decline in e-commerce profit is the market plunder coming from the competitors. However, in some cases, the opinions of experts may be too subjective; meaning, this revenue decreasement may be caused by internal factors rather than external ones, such as shopping cart abandonment on the website.

To improve the measurement of e-commerce success, the current study proposes a theoretically constructed and empirically viable way for predicting the relevant factors that may detract or foster e-commerce success. We apply web analytics to simultaneously, generalizably, and objectively measure the influential factors of e-commerce success. It should be noted that e-commerce success in our study is reasonably defined as an actual purchasing behavior performed by the online consumers, because consumer consumption should be one of necessary conditions for the success of e-commerce [7]. After our data analytics, an insightful strategic planning is provided to guide the improvement of e-commerce revenue.

2. Literature Review

2.1. E-Commerce Behavioral Process. Previous studies have divided consumer behavioral process into several steps. From the perspective of advertising, it is suggested that consumers' response to the receiving advertising conforms to the three steps, including thinking, feeling, and doing [8]. Thinking refers to consumers' cognitive aspects with regard to the receiving advertisements, such as beliefs, thoughts, or knowledge, while feeling pertains to their affectively experiences or emotions toward the brand or the content of the advertisement. As for doing, it describes consumers' willingness to proactively perform a certain behavior (e.g., intention to purchase). In the context of e-commerce, Kim et al. [9] proposed a three phase model that depicts consumer online purchase process, including prepurchase, purchase, and postpurchase. The scenario of these three phases can be that consumers often receive advertising over the Internet, and then they will determine whether the received message is trustworthy. If that message is worthy to be trusted, they are very likely to click the advertising URL that directs them into the shopping website (prepurchase phase). When the consumers are brought to the shopping website, they get involved in the next phase. In this purchase phase, consumers spend times on the website and try to scrutinize the products page by page. After surfing the product pages, some of consumers may immediately decide whether to purchase the product, and the others may take several days to consider their purchase decision. Regardless of immediate or hesitant consumers, there are only two results at this phase that is to buy or not to buy. Once they have actually purchased the product, the final phase is said to be emerged. In the postpurchase stage, consumers have converted their role from visitor to customer, and product quality, aftersales service, or product warranty should be emphasized by them. If customers are satisfied with these concerns, they are very likely to conduct their consumption again and share their positive shopping experiences with others (e.g., friends, family members, and colleagues).

Apparently, consumer online behavior can be concretely observed. We therefore theorize the journey of online transaction to be the three phases as well, including *acquisition*, *behavior*, and *conversion*. From the perspective of e-commerce proprietors, the proposed three phases here are relative to the phases mentioned earlier (see Table 1). For example, consumers performing certain online behaviors in the prepurchase phase can be attributed from the effectiveness of consumer acquisition strategy. We believe that this portraying can serve as a navigator to the guidance of e-commerce success.

2.2. Definitions of Big Data. Big data has been inconsistently defined by the previous studies, but most researchers believe that it consists of at least three characteristics, including volume, velocity, and variety [10]. Volume refers to the amount of data, while velocity denotes whether the data is collected in a real time manner and variety is the diversity of data. Although this 3Vs definition is well accepted, some researchers raise different views. For example, Mayer-Schonberger and Cukier [11] argued that big data denote exhaustively data collection rather than data sampling. Boyd and Crawford [12] defined big data as relationality data that contains common fields to enable the conjoining of a brand new dataset. Uddin and Gupta [13], on the other hand, considered that big data can be meaningfully shifting in relation to the context in which they are collected. Gandomi and Haider [14], however, defined big data as value-oriented data interpretation that many meaningful data insights can be extracted. In the current study, we postulate that the record of consumer online behavior is a kind of big data, because (1) as long as consumers visit the website, they generate footprints all the time until they leave the site. Thus, if the continuously generated data is collected, the amount of data should be very large (i.e., volume). (2) Consumers' footprints can be comprehensively collected as long as they happen on the website; meaning, the data is entirely captured rather than being sampled (i.e., exhaustivity). (3) Data value can be obtained through the interpretation of footprint data (i.e., value-oriented data). For example, consumers who stay on the website for more than five minutes showed a 10% purchase rate. Even though e-commerce proprietors possess consumer footprint data, most of them do not realize how to convert valuable information on hand into data insight [15], and we therefore introduce a feasible way in which consumer online behavior can be largely, exhaustively, and valuably collected, especially for those related to online transaction.

2.3. Big Data Analytics. Web analytics is one of big data techniques which can be used to observe the visitor's online journey and elicit the visitor's browsing and purchasing patterns from the website traffic [16]. Generally speaking, web analytics has two types: offsite web analytics and onsite web analytics. The former can be used regardless of the ownership of a website, while the latter can only be used if the

Table 1: The correspondence of the two process models.

Behavioral process model	Phase 1	Phase 2	Phase 3
Online transaction journey (managerial guidance)	Acquisition	Behavior	Conversion
Consumer shopping process (behavior observation)	Prepurchase	Purchase	Postpurchase

analyst owns a website or has the permission to access it. Onsite web analytics measures a visitor's behavior performed on the specific website especially for the measures that are unable to be captured by offsite web analytics (e.g., the number of a website's visitors who conduct a specific action to go beyond a casual content view or simply website visit). As the goal of our study is to predict online transaction, onsite web analytics is more applicable than offsite web analytics, because (1) setting a self-maintained e-commerce website is pervasive for almost all proprietors. This means that proprietors themselves are able to install tracking code into their websites without the need of access permission. (2) E-commerce website itself can be seen as a means by which proprietors apply to the persuasion of potential consumers, and onsite web analytics to this end helps proprietors to estimate consumer interaction with the website. For example, it can reconnoiter what pages potential consumers have visited, where they have been referred from, how much time they have spent on a website, and even clicking behaviors in their journeys of the website. (3) Due to the fact that consumers may withdraw a shopping website from their consideration sets if the product information provided by that website is not insufficient, with the aid of on-site web analytics, proprietors can understand the actual online behavior of potential consumers and accordingly conduct an improvement that highlights consumers' needs of product information. For instance, onsite web analytics can calculate the bounce rate (i.e., the percentage of visitors who enter the website and leave that website immediately) for a shopping website. If a high bounce rate is detected, this means that the landing page of the website is not capable of guiding potential consumers to continue browsing the website in depth. Accordingly, the current study applies onsite web analytics to collect behavioral big data on the e-commerce website which is essential to the prediction of online transaction.

3. Research Design

The current study attempts to identify the features of online behavior that encourage consumers to conduct a transaction on the website and then to develop a prediction model for e-commerce success. Figure 1 illustrates the research process. First, a researched e-commerce website was selected and then variables that are related to online transaction were derived from web analytics. Traffic data of the researched website were used to create a decision tree that predicts consumer transaction on the website. Finally, the accuracy of the prediction model was evaluated and compared to the results of the other prediction algorithms, such as logistic regression, support vector machine, and random forest.

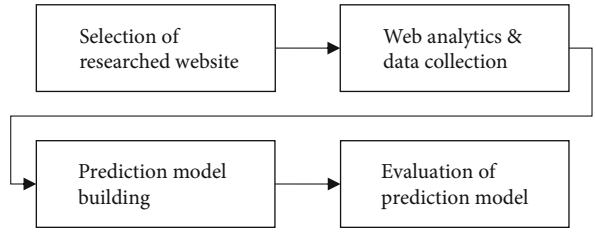


Figure 1: Research process.

3.1. Decision Tree. Decision tree is a supervised machine learning algorithm for the establishment of the discriminatory model [17]. It uses a tree-like model to extract valuable relationships from the information hidden in a data source and is frequently used for the purpose of classification or prediction [18]. There are two kinds of decision tree in general, including regression tree and classification tree. The former is developed when the goal factor is continuous, whereas latter is created when the goal factor is discrete. Regardless of which type of decision tree, they all go through a series of tests until the flow path reaches the terminal node. In the current study, we adopt classification tree rather than regression tree, because our goal is to predict whether or not consumers will conduct a transaction on the website which belongs to the concept of binary value.

3.2. Research Subject. For maintaining external validity, we collected the data from a real e-commerce website of Google (i.e., https://shop.googlemerchandisestore.com) which is a large-scale e-commerce website operated in the United States. Google e-commerce website sells many kinds of self-branded product, including backpack, apparel, electronics, and accessories. Collected data was elicited from the tool of web analytics developed by Google which is called Google Analytics (GA). Figure 2 outlines an overview of GA report. In default, this report includes data columns about users, new users, sessions, number of sessions per user, pageviews, pages/session, avg. session duration, and bounce rate.

3.3. Research Variables. In the context of online transaction, several variables are related to e-commerce success. Table 2 describes the variables that are used for the creation of decision tree. *Channel grouping*, *totals_bounces*, and *visitNumber* are the three factors relevant to the effectiveness of consumer acquisition (i.e., phase 1). Channel grouping (channelGrouping) is defined as the way of website traffic in which visitors are acquired, including organic search, direct, paid search, referral social, display, and affiliate. For example, if a visitor gets into the website by searching keywords on the search engine, this entrance traffic will be termed by GA as organic search. However, if a website visit is attributed by the search engine results that is the result of paid advertising, it can be termed as paid search. In addition, website

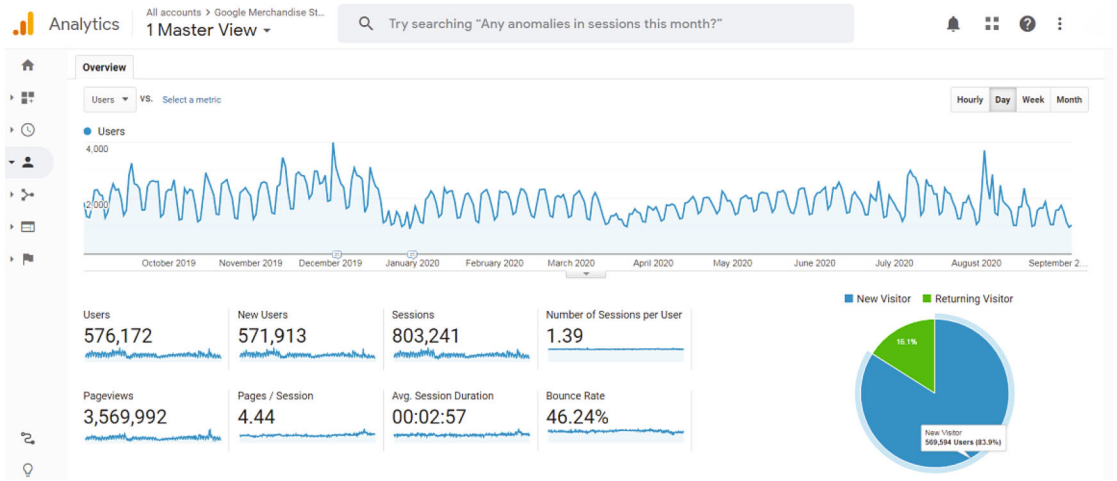


Figure 2: An overview of Google Analytics report.

proprietors may sometimes acquire wrong visitors by the wrong traffic channels. If this is true, the visitors who get into the website will very likely to leave that site without performing any further behaviors. Thus, the number of bounces (totals_bounces) can be used to monitor such a harmful behavior of the website. Furthermore, the number of times visitors have entered the website (visitNumber) plays as a necessary condition for the establishment of a complete shopping journey; as a result, it is impossible to make a purchase without the website visit. Due to the fact that website visitors attracted from different traffic channels exert differently on their website journeys, especially for those add-to-cart behaviors [19] and the appropriate traffic channel can effectively fix the problem of a landing page with high bounce rate [20], or even the number of website visits is highly related to website engagement of online consumers [21], the current study therefore includes these three measures into the decision tree and theorizes them as representative factors in the phase of consumer acquisition.

On the other hand, once visitors have actually entered the website, the number of pages they have viewed on the website (totals_pageviews) and the length of time the visitors have spent on the website (totals_timeOnSite) should be the other two important factors need to be addressed. Indeed, the more the webpages visitors have viewed, the more likely they will stay longer on the website, specifically for the visitors whose purpose is exploratory visit rather than goal-oriented visit [22]. Although it is difficult to measure the real success of e-commerce, technical measures here can serve as an alternative to indirectly validate the so-called success, because consumer website stickiness should be a consequence of high pageviews and long staying duration [23] which in turn affects consumers' intention to purchase online [24]. It is well known that purchase intention of a consumer does not necessarily correlated to her/his actual purchase behavior [25], for example, a consumer puts a product into the shopping cart but consequently regrets this temporary collection, and cart abandonment is said to have occurred. Thus, the number of transactions made by visitors on the website (totals_transactions) is particularly important to e-commerce success. In general, totals_pageviews and totals_timeOnSite are theorized in phase 2, while totals_transactions in phase 3 is characterized as the destination of online transaction journey.

4. Data Analysis

4.1. Summary of Research Data. In this study, the rationality of traffic data is regarded as an important data cleaning criterion, because unreasonable traffic will result in bias which is detrimental to our prediction model. In the context of web analytics, the so-called unreasonable traffic can be a webpage with very short time on site. If a website visitor stays on the page for very short period of time (e.g., 3 seconds), the possibility of completing a purchase in such short period of time is unlikely. Accordingly, webpages with time of site less than 10 minutes will be removed from our data set. Anesbury et al. [26] also reported that most consumers spend more than 10 minutes to complete their online transactions. After data cleaning, a total of 1,692 website traffic data were obtained in the current study. Table 3 summarizes the outcome of descriptive statistics for all the research variables, including three continuous variables and two categorical variables. For the continuous variables, which show that the average number of visits to Google Merchandise Store is 3.07 (visitNumber), the average pageviews received by this website is 14.20 pages (totals_pageviews), and the average time visitors spend on the website is 642.46 seconds (totals_timeOnSite). For the categorical factors, which show that 25 percent of data enter the website and bounce the websites, 51 percent of data has transacted. Referring to bounce rate, 25% of consumers leaved the website without additional interaction (totals_bounces), while nearly 51% of consumers conducted their transactions on the website (totals_transactions). It should be noted that totals_bounces and totals_transactions are categorical factors; thus, the values of them are extremely distributed.

4.2. Decision Tree Development. A decision tree to predict e-commerce success is developed by taking consumers' online

Table 2: Research variables.

Variables and corresponding phase	Description
Acquisition (phase 1)	
channelGrouping	The origins of the website traffic that commonly includes direct, referral, organic search, paid search, display, affiliate, and social.
totals_bounces	The number the visitors get into the website but immediately leave the site without additional activities.
visitNumber	The number of times visitors have entered the website.
Behavior (phase 2)	
totals_pageviews	The number of pages viewed by visitors on the website.
totals_timeOnSite	The length of time visitors spend on the website.
Conversion (phase 3)	
totals_transactions	The number of transactions made by visitors on the website.

Table 3: Summary of research data (sample size = 1692).

Research variables	Minimum	Maximum	Mean	Standard deviation
visitNumber	1	376	3.07	13.281
totals_ pageviews	1	165	14.20	15.320
totals_ timeOnSite*	600	6865	642.46	746.500
totals_bounces	0	1	0.25	0.434
totals_ transactions	0	1	0.51	0.500

*The metric of totals_timeOnSite is expressed in seconds by default.

transactions as the dependent variable (i.e., to buy or not to buy), and the four characteristics related to the phases 1 and 2 of online transaction journey are treated as the independent variables. Out of 1692 valid website traffic data, a total of 1184 were categorized as training data set (nearly 70% of all data), and the remaining of 508 was served as testing data set (nearly 30% of all data). Testing data set was used to validate the effectiveness of the proposed training model. SPSS Statistics with the Exhaustive CHAID (Chi Squared Automatic Interaction Detection) algorithm was applied to the tree construction. This algorithm is a modification of CHAID which evaluates all the possible splits for each independent variable. In addition, Exhaustive CHAID consists of three steps, including merging, splitting, and stopping. These three steps are repeatedly operationalized on each node from the top of tree until no additional nodes can be produced. In other words, if an independent variable has the strongest association with the root variable, it will be treated as the first branch in the tree with a leaf for each category that is significantly different from its parent variable. This process is continuous until no significant parent variables exist. It should be noted that the max_depth of the decision tree is not presupposed in our study, because we want to expand all the nodes in order to obtain pure leaves. This is the reason why we choose a modification of CHAID in SPSS so that all possible splits for each predictor can be measured. Yang and Shami [27] also adopted the same strategy to choose the best split node for the impure node. Accordingly, the hyperparameter of max_depth was determined to be three as it reveals the best accuracy for the prediction model and not the most complicated.

Figure 3 illustrates the outcome of decision tree development. The most important factor related to *totals_transactions* is *totals_pageviews*. Compared with consumers who are unwilling to perform purchase behavior and have website visit for less than 4 pages (nonpurchase: 99.6%, $n = 458$ vs. purchase: 0.4%, $n = 2$), consumers who visit the website for 4 to 11 pages are more likely to show purchase behavior (nonpurchase: 66.7%, $n = 93$ vs. purchase: 33.3%, $n = 47$). Moreover, when consumers visit the website for more than 11 pages, they exhibit higher number of transactions (nonpurchase: 5.4%, $n = 32$ vs. purchase: 94.6%, $n = 552$). This phenomenon is consistent with the finding of previous studies [28] that the higher the degree of website involvement, the more likely consumers are able to show their purchase intentions. Due to the fact that *totals_pageviews* plays a key role in purchase decision, it is necessary to clarify its facilitating factors. *totals_timeOnSite* and *visitNumber* function as main factors for the inducement of high *totals_pageviews* via respective traffic channels. For *totals_timeOnSite* (green squares in the figure), consumers who stayed on the website for over 1,461 seconds can perform higher transaction rate through organic and social channels in comparison with those who stayed on the website for less than 661 seconds.

This implies that consumers with higher stickiness are willing to spend more time on the website, because they get into the website voluntarily (i.e., channelGrouping: organic search) or they come to the website through the recommendation of their social friends (i.e., channelGrouping: social). Accordingly, organic search and social in *channelGrouping* are ingredients of customer acquisition, while *totals_timeOnSite* is prerequisite of customer retention, either the former or the later is essential to the success of customer value management [29]. As for *visitNumber* (blue squares in the figure), consumers who visit the website for over one time exhibit higher transaction rate through different channels (i.e., Direct, Referral, Paid Search, Display) than those who visit the website for merely one time. This finding confirms the importance of website visit that the more visits to the website, the more likely the purchase will

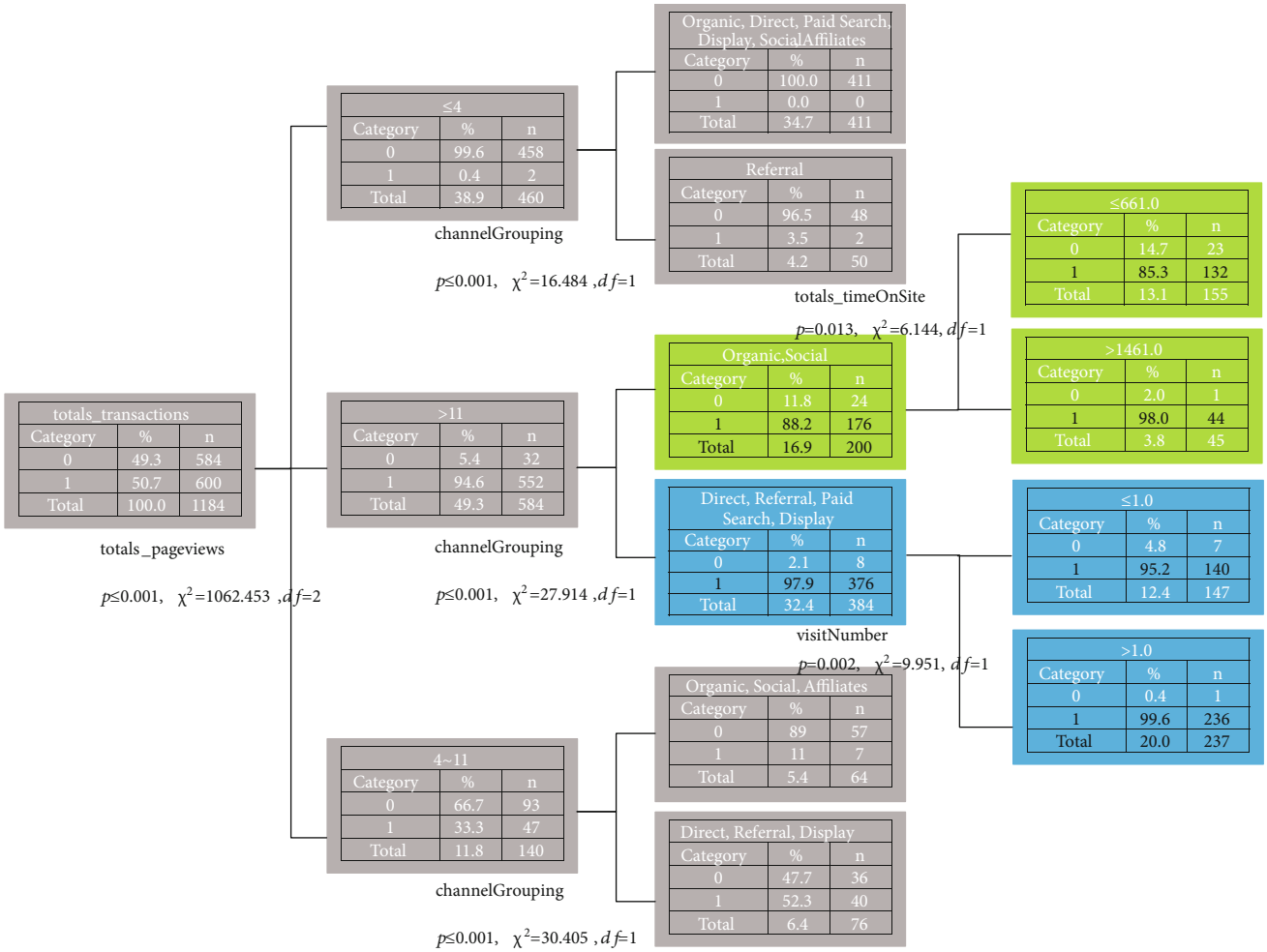


Figure 3: Decision tree.

be conducted. Despite a large number of website visits still end without any purchase, it can induce high pageviews when these traffics are triggered by the right channels which in turn derives purchase behavior. Generally speaking, if a visitor stays long enough on the website or if she/he frequently visits the website, a sufficient number of pageviews will be elicited, and the visitors should be inclined to have a purchase.

4.3. Model Evaluation. Table 4 outlines the evaluation results of the proposed mode. The prediction accuracy of the two data groups (i.e., non-purchase vs. purchase) is 93.9% and 99.2%; thus, the overall accuracy is 96.7%. Apparently, the prediction accuracy of the purchase group outperforms the nonpurchase group.

To further evaluate the effectiveness of the proposed model, we use different algorithms to develop other prediction models for the purpose of model comparison, including logistic regression, SVM, and random forest. Table 5 summarizes the results. Regardless of the overall prediction accuracy or the prediction accuracy of the purchase group, the performance of the proposed model outperforms the other three control models.

Table 4: Prediction accuracy of the proposed model (decision tree).

Observed	Predicted		
	Nonpurchase	Purchase	Percent correct
Nonpurchase	229	15	93.9%
Purchase	2	262	99.2%
Overall percentage	45.5%	54.5%	96.7%

Table 6 reveals the results of five-fold crossvalidation. The average accuracy and standard deviation of them are 94.4% and 0.009, respectively, confirming the stability of the model.

In summary, the decision tree algorithm in our study has better predictive ability than the other three algorithms, and it is therefore suitable for predicting the likelihood of consumer purchase and even to elicit under what circumstances consumers are willing to engage in online transaction [30].

5. Discussion

The current study found that pageviews is an important key for consumers to make transactions. In other words, as long as the number of pageviews reaches the threshold detected

Table 5: Prediction accuracy of the comparative models.

Observed	Predicted		
	Nonpurchase	Purchase	Percent correct
Logistic regression			
Nonpurchase	227	17	93.4%
Purchase	4	260	97.3%
Overall percentage	45.5%	54.5%	95.5%
SVM			
Nonpurchase	228	16	86.9%
Purchase	7	257	98.1%
Overall percentage	46.3%	53.7%	92.7%
Random forest			
Nonpurchase	227	17	93.0%
Purchase	4	260	98.5%
Overall percentage	45.5%	54.5%	95.9%

Table 6: The five-fold crossvalidation.

	1st fold	2nd fold	3rd fold	4th fold	5th fold
Accuracy	94.7%	94.4%	92.9%	94.7%	95.6%
Mean: 94.4% Std: 0.009					

from the web analytics, consumers are likely to engage in purchase behavior. Based on our findings, there can be two strategies to effectively increase the number of pageviews on the website, and they are *adhesion strategy* and *repeatability strategy*. The focus of adhesion strategy is to enable consumers increasing their stay on the website, especially when their website entrances are through voluntary search engine queries (i.e., organic search) or recommendations from friends (i.e., social). In fact, people can put more effort into voluntary things [31] or things recommended by friends [32]; hence, overly forcing consumers or the wrong source of information may reduce their website stickiness. As for repeatability strategy, a repeated visit to a website would be devoted to seek for information that was not covered by the first time visit [33], because not all consumers are willing to complete a purchase within one visit to the website. As a result, consumers may need to visit the website multiple times before they actually engaging in the purchase. The association between website repeated visit and actual consumption behavior can also be validated by our findings that consumers who have entered the website for more than once through various channels (i.e., direct, referral, paid search, and display) are more able to demonstrate actual purchase behavior. It is worth noting that the two strategies here may have been mentioned in literature, but they are frequently mentioned with other e-commerce success factors. However, too many factors are mentioned simultaneously will cause the operators have no idea about the priority of these influencing factors. Thus, the above two strategies specifically inform large-scale e-commerce companies to effectively improve online transactions. Furthermore, previous studies frequently adopt experimental design or questionnaire to be the way of data collection. However, these two methods have inherent limitations of research bias. In our study, we use web analytics to collect comprehensive data of online behavior that it has the potential to eliminate a certain degree of research bias. In other words, the two proposed strategies are based on the support of a large amount of behavior data. In sum, the current study is one of few empirical analytics that deal the prediction of website traffic with a managerial perspective of online transaction. It is also the preliminary study which merges the theoretical underpinning of online transaction journey into the decision tree model and therefore generates meaningful insights into online shopping characteristics that increase the likelihood of e-commerce success.

6. Conclusion

The current study focuses on the prediction of online purchase which is a prerequisite for e-commerce proprietors to succeed in the marketplace. We proposed a decision tree model for predicting whether or not consumers will actually conduct their online transactions. Variables in the proposed model were derived from e-commerce behavioral process, and the value for each variable is collected from web analytics which is real data performed by online consumers over the website. The prediction accuracy of the proposed model is higher than the other three control models which brings us to conclude that the decision tree model is a quite stable approach for the purpose of data classification, especially when the goal of analysis is to explore the data insights hidden in the dataset.

6.1. Academic Contribution. Several theoretical contributions can be drawn from our findings. *First*, unlike technical studies in general, the current study tries to merge theoretical underpinning into the decision tree under which the root node of the tree denotes conversion phase of online transaction journey, while the leaf nodes of the tree represent acquisition and behavior phases of that journey. We believe that such a theoretical mapping can shed light on data insight, especially when the research focus is to understand the uncertainty of consumer online transaction. As far as we know, this research plays a pioneering role in rationalizing the necessity of theory and technique integration. *Second*, we found that the bounce rate should not be a member factor of acquisition phase. Many previous studies have focused their research point on the decreasement of the bounce rate [34–36]; however, when the goal of the research is to clarify what causes the purchase rather than the unwillingness to purchase, paying attention of the bounce rate is somewhat unhelpful. Thus, the current study theoretically points out what causes consumers to engage in online purchase. *Finally*, online transaction journey is a theoretically inseparable concept even though it can be separately measured in practice. In other words, any phase in the journey should have certain impact on each other; therefore, it is unreasonable to treat them separately when the goal is to build a prediction model of e-commerce success. The decision tree model in our study is theoretically supported by online

transaction journey which has the potential to present the overall shopping behavior found on the website.

6.2. Managerial Implications. The current study proposes a prediction model that addresses the situations in which consumers would like to conduct online transactions. It is confirmed that the number of pageviews is a prerequisite to the success of e-commerce website. Proprietors can therefore clarify some ways that have the potential to promote consumer involvement of the website. For example, proprietors can invite their consumers to visit the website for more than once in order to exchange pageviews with them. Proprietors can also use channel identification to understand in what channels are able to guide valid traffic with both the longest stay and the highest pageviews on the website. All these strategic planning are relevant to online transaction which can be derived from the proposed prediction model. In other words, when there is any variance of independent variable found in the tree, proprietors can therefore predict its possible impact on consumer online transaction.

6.3. Limitations. Similar to most studies, the current study has several limitations without exception. *First*, our study applies the decision tree model to predict website traffic; however, there can be other classification methods that may produce different research findings. Future studies are recommended to used different classification algorithms and revalidate our research findings. *Second*, the research data in our study is limited to a specific date interval. It is suggested that future studies can try to expand the date range of the research data in order to get a complete picture of website traffic. *Finally*, the prediction model of our training data is based on the large-scale e-commerce website; thus, it may not be applicable to small- and medium-sized websites with insufficient traffic. Therefore, we suggest that future studies can apply our prediction model to similar scale e-commerce websites for the purpose of research revalidation.

Data Availability

The web traffic data used to support the findings of this study were supplied by researched websites, under license and so cannot be made freely available. Requests for access to these data should be made to Han-Ping Tsen (hanping311111@gmail.com).

Conflicts of Interest

This paper does not contain any conflict of interest.

References

[1] eMarketer, "Global Ecommerce 2020: E-commerce decelerate amid global retail contraction but remains a bright spot," 2020, 2020, https://www.emarketer.com/content/global-ecommerce-2020.

[2] W. T. Wang, Y. S. Wang, and E. R. Liu, "The stickiness intention of group-buying websites: the integration of the commitment-trust theory and e-commerce success model," *Information & Management*, vol. 53, no. 5, pp. 625–642, 2016.

[3] N. Kshetri, "Barriers to e-commerce and competitive business models in developing countries: a case study," *Electronic Commerce Research and Applications*, vol. 6, no. 4, pp. 443–452, 2007.

[4] H. Liu, V. V. Krasnoproshin, and S. Zhang, "Alogarithms for evaluating and selecting e-commerce websites," *Journal of Computational Optimization in Economics and Finance*, vol. 4, no. 2-3, pp. 135–148, 2012.

[5] W. W. Chin, "How to write up and report PLS analyses," in *Handbook of partial least squares*, pp. 655–690, Springer, Berlin, Heidelberg, 2010.

[6] W. J. Tibben, "Theory building for ICT4D: systemizing case study research using theory triangulation," *Information Technology for Development*, vol. 21, no. 4, pp. 628–652, 2015.

[7] A. Bodhani, "Getting a purchase on AR," *Engineering & Technology*, vol. 8, no. 4, pp. 46–49, 2013.

[8] T. E. Barry and D. J. Howard, "A review and critique of the hierarchy of effects in advertising," *International Journal of Advertising*, vol. 9, no. 2, pp. 121–135, 1990.

[9] D. J. Kim, D. L. Ferrin, and H. R. Rao, "Trust and satisfaction, two stepping stones for successful e-commerce relationships: a longitudinal exploration," *Information Systems Research*, vol. 20, no. 2, pp. 237–257, 2009.

[10] A. De Mauro, M. Greco, and M. Grimaldi, "A formal definition of big data based on its essential features," *Library Review*, vol. 65, no. 3, pp. 122–135, 2016.

[11] V. Mayer-Schonberger and K. Cukier, *Big Data: A Revolution That Will Transform how we Live, Work, and Think*, Houghton Mifflin Harcourt, London, 2013.

[12] D. Boyd and K. Crawford, "Critical questions for big DATA," *Information, Communication & Society*, vol. 15, no. 5, pp. 662–679, 2012.

[13] M. F. Uddin and N. Gupta, "Seven V's of big data understanding big data to extract value," in *Proceedings of the 2014 zone 1 conference of the American Society for Engineering Education*, pp. 1–5, Bridgeport, CT, USA, 2014.

[14] A. Gandomi and M. Haider, "Beyond the hype: big data concepts, methods, and analytics," *International Journal of Information Management*, vol. 35, no. 2, pp. 137–144, 2015.

[15] S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big data, analytics and the path from insights to value," *MIT Sloan Management Review*, vol. 52, no. 2, pp. 21–32, 2011.

[16] F. Shirazi and M. Mohammadi, "A big data analytics model for customer churn prediction in the retiree segment," *International Journal of Information Management*, vol. 48, pp. 238–253, 2019.

[17] H. Chen, R. H. Chiang, and V. C. Storey, "Business intelligence and analytics: from big data to big impact," *MIS Quarterly*, vol. 36, no. 4, pp. 1165–1188, 2012.

[18] J. P. Singh, S. Irani, N. P. Rana, Y. K. Dwivedi, S. Saumya, and P. Kumar Roy, "Predicting the "helpfulness" of online consumer reviews," *Journal of Business Research*, vol. 70, pp. 346–355, 2017.

[19] D. Drossos, M. Zacharioudakis, and G. Dionysiou, "Online traffic sources and persuasion techniques: how to change consumer behavior," in *Proceedings of the 2019 3rd International Conference on E-commerce, E-Business and E-Government*, pp. 80–84, New York, NY, USA, 2019.

[20] S. J. Turner, "Website statistics 2.0: using Google Analytics to measure library website effectiveness," *Technical Services Quarterly*, vol. 27, no. 3, pp. 261–278, 2010.

[21] Q. Jiang, C. H. Tan, C. W. Phang, J. Sutanto, and K. K. Wei, "Understanding Chinese online users and their visits to websites: application of Zipf's law," *International Journal of Information Management*, vol. 33, no. 5, pp. 752–763, 2013.

[22] J. Li, A. Abbasi, A. Cheema, and L. B. Abraham, "Path to purpose? How online customer journeys differ for hedonic versus utilitarian purchases," *Journal of Marketing*, vol. 84, no. 4, pp. 127–146, 2020.

[23] C. Tangmanee, "Comparisons of website visit behavior between purchase outcomes and product categories," *IJRBS*, vol. 6, no. 4, pp. 1–10, 2017.

[24] J. C. C. Lin, "Online stickiness: its antecedents and effect on purchasing intention," *Behaviour & Information Technology*, vol. 26, no. 6, pp. 507–516, 2007.

[25] V. G. Morwitz, "Methods for forecasting from intentions data," in *Principles of forecasting*, pp. 33–56, Springer, Boston, MA, 2001.

[26] Z. Anesbury, M. Nenycz-Thiel, J. Dawes, and R. Kennedy, "How do shoppers behave online? An observational study of online grocery shopping," *Journal of Consumer Behaviour*, vol. 15, no. 3, pp. 261–270, 2016.

[27] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, 2020.

[28] A. H. C. Hwang, J. Oh, and A. C. Scheinbaum, "Interactive music for multisensory e-commerce: the moderating role of online consumer involvement in experiential value, cognitive value, and purchase intention," *Psychology & Marketing*, vol. 37, no. 8, pp. 1031–1056, 2020.

[29] P. C. Verhoef and K. N. Lemon, "Successful customer value management: key lessons and emerging trends," *European Management Journal*, vol. 31, no. 1, pp. 1–15, 2013.

[30] D. Westreich, J. Lessler, and M. J. Funk, "Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression," *Journal of Clinical Epidemiology*, vol. 63, no. 8, pp. 826–833, 2010.

[31] P. Dekker, "Tocqueville did not write about soccer clubs: participation in voluntary associations and political involvement," in *Modernizing democracy*, pp. 45–47, Springer, New York, 2014.

[32] M. Zhao and J. Xie, "Effects of social and temporal distance on consumers' responses to peer recommendations," *Journal of Marketing Research*, vol. 48, no. 3, pp. 486–496, 2011.

[33] C. W. Choo, B. Detlor, and D. Turnbull, "Information seeking on the web–an integrated model of browsing and searching," in *Proceedings of the Annual Meeting of the American Society for Information Science (ASIS)*, pp. 3–16, Washington, 1999.

[34] A. Gofman, "Consumer driven multivariate landing page optimization: overview, issues and outlook," *The IPSI BgD Transactions on Internet Research*, vol. 3, no. 2, pp. 7–9, 2007.

[35] D. Schellong, J. Kemper, and M. Brettel, "Generating consumer insights from big data clickstream information and the link with transaction-related shopping behavior," in *Twenty-Fifth European Conference on Information Systems (ECIS)*, pp. 1–15, Guimaraes, 2017.

[36] J. Xun, "Revisiting the two-stage choice model: an empirical study of consumer choice on brand website visits," *Behaviour & Information Technology*, vol. 33, no. 11, pp. 1192–1207, 2014.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5518868, 10 pages
https://doi.org/10.1155/2021/5518868

WILEY | Hindawi

Research Article

Research on Business English Translation Architecture Based on Artificial Intelligence Speech Recognition and Edge Computing

Yunwei Xu

Nanchang Jiaotong Institute, Nanchang 330100, China

Correspondence should be addressed to Yunwei Xu; 07054@ecjtuit.edu.cn

Received 2 February 2021; Revised 15 April 2021; Accepted 5 July 2021; Published 18 August 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Yunwei 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 society, the continuous deepening of international cultural integration has become the background of the times. China has become more and more closely connected with the world, and many physical or online news media have become a platform for China to receive world information and spread Chinese culture. Business English translation is therefore valued by translation researchers and translators. Aiming at the shortcomings of current business English translation research, this paper designs and develops a business English translation architecture based on artificial intelligence speech recognition and edge computing. First of all, considering the relevance and complementarity between speech and text modalities, this paper uses the deep neural network feature fusion method to effectively fuse the extracted monomodal features and perform speech recognition. Secondly, adopt the edge computing method to establish the business English translation system architecture. Finally, the simulation test analysis verifies the efficiency of the business English translation framework established in this paper. Compared with the existing methods, our proposal improved the accuracy than others at least 10% and the time of model building also decreased obviously. The purpose of this research is to discuss how to deal with the many differences between the source language and the target language, and how to enhance the readability of the translation and meet the reader's cultural cognition and needs.

1. Introduction

Recently, international business activities in China have been increasing for many years. A deep understanding of the business translation theory has important practical significance and principles [1, 2]. From the actual situation of business translation, there are many difficulties in understanding and expression. International business translation has received increasing attention as communication tools. In the foreseeable future, businessmen should be able to communicate directly with their foreign partners in English [3, 4]. Government officials and decision makers, who are not businessmen, should be able to review business documents in English. In short, business translation in the era of globalization is very necessary and important [5, 6].

Although people are still too focused on specific translation skills and detailed functional applications, there is still a lack of systematic theoretical translation practices that can guide business [7, 8]. Business translation can provide a fertile ground for translation theory and provide a lot of help to enrich and develop it [9, 10]. In this sense, it is necessary to promote commercial translation to improve the translation theory and its improvement. Business translation plays an important role in various business situations. This requires translators to fully understand the relevant business knowledge and have the correct understanding of words, sentences, and paragraphs in the translation process, as well as to find out the exact meaning of words or sentences in different business situations [11, 12]. Since 1980, research on translation versatility has been explosive. Many linguists have studied generality from different perspectives, including syntactic, pragmatic, cognitive, and functional perspectives [13, 14]. Jordan divides ESP into two types: one is the purpose of academic English, which refers to the English and communication used to complete study and academic research. The other is professional English, which refers to English suitable for a certain industry, which is more practical and professional. Nick Briefer put forward a famous theory about the category of business English [15, 16]. He claimed that business English includes language knowledge, communication

skills, professional content, management skills, and cultural awareness, in other words, marketing and investment, foreign insurance, international finance, publicity, etc. The universality of translation is an inherent feature, expressed in terms of simplified explicitness and standardization, and has nothing to do with language and cultural differences and the translator's personal preferences [17, 18]. Compared with English texts, based on English translation texts, we can find that almost all English translation texts have several characteristics. These universal theorists have given simplification, explicitness, and normalization.

With the development of descriptive translation studies (DTS), translation is no longer considered a static text transmission process language from the source language to the target language for a long time. Descriptive translation studies (DTS) broadens the horizons of translation studies by shifting the focus from goal orientation to target culture rather than prescriptive translation studies [19, 20]. It provides a theoretical framework for corpus-based translation. China has become more and more closely connected with the world, and many physical or online news media have become a platform for China to receive world information and spread Chinese culture. Business English translation is therefore valued by translation researchers and translators. Aiming at the shortcomings of current business English translation research, this paper designs and develops a business English translation architecture based on artificial intelligence speech recognition and edge computing.

In our work, we design and develop a business English translation architecture based on artificial intelligence speech recognition and edge computing to deal with the many differences between the source language and the target language, and to enhance the readability of the translation and meet the reader's cultural cognition and needs. The main contributions of our proposal are as follows:

(1) It considers the relevance and complementary between speech and text modalities

(2) It uses the deep neural network feature fusion method to effectively fuse the extracted monomodal features

(3) It adopts the edge computing method to establish the business English translation system architecture

The rest of this paper is organized as follows. Section 2 discusses the business English translation framework, and edge computing technology is induced in Section 3. Business English translation based on speech recognition methods are discussed in Section 4. Section 5 shows simulation result analysis and discussion, and Section 6 concludes the paper with summary and future research directions.

2. Business English Translation Framework

With the help of relevant background and language knowledge, business translators should not only have reliable language knowledge but also should have economic, legal, management, trade, and other aspects [21–30]. Then, all this knowledge will be applied to provide the background of the content involved in the translation and help the translator to complete the task better and faster. Business theory and practice are two important parts of business English, which means that the language involved should be professional. Without the following foundation, you cannot fully understand business terms [24]. At the same time, due to the huge differences in Western culture between China and China, people have different values and ideologies; translators may encounter expressions that they have never seen before or are difficult to understand. In this case, we can apply the schema theory to cultural background, language environment activation, content structure, and other structure activation to help translators merge what they have learned and have a better understanding of the content [25]. The architecture of business English translation is shown in Figure 1.

With the globalization of the world economy, more and more companies are committed to participating in the international market competition. Therefore, more and more scholars are devoted to the research of business translation [26, 27]. A lot of content related to business translation also appeared. However, there is no uniform definition that is agreed by experts from all walks of life. In this part, the author tries to find out the characteristics and common points of specific previous research and gives the definition of business translation of this article. Business English is proposed under the theoretical framework of ESP, which is a kind of English business activities specially used for international professional subjects [28]. Jordan divides ESP into two types: one is the purpose of academic English, which refers to the English and communication used to complete study and academic research. The other is professional English, which refers to English suitable for a certain industry, which is more practical and professional. Nick Briefer put forward a famous theory about the category of business English [29, 30]. He claimed that business English includes language knowledge, communication skills, professional content, management skills, and cultural awareness. In other words, marketing and investment, foreign insurance, international finance, publicity, etc. are all covered in the scope of business English [31–37]. Therefore, it can be seen that business English is a kind of practical language, a medium of business activities, and business-related English needs and contains various business activities. In essence, business English is a combination of business activities and English. Therefore, business English can be defined as the English used in various business activities.

This society needs business English translators to communicate with each other, to have cross-cultural communication awareness, to grasp the characteristics of business translation, and to avoid contradictions and conflicts caused by cultural differences [38–41]. This paper explains the characteristics of cross-cultural factors in business English translation and study the method for cross-cultural phenomena in business English translation. This will contribute to strengthen awareness of cross-cultural in English translation and makes the ordinary people to have a better realization of various works around the word.

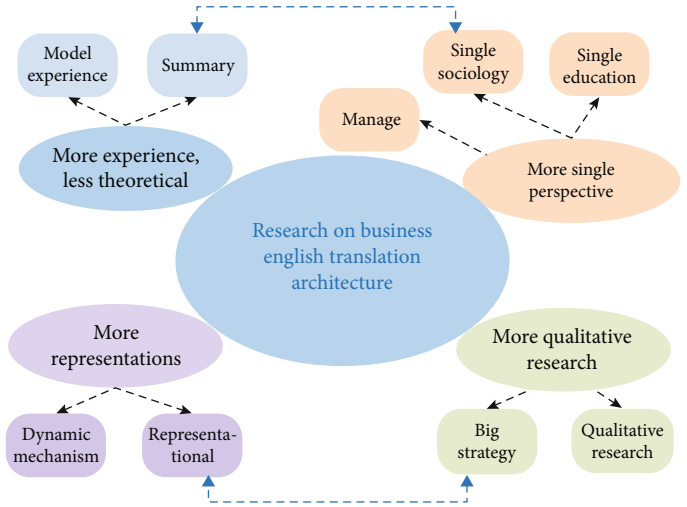


Figure 1: Schematic diagram of the architectural concept of business English translation.

3. Edge Computing Technology

The popularization of mobile terminals such as smartphones, tablet computers, and smart vehicles has had a huge impact on mobile and wireless networks, triggering challenges for global mobile networks. Mobile devices have certain computing and storage capabilities, but they still suffer from low storage capacity, low bandwidth, and high latency [31, 32]. Mobile cloud computing (MCC) integrates cloud computing and mobile computing, provides powerful functions for mobile devices, and provides them with storage, computing, and energy resources provided by centralized clouds. However, facing countless mobile devices, MCC faces huge challenges. For example, problems such as high latency, low coverage, and data transmission lag. Especially in the case of next-generation mobile networks (such as SG), it is particularly difficult to deal with these challenges. In addition, MCC is not suitable for scenarios where real-time applications guarantee high-quality services [33, 34].

Edge computing is the combination of mobile edge network and Internet services in the future 5G network. Its biggest feature is to sink core services to the edge of the mobile network, thereby improving the quality of various types of services, reducing waiting time, and providing a better user experience. MEC provides cloud computing functions in the RAN, allowing direct mobile communication between the core network and end users, while connecting users directly to the nearest edge network enabled with cloud services [35]. Deploying on base stations can enhance computing power and effectively avoid bottlenecks and system failures.

4. Research on Business English Translation Based on Speech Recognition

Machine translation is a deep learning technology combined with NLP using one of the most active, the most hopeful one direction. The original fully rely on people compiling rule-based machine translation method, then based on statistical method of SMT now nerve NMT machine translation, machine translation technology over the past 60 years has been updated, especially after the 2012 deep learning technology into the field of vision; the accuracy of machine translation is constantly refreshed; today, it is mainly taking stock of all kinds of deep learning the present situation of the application of machine translation on the inside, giving some representative papers to learn.

The biggest advantages of neural translation (NMT) based on deep learning technology are as follows: (1) It adopts an end-to-end structure and does not need to extract features artificially. (2) The network structure design is simple, and there is no need to carry out complex design work such as word segmentation, word alligation, and syntactic tree design. Meanwhile, the disadvantages of this method are also obvious: (1) Deep learning training often has hundreds of millions of samples. (2) Training a model requires a special GPU cluster, and it takes a few days or even a week to get a result. The iterative update of the model is very slow. Nal Kalchbrenner and Edward Grefenstette put forward in 2013, based on coding decoding structure framework, a new kind of machine translation in the source language sentence, using a convolution neural network to map it into a continuous dense implicit vector; using a recursive neural network as a decoder, the implicit vector decode sentences into the target language. Doing a little bit is that you can use the RNN to deal with different lengths of the input sentence, which try to capture all of its historical information, but a gradient disappeared and gradient explosion occurred because of RNN. There is no way to capture long-term dependencies.

Speech recognition is an interdisciplinary subject, involving many fields, such as artificial intelligence, pattern recognition, signal processing, and information theory. In information theory, the statistical method of speech recognition is modeled as a noise channel model [36, 37]. Each component is associated with the component of the general noise channel model at the bottom of the figure. The speech recognition process includes the speaker and the speech

recognizer. The speaker will consciously generate a sequence of words. A schematic diagram of the architecture of the convolutional neural network is shown in Figure 2.

Then, the word sequence is sent to the acoustic (corresponding to the noise channel) channel composed of a voice generator and an acoustic processing unit. The voice generator generates a voice signal of the word sequence in the acoustic environment, and then, the acoustic processing unit signals the voice signal processing to generate acoustic features; the speech decoder receives a sequence of feature vectors [42] and infers the word sequence that is closest to the original word sequence; the acoustic processing unit and the language decoder together form the speech recognizer.

An artificial neural network (ANN) is an information processing model inspired by the complex neural network of the brain. Simply put, the model consists of multiple layers of neurons. In order to obtain the most ideal model output results, it is necessary to constantly adjust the relevant parameters between neurons [38, 43]. A schematic diagram of business English translation based on speech recognition is shown in Figure 3.

Therefore, the artificial neural network model needs to be combined with the corresponding algorithm to solve the problem of how to adjust the weight parameters. The BP (back propagation) algorithm solves this problem well [39]. The basic principle of the BP algorithm includes the forward propagation of sample features at all levels. At the same time, the learning rules of the steepest descent method are used to adjust the weights of related neurons in the model through the feedback value of back propagation, so as to minimize the output error of the model. The BP algorithm is based on neurons, and its composition is divided into three parts: the input layer, hidden layer with different numbers, and output layer. The specific structure and algorithm are shown in Section 2. In recent years, neural networks have been rapidly developed and used, and they have performed well in the fields of pattern recognition, artificial intelligence, and automatic control [40].

In the first few layers of the convolutional neural network, the data nodes are transformed into a three-dimensional matrix. Unlike the fully connected layer, only some nodes are connected to adjacent layers, which greatly reduce the number of parameters and reduces subsequent models' training time. Usually a convolutional neural network consists of the following structure:

$$\begin{aligned} w_i &= \sqrt[n]{M_i}, \\ \lambda_{\max} &= \sum_{i=1}^{n} \frac{(BW)_i}{n w_i}. \end{aligned} \tag{1}$$

The input layer is the input of the neural network. In the image-oriented convolutional neural network, it represents the pixel matrix converted after a picture is read by a computer.

$$Xt = \sum_{i=1}^{k} \alpha_i X_{t-i} + \sum_{i=1}^{k} \beta_i Y_{t-i} + u_{lt}. \tag{2}$$

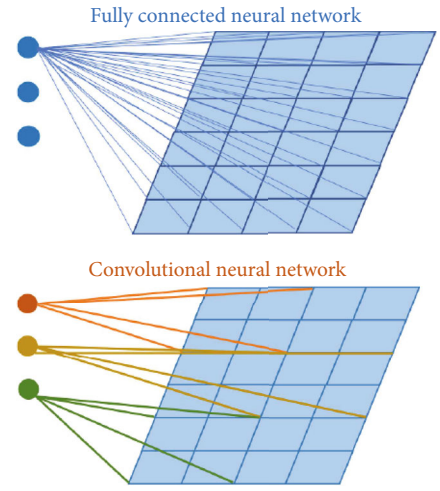


Figure 2: Schematic diagram of the architecture of the convolutional neural network.

The pixel matrix is a three-dimensional matrix. The length and width represent the size of the image, and the depth represents the number of color channels of the image. When the input picture is a black and white picture, the depth is 1; when the input picture is a color picture, the depth is 3.

$$\begin{aligned} F &= \frac{(\mathrm{SSE_r}\text{-}\mathrm{SSE_R})/\mathrm{k}}{SSE_R/(T-2k)} \sim F_{(k,T-2k)}, \\ \Delta f_r &= \left[\frac{\partial f_r}{\partial x_1}, \frac{\partial f_r}{\partial x_2}, \cdots, \frac{\partial f_r}{\partial x_n}\right]. \end{aligned} \tag{3}$$

Starting from the input layer, the three-dimensional matrix is transformed into another three-dimensional matrix through a different network structure until the final fully connected layer. The convolutional layer is the most important part of the convolutional neural network and the key to extracting features.

$$\begin{aligned} &f = f(x), \\ &x_j' \le x_j \le x_j'' \ (j = 1, 2, \cdots, n), \\ &g_j \le g_j'' \ (j = 1, 2, \cdots, n), \\ &h_j' \le h_j(x) \ (j = 1, 2, \cdots, n), \\ &w_j' \le w_j(x) \le w_j'' \ (j = 1, 2, \cdots, n). \end{aligned} \tag{4}$$

Generally, two-dimensional convolution is used in image recognition, that is, a discrete two-dimensional filter (also called a convolution kernel) and a two-dimensional image are convolved. The specific process is that the two-dimensional filter passes through all positions on the two-dimensional image, and do inner product with the pixel and its domain pixels at each position, and

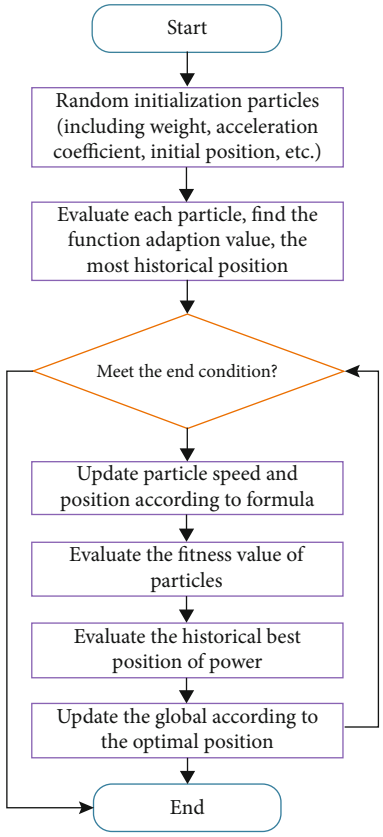


Figure 3: Schematic diagram of business English translation based on speech recognition.

different convolution kernels correspondingly extract different features.

$$x^{(j+1)} = x^{(j)} + s_j d^{(j)}. \tag{5}$$

Through the product of the above conditional probabilities, the joint probability of the word sequence is obtained. The above function can be further decomposed into two parts: feature mapping and calculating conditional probability distribution.

$$\begin{aligned} \left| f^{(j)} - f^{(j-1)} \right| &\le \tau, \\ \left| f^{(j)} - f^{(b)} \right| &\le \tau. \end{aligned} \tag{6}$$

In feature mapping, map each input word to a feature vector through the mapping matrix. C represents the feature vector corresponding to the Zth word in the dictionary, and Y72 represents the dimension of the feature vector. The process will be the vector obtained by feature mapping.

$$Q(x,q) = \frac{f}{f_0} + \sum_j p_x(x_j) + q\left[\sum_j p_g\left(g_j\right) + \sum_j p_h\left(h_j\right) + \sum_j p_w\left(w_j\right)\right]. \tag{7}$$

A previous work used RNN for temporal classification by combining RNN in a hybrid system. However, in addition to inheriting the shortcomings of HMM, this method does not make full use of the full potential of RNN for long-term sequence modeling. Therefore, it seems more feasible to directly train RNN for temporal classification tasks. CTC achieves this by allowing the network to perform label prediction at any point in the input sequence, as long as the overall order of the labels is correct. Since CTC does not care about the alignment of the target label with the input sequence and directly outputs the probability of the complete label sequence, this allows CTC to use the network as a timing classifier without data segmentation and external postprocessing [41].

A new method is proposed in this paper. The GAN translate network model, namely, CSGAN NMT model, is composed of two against models, one is the generation model G, which adopts the traditional attention-based neural network model, and where the source language sentence is translated into the target language; secondly, the discriminant model D, where the judgement sentence is translated by machine translation or people. Its design has two kinds of methods, namely, adopt the CNN or RNN design, and finally find that CNN's effect is better than LSTM, because a negative signal interference exists for LSTM training. In this paper, MLE is used to pretrain generator G, and then samples generated by G and real samples are used to pretrain D. When D reaches a certain accuracy rate, confrontational training will be entered. The part of GAN is basically the same as that of SeqGAN, which adopts the policy gradient method +MC search. Compared with these existing methods, our methods can do well in different reference standard environments though it may be overfitting. Compared with the existing methods, our proposal has three main advantages: firstly, it has a lower time consumed than others, which indicates that our proposal can translate as quickly as possible and in time. Secondly, our proposal has a higher accuracy than others, which shows that our proposal can perform well. Thirdly, our proposal can be adapt to any situations with English and business. However, the main limitation of our proposal is that it needs a huge computation space, which indicates that our proposal has a strong requirement for computation and it may not be easy to realize.

5. Case Study of Business English Translation

5.1. Data Case Source. The experiment in this section uses the deep learning model introduced in the paper to establish a business English speech translation model. In the training process of this experiment, the input of the deep learning model is input by frame, and the dimension of each frame of data is different according to the different feature parameters, and the corresponding output label of each frame of data is obtained after forward propagation. The dimension of the output layer is the same as the total number of labels, and the probability of each node output represents the probability of the corresponding label of the node in the frame of data. When the deep learning model is initially constructed, all weights will be initialized. After the input data enters the

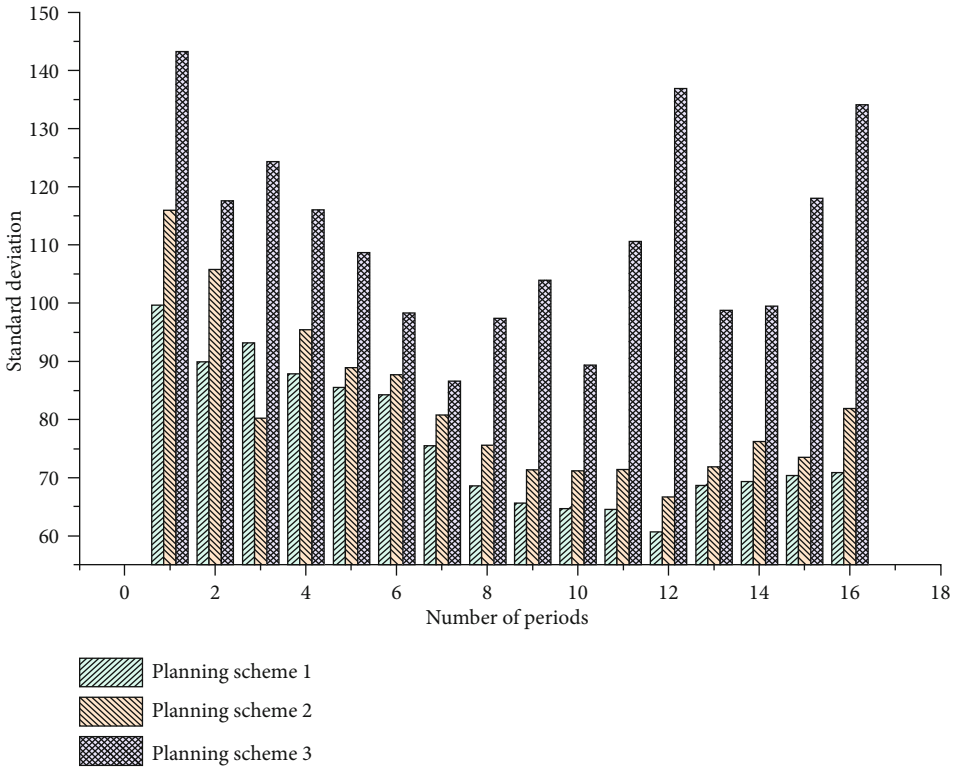


Figure 4: Schematic diagram of business English translation alignment label data distribution.

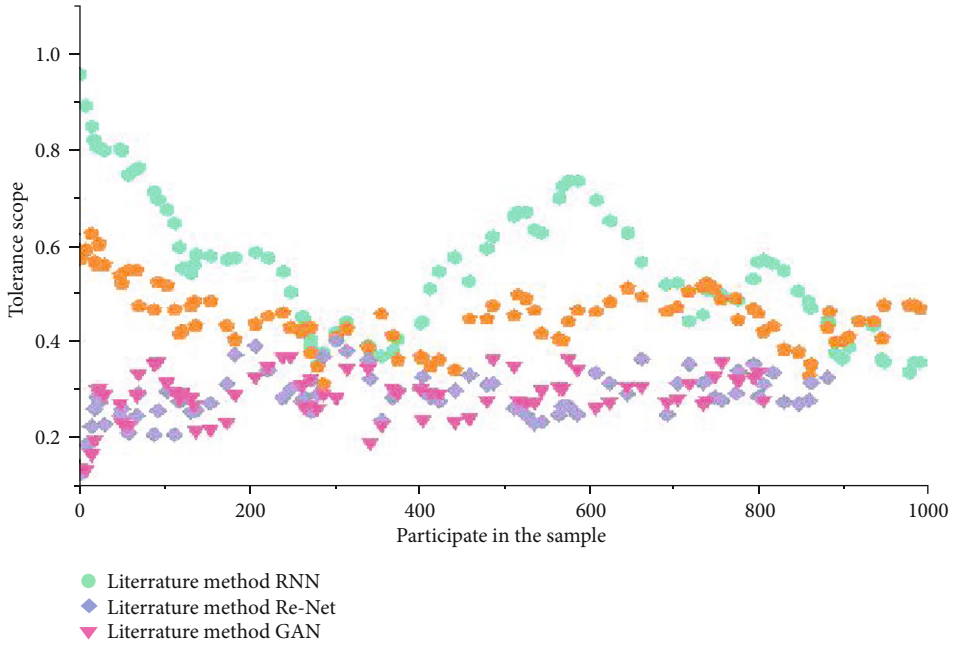


Figure 5: Comparison results of business English translation accuracy under different models.

network, it will be forward propagated according to the initial weight, and the output obtained is the predicted value.

The error between the predicted value and the true value is calculated through the loss function, and then, the time-based backpropagation algorithm is used to update the gradient in the direction where the loss function is falling, so that each weight and the parameters of each gate are updated. The weights are changed to complete a training process. The static optimization architecture implemented in this article and the static thermal path identification based on this optimization method have enabled the system to obtain a greater performance improvement. This chapter will specifically

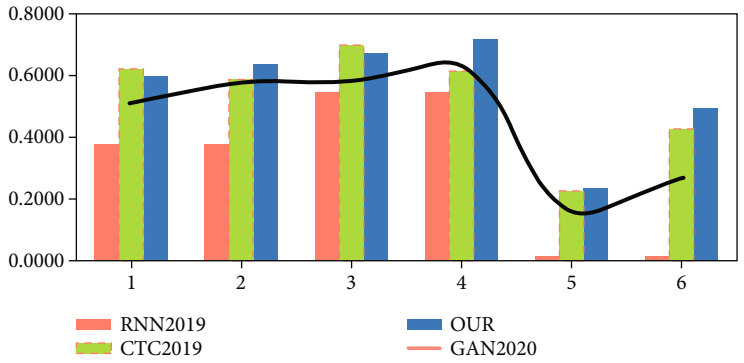


Figure 6: Comparison of accuracy of business English translation under different datasets.

introduce the resulting performance speedup and analyze the root cause of the performance improvement based on experimental data.

5.2. Application Effect Evaluation

5.2.1. Accuracy. Different feature parameters are extracted according to different experimental requirements. The feature parameters used in this experiment include MFCC, PLP, and FBANK features. For far-field voice data, we need to use microphone array technology to improve the quality of far-field voice. By beamforming the MDM data, the enhanced voice is obtained before subsequent processing. In order to obtain more accurate alignment label data, multiple GMM-HMM acoustic models are trained as described in the experiment, and label alignment is repeated, and finally, the alignment label data with higher accuracy is obtained. Figure 4 shows the distribution diagram of alignment label data for business English translation.

In planning scheme 1, the first alignment method to obtain the alignment label data is to averagely segment the training samples according to the number of sentence states and use the initial alignment label data to train a simple moon monophony model.

In planning scheme 2, after the acoustic model is obtained, the optimal path in the state diagram is selected according to the speech frame and the existing acoustic model, and each frame is matched to the state diagram to obtain the corresponding state of each frame.

In planning scheme 3, more accurate alignment label data was obtained.

According to this idea, we repeatedly train new models and repeatedly align them and finally get more accurate alignment label data. In order to prove the effectiveness of the deep neural network on the convergence speed of the model, we compared our model with existing methods. In order to ensure the reliability of the comparison, the same parameter settings are used for all models and run in the same environment. The error between the predicted value and the true value is calculated through the loss function, and then, the time-based backpropagation algorithm is used to update the gradient in the direction where the loss function is falling, so that each weight and the parameters of each gate are updated. The weights are changed to complete a training process. It can be seen from Figure 5 that the accuracy of different models in business English translation is still different. The accuracy of the validation set changes significantly. As the training progresses, the accuracy gradually increases and the range is more obvious. When the epoch is greater than 6 and continues to increase, the changes in the validation set and training tend to be flat and reach a stable state. Figure 5 shows the accuracy comparison of several models.

Table 1: Comparison table of results of performance conversions.

Real prediction	Positive	Negative
Positive	TP	FN
Negative	FP	TN

5.2.2. Time Consuming. If the epoch continues to increase, the accuracy changes on the training set and validation set are no longer obvious, which shows that there is still a certain overfitting. In order to ensure the reliability of the comparison, the model uses the same parameter settings and runs in different environments. Figure 6 shows the accuracy comparison under several environments and standards.

We use the confuse matrix and F1-score to evaluate the translate model performance; the matrix can be defined as follows:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{8}$$

For experimental results, the results of performance conversion are shown in Table 1.

$$\begin{aligned} \text{f1-score} &= \frac{2 * \text{precise} * \text{recall}}{\text{precise} + \text{recall}}, \\ \text{precise} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}. \end{aligned} \tag{9}$$

It can be seen from Figure 7 that the accuracy of the same model in business English translation under different environments and standards is still different. The reference

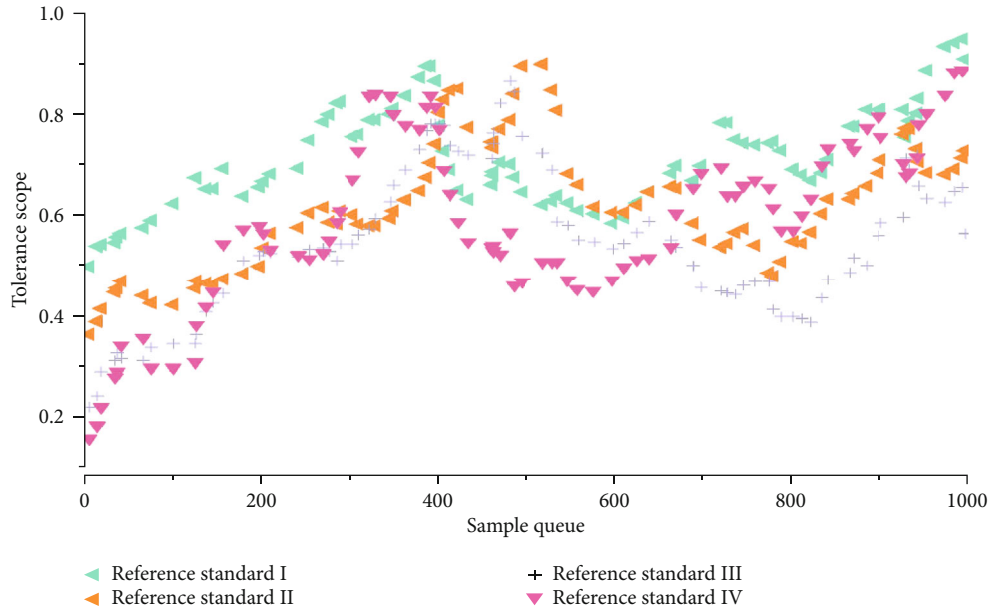


Figure 7: Comparison of accuracy of business English translation under different environments and standards.

standard is as follows: (1) conform to the expression habits of the target language, (2) when there are no words in the target language, keep the original words or create new words according to the meaning, (3) complete and accurate expression of the original information, without core semantic errors, and (4) the term is consistent with the common standards or practices of the industry and profession in the target language. Figure 6 shows the four methods results on different datasets. It indicated that our proposal obtained the best accuracy on 4 datasets of 6.This shows that there is still some overfitting. After the acoustic model is obtained, the optimal path in the state diagram is selected according to the speech frame and the existing acoustic model, and each frame is matched to the state diagram to obtain the corresponding state of each frame, thereby obtaining more accurate alignment label data. According to this idea, we repeatedly train new models and align them repeatedly and finally get more accurate alignment label data. During the initial construction of the deep learning model, all weights are initialized. After the input data enters the network, it will be forward propagated according to the initial weight, and the output obtained is the predicted value. The error between the predicted value and the true value is calculated through the loss function, and then, the time-based backpropagation algorithm is used to update the gradient in the direction where the loss function is falling, so that each weight and the parameters of each gate are updated. The weights are changed to complete a training process.

And we also compared three methods proposed in the last three years such as GNNE, RNNE, and CNNE with our proposal to investigate the effectiveness of our methods. The results can be shown in Figure 8:

As shown in Figure 8, the green line represents our proposal, and on all datasets, we can see that our proposal is better than others except D2, which all methods obtain the same

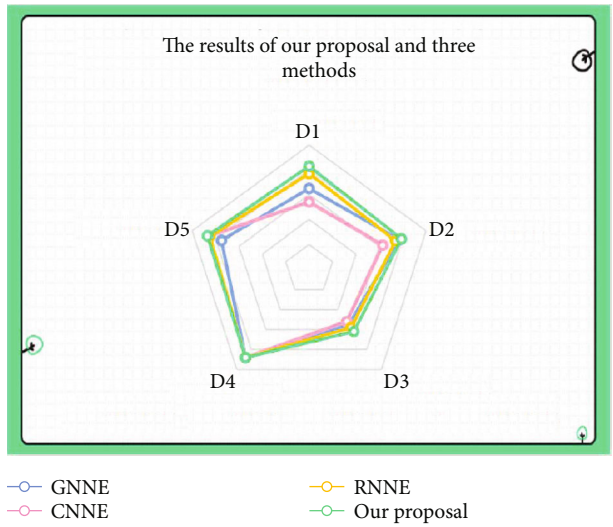


Figure 8: The results of our proposal and three methods on F1-score.

results. It indicated that our proposal can perform well than other three methods.

6. Conclusion

International business translation has received increasing attention as communication tools. In the foreseeable future, businessmen should be able to communicate directly with their foreign partners in English. Government officials and decision makers, who are not businessmen, should be able to review business documents in English. In short, business translation in the era of globalization is very necessary and important. In other words, marketing and investment, foreign insurance, international finance, publicity, etc. are all

covered in the scope of business English. Therefore, it can be seen that business English is a kind of practical language, a medium of business activities, and business-related English needs and contains various business activities. In essence, business English is a combination of business activities and English.

Therefore, business English can be defined as the English used in various business activities. Aiming at the shortcomings of current business English translation research, this paper designs and develops a business English translation architecture based on artificial intelligence speech recognition and edge computing. In this case, we can apply schema theory to the cultural background, the activation of the language environment, the content structure, and the activation of other structures to help translators merge the content they have learned and have a better understanding. We will also continue to devote ourselves to relevant research, hoping to provide a meager power for business English translation. Although our method has achieved good translation accuracy at present, it is still unable to achieve considerable accuracy in the face of complex speech environment, and the training time of the model is long, which cannot meet the purpose of real-time translation. In the future, we will further optimize the speech translation model to improve the training speed of the model while ensuring the accuracy.

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 conflicts of interest.

Acknowledgments

The research was supported by the Youth Project of University Humanities and Social Science Research of Jiangxi Province (No. JC19209) and the Project of 13th Five-Year Plan for Education and Science of Jiangxi Province (No. 19YB387).

References

[1] J. Zhang, L. Zhou, F. Zhou et al., "Computation-efficient offloading and trajectory scheduling for multi-UAV assisted mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2114–2125, 2020.

[2] K. Y. Zhan, J. H. Lewis, K. J. Vasil et al., "Cognitive functions in adults receiving cochlear implants: predictors of speech recognition and changes after implantation," *Otology & Neurotology*, vol. 41, no. 3, pp. e322–e329, 2020.

[3] X. Xiong, K. Zheng, L. Lei, and L. Hou, "Resource allocation based on deep reinforcement learning in IoT edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 6, pp. 1133–1146, 2020.

[4] T. Wang, Y. Liang, Y. Yang et al., "An intelligent edge-computing-based method to counter coupling problems in cyber-physical systems," *IEEE Network*, vol. 34, no. 3, pp. 16–22, 2020.

[5] P. Wagner, "Carnapian and Tarskian semantics," *Synthese*, vol. 194, no. 1, pp. 97–119, 2017.

[6] B. Trukeschitz, A. Hajji, J. Litschauer et al., "Translation, cultural adaptation and construct validity of the German version of the Adult Social Care Outcomes Toolkit for informal Carers (German ASCOT-Carer)," *Quality of Life Research*, vol. 30, no. 3, pp. 905–920, 2021.

[7] N. Tratnik, "Generalized cut method for computing the edge-Wiener index," *Discrete Applied Mathematics*, vol. 282, pp. 222–233, 2020.

[8] P. X. W. Zou, G. Zhang, and J. Wang, "Understanding the key risks in construction projects in China," *International Journal of Project Management*, vol. 25, no. 6, pp. 601–614, 2007.

[9] I. Speck, M. C. Ketterer, S. Arndt, A. Aschendorff, T. F. Jakob, and F. Hassepass, "Comparison of speech recognition and localization ability in single-sided deaf patients implanted with different cochlear implant electrode array designs," *Otology & Neurotology*, vol. 42, no. 1, pp. e22–e32, 2021.

[10] V. F. Uricchio, R. Giordano, and N. Lopez, "A fuzzy knowledge-based decision support system for groundwater pollution risk evaluation," *Journal of Environmental Management*, vol. 73, no. 3, pp. 189–197, 2004.

[11] Z. Song, "English speech recognition based on deep learning with multiple features," *Computing*, vol. 102, no. 3, pp. 663–682, 2020.

[12] S. P. Singh, A. Nayyar, R. Kumar, and A. Sharma, "Fog computing: from architecture to edge computing and big data processing," *Journal of Supercomputing*, vol. 75, no. 4, pp. 2070–2105, 2019.

[13] J. Shehorn, O. Strelcyk, and P. Zahorik, "Associations between speech recognition at high levels, the middle ear muscle reflex and noise exposure in individuals with normal audiograms," *Hearing Research*, vol. 392, article 107982, 2020.

[14] R. de Vries, F. Haas, and working group for revision of the Dutch Blood Transfusion Guideline 2011, "English translation of the Dutch Blood Transfusion Guideline 2011," *Clinical Chemistry*, vol. 58, no. 8, pp. 1266-1267, 2012.

[15] D. Raynaud, S. Gessner, and B. Mota, "Andalò di Negro's De compositione astrolabii: a critical edition with English translation and notes," *Archive for History of Exact Sciences*, vol. 73, no. 6, pp. 551–617, 2019.

[16] Y. Qu and H. Liu, "The basic problem of contemporary water resources environment research: translation and analysis from the English practice report on water resources," *Journal of Coastal Research*, vol. 104, 2020.

[17] Y. R. Oh, K. Park, H. B. Jeon, and J. G. Park, "Automatic proficiency assessment of Korean speech read aloud by non-natives using bidirectional LSTM-based speech recognition," *Etri Journal*, vol. 42, no. 5, pp. 761–772, 2020.

[18] A. Mehrabi, M. Siekkinen, and A. Yla-Jaaski, "Edge computing assisted adaptive mobile video streaming," *IEEE Transactions on Mobile Computing*, vol. 18, no. 4, pp. 787–800, 2019.

[19] Z. Liu, "A study on English translation of tourism publicity in coastal cities from the perspective of cross-cultural communication," *Journal of Coastal Research*, vol. 115, no. sp1, 2020.

[20] H. Liao, Z. Zhou, W. Kong et al., "Learning-based intent-aware task offloading for air-ground integrated vehicular edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–13, 2020.

[21] H. Li, H. Xu, C. Zhou, X. Lu, and Z. Han, "Joint optimization strategy of computation offloading and resource allocation in

multi-access edge computing environment," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 10214–10226, 2020.

[22] M. Lee, J. Lee, and J. H. Chang, "Ensemble of jointly trained deep neural network-based acoustic models for reverberant speech recognition," *Digital Signal Processing*, vol. 85, pp. 1–9, 2019.

[23] C. Lee, S. Hong, S. Hong, and T. Kim, "Performance analysis of local exit for distributed deep neural networks over cloud and edge computing," *Etri Journal*, vol. 42, no. 5, pp. 658–668, 2020.

[24] A. M. Kubiak, J. Rennies, S. D. Ewert, and B. Kollmeier, "Prediction of individual speech recognition performance in complex listening conditions," *Journal of the Acoustical Society of America*, vol. 147, no. 3, pp. 1379–1391, 2020.

[25] D. Kim, H. Lee, H. Song, N. Choi, and Y. Yi, "Economics of fog computing: interplay among infrastructure and service providers, users, and edge resource owners," *IEEE Transactions on Mobile Computing*, vol. 19, no. 11, pp. 2609–2622, 2020.

[26] Y. H. Jung, S. K. Hong, H. S. Wang et al., "Speech recognition: flexible piezoelectric acoustic sensors and machine learning for speech processing (Adv. Mater. 35/2020)," *Advanced Materials*, vol. 32, no. 35, article 2070259, 2020.

[27] J. M. Hung, X. Li, J. Wu, and M. F. Chang, "Challenges and trends indeveloping nonvolatile memory-enabled computing chips for intelligent edge devices," *IEEE Transactions on Electron Devices*, vol. 67, no. 4, pp. 1444–1453, 2020.

[28] D. Hülsmeier, A. Warzybok, B. Kollmeier, and M. R. Schädler, "Simulations with FADE of the effect of impaired hearing on speech recognition performance cast doubt on the role of spectral resolution," *Hearing Research*, vol. 395, article 107995, 2020.

[29] Y. Huang, X. Song, F. Ye, Y. Yang, and X. Li, "Fair and efficient caching algorithms and strategies for peer data sharing in pervasive edge computing environments," *IEEE Transactions on Mobile Computing*, vol. 19, no. 4, pp. 852–864, 2020.

[30] X. Huang, R. Yu, S. Xie, and Y. Zhang, "Task-container matching game for computation offloading in vehicular edge computing and networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–14, 2020.

[31] B. Hu and J. Li, "An edge computing framework for powertrain control system optimization of intelligent and connected vehicles based on curiosity-driven deep reinforcement learning," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 8, pp. 7652–7661, 2021.

[32] M. P. Hosseini, T. X. Tran, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Multimodal data analysis of epileptic EEG and rs-fMRI via deep learning and edge computing," *Artificial Intelligence in Medicine*, vol. 104, article 101813, 2020.

[33] H. Xing, L. Liu, J. Xu, and A. Nallanathan, "Joint task assignment and resource allocation for D2D-enabled mobile-edge computing," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4193–4207, 2019.

[34] J. T. Holder, N. C. Dwyer, and R. H. Gifford, "Duration of processor use per day is significantly correlated with speech recognition abilities in adults with cochlear implants," *Otology & Neurotology*, vol. 41, no. 2, pp. e227–e231, 2020.

[35] R. Du, Y. Liu, L. Liu, and W. Du, "A lightweight heterogeneous network clustering algorithm based on edge computing for 5G," *Wireless Networks*, vol. 26, no. 3, pp. 1631–1641, 2020.

[36] J. Dornhoffer, Y. F. Liu, E. E. Zhao, E. L. Camposeo, T. A. Meyer, and T. R. Mcrackan, "Factors that influence second-side cochlear implant speech recognition outcomes," *Otology & Neurotology*, vol. 42, 2021.

[37] S. Colby and B. Mcmurray, "The role of listening effort during degraded speech recognition: a comparison of the dual-task and pupillometry paradigms," *Journal of the Acoustical Society of America*, vol. 146, no. 4, pp. 3050-3051, 2019.

[38] D. Liu, J. Xu, P. Zhang, and Y. Yan, "A unified system for multilingual speech recognition and language identification," *Speech Communication*, vol. 127, pp. 17–28, 2021.

[39] S. Yang, Z. Gong, K. Ye, Y. Wei, Z. Huang, and Z. Huang, "EdgeRNN: a compact speech recognition network with spatio-temporal features for edge computing," *IEEE Access*, vol. 8, pp. 81468–81478, 2020.

[40] J. Zhang and D. Tao, "Empowering things with intelligence: a survey of the progress, challenges, and opportunities in artificial intelligence of things," *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 7789–7817, 2021.

[41] M. S. Hossain and G. Muhammad, "An audio-visual emotion recognition system using deep learning fusion for a cognitive wireless framework," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 62–68, 2019.

[42] A. Boukerche and V. Soto, "Computation offloading and retrieval for vehicular edge computing," *ACM Computing Surveys*, vol. 53, no. 4, pp. 1–35, 2020.

[43] K. Borowiak, C. Maguinness, and K. V. Kriegstein, "Dorsal-movement and ventral-form regions are functionally connected during visual-speech recognition," *Human Brain Mapping*, vol. 41, no. 4, pp. 952–972, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6623251, 10 pages
https://doi.org/10.1155/2021/6623251

Research Article

A Blockchain-Based Vehicle Condition Recording System for Second-Hand Vehicle Market

You-Ting Jiang and Hung-Min Sun

Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan

Correspondence should be addressed to Hung-Min Sun; hmsun@cs.nthu.edu.tw

Received 26 October 2020; Revised 20 January 2021; Accepted 13 May 2021; Published 29 May 2021

Academic Editor: Laurie Cuthbert

Copyright © 2021 You-Ting Jiang and Hung-Min 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.

Due to the cheap price and decent performance of used cars, the second-hand vehicle market is quite active in Taiwan. However, the vehicle information in the market of second-hand vehicles is centralized at present. In other words, the vehicle information is still provided by the second-hand vehicle dealers, which are not transparent and have no tampering cost. As the result, the authenticity, accuracy, and fairness of these second-hand vehicle data are worth discussing. To solve the above problems, this paper proposed a trusted vehicle data source system based on blockchain technology. The system is built on the Ethereum platform with the private chain structure. With the help of our platform, a trusted third party (e.g., maintenance plant and government branch) can record vehicle information into Blockchain so the integrity of vehicle information can be maintained. Customers can easily query the relevant vehicle information through our system interface and avoid receiving fake vehicle information. This system is a Proof of Concept (PoC), and the feasibility is examined by evaluating transactions per second (TPS) of the proposed system.

1. Introduction

The second-hand vehicle market is quite active in Taiwan. The reason is that although the second-hand vehicles are relatively old, they still maintain good performance due to the technological advancement of the automobile industry. Coupled with their relatively inexpensive price, many people will still consider used vehicle when purchasing a car. According to statistics from the Ministry of Transportation and Communications, Taiwan, there are about 600,000 transactions of vehicles a year in Taiwan's second-hand vehicle market [1], which is 1.5 times more than the new vehicle market. However, in the second-hand vehicle market, the most common situation which customers encounter is that they do not know whether the information of vehicle condition provided by the second-hand vehicle dealer is consistent with the actual condition and result in buying vehicles with problems (e.g., odometer fraud and accident vehicle). Although there are some vehicle information providers like Carfax [2], the situation of data fraud and tampering is still very serious.

Blockchain technology is a novel way to manage data. Due to its characteristics, a group of people (nodes) can work together to store, manage, and protect the blockchain without anyone trusting anyone, thereby eliminating the possible adverse consequences of centralized data control by third party. Moreover, each node would replicate the whole blockchain data through peer-to-peer networks, and so the information can hardly be destroyed. Therefore, we believe that blockchain technology can be applied to record the condition of second-hand vehicles.

This paper uses the blockchain architecture provided by Ethereum [3] to record the information of vehicle condition. By the properties of transparency and tamper proof of blockchain, customers can use the system proposed in this paper to check the actual vehicle condition without being worried misinformed by second-hand vehicle dealers.

2. Background

2.1. Introduce to Blockchain. The first successful blockchain application is Bitcoin [4]. In Bitcoin, blockchain acts as a publicly verifiable open ledger, which has the following properties:

(1) Transparency: The data on the blockchain is public and can be read by anyone; so, the record and flow of each transaction can be easily queried

(2) Decentralization: Bitcoin does not depend on additional third party to control the transactions. Each node in the Bitcoin network realizes verification, transmission, and management of transactions

(3) Tamper proof: Each block will contain the hash value of the previous block. Therefore, if someone tries to modify the data in the block, it will be immediately detected and rejected by other nodes, and so the blockchain cannot be tampered with

2.2. Smart Contract and Decentralized Application. The evolution of blockchain mechanism begins with trading of cryptocurrency, then gradually expands to the application of smart contract [5], and then enters the field of decentralized application. Below, we will introduce the concept of smart contract and decentralized application which are used to construct our vehicle condition recording system.

(1) Smart contract: The smart contract featured in Ethereum is written by a programming language called Solidity [6]. Before deploying, Solidity codes need to compile into Ethereum bytecode, being added into a transaction and send it to the network. Afterward, miners of Ethereum will verify the transaction and record it in a block to complete the deployment of a smart contract. Users of Ethereum can call a function of the smart contract by sending a transaction. When miners receive the transaction, they will verify it and record it in a block and run the bytecode of the smart contract in the Ethereum virtual machine (EVM) to complete the contract call. (Figure 1)

(2) Decentralized Application (DApp): DApp is defined as an application running on a Peer-To-Peer (P2P) network. In this network, every computer has the same status and same data and can execute the code of DApp, which giving DApp, following advantages

  (i) DApp running in the network will not be interrupted as long as there is at least one computer running in the P2P network

  (ii) When DApp is deployed in the blockchain network, the data stored in the block cannot be tampered with, even those who wrote and deployed DApp

In this paper, our vehicle condition recording system belongs to nonfinancial DApp, because the system does not involve the trading of vehicles.

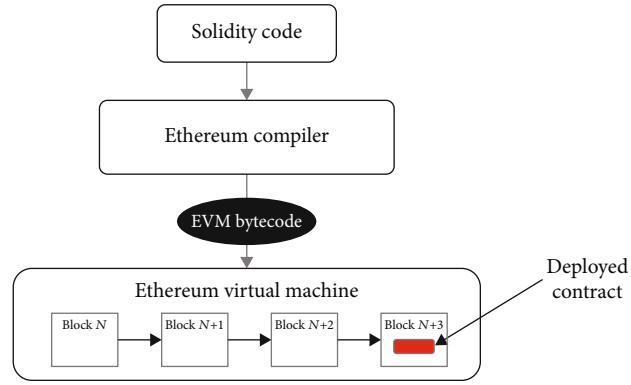


Figure 1: Ethereum smart contract execution flow.

2.3. Consensus Mechanism in Blockchain. Consensus is a mechanism to ensure that all the transactions occurring on the blockchain are genuine. It is often achieved by the contribution of other participants in the chain network. Below are three common types of consensus.

2.3.1. Proof of Work (PoW) [7]. PoW is the consensus mechanism used by Bitcoin and some other popular cryptocurrency networks. It requires participant nodes to solve a compute-intensive problem to receive the right to add new transactions to the blockchain. The participants who try to solve the problem are commonly called "miners" since they can get Bitcoins by verifying the blocks. However, this whole mining mechanism of Bitcoin needs high energy consumption and longer processing time.

2.3.2. Proof of Stake (PoS) [8]. PoS is another common consensus algorithm that serves as a low-energy consuming alternative to the PoW algorithm. Its algorithm is simple, and the node that has more cryptocurrency has more chance to verify transactions. The hypothesis behind this mechanism is that the node will have more motivation to maintain the genuineness of the system if they have more property on the chain. Since PoS does not include solving compute-intensive problems, it can significantly reduce the computing power than PoW. However, the scalability of transaction verification is still not improved in PoS.

2.3.3. Proof of Authority (PoA) [9]. PoA is a consensus algorithm proposed in 2017 by Ethereum cofounder and former CTO Gavin Wood. It introduces a practical solution for scale-up blockchain networks. Its concept is similar to PoS, but instead of using the number of cryptocurrencies, PoA decides the right of verifying transactions by nodes' reputation. In other words, transactions are verified by nodes that are selected as trustworthy entities. Since there are only a limited number of block validators, it makes PoA a highly scalable system. The PoA model enables companies to maintain their privacy while availing the benefits of blockchain technology. Microsoft Azure is an example where the PoA is being implemented.

2.4. Categories of Blockchain. At the present time, two types of blockchain are considered base on their publicity:

2.4.1. Public Chains. Public chains are open to anyone in the world and can be accessed, sent, received, and authenticated by anyone. It is a blockchain in which everyone can participate in a consensus process; thus, public chains are usually considered to be completely decentralized. However, the transaction on public chain is slow to update and iterate, and the initial set-up cost is very high if a developer wants to build his own chain with the current technology development framework [10]. If enterprises adopt public chain directly, they will be limited by the scaling problem, and the enterprise demand cannot be met.

2.4.2. Private Chains. A private chain is a completely private blockchain. Its write access is restricted to one organization. Access rights are either open to the public or restricted to a certain extent. The entire network is jointly maintained by the member organizations, and access to the network is generally through the member organizations' gateway nodes. With the consensus process controlled by preselected nodes, this type of blockchain is considered to be partially decentralized. However, at the expense of partial decentralization, private chains have characteristic of fast transaction, privacy-protecting, and having very low transaction costs and low set-up costs [11]. At the same time, it retains blockchain's tamper-evident nature.

3. Related Work

There are lots of DApps that have been developed since the emergence of blockchain [12]. Some of them deal with financial transaction issues, such as decentralized exchange [13]. Some deal with the issue of personal data ownership on the cloud storage platform [14, 15]. Other DApps include supply chain application [16, 17], insurance application [18], Internet of Things (IoT) with blockchain [19, 20], game application [21], and secondary radio spectrum trade [22]. Our system focuses on data recording, and this type of application is similar to the application of product biography or supply chain.

In [16], the authors propose a framework for blockchain-based supply chain quality management. This framework consists of blockchain, smart contracts, and various IoT sensors. Blockchain provides a safe distributed ledger with various quality information, assets information, logistic information, and transaction information. Smart contracts bring privacy protection, automation, and intelligence into this system, while IoT sensors gather various data from the real world. By using the framework, each piece of information in the production process of the product will be recorded on blockchain, including temperature, humidity, positioning, and other information in the process of product transportation. Record all information in the production process of the product in blockchain can help users quickly understand the real information of the product by means of the blockchain's tamper-proof characteristic. This way of creating a biography for a product can also be used in recording vehicle conditions.

Another research [23] is a big data-sharing system for the car industry in China based on blockchain. They have set up an automotive industry alliance in China and published a white paper, which mainly aims at establishing a blockchain platform for recording vehicle information. At the end of 2018, they officially released this blockchain platform called Engine Chain (EGCC), hoping to record the information of the whole life cycle of a car through this blockchain platform. This research is similar to what our system wants to do, but our system is slightly different from EGCC. First of all, to query car information or execute other operations on EGCC, users need to spend EGCC coin. On the other hand, our system would not charge for querying data. The second difference is the group that has record permission. EGCC allows all registered users to record data, but our system only allows trusted third party, which avoid malicious users trying to tamper data. The third difference is the consensus algorithm used. EGCC uses validated DPoS (VDPoS) [24], which is an algorithm proposed in the Insight Chain [24]. It is a modified PoS algorithm which can improve speed. Table 1 is a simple comparison between our system and EGCC.

Therefore, in this paper, we use blockchain technology to achieve the purpose that vehicle condition information cannot be tampered with, so as to help customers to have authentic information of vehicles for reference when choosing second-hand vehicles.

4. System Design

4.1. Design Goal. Our system is designed according to the following three main goals.

(1) Transparency of vehicle condition information: Our system is based on the PoA consensus of Ethereum so that the condition information of vehicles can be stored on the private blockchain of Ethereum by specific users, and the information recorded on Blockchain is transparent to everyone in this private network

(2) Access control of smart contract and tamper proof of blockchain information: Because our system must rely on the help of some specific users (e.g., government branch and maintenance plant) to record or update the information of vehicle condition to change the state of smart contract, we must only allow these specific users to update the state of smart contract; so, we use the conditional branch in a smart contract to judge whether the transaction sender has the permission to change the state or not, so as to achieve the access control of the smart contract

(3) User friendly of operation: We also design a web user interface (UI) for the system. Each user can use the web page to interact with the contract function to record or update the information of vehicle condition without having any background knowledge of smart contract. We design different web UI for users depending on each user's identity to record or read different information on the contract

4.2. System Model. There are many roles in the proposed system (Figure 2), and the duty of each different role will be explained in the following paragraphs.

Table 1: Comparison between the proposed system and EGCC.

	The proposed system	EGCC
Blockchain platform	Ethereum private chain	Engine Chain
Query fee	Free for query	Purchase for unlocking information
Record permission	Trusted third party	Registered user
Consensus algorithm	PoA	VDPoS

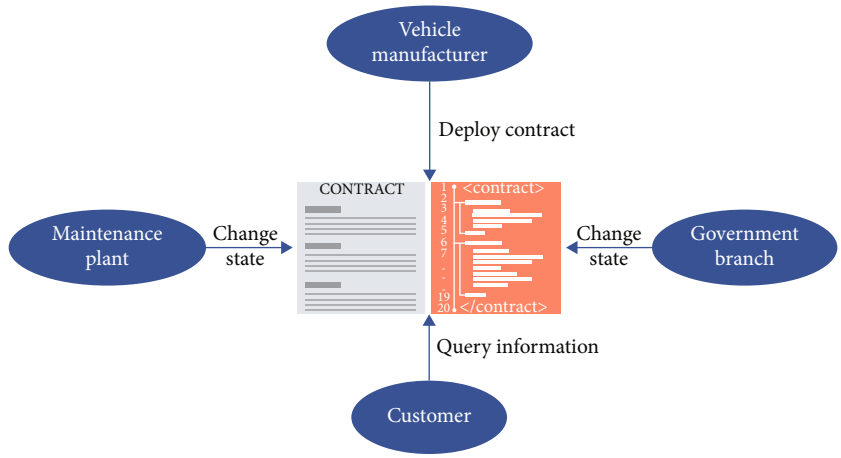


Figure 2: System model.

(1) Vehicle manufacturer: For vehicle manufacturers, the only function provided is the deployment of new contracts to the blockchain. After a manufacturer completes a car build, they can use the proposed system to deploy a smart contract that contains initial information about the vehicle. Initial information includes car model, year of release, mileage, and engine capacity.

(2) Maintenance plant: The maintenance plant can use the functions of the proposed system to record the maintenance history of vehicles. A maintenance record includes two types of information, one is the maintenance time, and another is the maintenance description. For instance, the maintenance plant can record the mileage of the vehicle to prevent customers from suffering from odometer fraud. In addition, before recording the information to blockchain, the owner ID of the vehicle is needed. Since the owner ID is a secret from the vehicle owner, this approach ensures that the maintenance plant can only record vehicle information with the vehicle owner's consent

(3) Government branch: The government branch can use the proposed system to record information. First, the Motor Vehicles Office, the government branch that is in charge of vehicles in Taiwan, can update the ownership history when the owner of a vehicle is changed. This information can make customers go through the ownership history of a vehicle and bring it into assessment. Motor Vehicles Office also has the permission to maintain the mapping table which maps the Vehicle Identification Number (VIN) to its Owner ID and only they can maintain this table. VIN is like an ID card number of a vehicle; thus, each vehicle has a different VIN. Second, when dealing with a traffic accident, the traffic police can record the time of the accident by our system, and this allows customers to understand the accident history of a vehicle and prevents them from buying an accident car. Third, when a vehicle is scrapped, the contract which represents the vehicle needs to be destructed by triggering the self-destruct function in the contract. Only Motor Vehicles Office has the permission to trigger the function

(4) Customer: Customers can use our system to query about the information of vehicles, including initial information, accident history, and maintenance records. Customers can simply enter the VIN of the vehicle and the token acquired from a second-hand vehicle dealer on the web page to query the information of the vehicle. Besides, when a general user wants to query the information of his vehicle, he can just input the VIN and owner ID to get his vehicle information rather than input a token. Customers do not need to have an account to query about the information of vehicles, as long as the customer inputs the correct token or owner ID, and he has the permission to read the information from blockchain. Moreover, they do not have the permission to change the state of smart contracts

(5) Second-hand vehicle dealer: The second-hand vehicle dealer in the proposed system do not have the permission to record or update the information of vehicles. They are only responsible for generating a

query token for the customer who wants to query the vehicle information. After inputting this token by the customer, our system will verify the token. If it is valid, then the system will return the vehicle information

(6) Authority node: As mentioned before, our system is based on the PoA consensus of Ethereum blockchain called Clique. Different from the PoW consensus, there has no miner in the PoA network. The accounting right of blocks is handled by predefined authority nodes. By using the PoA consensus Clique, transactions can be verified and accounted into the new block by the authority nodes without costing any fee

4.3. Life Cycle of Vehicle. In our design, the vehicle manufacturer is responsible for pushing the initial information of each vehicle to each contract, including car model, released year, mileage, and engine capacity. After the sale of the vehicle, the government branch (Motor Vehicles Office) needs to update the ownership history and change the owner ID of the vehicle on the contract. When using the vehicle, every using data that includes accident history, maintenance history, and mileage will be recorded or updated in the contract by the government branch and maintenance plant, respectively. When the owner of the vehicle sells the car to a second-hand vehicle dealer, Motor Vehicles Office also needs to update the ownership history. Customers can use our system to query the information of each vehicle by a valid token at any time in the life cycle of each vehicle until the vehicle is scrapped. When the vehicle is scrapped, the contract that represents the vehicle is also destructed by the self-destruct function of the contract (Figure 3).

5. System Implementation

In this chapter, we will focus on the details of the proposed system, including a full description of each function and the user interface of the system. In the first section, we explain the structure of the system and what software and tools are used. In the second section, we describe the functions of the proposed system in detail.

5.1. System Structure. The proposed system is performed on a private blockchain environment using Go-Ethereum (Geth) [25] client implementation. Clique, a PoA Geth implementation, is used as the back-end blockchain operating system, and Solidity is used as the programming language to write the smart contract. The system structure includes the following two parts:

(1) Private chain setup: Puppeth [26] is used to set up the configuration of the private chain. After the setup, Puppeth will generate a file that includes information of the configuration such as block time, gas limit, and authority node setup. Geth will later use this file to build up the private chain environment

In the configuration, we create different nodes for different identities of participants in the proposed system including vehicle manufacture, maintenance plant, and government branch for account management. In addition to the above nodes, three authority nodes are set up for system simulation.

(2) User interface (UI): The UI of the proposed system is a web page. The server side is constructed by the http-server suite provided by Node.js [27]. Web3.js [28] is used to link the front-end UI to the back-end smart contract. After the setup, private chain and account information can be connected

The system structure is shown in the following figure (Figure 4).

5.2. System Operation

(1) Log-in process: Since there are different roles of users in the proposed system, each user must choose which role to log in. Every user's account is connected to the accounts in Geth nodes. Different roles lead to connect to different Geth nodes. Because of the serial number of the account list in the Geth node, users can just input the serial number as the account number instead of inputting the account address (Figure 5)

(2) Operation of customer: Because of avoiding anyone that can query the vehicle information without any permission, the customer needs to input a token which is generated by second-hand vehicle dealer for token authentication. After passing the authentication, the customer can get the information of the vehicle. For the authentication purpose, we keep an array in smart contract to record these valid tokens for authentication. The customer needs to input the VIN of the vehicle and a token generated by a second-hand vehicle dealer to query its information. We use a smart contract to keep a mapping table to store each VIN and corresponding contract address; thus, the customers do not need to input complicated contract address string. Our system provides information query function in smart contract which can return car model, released year, engine capacity, mileage, ownership history, accident history, maintenance history, and so on

(3) Operation of vehicle nanufacturer: In the proposed system, the vehicle manufacturer is responsible for deploying a new contract that represents a vehicle. After logging in the system, the vehicle manufacturer can input the initial information of the vehicle and deploy the contract (Figure 6). After the deployment, the vehicle manufacturer will update the mapping table which stores each VIN and corresponding contract address. In the proposed system, only they can deploy a contract and update the mapping table

(4) Operation of maintenance plant: The maintenance plant in our system is responsible for recording maintenance history and mileage of vehicles. Before recording the information, our system will require that the maintenance plant should input the owner ID of the vehicle at first. If the owner ID is correct,

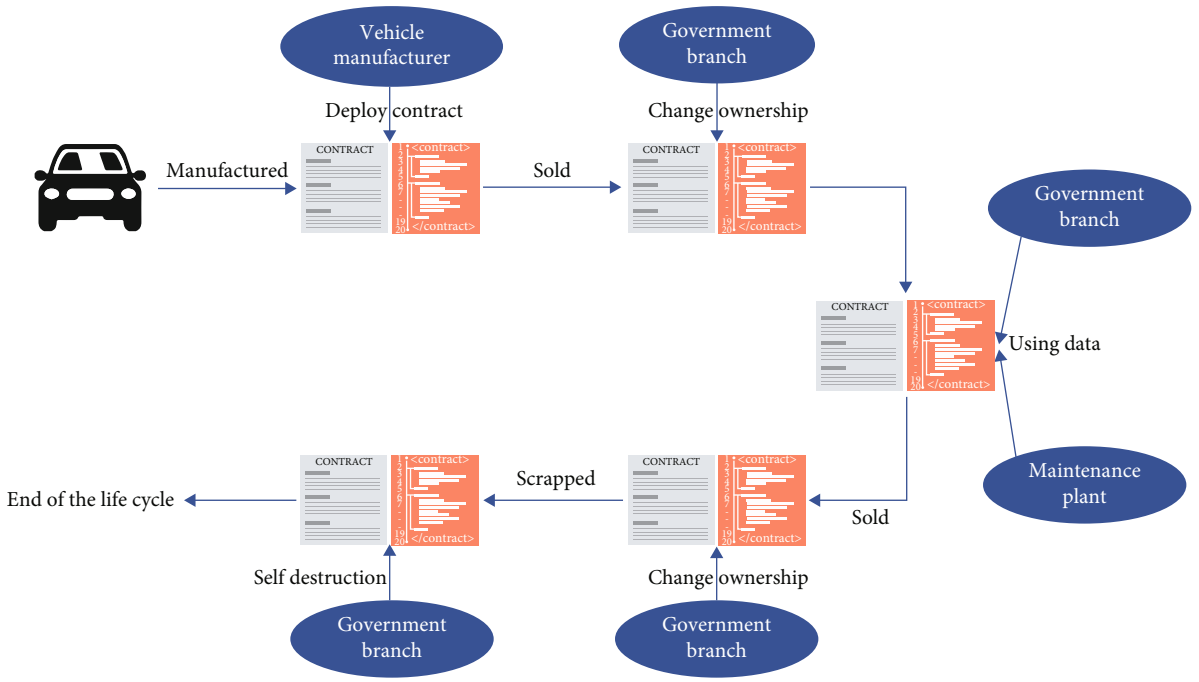


Figure 3: The life cycle of a vehicle in our system.

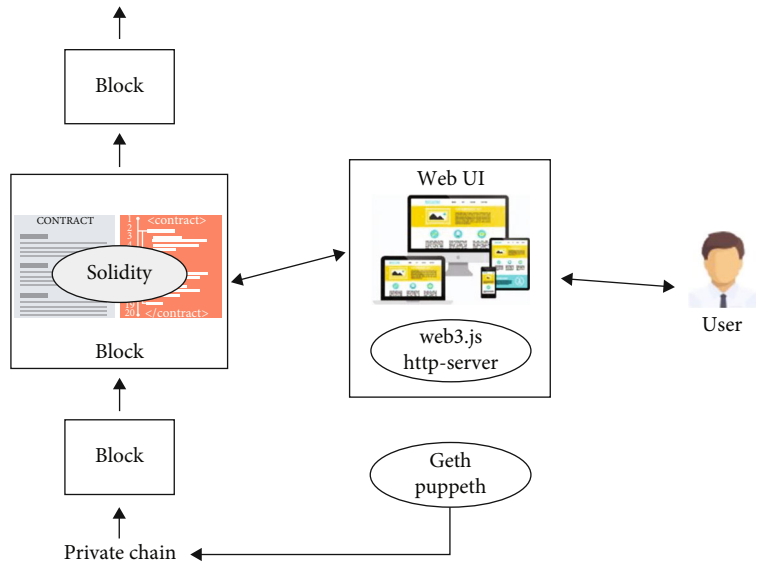


Figure 4: The overview of the proposed system structure.

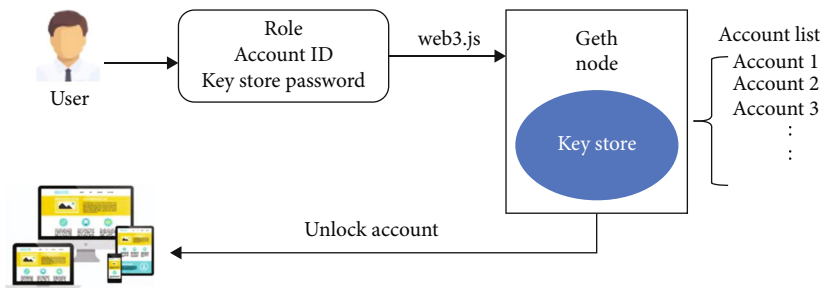


Figure 5: The log-in process of the proposed system.

```
var car = carContract.new(
        /* input of constructor */
        {
                from: web3.eth.accounts[accountNum],
                data: 'byte code of the contract',
                gas: '4700000'
        }, function (e, contract){
                console.log(e, contract);
                if (typeof contract.address !== 'undefined') {
                    alert("success! address: " + contract.address.toString);
                    document.getElementById("address").innerText = contract.address.toString();
                }
            }
        )
```

Figure 6: The code of deploying a contract.

then they can start recording the vehicle information. We use the mapping table *VIN-to-OwnerID* to help us for authentication.

When recording the information, the program in smart contract will check whether the role of the transaction sender is the maintenance plant that exists in the account list of the node. If the identity of the transaction sender is authentic, the program will append the time and the description of maintenance history to the Struct array *maintainHistories* and update the mileage.

(5) Operation of government branch: The operation of the government branch is similar to maintenance plant. For the ownership history, the program in smart contract which is triggered by Motor Vehicles Office will update the ownership number when the owner of the vehicle is changed. Motor Vehicles Office is also responsible for maintaining the mapping table *VIN-to-OwnerID*, and when the owner of a vehicle is changed, only Motor Vehicles Office can update the mapping table. We use SHA256 to store the hash value of owner ID in the mapping table in smart contract rather than store the plaintext of owner ID.

For the accident history, the program in smart contract which is triggered by police will append the time of the accident to the String array *accidentHistory*. Finally, when the vehicle is scrapped, Motor Vehicles Office needs to trigger the self-destruct function of the contract, and we use the instruction *selfdestruct()* in Solidity to destruct a smart contract. Before all of the above operation, the program in smart contract still first checks whether the role of the transaction sender has the permission to change the state of the contract or not.

(6) Operation of second-Hand vehicle dealer: The second-hand vehicle dealer does not have the permission to record or update the information of vehicles. They are only responsible for generating a query token for the customer who wants to query the vehicle information. For the token generation, our system concatenates the password of the second-hand vehicle dealer and current time as the input string and output the hash value (token). Then, the dealer will write this token into an array in the smart contract

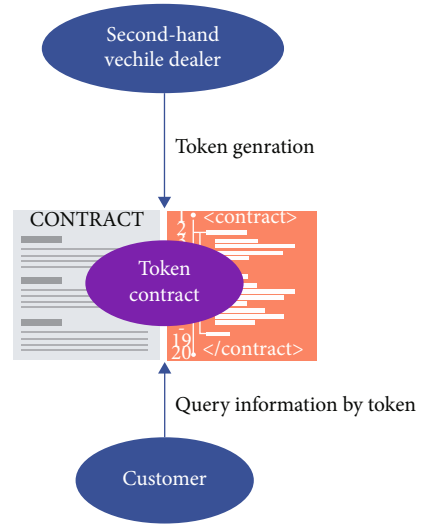


Figure 7: The operation of second-hand vehicle dealer.

Every token generated by a second-hand vehicle dealer has its lifetime. The proposed system will delete it by sending a transaction to delete the token in three minutes after its generation. This ensures that every token is timeliness and difficult to reuse (Figure 7).

6. System Evaluation

6.1. Gas Limit and Transaction Throughput. The analysis is performed on a PoA private Blockchain development based on the Ethereum platform, using Go-Ethereum (Geth) client implementation. Since there is no fee for sending transactions in a private Blockchain environment, we focus on the throughput of transactions. The proposed system is feasible only if the throughput of transactions is greater than the daily transactions.

Ethereum platform prevents transaction spamming and rewards block miners by charging a gas fee on transactions. There are three important concepts in the Ethereum gas mechanism, gas, gas price, and gas limit.

(1) Gas: Each bytecode instruction in the EVM execution process requires a corresponding amount of gas, and the more complex the logic, the more gas is required.

Table 2: The TPS of each function execution using block time = 5 seconds.

Function description	Transaction cost	TPS	Transactions per day
Deploy the contract	2,738,400	0.58	50,482
Record the accident time	63,522	25.18	2,176,253
Record the maintenance history	88,448	18.08	1,562,952
Update the mileage	27,897	57.35	4,955,371
Update the ownership history	41,814	38.26	3,306,069
Average (without deploy contract)	55,420	34.71	3,000,161

Table 3: The TPS of each function execution using block time = 10 seconds.

Function description	Transaction cost	TPS	Transactions per day
Deploy the contract	2,738,400	0.29	25,241
Record the accident time	63,522	12.59	1,088,126
Record the maintenance history	88,448	9.04	781,476
Update the mileage	27,897	28.67	2,477,685
Update the ownership history	41,814	19.13	1,653,034
Average (without deploy contract)	55,420	17.36	1,500,080

Table 4: The TPS of each function execution using block time = 10 seconds.

Function description	Transaction cost	TPS	Transactions per day
Deploy the contract	2,738,400	0.19	16,827
Record the accident time	63,522	8.39	725,417
Record the maintenance history	88,448	6.03	520,984
Update the mileage	27,897	19.12	1,651,790
Update the ownership history	41,814	12.75	1,102,023
Average (without deploy contract)	55,420	11.57	1,000,054

The table of gas spend to corresponding bytecode is defined in Ethereum yellow paper [3]

(2) Gas price: Gas price means how much you are willing to pay for each gas, in unit Gwei, which equals to 10^{-9} Ether, the basic unit of coin in Ethereum

(3) Gas limit: The gas limit is the maximum amount of gas you are willing to buy for the transaction. The reason for the gas limit is that no one knows how much gas will be used in the execution of a smart contract until it is actually executed, and so the gas limit is meant to cap the fee for each executed smart contract

Transactions per second (TPS) is a common criterion to evaluate how fast can a transaction be verified on the blockchain. The theoretical maximum TPS can be calculated using the following equation [29] (Eq. (1)), where *GasLimit* is the maximum number of units of gas user willing to spend on a transaction. *TxGas* is the gas needed to compute the simplest transaction, and *BlockTime* is the time required to create the next block in a chain.

$$\mathrm{TPS}_{\max} = \frac{\mathrm{GasLimit}}{Tx\mathrm{Gas} \times \mathrm{BlockTime}}. \quad (1)$$

6.2. Evaluations. The gas limit and the block time could be set as a configuration parameter. We set the gas limit to 8,000,000 gas by default configuration in the proposed system, which is the same as the gas limit of Ethereum main network [3].

We test the throughput using different block times, 5 s, 10s, and 15 s, respectively. For the gas cost of each transaction, we use Remix [30] to calculate the gas cost of function execution in the system. Remix is a web browser integrated development environment (IDE) for developers in developing Solidity smart contract and DApp. The TPS is calculated using Eq. (1). Transactions per day are then calculated by TPS multiplying seconds within a day. Tables 2–4 report the TPS of each function execution using different block times in the proposed system. For display conciseness, the TPS column is rounded to two decimal places, while the transactions per day column were computed using unrounded TPS.

According to Tables 2–4, it can be observed that the acceptable number of transactions per day of deploying contract is lower than the others since it needs higher gas cost. Besides the deploy contract function, the TPS of other functions are good enough. Therefore, we only focus on the TPS of deploying contract, since lower block times were found to be strongly related to the higher number of lost blocks per valid block [29], which leads to a bad performance of the Blockchain. For avoiding this problem, we set the block

time to 15 seconds by default configuration in our system. Assume that we have 600,000 vehicles (yearly second-hand vehicle transaction amounts in Taiwan) to deploy to the blockchain at the beginning, the required time of deploying these contracts is about 35 days, and it is acceptable.

7. Limit and Challenges

This paper's implementation relies on the existing platform (Ethereum). Thus, the functionality is restricted by the functions their system provides. In addition, this paper's method is a proof of concept; thus, it might still have problems when it runs online, such as how to choose a proper trusted third party to record the chain. Also, PoS is not a fully centralize consensus algorithm; thus, it is not as decentralized as Bitcoin which uses PoW. However, it still could avoid second-hand traders hold car information by themselves. So, it is still could solve this problem.

8. Conclusions

This paper builds a trusted vehicle data source system based on blockchain technology. With the help of trusted third parties to record vehicle information to our platform instantly, customers no longer need to worry about the problem of getting untrustworthy vehicle condition information from second-hand vehicle dealers. Moreover, because of the property of blockchain, we can ensure that the information of vehicle condition in the system is transparent to everyone and cannot be tampered with.

Vehicle manufacturers can use the system to deploy a smart contract that represents a vehicle. The maintenance plant and government branch can record or update the vehicle condition information by using the proposed system. Customers can easily query the relevant vehicle information through the system interface. All of the above operations can be done without costing any fee.

In the evaluation of the proposed system, the result shows that the transaction throughput is enough for daily transactions.

Data Availability

The code implementation used to support the findings of this study were supplied by Hung-Min Sun under license and so cannot be made freely available. Requests for access to these data should be made to Hung-Min Sun: hmsun@cs.nthu.edu.tw.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This study was funded by the Ministry of Science and Technology, Taiwan [grant numbers MOST 106-2221-E-007-026-MY3, MOST 107-2221-E-007-015-MY3, and MOST 108-2218-E-001-001].

References

[1] "Ministry of Transportation and Communications, Taiwan. Commonly Used Transportation Statistics," 2020, https://stat.motc.gov.tw/mocdb/stmain.jsp?sys=100&funid=emenu.

[2] "Carfax," 2020, https://www.carfax.com/.

[3] G. Wood, "Ethereum: a secure decentralised generalized transaction ledger," *Ethereum project yellow paper*, vol. 151, pp. 1–32, 2014.

[4] S. Nakamoto and A. Bitcoin, "A Peer-to-Peer Electronic Cash System," 2008.

[5] "Smart contract," 2020, https://ethereum.org/en/developers/docs/smart-contracts/.

[6] "Solidity," 2020, https://docs.soliditylang.org/en/v0.7.4/.

[7] M. Jakobsson and A. Juels, "Proofs of work and bread pudding protocols(extended abstract)," in *Secure information networks*, pp. 258–272, Springer, Boston, MA, 1999.

[8] J. Kang, Z. Xiong, D. Niyato, P. Wang, D. Ye, and D. I. Kim, "Incentivizing consensus propagation in proof-of-stake based consortium blockchain networks," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 157–160, 2019.

[9] "Proof of authority," 2020, https://github.com/PoAnetwork/wiki/wiki/PoA-Network-Whitepaper/.

[10] D. Guegan, *Public Blockchain Versus Private Blockhain*, 2017.

[11] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An overview of blockchain technology: architecture, consensus, and future trends," in *2017 IEEE International Congress on Big Data (BigData Congress)*, pp. 557–564, Boston, MA, USA, 2017.

[12] "State of the DApps," 2020, https://www.stateofthedapps.com/zh/stats/.

[13] "Idex - decentralized ethereum asset exchange," 2020, https://idex.market/eth/idex/.

[14] N. Kaaniche and M. Laurent, "A blockchain-based data usage auditing architecture with enhanced privacy and availability," in *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)*, pp. 1–5, Cambridge, MA, USA, 2017.

[15] G. Zyskind and O. Nathan, "Decentralizing privacy: using blockchain to protect personal data," in *2015 IEEE Security and Privacy Workshops*, pp. 180–184, San Jose, CA, USA, 2015.

[16] S. Chen, R. Shi, Z. Ren, J. Yan, Y. Shi, and J. Zhang, "A blockchain-based supply chain quality management framework," in *2017 IEEE 14th International Conference on e-Business Engineering (ICEBE)*, pp. 172–176, Shanghai, China, 2017.

[17] A. Iftekhar, X. Cui, M. Hassan, and W. Afzal, "Application of blockchain and Internet of Things to ensure tamper-proof data availability for food safety," *Journal of Food Quality*, vol. 2020, Article ID 5385207, 14 pages, 2020.

[18] I. Nath, "Data exchange platform to fight insurance fraud on blockchain," in *IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pp. 821–825, Barcelona, Spain, 2016.

[19] P. Urien, "Blockchain iot (biot): a new direction for solving internet of things security and trust issues," in *2018 3rd Cloudification of the Internet of Things (CIoT)*, pp. 1–4, Paris, France, 2018.

[20] Y. Wu, L. Song, L. Liu, J. Li, X. Li, and L. Zhou, "Consensus mechanism of IoT based on blockchain technology," *Shock and Vibration*, vol. 2020, Article ID 8846429, 9 pages, 2020.

[21] "Etheremon," 2020, https://www.etheremon.com/.

[22] M. A. Khan, M. M. Jamali, T. Maksymyuk, and J. Gazda, "A blockchain token-based trading model for secondary spectrum markets in future generation mobile networks," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 7975393, 12 pages, 2020.

[23] "Enginechain white paper," 2020, https://white-paper.oss-cn-beijing.aliyuncs.com/v.3.0.0/Engine%20V.3.0.0%20-%20%20English%20Version.pdf/.

[24] TEAM, "Insight Chain Founding," *Insight Chain Technical White Paper V1. 0*, 2019.

[25] "Geth," 2020, https://github.com/ethereum/go-ethereum/wiki/.

[26] "Puppeth," 2020, https://github.com/puppeth/.

[27] "Node.js," 2020, https://nodejs.org/en/.

[28] "Web3.js," 2020, https://web3js.readthedocs.io/en/1.0/.

[29] "On ethereum performance evaluation using PoA," 2020, https://blog.coinfabrik.com/on-ethereum-performance-evaluation-using-PoA/.

[30] "Remix," 2020, https://remix.ethereum.org/#optimize=false&evmVersion=null/.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5590754, 13 pages
https://doi.org/10.1155/2021/5590754

Research Article

Medical Knowledge Graph in Chinese Using Deep Semantic Mobile Computation Based on IoT and WoT

Wanheng Liu,[1] Ling Yin,[2] Cong Wang,[1] Fulin Liu,[3] and Zhiyu Ni[3]

[1]*Beijing University of Posts and Telecommunications, China*
[2]*National Population Health Data Center, China*
[3]*Affliated Hospital of Hebei University, China*

Correspondence should be addressed to Ling Yin; lwhlcf@163.com

Received 4 February 2021; Accepted 13 May 2021; Published 28 May 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Wanheng Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a novel medical knowledge graph in Chinese approach applied in smart healthcare based on IoT and WoT is presented, using deep neural networks combined with self-attention to generate medical knowledge graph to make it more convenient for performing disease diagnosis and providing treatment advisement. Although great success has been made in the medical knowledge graph in recent studies, the issue of comprehensive medical knowledge graph in Chinese appropriate for telemedicine or mobile devices have been ignored. In our study, it is a working theory which is based on semantic mobile computing and deep learning. When several experiments have been carried out, it is demonstrated that it has better performance in generating various types of medical knowledge graph in Chinese, which is similar to that of the state-of-the-art. Also, it works well in the accuracy and comprehensive, which is much higher and highly consisted with the predictions of the theoretical model. It proves to be inspiring and encouraging that our work involving studies of medical knowledge graph in Chinese, which can stimulate the smart healthcare development.

1. Introduction

In recent years, healthcare was attached much importance in people's daily life, and semantic mobile computing applied in healthcare has become a hot topic and received much attention from researchers and academia from all over the world recently. To make it more convenient and easy for both doctors and patients, a comprehensive medical knowledge graph in Chinese using the deep learning method on a mobile device based on IoT and WoT is of great significance. Also, it can serve for smart healthcare which can be applied in remote consultation platform and mobile devices in hospital to implement disease classification and treatment recommendations. It is more time-saving, economic, and efficient.

It is well known that technological revolutions have hit the industrial world and semantic mobile computing in Internet of Things (IoT) and Web of Things (WoT) has emerged from intuitive, which can make it more convenient for people to get the information more faster, accurately, and intelligently from mass amount of Internet information. The knowledge graph was firstly proposed by Google in 2012 [1], which is the knowledge base that Google utilized to enhance its search engine. There are mainly three eras in the knowledge graph, including the ontology era, semantic web, and knowledge graph era. In previous studies, knowledge graphs in the ontology era were constructed manually by experts in relevant fields, including Cyc [2] in 1984, WordNet [3] proposed by Miller G A. in 1985, and Chinese dictionary HowNet. When the World Wide Web was set up in 1989, it resolves the previous restriction that the query tool could only look up information step by step in a specific path. To address the limitations in application, the Semantic Web [4] era occurred when Linked Data [5] was presented by Berners-Lee T. in 2006, and DBpedia [6, 7] was the representative one. Then in 2001, Wikipedia was launched to provide a free encyclopedia for people, for example, DBpedia and Yago [8]. Then, Freebase [9] was built in 2007, which can be manually and actively imported. However, to conduct effective search and analysis of knowledge, the knowledge graph era emerged. The typical examples are Knowledge

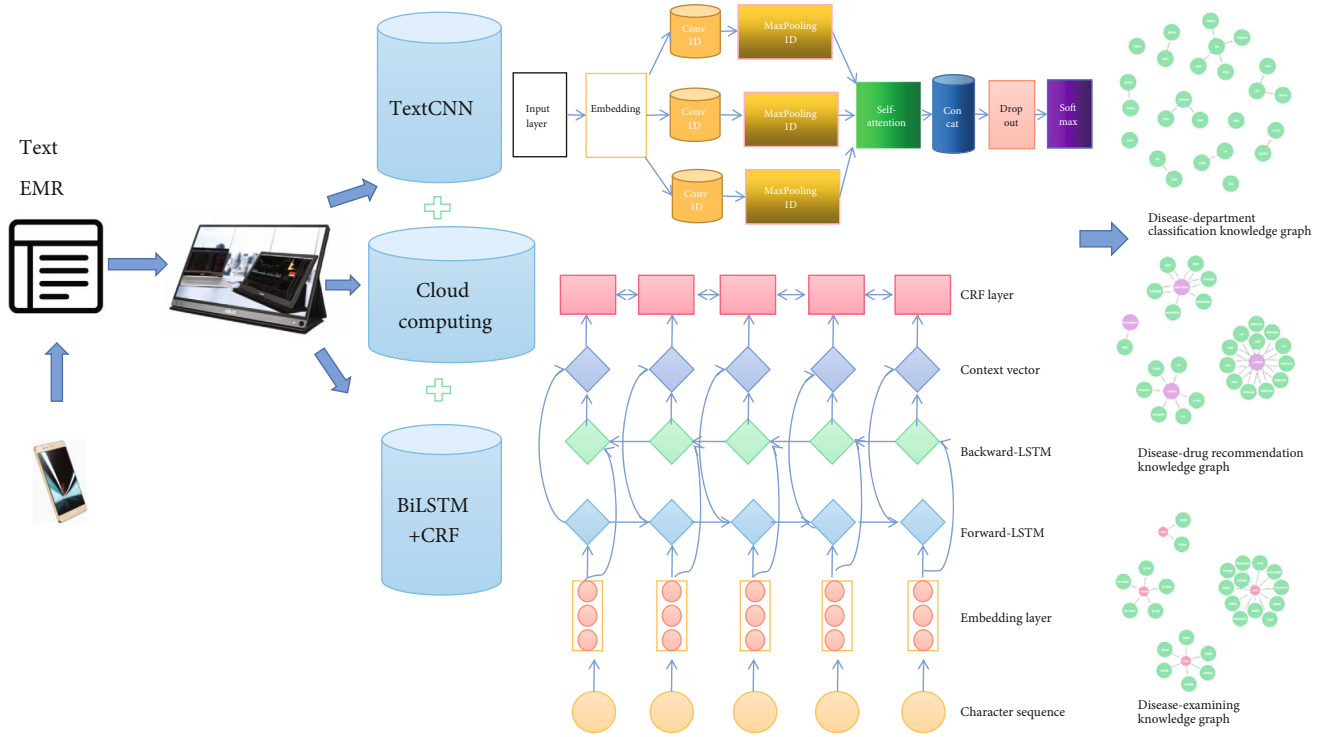


Figure 1: The architecture of our medical knowledge graph in the Chinese system.

Vault [10], Google Knowledge graph, Facebook graph search, and Baidu Knowledge.

In recent years, as the maturity of technology and more emphasis on healthcare, it has become more and more popular to apply knowledge graph in the medical field and has attracted much attention from researchers in computer and medical to combine these two fields [11–26]. Rotmensch et al. [11] in 2017 proposed learning a health knowledge graph from electronic medical records by using probabilistic. Longxiang et al. [12] presented a framework that can retrieve knowledge in a knowledge graph automatically in 2017. Wen et al. [14] proposed an attention-aware path-based relation extraction for medical knowledge graph. Sun et al. [21] in 2018 incorporate description embeddings into medical knowledge graphs representation learning. Kai et al. [22] presented a novel semantic similarity measure in biomedical knowledge graphs in 2018. Wang et al. [23] based on knowledge graph design and implement personal health record systems using semantic integration methods in 2018. Wang et al. [25] made use of a knowledge graph to process medical archives in 2019. In addition, Song et al. [13] put forward the Chinese pediatric medical knowledge graph in 2019. Ying et al. [15] proposed a system for drug-drug similarity measure based on a knowledge graph in 2019. Xiao et al. [26] in 2019 based on medical knowledge graph to enhance the fraud, waste, and abuse detection on claim data using the deep learning-based method to extract the entities and relationships from the knowledge sources. Li et al. [27] used an imperfect knowledge graph to classify rare diseases in 2019. Chen et al. [16] in 2020 proposed robustly extracting medical knowledge from EHRs, in which nonlinear functions are adopted in building the causal graph to better understand exiting model assumptions. Zheng et al. [17] put forward a dedicated knowledge graph benchmark for biomedical data mining in 2020. Also, medical knowledge graph can be applied in specific departments, for example, Xie et al. [24] in 2018 presented a data-driven traditional Chinese medicine knowledge discovery method based on a knowledge graph. Li et al. [18] came up constructing a knowledge graph of knee osteoarthritis in Chinese using an automatic approach in 2020. Chai et al. [19] in 2020 proposed the diagnosis method of thyroid disease combining knowledge graph and deep learning. Alawad et al. [20] utilize medical knowledge graph combined with a deep learning method for cancer phenotyping in 2021. Great progress has been made in the knowledge graph applied in the medical field above; however, it still suffers from problems as follows: (1) owning to data lacking and inadequate, it is hard for the model of knowledge graph to achieve the best performance. (2) On account of limitations of the medical field, the tools for building knowledge graph in the healthcare industry are immature and domain knowledge to be integrated in the same medical system has some limitations. (3) There are still some difficulties in updating medical knowledge. (4) Due to the most medical knowledge graph focusing on a single disease, various disease classifications and treatment recommendations have been overlooked. (5) Because of complex medical knowledge, it is hard to find the balance in the process of the knowledge graph structure.

To overcome the limitations above, a novel Chinese healthcare knowledge graph using deep learning method which can be utilized in the mobile device based on IoT and WoT is proposed in this paper. In order to address the lack of data, 600 thousand pieces of data were processed in our work to support the training of the model for good

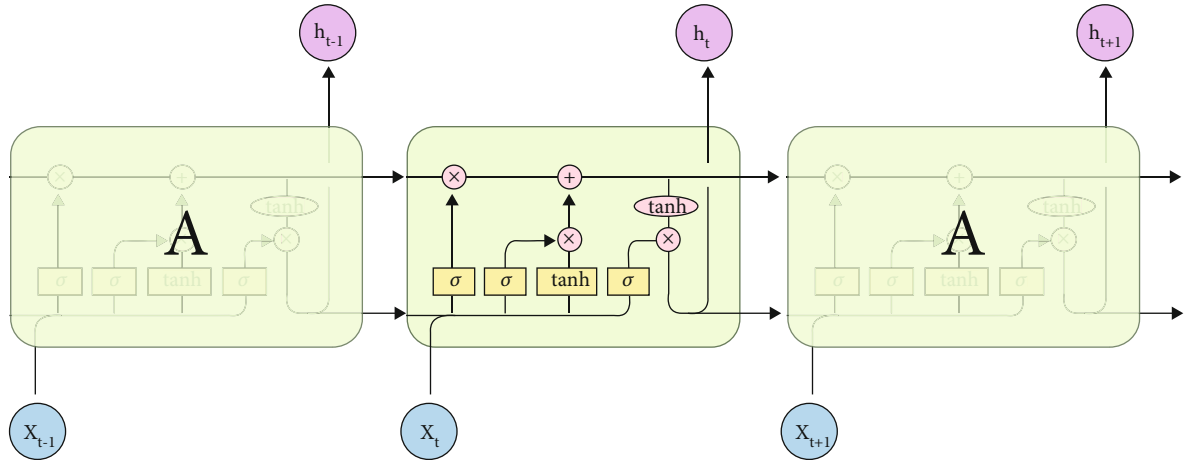


Figure 2: The structure of LSTM.

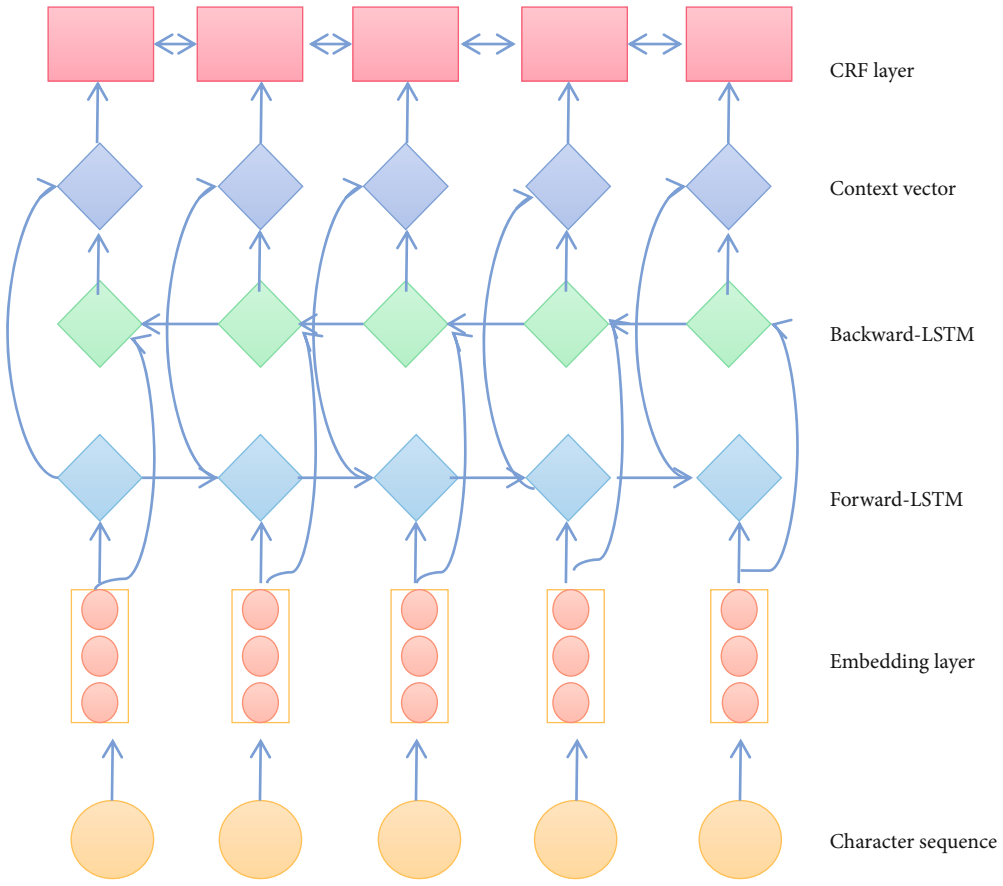


Figure 3: The architecture of BiLSTM+CRF.

performance achieved. Furthermore, the deep learning method instead of the traditional method is adopted in our study to realize an intelligent medical knowledge graph. And the self-attention is used in our deep learning network to reduce the amount of computation cost. Besides, cloud computing is also utilized to make it in practice which is convenient for both doctors and patients through a mobile device based on IoT and WoT. New data would be computed and stored on the cloud, in which text and image computing, processing, and storage capabilities are involved. It counts that not only can the diseases be classified but also treatment recommendations can be made in the intelligent knowledge graph we presented. What is more, our medical knowledge system can be applied in practice in remote consultation platform served for both patients and doctors. And the subject is concerned chiefly with the study of a knowledge graph. The contributions in our study are the following:

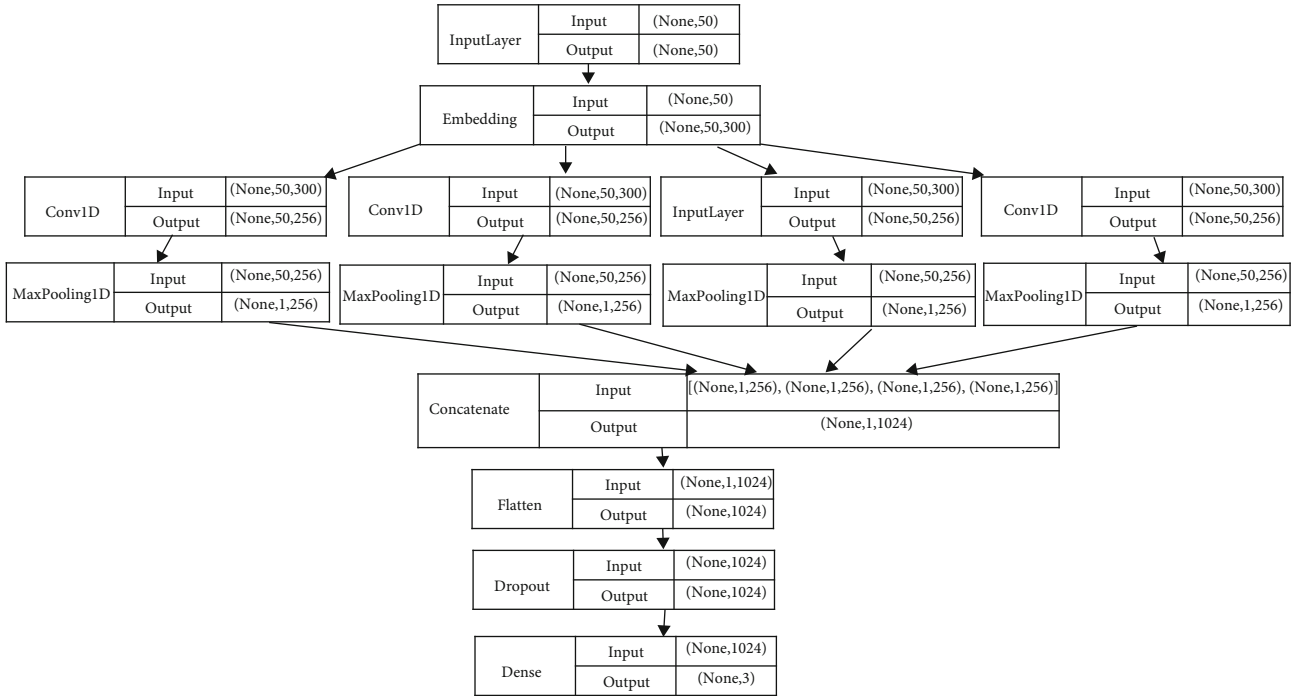


Figure 4: The architecture of Text-CNN.

(i) A novel and more comprehensive medical knowledge graph in Chinese based on deep learning method is proposed, which can be served for doctors and patients

(ii) The self-attention is utilized instead of the traditional attention mechanism in our deep learning network, which is more simple and the computational cost is relatively low

(iii) A new dataset is labelled and processed to train and obtain a well model

(iv) To make it can be in practice based on IoT and WoT, the cloud computing is combined to support our study

In this paper, it is divided into 4 sections as follows: in Section 2, it introduces the materials and methods in which the system and network architecture contained. And in Section 3, we show the results and discussion of our work in this part. Finally, it is concluded in this paper in Section 4.

2. Materials and Methods

The architecture of the medical knowledge graph in Chinese based on deep neural networks applied in telemedicine as can been seen in Figure 1. The cloud computing is adopted to process and store the data, which makes it in practice based on IoT and WoT when using a mobile device. When new data is acquired, it can be computed by the cloud, and then, new data would be stored on the cloud to improve the medical knowledge graph, which is more efficient and convenient. Firstly, text information which can be recognized and

Table 1: The structure of TextCNN.

Embedding		
Conv(3)	Conv(4)	Conv(5)
Bn	Bn	Bn
Relu	Relu	Relu
Conv(3)	Conv(4)	Conv(5)
Bn	Bn	Bn
Relu	Relu	Relu
Max-pool	Max-pool	Max-pool
Concat		
Dropout		
Softmax		

EMR on the mobile device of the patients is input to the mobile device of the doctor. When the EMR and text information can be recognized is obtained, it would be transmitted to our model to compute, which BiLSTM, CRF, TextCNN, and self-attention contained to accomplish the feature representation and sequence labeling. What is more, our model is based on 600 thousand of training data. Therefore, it has better performance in accuracy and comprehensive of our medical knowledge graph generated. It can not only perform department classification according to the disease and also can recommend food and drug for the patients. In addition, it can advise the patients to examine the major project of the disease to reduce the time and economy consuming.

2.1. Network Architecture. The neural network we adopted in our study is BiLSTM (Bidirectional Long Short-Term

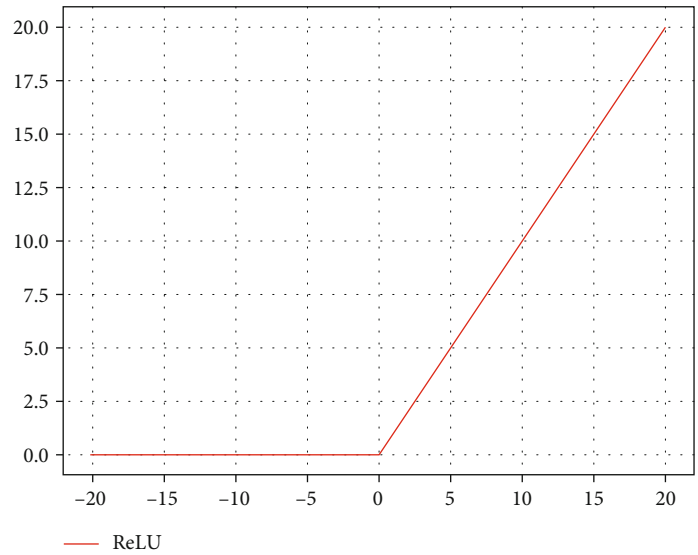


Figure 5: Representation of Relu Activation Function.

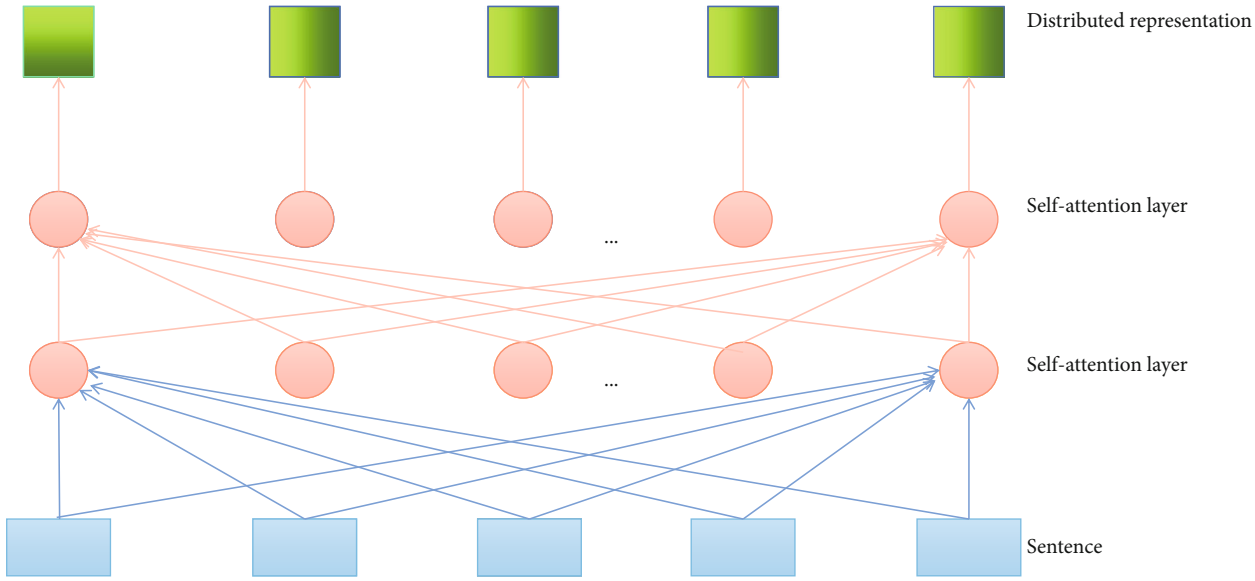


Figure 6: Self-attention.

Memory), which was consisted of the forward LSTM and the backward LSTM, and it is often used to model contextual information in natural language processing tasks. It is a fact that the LSTM is a special type of RNNs. And RNN is a special neural network structure made up of input layer, hidden layer, and output layer three layers, which is proposed according to the view that human cognition is based on past experience and memory. Not only the input from the previous moment is taken into account but also the memory function of the network for the previous content is also added in RNN. It works well for time series problems; however, it is can result in a small memory value when costing too much time. Therefore, the LSTM is proposed to address the problem of gradient vanishing and gradient explosion in long sequence training, which is the extension of the traditional model and has a better performance in a long sequence. As shown in Figure 2, it can be seen that in LSTM proposed by Sepp Hochreiter in 1997 [28], which is not a single layer of neural networks, but there are four layers and a lot of gates added that interact in a specific way. In LSTM, the first step is to decide what information should be discarded from the cellular state, which was made by the sigmoid in the forget-gate layer. Then, to decide what new information ought to be stored in the cell state, where the sigmoid layer in the input gate determines which values to be updated and the tanh layer creates a new candidate value vector to be added to the state. And a status will be updated when the above steps are combined, and the old cell state will be updated to a new cell state. Finally, it needs to decide what to output, in which the sigmoid layer determines to output which parts

of the cell state. In conclusion, under the control of the gate, the state vector of an LSTM unit and the output can be obtained. Based on LSTM, in which the state vector of each cell is input to the cell at the next moment and the process is unidirectional, BiLSTM can be considered as two layers of LSTM. It is found that the state inputs of the two layers are in opposite directions, which are forward and backward propagation, respectively. In BiLSTM, it has better performance and better capture bidirectional semantic dependencies which can solve the problem of unable to encode information from back to front in LSTM.

2.2. BiLSTM+CRF. CRF (Conditional Random Field) [29] is often utilized to achieve better performance in sequence labeling tasks, which the whole sequence can be taken into account. In BiLSTM, due to no conditional constraint of label transitions, a completely wrong sequence of annotations may be output. Therefore, in our architecture, BiLSTM was combined with CRF is adopted in our architecture. Due to the independent of each other in the output of the softmax layer of the LSTM, it cannot affect each other in output when the contextual information learned by BiLSTM. Nevertheless, the label output of the maximum probablity value can be picked at each step. And the transition feature function is contained in CRF which the context of the label can be learned. Then CRF is used to do the output layer of BiLSTM when the order between output labels considered. It can regard the output of BiLSTM as the probability of each tag's value, and the function of CRF as the results adjusted according to the relationship between labels. And the architecture of BiLSTM combined with CRF is shown in Figure 3, which strengthens the relevance of information between texts and the optimal sequence result can be obtained through the probability of the feature to the tag.

2.3. TextCNN. The TextCNN is added in our model to extract local features of the current word. On account of the capacity of the model, some important information may discard when long sentences are learned by BiLSTM. Also, for the word segmentation task, the label of the current word is basically only associated with the previous several words. As shown in Figure 4, when a sample of some sentences input and needs to be classified, the convolution operation will be conducted by CNN. Each word in the sentence is composed of n -dimensional word sentence, so the size of the input matrix is $m \times n$, where m is the sentence length. It is counted that for text data, the filter moves down no longer swipes horizontally. Then the max-pooling operation of each vector would be carried out and the pooling value be stitched. Finally, the feature representation of the sentence is obtained and the sentence vector is classified by the classifier.

There are four parts in Text-CNN, embedding layer, convolution layer, pooling layer, and fully connected layer. In the embedding layer, the words can be mapped into corresponding vectors. And the main purpose of this layer is to encode the natural language input into distributed representation. In the convolution layer, the one-dimensional convolution operation is performed for the word mapped features. And it is mainly used to extract different N-gram features through convolution. The input statement or text can be converted to a two-dimensional matrix when passed through the embedding layer. It is noted that in TextCNN, various types of kennels should be utilized at the same time, and each size of kernels is also multiple. Then, in the pooling layer, the pooling operation is conducted on the convolutioned results, which aims at extracting the features with the highest activation degree to N-gram features extracted by convolution. The maximum value is taken when a number of one-dimensional vectors have been convolutioned; then, they are pieced together as output in the max-pooling layer. Finally, in a fully connected layer, softmax is added to output the probabilities for each category.

It is clear that the basic structure of the TextCNN is shown in Table 1. The batch normalization is added in the structure. When the natural language is normalized, it can reduce randomness and bring it closer to predefined standards. What is more, it can contribute to reducing the amount of different information that would be processed and increase efficiency. In other words, if batch normalization is performed, the text distribution is more closer to normal distribution.

2.4. Loss Function. The sequence tagging at the sentence level can be executed in the CRF layer. And the parameter in the CRF layer can be treated as a matrix A, which is $(k+2) \times (k+2)$. As shown in Formula (1), it means that for the whole sequence x, the score of the whole sequence annotation is equal to the sum of the scores of all positions, which is divided into two parts. The first part in the formula can be affected by the pi, which is the output of BiLSTM. And the second part was determined by the CRF transfer matrix A, where A_{yi-1} and y_i represent the transfer score from the y_{i-1} tag to the y_i tag. In conclusion, front-and-back tag dependency constraints were taken into account in CRF, and the mark state transition probability was used as the score.

Table 2: The comparison of F1 scores (state-of-the-art) of different models for NER.

System	Accuracy
CRF	90.9
Conv-CRF	89.59
BiLSTM-CRF	90.1
BiLSTM-CNNs	90.77
Attention-BiLSTM-CRF	90.84
Self-attention-BiLSTM-CRF+TextCNN	90.93

Table 3: The comparison of accuracy(state-of-the art) in different models for our dataset.

Model	Accuracy
BiLSTM-CRF	84.97
Attention-BiLSTM-CRF	85.1
Self-attention-BiLSTM-CRF+TextCNN	85.5

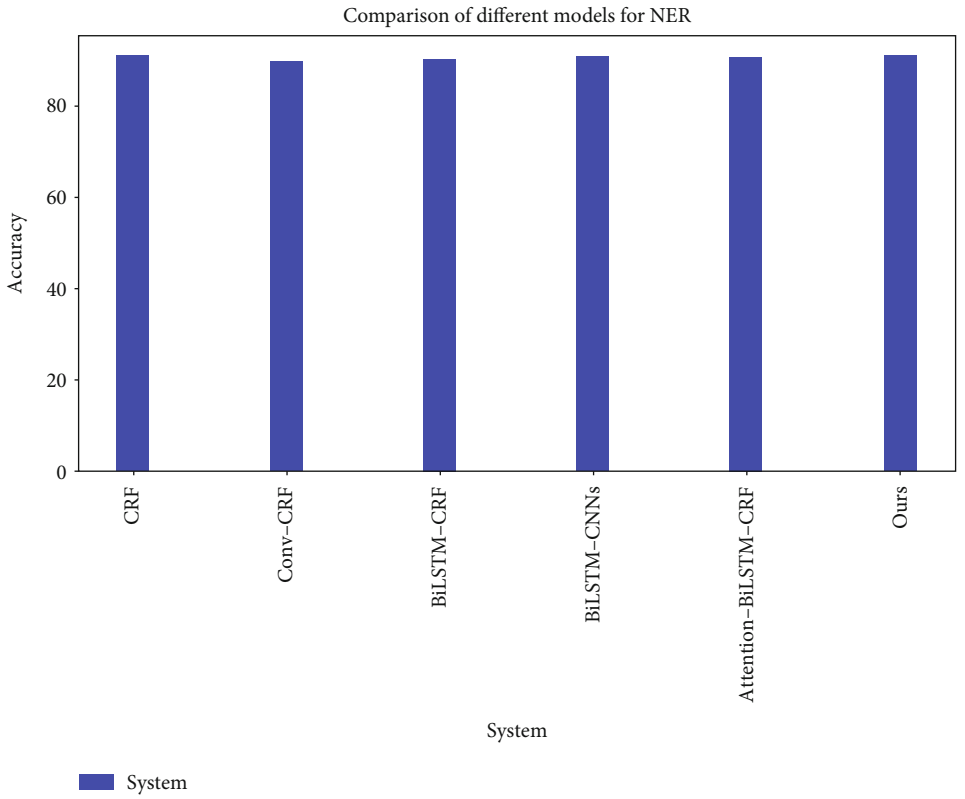


Figure 7: The accuracy in different models for NER.

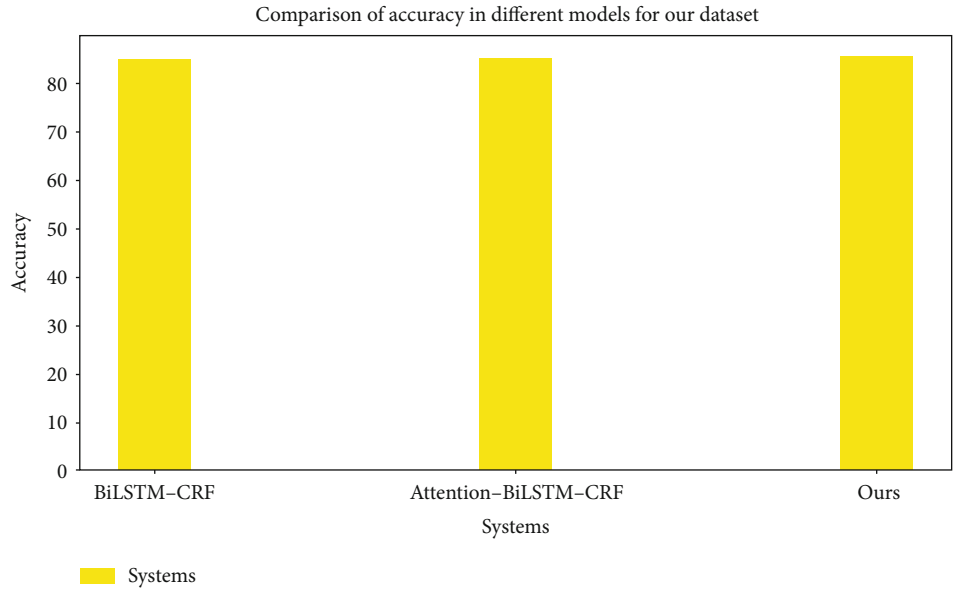


Figure 8: The accuracy in different models for our dataset.

$$\text{score}(x, y) = \sum_{i=1}^{n} Pi, yi + \sum_{i=1}^{n+1} Ayi - 1, yi. \quad (1)$$

Then, the normalized probability can be acquired by the softmax, which can be presented as in (Formula) (2).

$$P(y \mid x) = \frac{\exp\left(\text{score}(x, y)\right)}{\sum_{y'} \exp\left(\text{score}\left(x, y'\right)\right)}. \quad (2)$$

When model training, the maximizing logarithm likelihood function would be utilized. It can be defined in Formula

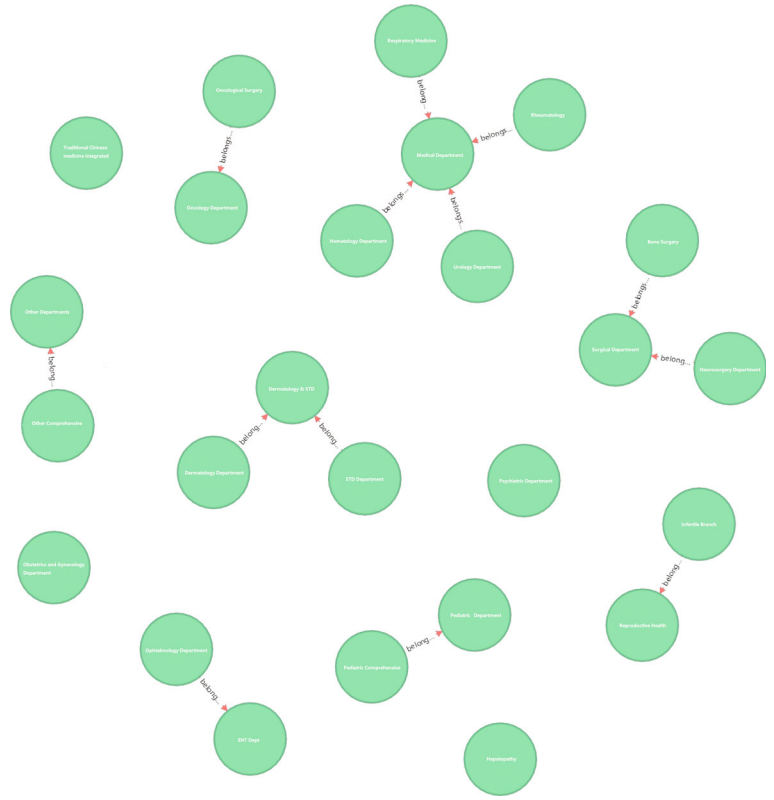


Figure 9: Disease-department knowledge graph.

(3), where represents the logarithmic likelihood of a training sample (x, y^x).

$$\log P(\mathrm{y}^x \mid x) = \mathrm{score}(x, y^x) - \log\left(\sum_{y'} \exp\left(\mathrm{score}\left(x, y'\right)\right)\right). \tag{3}$$

The Viterbi algorithm can be defined as a dynamic programming algorithm for selecting an optimal path, every step from the beginning, which can record the maximum probability of all paths to this state and finally keeping up pushing backward based on the maximum value. And then, finally, backing up from the end to the maximum probability is most probably the optimal path. In addition, dynamic programming is a method of solving complex problems by decomposing the original problem into relatively simple subproblems, which is suitable for overlapping subproblems and the problem of optimal substructural and less time-consuming. Therefore, the Viterbi algorithm which is a dynamic programming is adopted when the model predicts the process or decodes to find the optimal path, as shown in Formula (4).

$$\mathrm{y}^* = \arg\max_{y'} \mathrm{score}\left(x, y'\right). \tag{4}$$

In the CFR loss function, it consisted of two parts, the score of the real path and the total of all paths. What is more, the real path is the highest score of all paths, which can be defined in Formula (5), where pi represents the score of each possible path. And the logarithmic loss function can be represented as in Formula (6).

$$\mathrm{LossFunction} = \frac{P_{\mathrm{RealPath}}}{P_1 + P_2 + \cdots + P_N}, \tag{5}$$

$$\mathrm{LogLossFunction} = \log \frac{P_{\mathrm{RealPath}}}{P_1 + P_2 + \cdots + P_N}. \tag{6}$$

It is obvious that the loss function is usually minimized in training, so the negative sign is also added in the above formula. As shown in Formula (7), where e is a constant and Si is the sum of the EmissionScore and TransitionScore.

$$\mathrm{LogLossFunction} = -\log \frac{e^{S_{\mathrm{RealPath}}}}{e^{S_1} + e^{S_2} + \cdots + e^{S_N}}. \tag{7}$$

Similar to the Viterbi decoding algorithm, each node records the sum of the path from all previous nodes to the current modes, and the final step is to get the sum of all paths, which is shown in Formula (8) and $\omega 0$ to $\omega 1$ is taken as an example.

$$\mathrm{TotalScore}(\omega_0 \longrightarrow \omega_1) = \log \cdot \left(e^{x_{01}+x_{11}+t_{11}} + e^{x_{02}+x_{11}+t_{21}} + e^{x_{01}+x_{12}+t_{12}} + e^{x_{02}+x_{12}+t_{22}}\right). \tag{8}$$

Besides, the prediction is also made by Viterbi decoding, when each node can record the optimal path from all previous

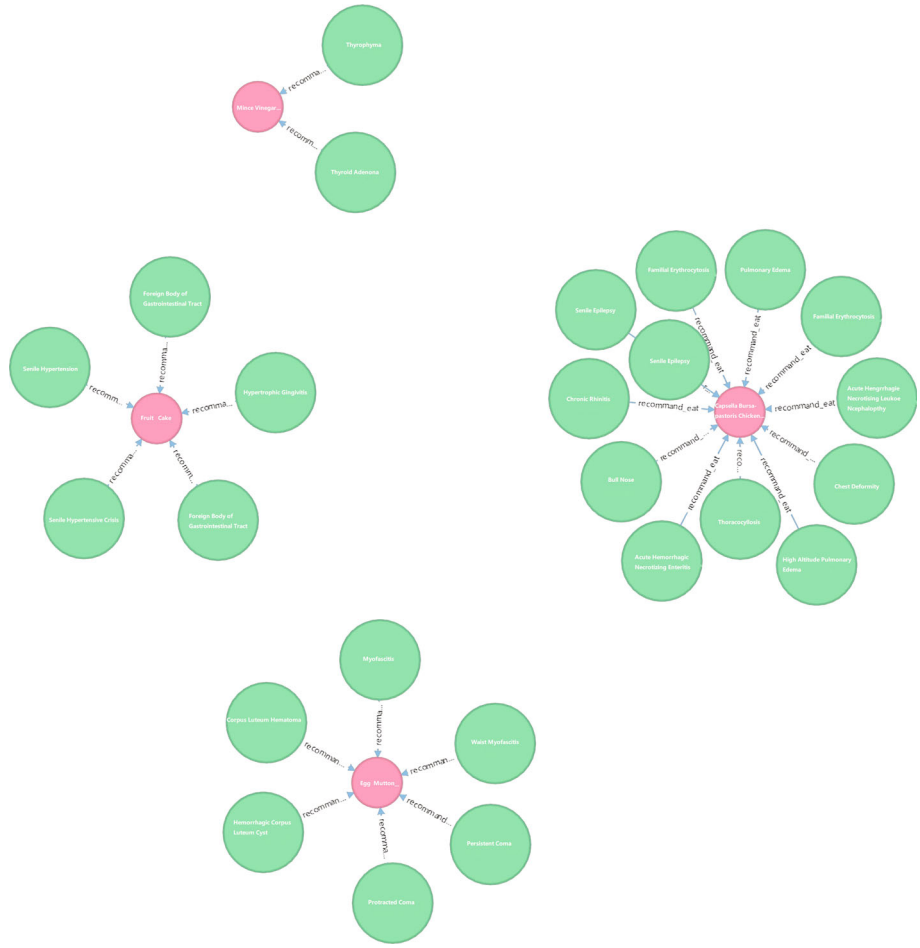


Figure 10: Disease-food recommendation knowledge graph.

nodes to the current node. And then, an optimal path can be obtained by backtracking, as shown in Formula (9).

$$\text{previous} = [\max(\text{scores}[00], \text{scores}[10]), \max(\text{scores}[01], \text{scores}[11])]. \tag{9}$$

2.5. Activation Function. As shown in Table 1, the activation function utilized is Relu in TextCNN structure, which can perform a nonlinear transformation to the output of the convolution layer. It is defined in Formula (10), and figure representation is shown in Figure 5.

$$f(x) = \max(0, x). \tag{10}$$

Combined Formula (10) with Figure 5, it is seen that the convergence rate is fast and no gradient disappearance in activation function. Besides, it only evaluates the max function, not the exponential function like Sigmoid, which leads to less forward calculation. Furthermore, due to simple derivative without exponential calculation, the back propagation calculation is very fast. On account of the value of some neurons is 0, which makes the network sparse properties which can reduce overfitting. Compared with sigmoid and tanh, the gradient vanishing problem in a positive interval is addressed, and it is much faster in convergence rate and computation than that. In a word, in Relu, it is suitable for backward propagation and has lower computation.

2.6. Attention Mechanism. Attention Mechanism is also adopted in our work, as shown in Figure 6, which is a technique that enables the model to focus on the important information and learn to absorb it sufficiently. Also, it can work on any sequence model and is linked to a specific purpose or task for the text. When model training, the information in a decoder is defined as a Query, and all the possible words are contained in the encoder, which is used as a dictionary. So the key and the value of this dictionary are usually the sequence information of all encoders. In our work, self-attention is utilized as an attention mechanism, which is simple and the computational cost is relatively low. The input sentence sequence information is after the word vector lookup or processed sequence information obtained after a sequence encoding with TextCNN. Firstly, each word qi in the sentence sequence and the attention weight ai of other words k in the sentence is established. Then, softmax normalization would be carried out on the attention weight vector, and all the moment information of the sentence sequence would also be linearly weighted. In a word, it can raise the limit of long-distance semantic dependency modeling capabilities.

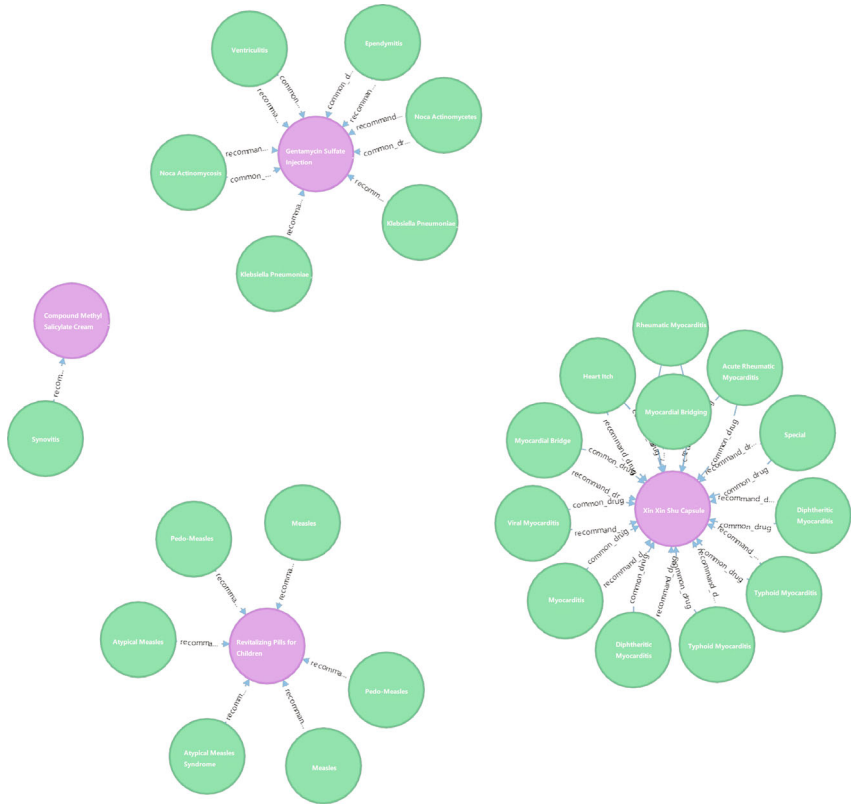


Figure 11: Disease-food not recommended.

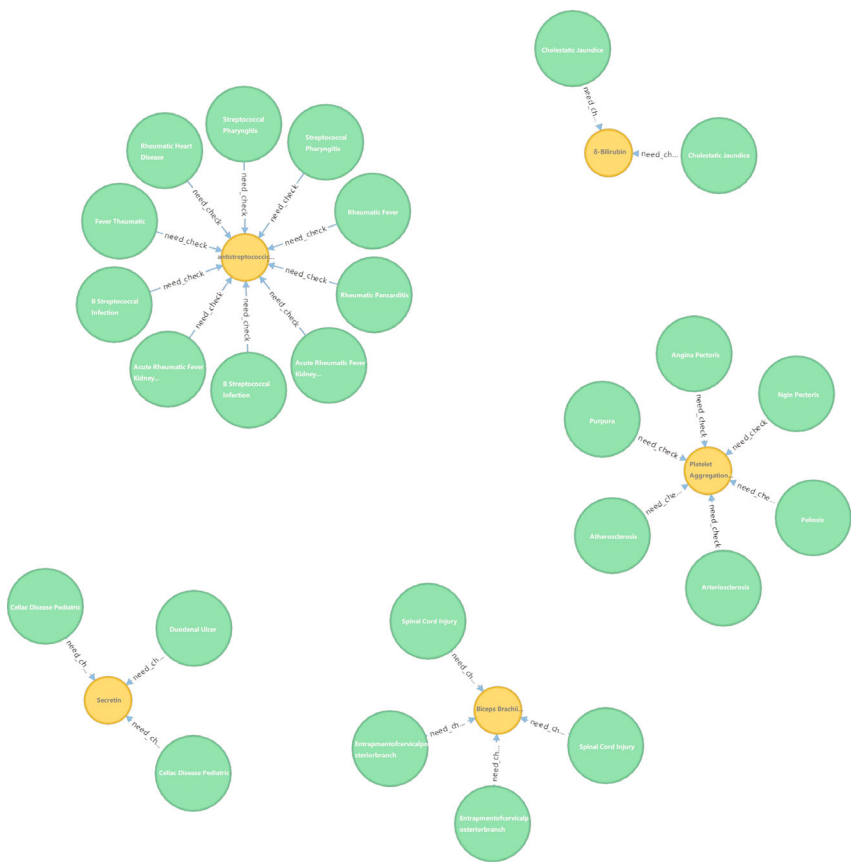


Figure 12: Disease drug recommendation knowledge graph.

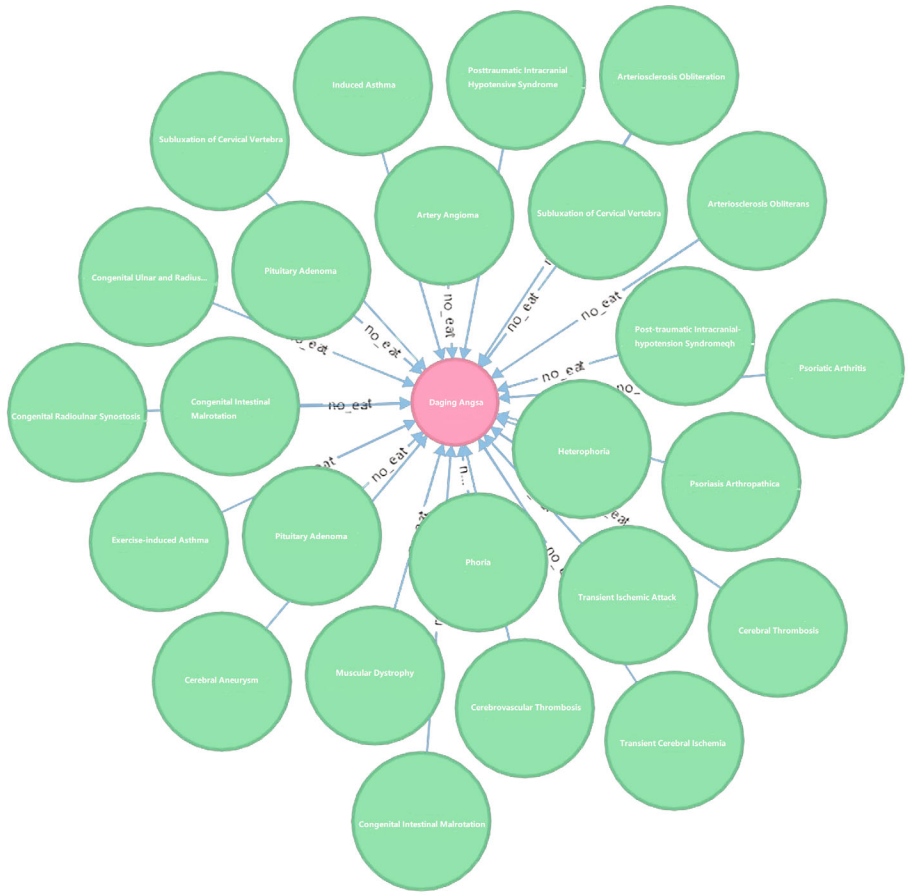


Figure 13: Drug-producer.

2.7. Cloud Computing. To make our model in practice, cloud computing is also adopted in our system. Cloud computing is an Internet-based computing in which the shared hardware and software resources and information are made available to computers and other devices on demand. When it comes to our system, to make it more efficient, the amount of data that patients and doctors are working with is not stored locally, but in the medical data center. When new data is acquired, it can be computed by the cloud, and then, new data would be stored on the cloud, in which text and image computing, processing, and storage capabilities are involved. What is more, the cloud computing is low-price and more convenient. Due to its automated centralized management, it can free engineers in the hospital from the increasing cost of data center management and enjoy the advantages of low cost. In addition, the terminal equipment requirements are very low in cloud computing, so it is no need to purchase expensive terminal equipment. Furthermore, it makes it easy to share medical text data and applications between different devices when cloud computing is applied. Therefore, it can help doctors with remote diagnosis and remote treatment.

3. Results and Discussion

Our experimental hardware device is based on Windows 10 operating system and intel i7 CPU processor, and the training and testing process is based on python, tensorflow, and neo4j. In Table 2, it shows the state-of-the-art F1 scores of different models for NER. And it can be seen that our system has a better performance than common systems for NER. When testing on our dataset, our model still has higher accuracy and performs better in Table 3. And the standard evaluation method we used is based on the macromean F1 value. Then, the bar chart shows in Figure 7 and in Figure 8 to make it more intuitive and clear. Meanwhile, it is shown the results of the medical knowledge graph in Chinese obtained from the experiments of our model. Firstly, it demonstrated that the disease department classification can be formed a knowledge graph in Figure 9, which the disease can be recommended to the correct and corresponding disease department. Secondly, the recommendation food of the disease can also be generated knowledge graph in Figure 10, which the reasonable diet can be advised according to the corresponding disease while the unsuitable food is not advised to the disease which is shown in Figure 11. Also, according to the diseases, the drug recommendation knowledge graph is produced as shown in Figure 12, in which the reasonable and effective medicine can be suggested to the corresponding diseases for the patients. And the drug producer is shown in Figure 13, which can be seen in the medicine factory. What is more, our model can suggest the patients to check the main factor of the disease in

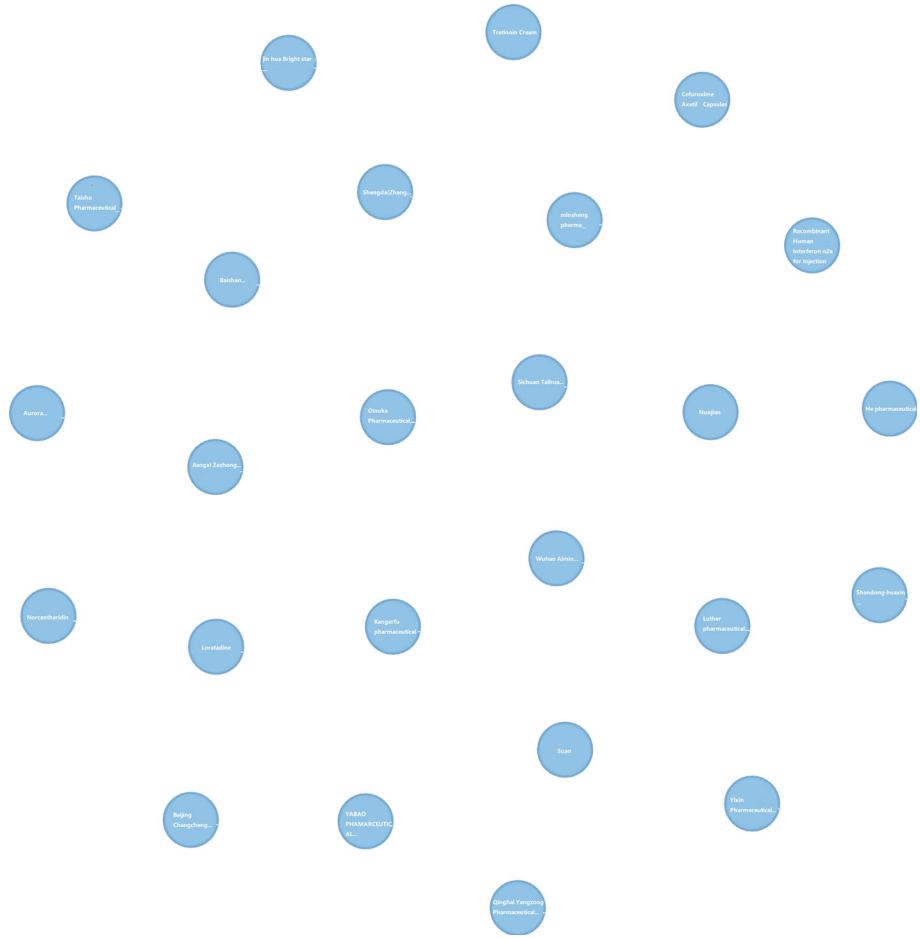


Figure 14: Disease-examining recommendation.

Figure 14, namely, to advise the patients to examine the main physical items matched with the diseases generated by the medical knowledge graph to perform further diagnosed. These results suggest that medical knowledge graph in Chinese in deep learning method on the mobile device based on IoT and WoT is essential and promising. And it works well and has better performance when generated various types of knowledge graph for both the doctors and patients. In addition, the accuracy and comprehensiveness of our knowledge graph are much higher and highly consistent with the predictions of the theoretical model. However, it is clear that there are still some limitations in this study that the suitable doctor recommended for the patient has not fused in our model, which remains to be addressed and further study.

4. Conclusions

To sum up, it is revealed that the medical knowledge graph in Chinese based on semantic computing using neural networks in IoT and WoT is considerable, which can make it convenient and intelligent for both doctors and patients. What is more, it is more time-saving, economic, and efficient which can contribute to smart healthcare applied in remote consultation platforms and mobile devices in hospital to implement disease classification and treatment recommendations.

However, it can be found that the suitable doctors recommended for the corresponding patients are still a problem that should be addressed. Besides, a more adaptive cloud computing method for the deep learning approach in practice ought to be adopted in the future work. Meanwhile, image data should be added to our model to achieve a more comprehensive model to serve for patients and doctors. What is more, it is considered that the patient's daily information data can be automatically recorded in our system to make it easier for the doctors to analyze the physical condition of the patients in real time, and it can also act as an early warning to the patients. Therefore, in the future, appropriate doctors recommended to the corresponding patients, image and text data fusion and patient's daily information automatically recorded should be paid much attention in our future work.

Data Availability

The data we used is available and can be accessed to perform medical knowledge graph in Chinese based on semantic computing study. And part of them are available to you from the corresponding author upon request (liuwanheng301@163.com).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the study of this work and publication of this paper.

Acknowledgments

Our work is supported by the Key R&D Program of Hebei Province, People's Livelihood Science and Technology Project, Project No.: 20277716D, Project name: Study on the pathological characteristics and pathogenic mechanism of Covid-19.

References

[1] A. Singhal, "*Introducing the Knowledgegraph: Things, Not Strings*," *Official google blog*, vol. 5, p. 16, 2012.

[2] D. B. Lenat, M. Prakash, and M. Shepherd, "CYC: using common sense knowledge to overcome brittleness and knowledge acquisition bottlenecks," *AI Magazine*, vol. 6, no. 4, p. 65, 1985.

[3] G. A. Miller, "Wordnet: Dictionary Browser' in Information in Data," *Proceedings of the First Conference of the UW Centre for the New Oxford Dictionary*, University of Waterloo, Waterloo, Canada, 1985.

[4] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific American*, vol. 284, no. 5, pp. 28–37, 2001.

[5] T. Berners-Lee, "Linked data-design issues," 2006, http://www.w3.org/Designlssues/LinkedData.html.

[6] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, "Dbpedia: a nucleus for a web of open data," *The semantic web*, pp. 722–735, 2007.

[7] R. R. Swick, *Resource Description Framework (RDF) Model and Syntax Specification W3C Recommendation*, W3c Recommendation World Wide Web Consortium, 1999.

[8] F. M. Suchanek, G. Kasneci, and G. Weikum, "Yago: a core of semantic knowledge," in *Proceedings of the 16th international conference on World Wide Web*, pp. 697–706, Banff, Alberta, Canada, 2007.

[9] K. Bollacker, R. Cook, and P. Tufts, "Freebase: a shared database of structured general human knowledge," *AAAI*, vol. 7, pp. 1962-1963, 2007.

[10] X. Dong, E. Gabrilovich, G. Heitz et al., "Knowledge vault: a web-scale approach to probabilistic knowledge fusion," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 601–610, New York, USA, 2014.

[11] M. Rotmensch, Y. Halpern, A. Tlimat, S. Horng, and D. Sontag, "Learning a health knowledge graph from electronic medical records," *Scientific Reports*, vol. 7, no. 1, p. 5994, 2017.

[12] L. Shi, S. Li, X. Yang, J. Qi, G. Pan, and B. Zhou, "Semantic health knowledge graph: semantic integration of heterogeneous medical knowledge and services," *BioMed Research International*, vol. 2017, 12 pages, 2017.

[13] Y. Song, L. Cai, K. Zhang, H. Zan, T. Liu, and X. Ren, *Construction of Chinese Pediatric Medical Knowledge Graph [C]//Joint International Semantic Technology Conference*, Springer, Singapore, 2019.

[14] D. Wen, Y. Liu, K. Yuan, S. Si, and Y. Shen, "Attention-aware path-based relation extraction for medical knowledge graph," in *International Conference on Smart Computing and Communication*, pp. 321–331, Springer, Cham, 2017.

[15] Y. Shen, K. Yuan, J. Dai, B. Tang, M. Yang, and K. Lei, "KGDDS: a system for drug-drug similarity measure in therapeutic substitution based on knowledge graph curation," *Journal of Medical Systems*, vol. 43, no. 4, pp. 1–9, 2019.

[16] I. Y. Chen, M. Agrawal, S. Horng, and D. Sontag, *Robustly Extracting Medical Knowledge from EHRs: A Case Study of Learning a Health Knowledge Graph*, Pacific Symposium on Biocomputing, 2020.

[17] S. Zheng, J. Rao, Y. Song et al., "PharmKG: a dedicated knowledge graph benchmark for bomedical data mining," *Briefings in Bioinformatics*, 2020.

[18] X. Li, H. Liu, X. Zhao, G. Zhang, and C. Xing, "Automatic approach for constructing a knowledge graph of knee osteoarthritis in Chinese," *Health Information Science and Systems*, vol. 8, no. 1, pp. 1–8, 2020.

[19] X. Chai, "Diagnosis method of thyroid disease combining knowledge graph and deep learning," *IEEE Access*, vol. 8, pp. 149787–149795, 2020.

[20] M. Alawad, S. Gao, M. C. Shekar et al., *Integration of Domain Knowledge Using Medical Knowledge Graph Deep Learning for Cancer Phenotyping*, 2021, arXiv:2101.01337.

[21] X. Sun, Y. Man, Y. Zhao, J. He, and N. Liu, "Incorporating description embeddings into medical knowledge graphs representation learning," in *International Conference on Human Centered Computing*, pp. 188–194, Springer, Cham, 2018.

[22] K. Lei, K. Yuan, Q. Zhang, and Y. Shen, "MedSim: a novel semantic similarity measure in bio-medical knowledge graphs," in *International Conference on Knowledge Science, Engineering and Management*, Springer, Cham, 2018.

[23] H. Wang, X. Miao, and P. Yang, *Design and Implementation of Personal Health Record Systems Based on Knowledge Graph [C]//2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, IEEE (Institute of Electrical and Electronics Engineers), 2018.

[24] Y. Xie, C. Yan, and D. Zhang, *Personalized Diagnostic Modal Discovery of Traditional Chinese Medicine Knowledge Graph [C]// 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, IEEE, 2018.

[25] X. Wang, R. Wang, Z. Bao, J. Liang, and W. Lu, "Effective medical archives processing using knowledge graphs," in *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1141–1144, 2019.

[26] J. Xiao, H. Sun, W. Zhu et al., "Building a medical knowledge graph to enhance the fraud, waste and abuse detection on claim data (preprint)," *JMIR Medical Informatics*, vol. 8, no. 7, 2019.

[27] X. Li, Y. Wang, D. Wang, W. Yuan, D. Peng, and Q. Mei, "Improving rare disease classification using imperfect knowledge graph," *BMC Medical Informatics and Decision Making*, vol. 19, no. 5, pp. 1–10, 2019.

[28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[29] J. Lafferty, A. McCallum, and F. Pereira, *Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data*, Proceedings of ICML, 2001.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6652246, 14 pages
https://doi.org/10.1155/2021/6652246

Research Article

A Security Wireless Monitoring and Automatic Protection System for CCEL

Qiao-yan Sun,[1] Xue-jiao Liu,[2] Yu-mei Sun,[1] Mei-chun Wang,[1] Xiang Han,[2] and Xiang-guang Chen[1,2]

[1]*Department of Electrical and Electronic Engineering, College of Engineering, Yantai Nanshan University, Longkou 265713, China*
[2]*School of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing 100081, China*

Correspondence should be addressed to Xiang-guang Chen; xgc1@bit.edu.cn

Received 9 December 2020; Revised 15 March 2021; Accepted 1 May 2021; Published 15 May 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Qiao-yan Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A wireless monitoring and automatic protection system for laboratory safety is proposed to solve the security problems in chemistry and chemical engineering laboratories (CCEL). The system is based on a wireless sensor network (WSN), including indoor wireless security monitoring node, information aggregation and management node, portable handheld terminal, and laboratory security information management system. Each indoor safety wireless monitoring node can monitor the laboratory security information (for example, fire information, toxic gas content, abnormal sound in chemistry, and chemical engineering laboratory) and control the corresponding protection device. The portable handheld terminal can remind the safety responsible person to check the safety status of the laboratory remotely and can issue corresponding operation and control instructions. According to the instructions it sends, the indoor security wireless monitoring and control node can start the corresponding automatic protection device. The experimental results demonstrate that the system can realize all-weather online monitoring of indoor safety information, ensuring that some important scientific research experiments can be conducted continuously and safely, significantly reduced the workload of laboratory managers, and precisely reported the location and conditions of accidents; the loss caused by the accident can be minimized based on the automatic protection device. The system can also be used in dangerous goods warehouse and inflammable and explosive places.

1. Introduction

At present, the CCELs in many universities and colleges in China are used for both teaching and scientific research, and the number of laboratories for teaching and scientific research is increasing rapidly, and the safety problems of laboratories are worthy of attention [1]. Based on the current status, doing a good job in laboratory safety management, studying the early-warning system of laboratory safety accidents, etc., are all important contents of laboratory safety management in colleges and universities [2]. Effective laboratory safety management measures can ensure the smooth progress of school teaching and scientific research, which is of great significance for students to form correct concepts of the world, life, and values [3].

Reference [4] summarized the laboratory safety education and training, safety management system and guarantee facilities, safety management level, and so on and proposed a number of practical measures to strengthen laboratory safety management. Reference [5] made a brief comparison between China and the United States in the aspects of student safety education, laboratory construction, chemical management, laboratory management, and establishment of management institutions. Reference [6] introduced that some universities in China had achieved good results in the safety management of chemistry and chemical laboratory by referring to the advanced management mode and management system of foreign countries. In order to improve the reliability and practicability of the wireless safety monitoring system, reference [7] had diagnosed the fault of wireless safety monitoring system in chemistry and chemical engineering laboratory.

At present, the management information of the laboratories in some research institutes and universities in China

cannot be released in time. It is indeed difficult to assess laboratory conditions and environments [8]. Unadvanced monitoring and management system will occupy more resources, and the facilities of the system are limited. The unadvanced databases cannot meet the requirement of most of the experimental facilities and cannot be used for the relevant information management system. In order to ensure the safety of workers and technicians, some students at Chongqing University have designed a system that may not only use sensors to monitor changes in the concentration of gases on site but also automatically select suitable ventilation facilities for different levels of danger to reduce the content of flammable and explosive gases [9, 10]. Reference [11] designed a laboratory remote monitoring device, which achieves the real-time monitoring and gives the automatic voice alarm and alert message to the remote area or host area by using wireless network technology. This work is designed using a PIC16F877 microcontroller and focuses on sensing the existence of the critical parameters and issuing the alarm if those parameters are sensed. Reference [12] proposed a safety information monitoring and data management system, the purpose of which is to monitor the laboratory and obtain relevant data, judge experimental environment data, and transmit relevant sound-light alarm information. The system integrates some intelligent methods in the management of the laboratory and has certain reference value in the safety and telemetry of the laboratory. Reference [13] proposed a real-time detection and automatic alarm system for laboratory status information, which its purpose is to monitor real-time data and laboratory status, automatically determine the relevant environmental indicators, detect whether there is an external intrusion, and configure the sound-light alarm device. The monitoring system can communicate wire line or wirelessly with PC and can realize remote monitoring, which can meet the needs of unattended duty in related laboratories. Reference [14] presented an embedded laboratory security monitoring system. The system uses a camera to obtain video information and a sensor to obtain environment-related information and sends the data to the network server through the wireless sensor network. When an emergency occurs, the alarm information can be displayed on the web, and the sound and light alarm information can be sent. The system can also be used to monitor laboratories or places where valuables are stored. Besides, a wireless remote monitoring system for the laboratory was designed and developed in reference [15].

Laboratories are the very important places for personnel training and scientific and technological innovation in a university. Reference [16] systemically discussed telecommunications, standard management of chemistry laboratory security based on three aspects of logistical support system, management responsibility system, and education security system. The safety operation of CCEL is the key of the CCEL management system. Reference [17] analyzed safety system construction, safety hardware guarantee, and safety protection skills training of CCEL. Reference [18] comprehensively discussed the importance of experimental safety education, whose purpose is to require relevant personnel to improve CCEL safety management ability and ensure the safety of the CCEL environment. Reference [19] indicated to pressingly study laboratory safety management measures, because the prolonged teaching experiment time and the unproperly disposed wastes in the laboratory are all laboratory safety problems.

The authors of reference [20], aiming at the frequent occurrence of laboratory accidents in Chinese universities at present, took chemical engineering laboratory as an example and put forward corresponding safety management methods from four aspects: equipment, reagents and consumables, water and electricity, and rules and regulations. The authors of reference [21] analyzed the main problems existing in chemical engineering laboratories in China, summarized the types of safety accidents in chemical engineering laboratories in Chinese universities in recent years, and put forward a set of effective laboratory safety management mechanism. The authors of reference [22] studied the safety management of chemical engineering laboratories in China; normalized the management of safety education mechanism, safety supervision and evaluation mechanism, prevention and control mechanism, and accident emergency treatment mechanism; and proposed a safety access management system for chemical engineering laboratories. The authors of reference [23] discussed the methods and measures of chemical engineering laboratory safety management, such as the implementation of student safety education, strengthening the construction of chemical engineering laboratory safety rules and regulations, construction and management of safety facilities, and strengthening the management and use of hazardous chemicals, and put forward the hierarchical management of chemical engineering laboratories and the establishment of a "three-level safety inspection" system.

The main references in recent years are the following: the authors analyzed some problems existing in the management of chemistry and chemical engineering laboratory and put forward some improvement schemes [24]: (1) implement scientific management and further improve rules and regulations, (2) establish reward and punishment mechanism to further strengthen safety management, (3) use the Internet to strengthen instrument management and resource utilization, and (4) constantly improve the quality of the laboratory management team.

According to the problems commonly existing in university laboratories in China, the authors put forward several improvement methods [25]: (1) formulate a reasonable and effective safety management system, (2) strengthen the construction of laboratory safety equipment, and (3) strengthen laboratory safety education and training.

The authors pointed out the types of accidents in university chemical engineering laboratories [26]: for example, (1) fire accident, (2) explosion accident, and (3) chemical poisoning accidents. And put forward several improvement measures: (1) strengthen the safety education of students, eliminate people's unsafe behavior, (2) to do a good job in quality management and eliminate the unsafe state of materials, and (3) strengthen the laboratory safety management and conduct the laboratory safety management according to the rules.

The authors compared Chinese and Norwegian laboratories in five aspects: safety education, safety facilities and management, chemicals management, operation management, and equipment management [27]. It was pointed out that safety education, experimental facilities, drug management, and management should be strengthened and improved.

According to the author's review of relevant statistical data [28], in the past 10 years, there have been 14,539 school fires in China, killing 29 people and injuring 50 others. At the same time, the authors pointed out the problems existing in laboratory safety management in colleges and universities, such as (1) relaxation of safety education process, (2) the dynamic supervision of hazard sources lags behind, and (3) lack of emergency plan. Several improved schemes were proposed, such as (1) introducing online safety training courses, (2) realize dynamic supervision of hazard sources, and (3) virtual simulation of an emergency drill.

In recent years, our research group has completed several research projects in the fields of research and application of WSN and published some articles; the articles [29–34] are representative.

At present, the main safety problems in chemistry and chemical engineering experiments are as follows:

(1) *Fire Accident.* Since chemistry and chemical engineering laboratories are places for chemical reactions, which are easy to cause fires due to heat release, flammable gas leakage, or aging of electrical equipment, etc. If not timely alarm and take effective firefighting measures, it will cause casualties and property losses

(2) *Accident of Chemical Poisoning.* In the process of chemistry and chemical engineering experiments, chemicals are needed. There are many toxic substances in these chemicals and they are frequently used. If students lack safety awareness and fail to carry out effective protection in the experiment, they are prone to poisoning. If they fail to report to the police in time and take effective protection measures, death accidents will occur

(3) *Explosion Accident.* Because there are so many chemicals in a chemical engineering laboratory, the reactions of the chemicals give off heat and produce energy. If the amount of chemical substances used in the experiment is not used properly, it will easily explode. If the alarm is not timely and effective protection measures are not taken, it will cause heavy casualties and property losses

For some chemistry or chemical engineering experiments, it often takes a day or days to complete. During the experiment, if a fire (or poisoning, or explosion) occurs in the laboratory, how can the remote monitoring center be notified in time? How to report the exact location of the accident in time? How can effective measures be taken to protect the accident site? These problems have not been well solved in the chemistry and chemical engineering laboratories of Chinese universities.

In this paper, we proposed a wireless monitoring and automatic protection system for chemistry and chemical engineering laboratory safety. The system is based on a wireless sensor network (WSN), including indoor wireless security monitoring node, information aggregation and management node, portable handheld terminal, and laboratory security information management system. Each indoor safety wireless monitoring node can monitor the laboratory security information (this is fire, poisonous and harmful gas, abnormal explosion sound, and so on) and control the corresponding protection device. When the safety wireless monitoring node detects the unsafe information in the laboratory, it will upload the field information in time and start the indoor safety automatic protection device. The portable handheld terminal can remind the safety responsible person to check the safety status of the laboratory remotely and can issue corresponding operation and control instructions. According to the instructions it sends, the indoor security wireless monitoring and control node can start the corresponding automatic protection device. The laboratory security information management system receives the information transmitted by all wireless monitoring nodes in the laboratories, and processes, analyzes, and stores the data of each wireless monitoring node in the laboratories. The experimental results demonstrate that the system can realize all-weather online monitoring of indoor safety information, ensuring that some important scientific research experiments can be conducted continuously and safely, significantly reduced the workload of laboratory managers, and precisely reported the location and conditions of accidents; the loss caused by the accident can be minimized based on the automatic protection device. The system proposed in this paper can also be used in dangerous goods warehouse and inflammable and explosive places in the process of industrial production.

The rest of this paper is structured as follows. Section 2 provides monitoring and control system structure design in CCEL, including a schematic structure of wireless monitoring and protection system, design of indoor security wireless monitoring node, design of portable handheld terminal, and structural design of laboratory safety information management system. Section 3 proposes a security wireless monitoring and automatic protection method in CCEL. Section 4 presents the safety judgment procedures of CCEL. Section 5 discusses the application examples, including the structure of the experimental system, testing of security monitoring node in CCEL, and remote testing based on handheld terminal. Section 6 summarizes the research work of the whole paper and puts forward some constructive proposals for further and in-depth research work in the future.

2. Monitoring and Control System Structure Design

Figure 1 shows a schematic structure of a CCEL wireless monitoring and protection system. The system comprises several indoor security wireless monitoring nodes, an information aggregation management node, several portable

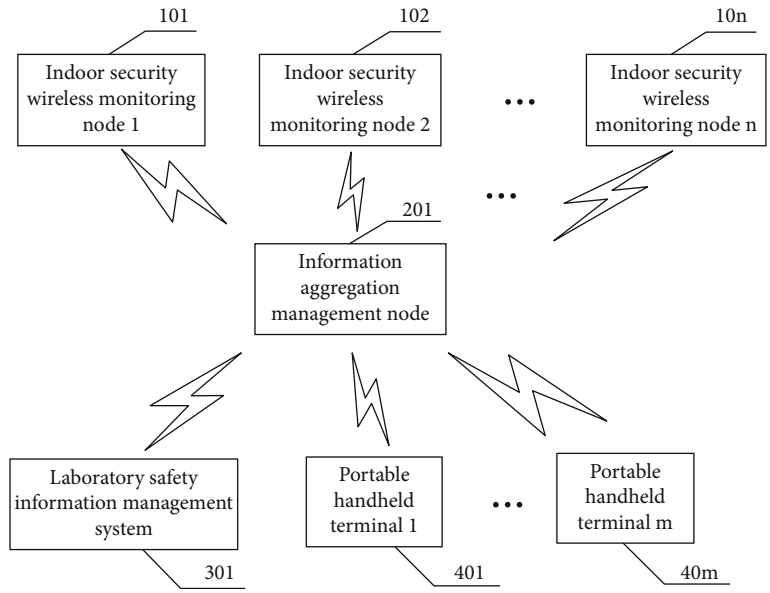


Figure 1: A schematic structure of wireless monitoring and protection system in the CCEL.

handheld terminals, and the laboratory safety information management system.

Each indoor security wireless monitoring node in Figure 1 is used to monitor the CCEL security state and to control the automatic protection equipment corresponding to the indoor monitoring node according to the monitoring situation or the instructions of the portable handheld terminal or CCEL information management system. The monitoring CCEL security state includes monitoring the fire information, the poisonous and harmful gas information, and the explosion sound information in the CCEL environment.

The information aggregation management node in Figure 1 is used to support each CCEL security wireless monitoring node to communicate with the corresponding handheld terminal and the laboratory safety information management system. The CCEL safety information management system is used for receiving the monitoring data of the CCEL security wireless monitoring node which is transferred by the information aggregation management node and processing, analyzing, and storing the monitoring data. The handheld terminal in Figure 1 is used for receiving the monitoring data of the CCEL security wireless monitoring node corresponding to the handheld terminal which is transferred by the information aggregation management node and remotely controlling the automatic protection device corresponding to the CCEL security wireless monitoring node.

The CCEL wireless monitoring and protection system may be provided with an indoor security wireless node in each laboratory, such as the indoor security wireless nodes 101~10n in Figure 1, respectively, correspond to n laboratories; then, you can set an information aggregation management node (201) and a laboratory safety information management system (301), the laboratory safety information management system (301) is a subsystem of the laboratory wireless monitoring system, which can be used to record the whole indoor security wireless node data for backup, analysis, and other operations. The system would also provide several portable handheld terminals, that is, m portable handheld terminals 401 to 40 m in Figure 1. In this system, n indoor security wireless nodes correspond to m portable handheld terminals, and the corresponding relationship can be one-to-one or many-to-one, which means one portable handheld terminal can manage one CCEL security wireless node, or several portable handheld terminals manage several multiple indoor security wireless nodes.

CCEL security wireless monitoring nodes (101~10n) are applied to monitor the indoor environment temperature, smoke and flame, poisonous and harmful gases, and explosion sound.

The information aggregation management node (201) receives the monitoring information from each laboratory and transmits the monitoring information to the CCEL safety information management system (PC host computer) and the portable handheld terminal, respectively.

The laboratory safety information management system (301) receives the monitoring information from each laboratory transmitted by the information aggregation management node, and processes, analyzes, and saves the information.

The portable handheld terminal (401~40 m) receives the monitoring information from each laboratory transmitted by the information convergence management node, monitors the situation in the laboratory through the handheld terminal, communicates with the indoor security wireless monitoring node (101~10n), and realizes remote monitoring of the laboratory automatic protection devices.

We can divide the system shown in Figure 1 into four levels. Level 4 is portable handheld terminal. Typically, a laboratory manager corresponds to a terminal or several terminals. That is, he or she is responsible for the safety of a laboratory or several laboratories. Level 3 is a laboratory safety information management system; it is a safety monitoring center for laboratories. It manages all laboratory safety in a building or several buildings. Once a laboratory has an

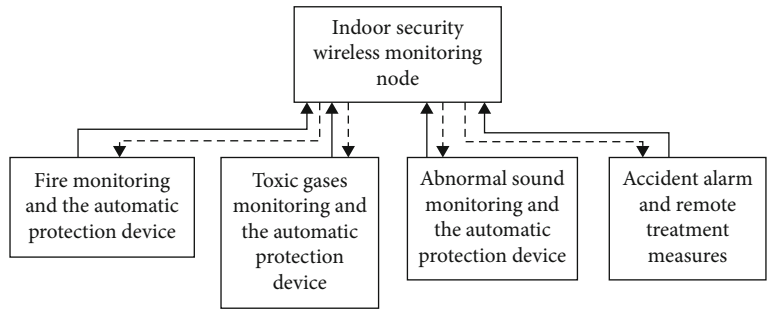


Figure 2: Overall functional structure of each indoor security wireless monitoring node.

accident, the monitoring center will contact the person in charge of the handheld terminal in time. They will take the solution as soon as possible.

2.1. Design of Indoor Security Wireless Monitoring Node. Figure 2 illustrates the overall functional structure of the indoor monitoring node. Including the indoor security wireless monitoring node control modes: fire, toxic gases, abnormal sound monitoring and the corresponding automatic protection device, and the accident alarm and remote treatment measures, such as sending alarm information to the appropriate portable handheld terminal, receiving remote control commands to activate the automatic protection device.

(1) *Sensors.* For the detection of temperature, smoke, and flame in the indoor environment, for the detection of toxic and harmful gases in the indoor environment, and for the detection of abnormal sounds in the indoor environment (such as light detonation, explosion, and detonation)

(2) *Controller Module.* It is used to collect the field data of CCEL, analyze and judge the data, and control the indoor safety automatic protection device and send alarm information

(3) *Communication Module.* Transmit the parameter information and image information obtained by CCEL on site to the security information management system in CCEL and handheld terminal. Moreover, the command or control information of the handheld terminal and the security information management system in CCEL can be transmitted to the indoor security wireless monitoring node

Each indoor safety wireless monitoring node in CCEL can monitor the laboratory safety status online, can analyze and judge the obtained node monitoring data, and can control the automatic protection equipment through the safety wireless monitoring node. At the same time, the monitoring data will be sent to the security information management system in CCEL, and the corresponding handheld terminal will be sent to remind the security responsible person to conduct remote viewing and determine the corresponding measures, and instructions can be sent to the CCEL wireless monitoring node through the handheld terminal, so as to timely start the corresponding indoor automatic protection device.

In the practical application of this system, each CCEL indoor security wireless monitoring node can transmit the data to the information aggregation management node and, then, send the data from the information aggregation management node to each portable handheld terminal and the security information management system in CCEL. Each security wireless monitoring node in CCEL can also communicate and transmit information with the corresponding portable handheld terminal. When a node fails, it will not affect the normal operation of other nodes, information aggregation management nodes, portable handheld terminals, and laboratory security information management system. Moreover, the information aggregation management node and the portable handheld terminal can be used to determine whether a node has a fault, which can obviously improve the reliability and security of the whole system.

2.2. Design of Portable Handheld Terminal. Figure 3 shows the schematic diagram of the handheld terminal. The handheld terminal includes processor module, communication module, power supply module, liquid crystal display, and keyboard (or touch keyboard or touch screen). It obtains the indoor detection information and image information through the communication module, analyzes the detection information based on the judgment algorithm, determines the type of the accident, and remotely controls the corresponding safety protection device, by looking at the scene images of the accident and timely reporting of the accident site and field conditions.

The solid line shown in Figure 3 indicates the information transmitted to the handheld terminal by the field wireless monitoring node, and the broken line indicates the control information sent out by the handheld terminal to the indoor security wireless supervisory node or the accident type alarm information sent from the handheld terminal to the real field.

2.3. Structural Design of Laboratory Safety Information Management System. The laboratory safety information management system shown in Figure 4 includes a laboratory security wireless monitoring node (301-1), a PC of security information management system (301-2), and a laboratory safety information management system server (301-3).

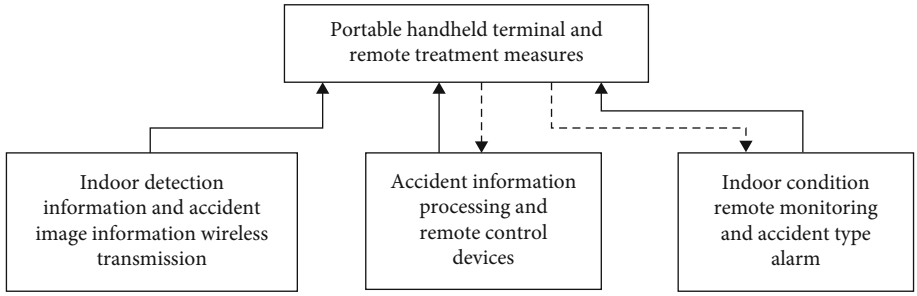


Figure 3: The schematic diagram of the portable handheld terminal.

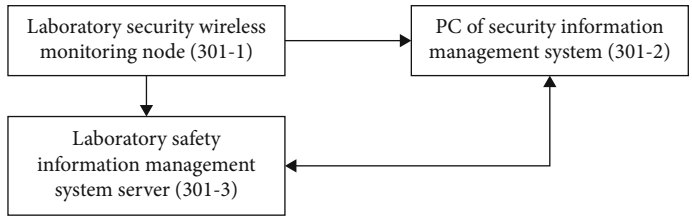


Figure 4: Structure of laboratory safety information management system.

The indoor information transmitting and receiving node (301-1) is used for receiving the monitoring data of the indoor monitoring node which is sent by the information aggregation management node and sending the monitoring data to the server (301-3) or the PC host computer (301-2) and sending the instruction of the PC host computer (301-2) to the information aggregation management node.

The server (301-3) is used for storing the monitoring data received by the indoor information transmitting and receiving node and calls the corresponding stored data according to the instruction of the PC host computer (301-2).

The PC host computer (301-2) is used for providing an exhibition and management interface for displaying the monitoring data and receiving the operation instruction of the manager.

In the system, the indoor information transmitting and receiving node is also called a gateway. It is composed of a processor module, communication module, and node power supply module. It is connected with a PC. Its main function is data receiving and sending.

In the embodiment, a set of management software is provided in the PC host computer for providing the display and the management interface for displaying the monitoring data and receiving the operation instruction of the manager.

The portable handheld terminal cannot store the nodes' data for a long time; in order to analyze the cause of the accident and to back up the remote control, the laboratory safety information management system is designed for backup and analysis of the monitoring data. The information management system receives the detection information from each laboratory transferred from the information aggregation management node, and processes, analyzes, and saves the information.

In addition, in this system, the control modes of automatic protection device also will be designed in the node, and the automatic protection device includes the fire extinguishing device, the ventilator, and the dangerous goods isolation device.

3. Security Wireless Monitoring and Automatic Protection Method in CCEL

This paper proposes a method of wireless monitoring and automatic protection for safety in CCEL. The realization process of this method includes the following steps:

(1) When the fire situation is detected by the safety wireless monitoring node installed in the CCEL room, the fire situation information obtained is analyzed, and the first automatic protection device is determined to be activated according to the analysis results

(2) When the toxic gas information is detected by the safety wireless monitoring node installed in the CCEL room, analyze the toxic and harmful gas information obtained, and determine whether to start the second automatic protection device based on the analysis results

(3) When abnormal sound information is detected by the security wireless monitoring node installed in the CCEL room, the abnormal sound information obtained shall be analyzed, and the third automatic protection device shall be activated according to the analysis results

(4) When the above information is detected by the security wireless monitoring node installed in the CCEL room, the obtained information is transmitted to the security information management system in CCEL through the information aggregation management node and to the handheld terminal corresponding to the indoor security wireless monitoring node

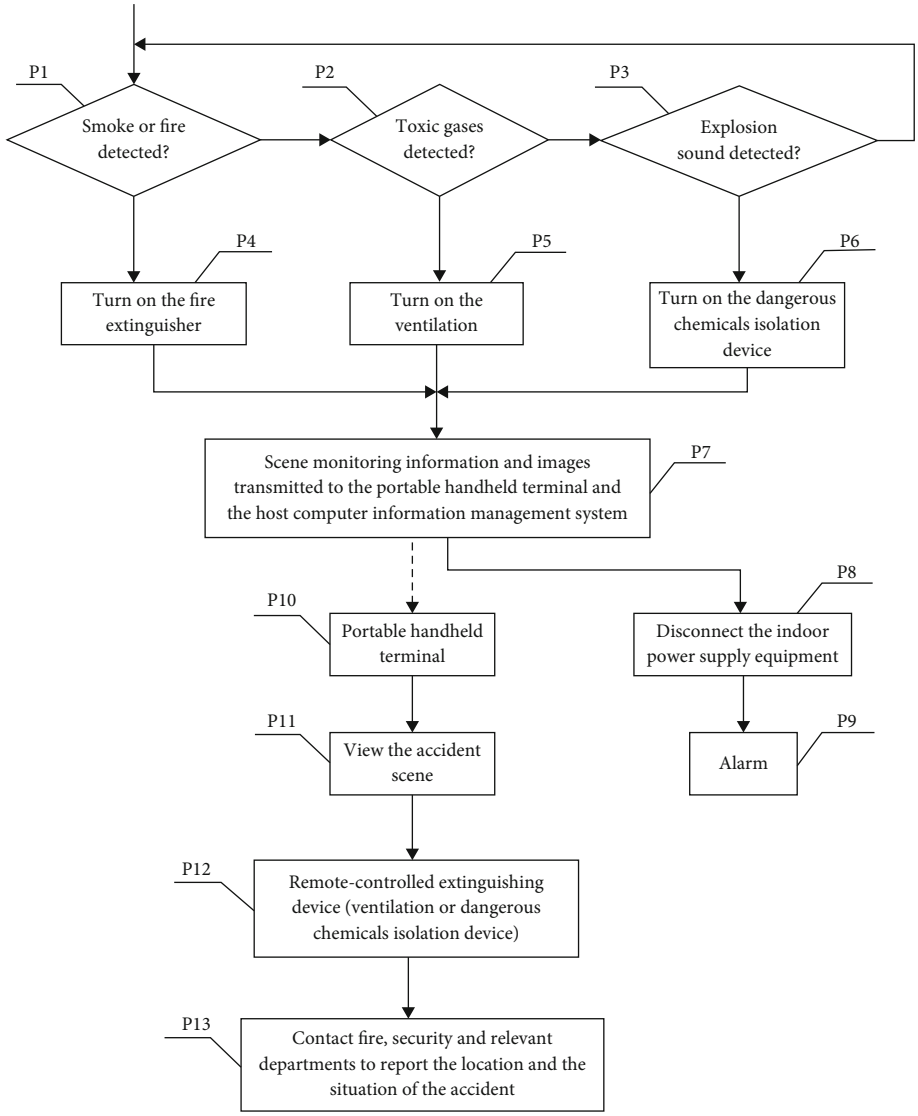


Figure 5: Flow chart of laboratory safety wireless monitoring and automatic protection method.

(5) Indoor safety can show by indoor security of wireless data transmission of video monitoring node, the researchers can also use a portable handheld terminal to send the instructions to the indoor security wireless monitoring nodes and can also send functional instructions to the corresponding indoor security wireless monitoring node through information gathering management node. The node controls the relevant automatic protection device and sends the relevant alarm information at the same time

In this system, when a security monitoring node in the CCEL enters the working state, it can simultaneously or successively perform steps (1) to (3), and the order of steps (1) to (3) can be changed, not be restricted.

Figure 5 illustrates the flow chart of a secure wireless monitoring and automatic protection method proposed in this paper. Each step function is described as follows:

P1 Description: determine whether smoke or flame information is detected? If there is, then turn on the fire extinguishing device; if not, then P2 judgment

P2 Description: determine whether toxic gases information is detected? If there is, then turn on the ventilation; if not, then

P3 Description: determine whether abnormal sound information is detected? If there is, then turn on the dangerous chemicals isolation device; if not, then P1 judgment

P4 Description: after confirming by the judgment algorithm, the fire extinguishing device is controlled by indoor security wireless monitoring node, which includes high pressure water and fire extinguishing agent

P5 Description: after confirming by the judgment algorithm, the indoor security wireless monitoring node controls the ventilation device, the ventilation device includes an exhaust fan and a fume hood

P6 Description: after confirming by the judgment algorithm, the dangerous chemicals isolating device is controlled by the indoor security wireless monitoring node, the isolating

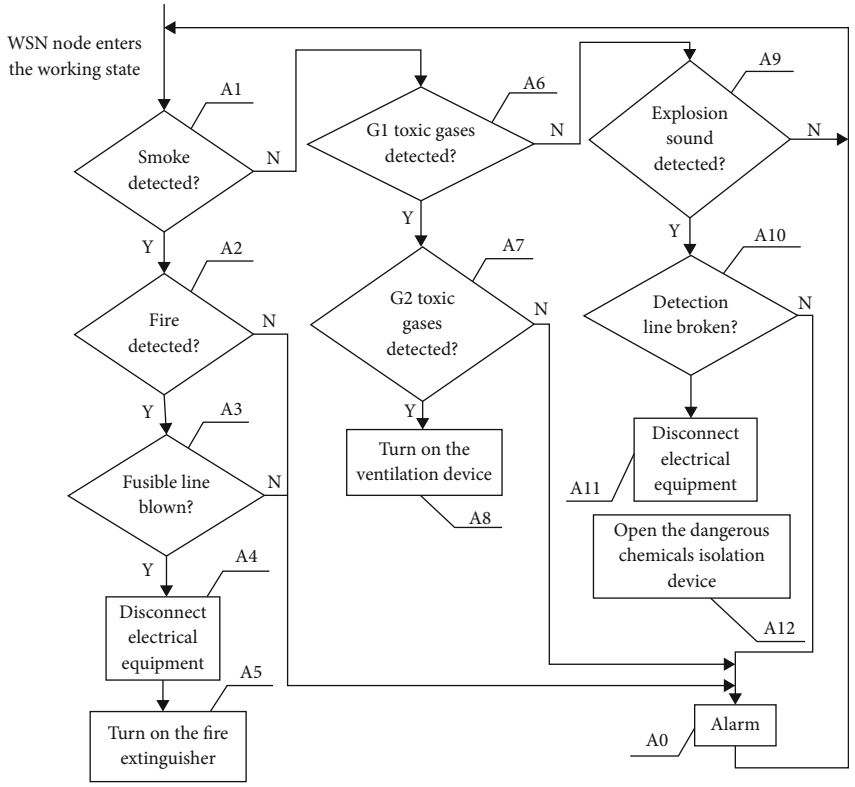


Figure 6: The flow chart of the security judgment method in CCEL.

device includes fire-proof material, flameproof material, and explosive-proof material

P7 Description: when the accident occurs, on-site inspection information and images will be transmitted to the handheld terminal and laboratory safety information management system for timely alarm, the effective handling of accidents and analyzing accident cause

P8 Description: disconnect the indoor power supply equipment, to avoid causing other accidents

P9 Description: send out an alarm signal, to facilitate the rapid departure of laboratory staff

P10 Description: quickly view digital and images information when a laboratory alarm is received

P11 Description: the accurate location of the accident can be determined by data and alarm information, and the image information of the accident site can be viewed to determine the remote control scheme

P12 Description: minimize accident losses by remotely operating fire extinguishing devices (ventilation devices or isolating devices)

P13 Description: contact fire department, security department, and other departments in time, and send the exact location of the accident and the relevant information of the accident site

In Figure 5, judgment box P1 judges the temperature, smoke, and flame detected in the indoor environment; judgment box P2 makes judgments on toxic and harmful gases (such as carbon monoxide, nitric oxide, hydrogen sulfide, sulfur dioxide, and chlorine) detected in the indoor environment; judgment box P3 is used to judge abnormal sounds (such as light detonation, explosion, and detonation) detected in the indoor environment.

In Figure 5, (1) through P4 judgment, it is confirmed that when an indoor fire occurs, different fire-extinguishing devices, including high-pressure water, foam fire-extinguishing agent, and dry powder fire-extinguishing agent, shall be controlled by indoor safety wireless monitoring node, etc. (2) Judging from P5, when the indoor toxic gas leakage is confirmed, different ventilation devices are controlled by the indoor safety wireless monitoring node, including the exhaust fan, the fume hood, and the automatic opening of Windows. (3) According to P6 judgment, when the indoor explosion sound is confirmed, different dangerous chemical isolation devices shall be controlled by the indoor security wireless monitoring node, including flame-retardant materials, flameproof materials, and explosive-proof cover.

4. Safety Judgment Procedures of CCEL

The flow chart of the CCEL indoor safety judgment algorithm proposed by us is shown in Figure 6. In order to improve the accuracy of node judgment and the effectiveness of node control, the specific steps and instructions of the algorithm are as follows:

A1 Description: to determine whether smoke information has been detected? If so, enter A2; if not, enter A6 judgment

A2 Description: to determine whether flame information has detected? If so, enter A3 judgment; if not, switch on P0 to alarm and transfer to A1 for judgment

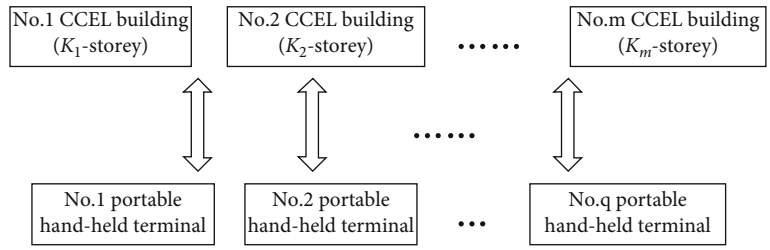


Figure 7: The structure chart of handheld terminal monitoring laboratory buildings.

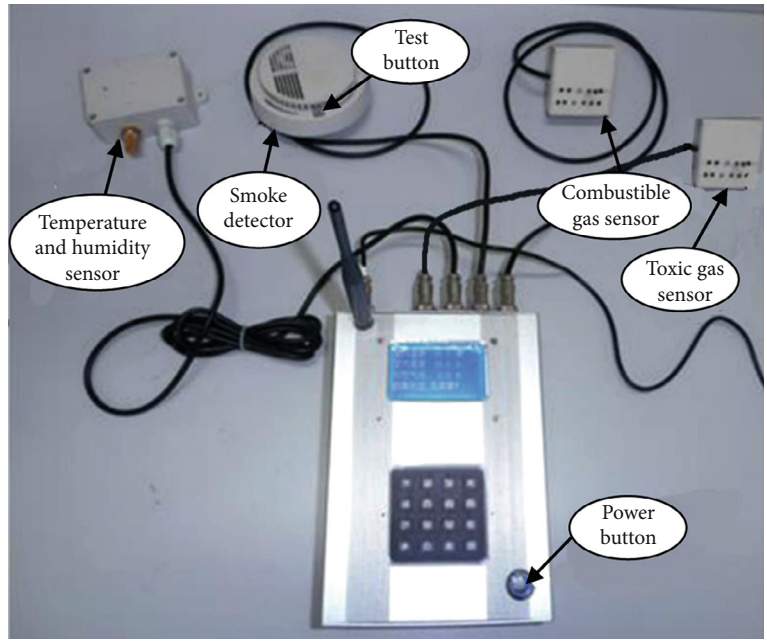


Figure 8: The experimental device of an indoor security monitoring node connected to several sensors.

A3 Description: is it detected that the fusible line has been burnt? If it is burnt out, switch to A4; if it is not burnt out, connect P0 alarm and transfer to A1 for judgment

A4 Description: disconnect the power supply of indoor electrical equipment in the CCEL and switch to A5

A5 Description: activate the fire extinguishing device in the CCEL and transmit the field information and images to the portable handheld terminal

A6 Description: to determine whether G1 has detected toxic and harmful gas? If so, enter A7 judgment; if not, enter A9 judgment

A7 Description: determine whether G2 has detected toxic and harmful gas? If so, enter the A8; if not, A0 alarm is turned on and A1 judgment is transferred

A8 Description: open the ventilation device in the CCEL room and send the field information and images to the portable handheld terminal

A9 Description: to determine whether explosion sound information has been detected? If so, enter the A10; if not, transfer to A1

A10 Description: judge whether the detection line is broken or not? If broken, enter A11; If it is not broken, A0 is connected to alarm and A1 is transferred to judgment

A11 Description: disconnect the power supply of indoor electrical equipment in the CCEL and switch to A12

A12 Description: activate the isolation device in the CCEL and transmit the field information and image to the portable handheld terminal

A0 Description: connect the alarm device of CCEL, namely sound alarm, voice alarm, etc.

The above is a description of each component in Figure 6. Of course, relevant steps for judging fire situation, relevant steps for judging toxic and harmful gas, and relevant steps for judging abnormal sound are at the same level, and the sequence can be changed.

5. Application Examples

5.1. Structure of the Experimental System. The schematic diagram of handheld terminal monitoring laboratory buildings is shown in Figure 7, and it allows multiple managers to monitor all laboratories at the same time. Assuming that there are mk_i-storey ($i = 1, 2, 3 \cdots, m$) CCEL buildings and q handheld terminals, every terminal can receive all experiment information in m buildings. That means the WSN

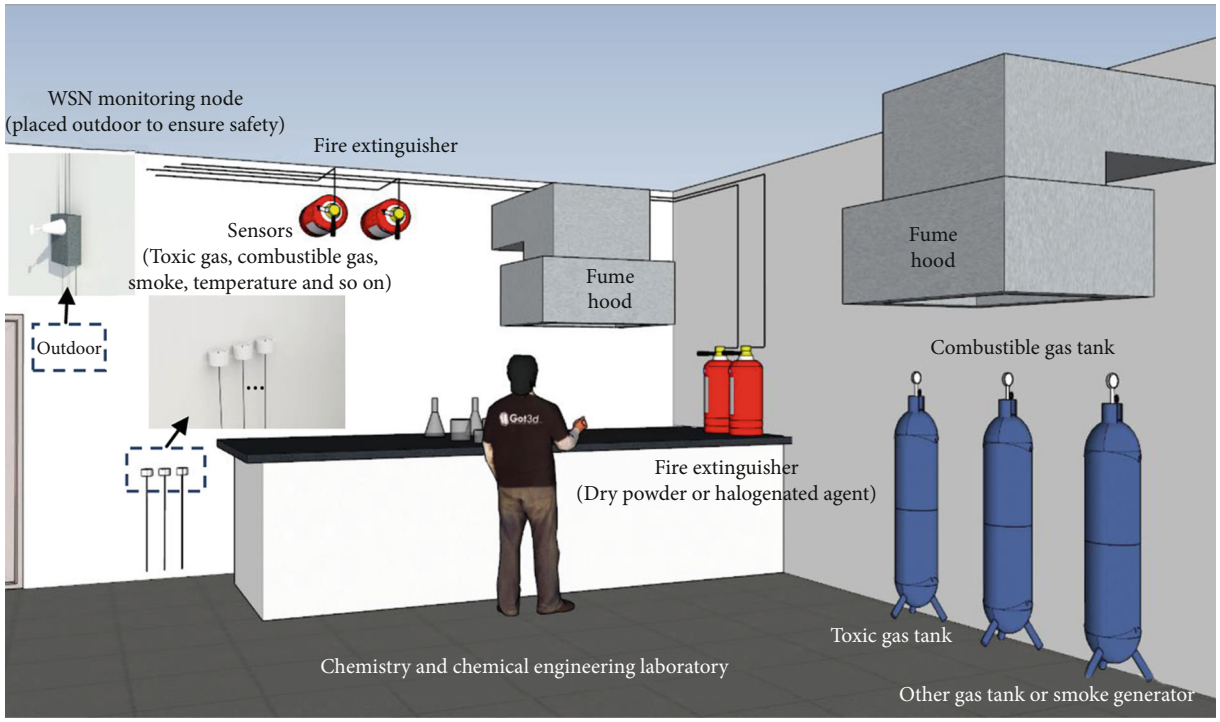


Figure 9: The experimental wireless monitoring system diagram for CCEL.

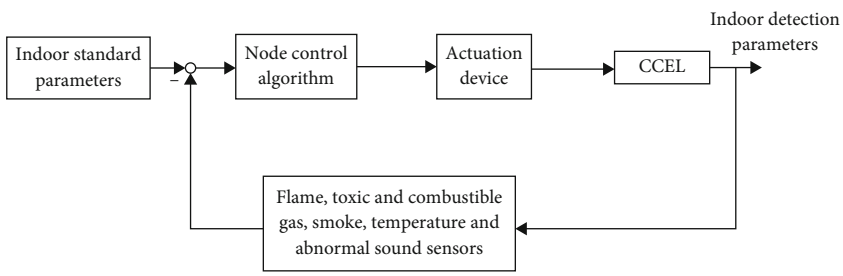


Figure 10: Structure diagram of measurement and control system in CCEL.

nodes (including the external sensors) and corresponding protective devices are fixed in m buildings.

5.2. Testing of Security Monitoring Node in CCEL. For experimental verification, m monitoring nodes are, respectively, installed in m laboratories, the WSN node consists of several sensors, corresponding protective device, alarm system, and so on. The WSN monitoring node used in this experiment is shown in Figure 8.

The wireless monitoring node has internal and external sensors. The internal sensors are integrated on the monitoring node, such as temperature sensor and humidity sensor; they are applied to detect temperature and humidity inside this node. These external sensors are installed in suitable positions around the experimental devices. These sensors are connected to a wireless monitoring node via signal wires.

We put combustible gas, toxic gas, and smoke generators and the security monitoring node into one of the laboratories, then open the generators, as shown in Figure 9, which is the experimental wireless monitoring system diagram for CCEL. The system includes extinguishers, fume hood, WSN nodes, and gas generator. The WSN node is placed out of the laboratory to prevent the high temperature and fire damage. The signal line between sensors, extinguishers, and WSN node is coated with fireproof material. The fume hood can be used to release toxic and combustible gas when the gas concentration in the room exceeds the standard. The WSN node would alarm at the same time when the CCEL is in danger. Figure 10 illustrates the structure diagram of the measurement and control system in CCEL.

When we activate the monitoring system, the sensors start to monitor the laboratory data. And WSN node would transmit the data to and the handheld terminal and feed it back to the control system. If the detection parameters exceed the indoor standard parameters, the WSN node would control the actuation device to take a series of measures for the CCEL. For example, when there is a fire in the laboratory, the safety monitoring node turns on the fire extinguisher. The extinguisher sprays fire agent to the experiment device

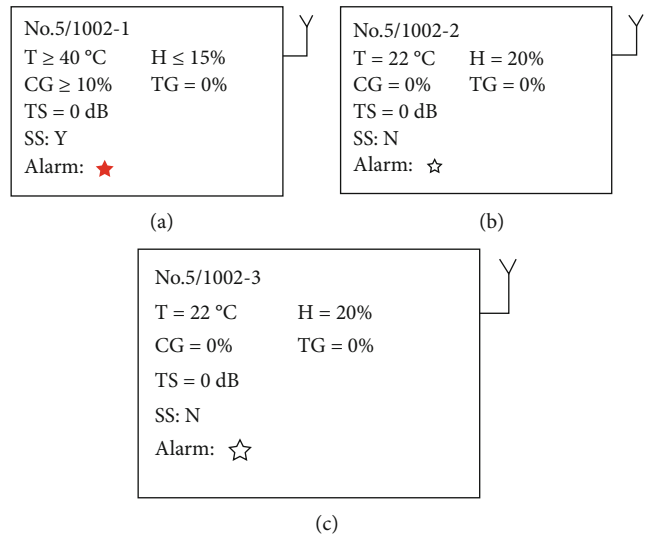


Figure 11: Testing experimental example 1 on the interface of the handheld terminal.

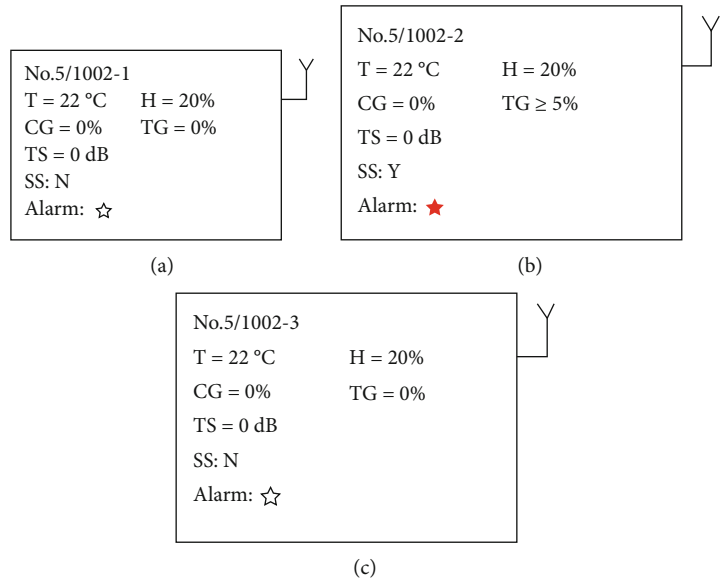


Figure 12: Testing experimental example 2 on the interface of the handheld terminal.

in order to minimize the loss of laboratory equipment and to protect the safety of indoor staff.

5.3. Remote Testing Experiments Based on Handheld Terminal. This section introduces remote testing experiments based on handheld terminals. The remote test experiments in this paper were carried out in three laboratories, respectively. They are numbered 1002-1, 1002-2, and 1002-3, and each laboratory covers an area of 50 square meters. When the indoor safety monitoring node detects smoke, flame, toxic gas leakage, and abnormal sound, the field information will be sent to the handheld terminal. The terminal will display the specific location of the alarm (i.e., which building? Which room? And the parameters indicated by the alarm (i.e., smoke, fire, gas leakage, and abnormal sound). Alarm information will also be displayed on the handheld terminal, as shown in Figures 11–13.

5.3.1. Test Experiment Example 1. This experiment was conducted in the laboratories of 1002-1, 1002-2, and 1002-3, Teaching building No. 5, respectively. Let T stands for temperature, H for relative humidity, CG for combustible gas concentration, TG for toxic gas concentration, and TS for decibel value. SS stands for confirmation status (i.e. SS: N means normal; SS: Y means there is a dangerous accident). Alarm indicates the state of the Alarm (i.e., no sound on the Alarm indicates normal; a loud or flashing Alarm indicates a dangerous accident). When the concentration of combustible gas is measured to be greater than or equal to 10%, and the indoor temperature rises and the relative humidity

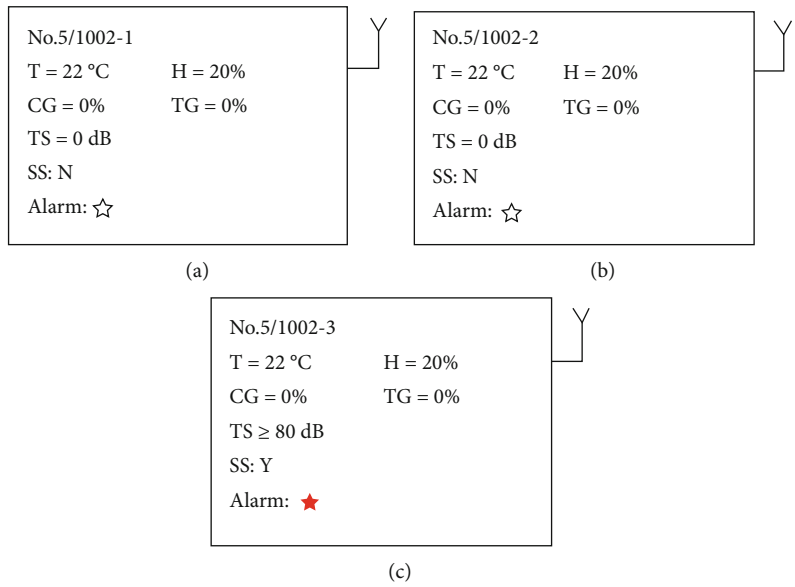


Figure 13: Testing experimental example 3 on the interface of the handheld terminal.

drops, the indoor combustion accident is considered to have occurred. The display information of the handheld terminal is shown in Figure 11, and its alarm information is shown in Figure 11(a).

(1)Automatic Protection Measures. Turn on the indoor camera, point the extinguisher at the ignition point, and spray dry powder extinguishing agent (or safety extinguishing agent) to quickly extinguish the ignition point. At the same time, send out an alarm signal to the indoor staff or nearby staff to evacuate the scene quickly.

5.3.2. Test Experiment Example 2. This experiment was conducted in the laboratories of 1002-1, 1002-2, and 1002-3, Teaching building No. 5, respectively. Let T stands for temperature, H for relative humidity, CG for combustible gas concentration, TG for toxic gas concentration, and TS for decibel value. SS stands for confirmation status (i.e., SS: N means normal; SS: Y means there is a dangerous accident). Alarm indicates the state of the Alarm (i.e., no sound on the Alarm indicates normal; a loud or flashing Alarm indicates a dangerous accident). When the toxic gas concentration is measured to be greater than or equal to 5%, it is considered that a toxic gas leak has occurred in the room. The display information of the handheld terminal is shown in Figure 12, and its alarm information is shown in Figure 12(b).

(1)Automatic Protection Measures. Open the indoor fume hood or ventilator and at the same time send an alarm signal for indoor or nearby workers to evacuate the scene quickly.

5.3.3. Testing Experimental Example 3. This experiment was conducted in the laboratory of 1002-1, 1002-2, and 1002-3, Teaching building No. 5, respectively. Let T stands for temperature, H for relative humidity, CG for combustible gas concentration, TG for toxic gas concentration, and TS for decibel value. SS stands for confirmation status (i.e., SS: N means normal; SS: Y means there is a dangerous accident). Alarm indicates the state of the Alarm (i.e., no sound on the Alarm indicates normal; a loud or flashing Alarm indicates a dangerous accident). When the decibel value of the measured abnormal sound is greater than or equal to 80 dB (in this paper, it is assumed that the noise of more than 80 dB produced in the laboratory is an abnormal sound), it is considered that an explosion or an abnormal accident has occurred in the room. The display information of the handheld terminal is shown in Figure 13, and its alarm information is shown in Figure 13(c).

(1)Automatic Protection Measures. Turn on the indoor camera and activate the isolation device near the accident point. At the same time, send an alarm signal for the indoor staff or nearby staff to evacuate the scene quickly.

When doing experiments in the chemistry and chemical engineering laboratory, safety is of the first importance. Here, we give some safety measures that must be paid attention to during the experiment:

(1) When conducting fire experiment, it is necessary to have reliable fire extinguishing equipment and measures

(2) When conducting toxic gas experiments, the experiments can be carried out in a small space first, with reliable ventilation equipment and protective measures

(3) The isolation device and protective measures should be provided when conducting abnormal sound experiments

All in all, researchers should avoid major accidents during the experiment and must have protective equipment and safety measures.

6. Conclusion

In view of the existing safety problems in chemical and chemical engineering laboratories, we propose a CCEL indoor safety wireless monitoring and automatic protection system. The testing results demonstrate that the main advantages of this system can be summarized as follows:

(1) The system proposed can achieve all-weather real-time monitoring of indoor safety information in CCEL, ensure that teaching experiments and scientific research experiments can be conducted continuously and safely, and significantly reduce the work intensity of laboratory managers. When the accident occurred, they were able to take effective measures to deal with it quickly and report relevant information to relevant departments in time, so as to reduce the damage caused by emergencies

(2) Because this system can in real-time monitor the safety status of CCEL, timely and accurately obtain the fire, toxic and harmful gas, abnormal sound, and other danger information of CCEL. Based on the acquisition of on-site monitoring data and image message, and it can remotely, wirelessly, and rapidly control the automatic protection device in CCEL. When an accident occurs, it will do its best to protect students and teachers working in the lab from harm and minimize damage to laboratory equipment

(3) The application of the system will solve a difficult problem in CCEL, and the system can be used for CCELs (such as toxic and hazardous materials warehouses and flammable and explosive places) of colleges, research institutes, vocational colleges, and colleges. Therefore, the system has a good application prospect

The limitations of the method proposed in this paper and the prospect of future research work are as follows:

(1) The specification standard of automatic fire extinguishing equipment applicable to the "gunpowder and explosives" chemistry and chemical engineering needs further study. A variety of automatic fire extinguishing methods are further studied, and the experimental results are analyzed and compared. The selection of fire-extinguishing materials (i.e., fire-extinguishing agents) suitable for use in "gunpowder and explosives" chemistry and chemical engineering laboratories is required

(2) It is required to select a sensor that can quickly detect the toxic gas content so as to accurately find the leakage point of toxic gas. Once the toxic gas is detected, it is required to have a quick method to cut off the source of toxic gas. The effective methods and measures of automatic protection of experimental personnel and experimental equipment need to be further studied and improved

(3) When an explosion occurs in the laboratory, the system can automatically activate the protective device and quickly isolate the explosion site, so as to minimize casualties and property losses. These works still need to be further studied and improved

Data Availability

The effectiveness and availability of the proposed system are verified by the experimental results. Therefore, the experimental data supporting the results of this study (the experimental data type is number, i.e., including integer and decimal numbers) can be found in Section 5.3 of the manuscript or obtained from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Xiang-Guang Chen and Xue-Jiao Liu constructed the experimental system and designed the experimental scheme and carried out many experiments on the system in the laboratory. Qiao-Yan Sun and Xiang Han participated in several experiments of the system and studied the scheme of the automatic protection system. Yu-mei Sun and Mei-chun Wang reviewed references and materials (including recent years) related to the experiments of this system.

Acknowledgments

This scientific work is financially supported by the Yantai Double Hundred Plan Talent Project (YT201803) in 2018 and the Natural Science Foundation (No. ZR2016FM28) of Shandong Province in 2016.

References

[1] M.-Q. Bao, Y. Zhang, and S.-C. Zhang, "Analysis of college chemistry laboratory safety problems and exploration of safety management measures," *Experimental Technology and Management*, vol. 29, no. 1, pp. 188–191, 2012.

[2] Y.-q. Liu, H.-g. Bao, and S.-j. Song, "Research on University Laboratory Safety Management," *Experiment Science & Technology*, vol. 10, no. 1, pp. 173–175, 2012.

[3] N.-w. Qiu, X.-c. An, J.-w. Jia, and Z. Kun, "Challenges and solutions confronted with for laboratory safety management in colleges and universities during new period," *Experimental Technology and Management*, vol. 29, no. 1, pp. 181–185, 2012.

[4] F. Jian, X. Wei, Q. Huidong, and L. Yanhui, "Exploration on strengthening the safety of chemistry and chemical engineering laboratory," *Guangdong Chemical Engineering*, vol. 43, no. 6, pp. 209-210, 2016.

[5] Q. Huidong, L. Guang, W. Wei, and L. Yanhui, "Comparison 0n safety management of chemistry and chemical engineering laboratories between universities of China and United States," *Experimental Technology and Management*, vol. 31, no. 4, pp. 2013–2015, 2014, 209.

[6] Q. Qiu, Z. K. Luo, W. Z. Lu, and W. Fang, "Research of chemical laboratory safety management in domestic and abroad universities," *Guangzhou Chemical Industry*, vol. 38, no. 5, pp. 272–274, 2010.

[7] L. Zhi-Ming, C. Xiang-Guang, W. Lei, and W. Yanfu, "Fault diagnosis method for wireless security monitoring system in chemistry and chemical engineering laboratory," *Transactions of Beijing Institute of Technology*, vol. 35, no. 10, pp. 1062–1066, 2015, 1073.

[8] W. Ping and W. Zheng, "Design and implementation of open computer lab monitoring and management system," *Computer and Modernization*, vol. 11, pp. 125–128, 2007.

[9] W. Nie, J. Xu, and Y. Liu, "Design of automatic closed system about gas outburst monitor and its role of the reducing security cost," *Proceedings of the 2nd International Conference Mine Hazards Prevention and Control*, , pp. 538–543, Atlantis Press, Paris, France, 2010.

[10] W. Nie, Y. Liu, and J. Xu, "A gas monitoring and control system in a coal and gas outburst laboratory," *Journal of Sensors*, vol. 3, 11 pages, 2014.

[11] V. Ramya, B. Palaniappan, and V. Sumathi, "Gsm based embedded system for remote laboratory safety monitoring and alerting," *International Journal of Distributed and Parallel Systems (IJDPS)*, vol. 3, no. 6, pp. 31–49, 2012.

[12] J. Linyiang, L. Hailong, and G. Zhenhua, "Embedded laboratory environment monitor system," in *WASE International Conference on Information Engineering*, pp. 197–201, Yaiyuan, China, 2009.

[13] L. Yang, L. Jiang, K. Yue, and P. Heming, "Design and implementation of the lab remote monitoring system based on embedded web technology," *International Forum on Information Technology & Applications*, vol. 2, pp. 172–175, 2010.

[14] G. Zhao, Z. Wang, W. Li, and K. Wang, "An embedded laboratory security monitoring system," in *Third International Conference on Measuring Technology and Mechatronics Automation*, pp. 395–398, Shanghai China, 2011.

[15] G. Ying, C. Yan, and H. Fei, "Research of laboratory remote monitoring and early warning system based on wireless sensor network," *Research & Exploration in Laboratory*, vol. 35, no. 3, pp. 105–109, 2016.

[16] L. Yu-yu, "Exploration of safety management for chemical laboratories in universities," *Laboratory Science*, vol. 18, no. 4, pp. 203–205, 2015.

[17] P. Yanling, "Construction study of the safety management system for chemical engineering laboratory," *Modern Chemical Research*, no. 2, pp. 11-12, 2017.

[18] F. E. N. G. Shou-chun, "Chemistry and chemical engineering laboratory safety management and safety education," *Laboratory Science*, vol. 17, no. 3, pp. 196–198, 2014.

[19] H. Xue, L. Yingling, Y. Guoqiang, and F. Guangzhu, "Exploration on the safety construction of chemistry and chemical engineering laboratory," *Guangdong Chemical Industry*, vol. 43, no. 13, p. 296, 288, 2016.

[20] L. Hai-Bin and J. Jie, "Safety management of chemical engineering laboratory in colleges and universities," *Anhui Chemical Industry*, vol. 49, no. 2, pp. 142–144, 2019.

[21] L. Yan, R. Shun-Lin, and H. Dong-Xian, "Exploration on safety management mechanism of chemical laboratories in universities," *Guangzhou Chemical Industry*, vol. 48, no. 5, pp. 198-199+211, 2020.

[22] L. Shiping, X. Aijuan, X. Bing et al., "Construction and practice of safety access system for laboratories of chemical engineering majors," *Experimental Technology and Management*, vol. 36, no. 11, pp. 11–14, 2019.

[23] G. Quanhui, T. Yanting, L. Juan, Z. Hui, and M. Xinqi, "Exploration on safety management of university chemical engineering laboratory under new normal situation," *Experimental Technology and Management*, vol. 34, no. 10, pp. 249–251, 2017.

[24] Z. Jìng, S. Qiao-Fang, J. Wu, and L.-X. Wang, "Construction and management of chemistry and chemical engineering laboratory in colleges and universities under the new situation," *Chemical Industry Times*, vol. 34, no. 2, pp. 39–41, 2020.

[25] W. Wu and C. Jian-Feng, "Exploration on the safety management system chemical laboratory in colleges and universities under the new situation," *Chemical Enterprise Management*, vol. 10, pp. 21-22, 2020.

[26] B.-J. Wan, "Analysis and preventive measures of safety hidden danger in chemical and chemical laboratories of local universities," *Industrial & Science Tribune*, vol. 18, no. 11, pp. 234-235, 2019.

[27] Z. Bao-han, T. Qiang, W. Wei et al., "Comparison of safety management between Chinese and Norwegian chemical laboratories," *Experimental Technology and Management*, vol. 27, no. 4, pp. 49-50, 2019.

[28] Z. Dong, M. Kai-Yuan, J. Nan, L. Ting-Ting, and L. Sen-Lín, "A brief analysis on the construction of university laboratory safety management system based on information platform," *China Inspection Body & Laboratory*, no. 4, pp. 49-50, 2019.

[29] Y.-m. Sun, X.-j. Liu, X.-g. Chen, Q.-y. Sun, and Z. Jun, "Research and application of a fault self-diagnosis method for roots flowmeter based on WSN node," *Wireless Personal Communications*, vol. 59, no. 3, pp. 2315–2330, 2017.

[30] Y.-M. Sun, X.-J. Liu, Q.-Y. Sun, and X.-G. Chen, "Research and application of wireless security monitoring system in chemistry and chemical engineering laboratory," *Wireless Personal Communications*, vol. 59, no. 3, pp. 2331–2344, 2017.

[31] L. Zhi-Min and C. Xiang-guang, "Research and application of WSN node communication energy consumption model based on asynchronous MAC protocol," *Transactions of Beijing Institute of Technology*, vol. 35, no. 2, pp. 171–175, 2015.

[32] L. Zhimin and X. Cheng, "Research on fault diagnosis of wireless sensor node on module level," *Chinese Journal of Scientific Instrument*, vol. 34, no. 12, pp. 2763–2769, 2013.

[33] X. Ying-Xin, C. Xiang-Guang, and Z. Jun, "A double weighted LS-SVM model for data estimation in wireless sensor networks," *Journal of Beijing Institute of Technology*, vol. 21, no. 1, pp. 134–139, 2012.

[34] Z. Jun, C. Xiang-Guang, L. Chun-Tao, Y. Xiang-Ming, and Y. Bin, "Study on application of distributed multi-channel MAC protocols in WSN," *Acta Armamentarii*, vol. 33, no. 6, pp. 695–701, 2012.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6639894, 12 pages
https://doi.org/10.1155/2021/6639894

Research Article

A Novel Subcarrier-Level Spectrum Sensing Method by Utilizing Fine-Grained Channel State Information in Wireless Networks

Zheng Lu[1,2] **Handong Wang**,[1,2] **Hongyu Sun**[1,2] **Chin-Ling Chen**,[3,4,5] **and Zhenjiang Tan**[1,2]

[1]*Department of Computer Science, Jilin Normal University, Siping 136000, China*
[2]*State Key Laboratory of Numerical Simulation, Jilin Province, Siping 136000, China*
[3]*Department of Computer Science and Information Engineering, Chaoyang University of Technology, Taizhong 41349, Taiwan*
[4]*School of Information Engineering, Changchun Sci-Tech University, Changchun 130000, China*
[5]*School of Computer and Information Engineering, Xiamen University of Technology, Xiamen 360000, China*

Correspondence should be addressed to Chin-Ling Chen; clc@mail.cyut.edu.tw and Zhenjiang Tan; tanzj@jlnu.edu.cn

Received 31 December 2020; Revised 23 February 2021; Accepted 12 March 2021; Published 7 May 2021

Academic Editor: Feng-Jang Hwang

Copyright © 2021 Zheng 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.

Traditionally, the channelization structures of wireless technologies (802.11/ZigBee/BLE) have been fixed. Each node content for the spectrum is assigned one channel with a specific bandwidth. However, classical channel-based spectrum sensing and sharing algorithms have great limitations to further optimize spectrum utilization when multiple IoT with different wireless technologies coexisting in the same environment. Therefore, exploring the fine-grained spectrum sensing algorithm becomes an essential work to further improve the spectrum utilization efficiency, especially in the Industrial Scientific Medical (ISM) band. This paper proposes *Subcarrier-Sniffer*, a novel subcarrier-level spectrum sensing and sharing method, which utilizes channel state information (CSI) to sense the fine-grained status of each subcarrier of the traditional channel. To evaluate the performance of *Subcarrier-Sniffer*, we implemented *Subcarrier-Sniffer* by USRP B200min, and the experimental results show that the accuracy of subcarrier-level spectrum sensing could achieve 100% in our settings that the distance between *Subcarrier-Sniffer* and the monitor is not greater than 7 m. *Subcarrier-Sniffer* could be applied in WiFi and ZigBee, WiFi and BLE, and WiFi and LTE-U coexisted environments for better spectrum utilization.

1. Introduction

Most of the wireless technologies today operate with the default bandwidth, for example, the default bandwidth of 802.11n is 20 MHz, while the default bandwidths of ZigBee and BLE are 3 MHz and 2 MHz, respectively. However, the channels are partially overlapped in ISM 2.4 GHz which hampers further optimize the limited spectrum resource. Moreover, the ever-increasing of the heterogeneous IoT devices coexisting in interfered environments also motivated the developing spectrum sharing and management techniques [1]. The spectrum sensing and sharing algorithms are widely used in a variety of wireless technologies, for example, DeepWiFi [2] for WiFi spectrum sensing, OutSense [3] for ZigBee spectrum sensing, and RealSense [4] for TV spectrum sensing. Spectrum sensing and sharing methods are also used in cross-technology coexisted environments, for example, literature [5–11] conducts the spectrum sensing and management between WiFi and LTE-U technologies, literature [12–15] introduces the spectrum sensing between ZigBee and WiFi technologies, [16–21], uses spectrum sensing method for OFDM-based heterogeneous wireless devices.

However, most of the current work uses a channel-based spectrum sensing algorithm to share ISM band resources [22]. For example, WiFi defines 11 channels with 20 MHz bandwidth, ZigBee defines 26 channels with 3 MHz bandwidth, and Bluetooth 4.0 defines 30 channels with 2 MHz bandwidth in the 2.4 GHz ISM bands. Therefore, the channel bandwidths

of WiFi, ZigBee, and BLE are 20 MHz, 3 MHz, and 2 MHz, respectively. It is obvious that channel-based contention methods between heterogeneous wireless technologies deployed in the interfered environment could not further optimize spectrum utilization, Figure 1 illustrates the spectrum wasting between the narrow-wide band and wide-wide band technologies coexisting scenarios.

Figure 1(a) shows the disadvantage of channel-based spectrum sharing methods when narrow-wide band wireless technologies coexisted in the same interfered environment since both of them trying to access the idle channel which practically overlapped in the spectrum, and only one of them could obtain the channel. Suppose the bandwidth of the wideband technology is x MHz and the bandwidth of the narrowband technology is y. Thus, the wasted spectrum is $(x - y)$ MHz when the narrowband technology allocating the channel. Figure 1(b) shows the spectrum wasting when two different wideband technologies content the same channel, and assume that there n MHz of the subcarriers are idle in the first wideband technology, and m MHz subcarriers are idle in the second wideband technology. If the idle subcarriers are not overlapping, the total maximum spectrum wasting is $(m + n)$ MHz. Therefore, it is obvious that the subcarrier-level spectrum sensing and sharing method could further optimize the spectrum efficiency in ISM bands when massive heterogeneous wireless technologies are deployed in the same interfered environments.

Since the channel-level spectrum sensing methods determine whether the channel is idle by measuring the signal-to-noise ratio (SNR) changes (such as SNR and RSS) of the channel's frequency center, the subcarrier-level spectrum sensing methods measure fine-grained SNR in each subcarrier to determine the status of the fine-grained spectrum. Therefore, the design challenges introduced by the fine-grained subcarrier-level spectrum sensing and sharing methods are at least as follows: (i) how to perform energy-efficient subcarrier-level spectrum sensing algorithm without affecting normal communications. The traditional method uses full-spectrum scanning to sense the energy variation of each subcarrier, which affects normal communication and requires an expensive device with a relatively high sampling rate. To solve this challenge, this paper proposes to use channel state information (CSI) to sense the subcarrier-level SNR sensing, (ii) how to remove the noise jitters of each subcarrier caused by the frequency selective characteristics of the signal; therefore, for the wideband end, how to obtain the pure SNR jitters caused by the narrowband signal interference is a significant challenge; to solve this problem, this paper collects the initial state of CSI first to fetch the environmental information without cross-technology interference, and (iii) how to determine which subcarriers do have serious interference according to the SNR jitters generated by the interference is another design challenge of subcarrier-level spectrum sensing algorithm. The main reason for the problem is that if the narrowband signal is 3 MHz, the interference range (IR) caused by harmonics usually satisfies IR = $\alpha + 3 + \beta$MHz, where α and β are, respectively, the influence on the left and right spectrum range of 3 MHz. To solve this problem, a method for delimiting the range of interfered subcarriers is proposed.

Specifically, the contributions of this paper are detailed as follows:

(1) We first propose to use channel state information collected from the off-the-shelf wireless devices to evaluate the subcarrier interference and guide the spectrum sensing and sharing algorithm for optimized spectrum utilization

(2) We propose a denoise method base on the initial state of the CSI profile to recognize the narrowband utilization in a more precise way

(3) We propose a threshold-based interference range determination method to identify the interfered subcarriers when wideband and wideband coexist in the propagation scenarios

(4) We implemented the system using USRP B200min and analyze the performance of *Subcarrier-Sniffer*, and the experimental results show that the accuracy of *Subcarrier-Sniffer* could achieve 100% when the distance between *Subcarrier-Sniffer* and the monitor is not greater than 7 m

The rest of this paper is organized as follows: Section 2 introduces the related works, Section 3 introduces the theory of our method, Section 4 shows the evaluation results, and the conclusion is in Section 5.

2. Related Works

To further optimize the spectrum utilization in the ISM band, researchers have explored various nonchannel-based spectrum sensing and allocation methods. Subcarrier-Sniffer is closely related to three broad areas: (i) fine-grained channel access methods, (ii) adapted width channel access methods, and (iii) narrow-wide band coexistence methods.

2.1. Fine-Grained Channel Access Methods. FICA [23] splits a channel into multiple subchannels and allows devices content for subchannels, which could reduce the MAC-layer overhead of high-rate WLANs. It uses a frequency-domain back-off algorithm that cannot coexist directly with current 802.11 WLANs. Moreover, FICA requires a tight synchronization between all nodes contend for spectrum, which could not retain the distributed mechanism in CSMA/CA protocol. FSS [24] proposes a fine-grained spectrum sharing method for fair and efficient spectrum utilization between existed and emerging 802.11 protocols with various channel widths. It allows users to contend for each spectrum chunk and opportunistically split a wideband channel or bond multiple discontinuous chunks to ensure fair and efficient access to the available spectrum. The authors of [25] propose to divide a frame into multiple miniframes and divide the channel into multiple minichannels; they use a greedy spectrum assignment algorithm to determine where to place the miniframes, without considering the guard band cost of noncontiguous minichannels.

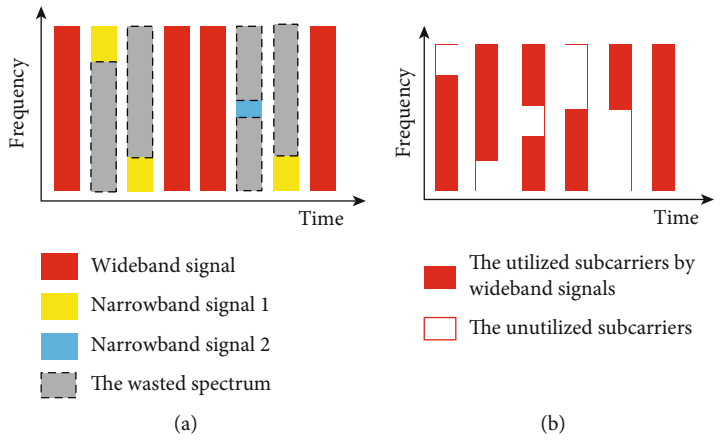


Figure 1: Spectrum wasting under channel-based sensing methods: (a) narrow-wide band technologies coexisting scenarios and (b) wide-wide band technologies coexisting scenarios.

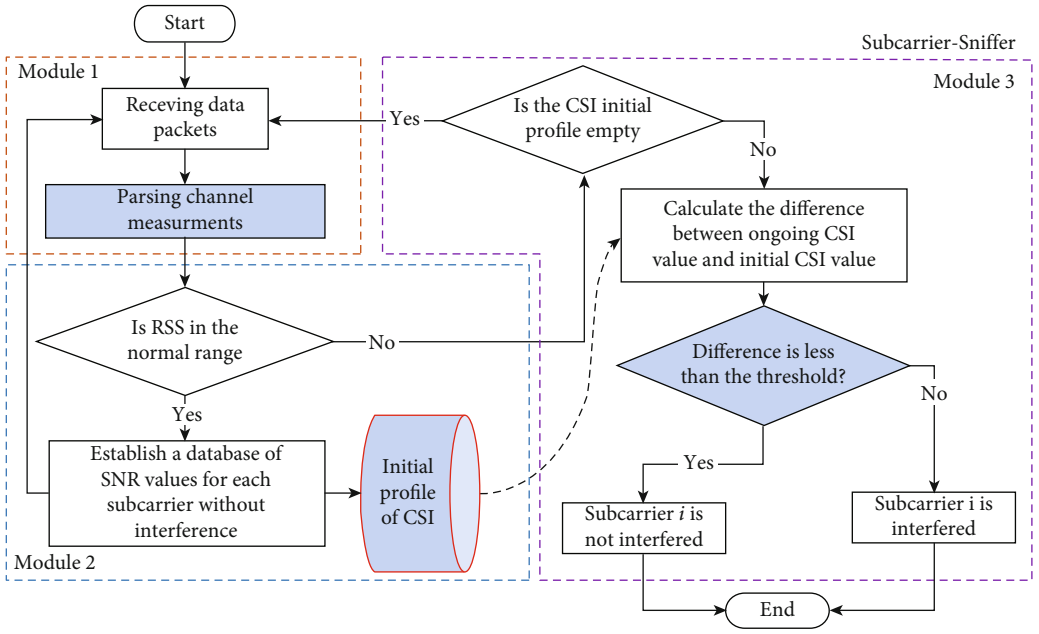


Figure 2: Design overview of Subcarrier-Sniffer.

2.2. Adaptable Width Channel Access Methods. The proposal of IEEE standards such as 802.11n [26] and 802.11ac [27] suggested changing channel width adaptively for spectrum utilization. FLUID [28] uses flexible channelization, which allows each transmitter to choose the appropriate channel width and center frequency for transmission. The experimental results show that flexible channelization can improve the throughput of the system. The authors of [29] modeled the adaptive width channel allocation from a game-theoretic point of view, in which the node is rational and always pursue their objectives. The authors of [30] explored the benefits of adapting channel width to balance the trade-off between through and energy efficiency. Article [31] proposes to share the spectrum according to throughput requirements and signal strength values of a specific node. [32] proposed to change both the frequency center and bandwidth to match network traffic loads.

2.3. Narrow-Wide Band Coexistence Methods. As the deployment of heterogeneous IoT devices in an indoor environment, researchers proposed several narrow-wide band coexistence methods to optimize spectrum utilization in ISM bands. SWIFT [32] proposes to enable the coexistence between the OFDM-based ultrawideband (UWB) system and the WiFi WLANs that have a relatively narrower bandwidth. LASI [11] enables simultaneous transmission between LTE and WiFi in ISM 5G bands by low-amplitude injection methods. B^2W^2 [14] designed N-way concurrent cross-technology communications while concurrently support the original WiFi to WiFi and BLE to BLE communications among multiple WiFi and BLE devices. ECT [33] proposed a network layer design for WiFi and ZigBee coexisted networks for reducing packet delivery delay in IoT networks.

In summary, algorithms have been proposed to solve the spectrum crisis problems. However, most of the work

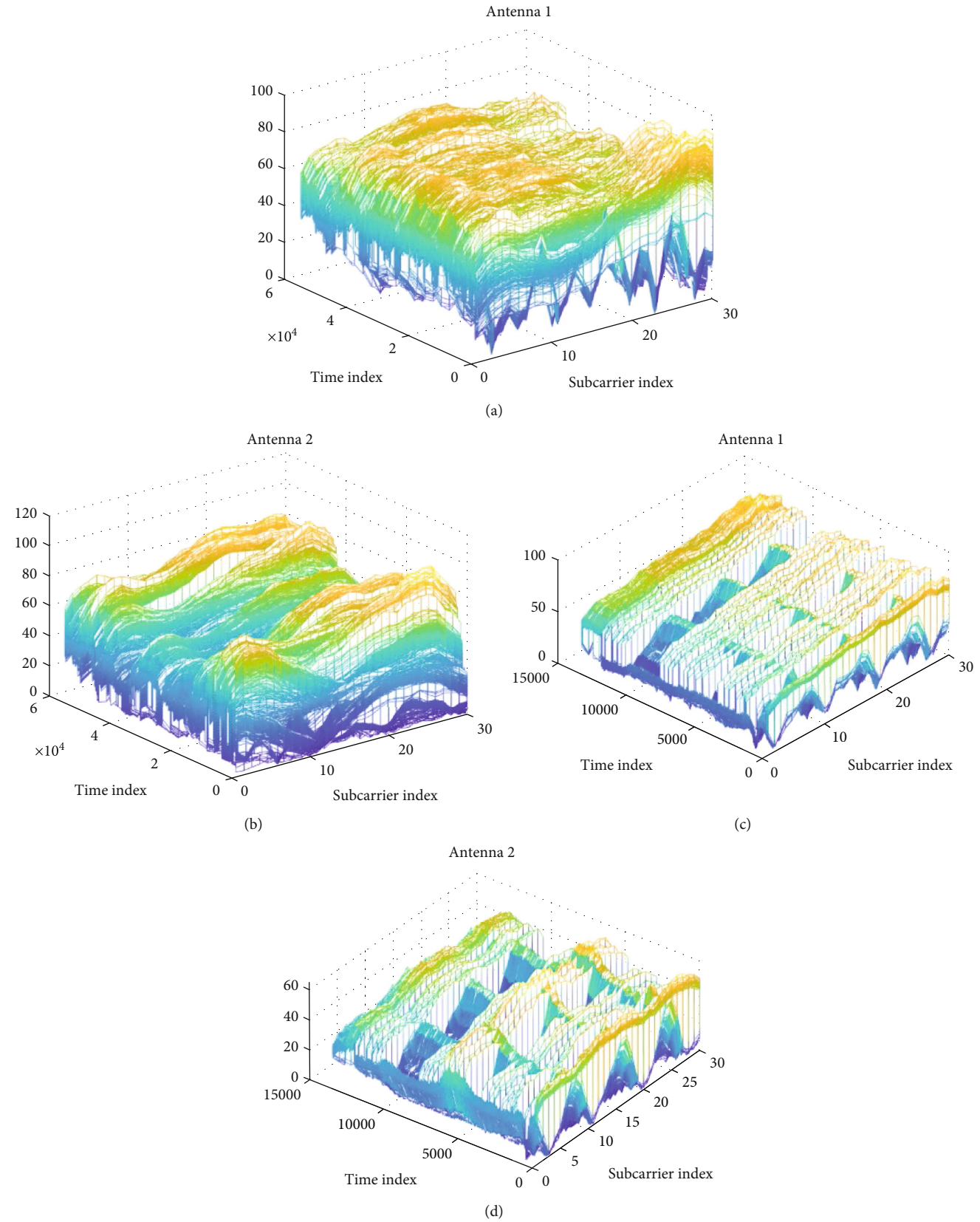


Figure 3: Examples of using CSI to sense subcarrier-level interference: (a) SNRs of antenna 1 calculated by CSI without interference; (b) SNRs of antenna 2 calculated by CSI without interference; (c) SNRs of antenna 1 calculated by CSI with three ZigBee interferences; (d) SNRs of antenna 2 calculated by CSI with three ZigBee interferences.

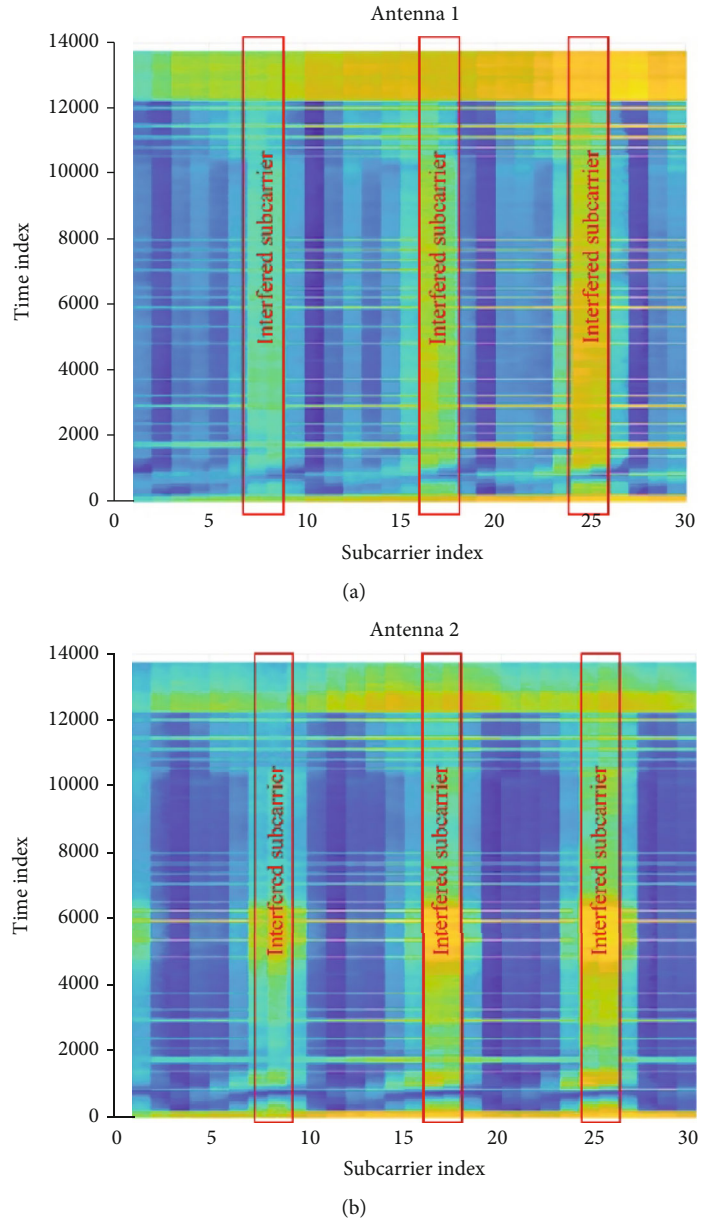


Figure 4: The difference between noninterfered subcarriers and interfered subcarriers: (a) data collected from antenna 1 and (b) data collected from antenna 2.

introduces additional monitoring systems or is not compatible with current IoT protocols. In this paper, we propose Subcarrier-Sniffier, which could utilize the off-the-shelf devices and network measurements to enable subcarrier-level spectrum sensing among the narrow-wide band or wide-wide band spectrum sharing scenarios.

3. Design of Subcarrier-Sniffer

According to the design challenges and our solutions, *Subcarrier-Sniffer* contains three parts: (i) channel state information collection and analysis module, (ii) CSI initial states establish and denoising module, and (iii) interfered range of target subcarrier determination module. The design overview of *Subcarrier-Sniffer* and the data stream among the three are shown in Figure 2.

The main functions of the three modules are shown as follows:

(1) *Channel State Information Collection and Analysis Module (Module 1)*. This module is responsible for collecting the channel measurements from off-the-shelf wideband wireless devices, such as the device that runs WiFi or LTE technologies. Currently, the important measurements of wireless technologies such as WiFi, ZigBee, and BLE are RSS (received signal strength) and CSI (channel state information), in which CSI is a complex number that describes the

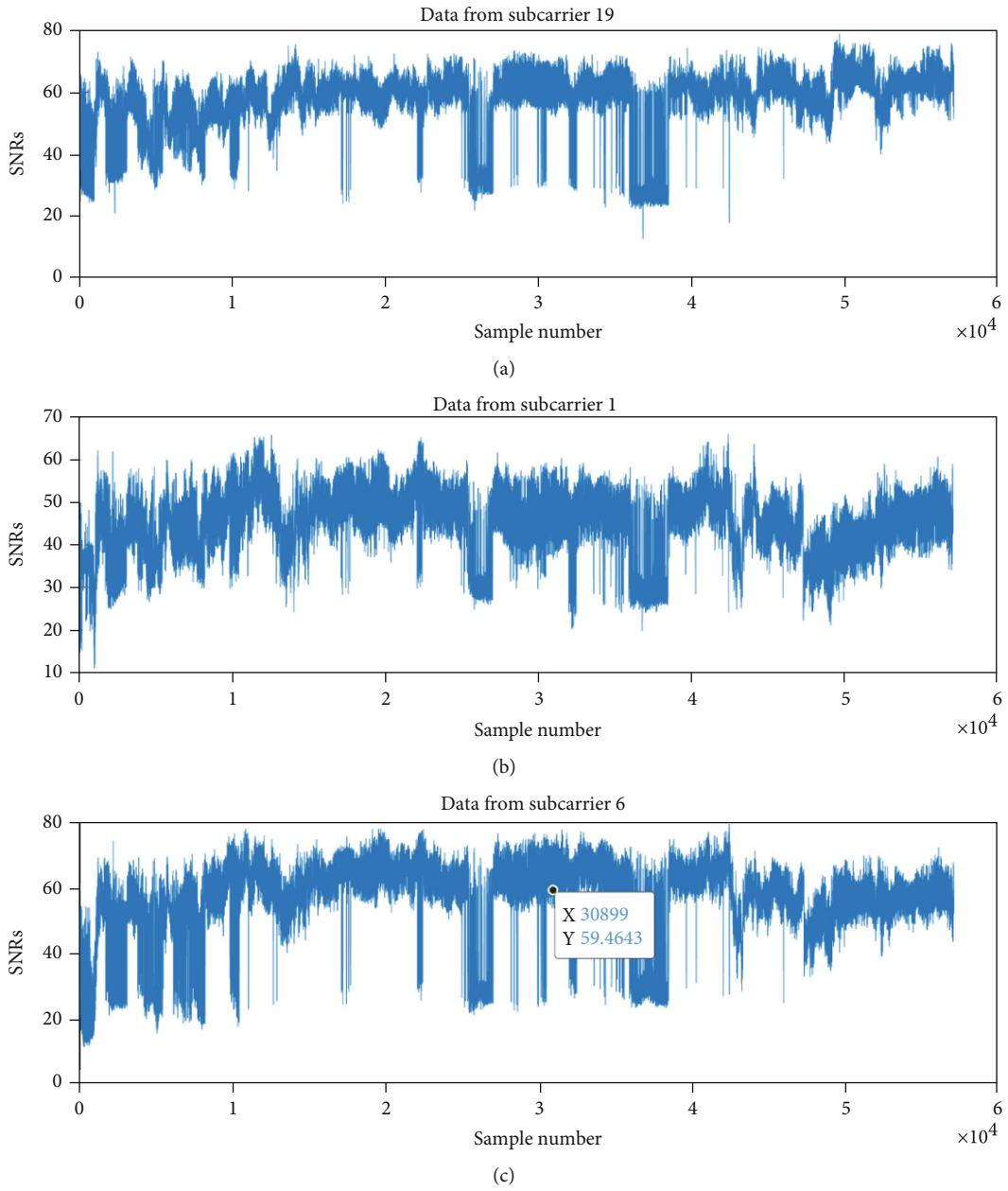


Figure 5: The variations of different subcarriers without any interference: (a) data from subcarrier 1; (b) data from subcarrier 6; (c) data from subcarrier 19.

phase and amplitude shift of arriving signals. The main procedures are detailed in Section 4 (1).

(2) *CSI Initial States Establish and Denoising Module (Module 2)*. This module is used for the initial CSI database establish by estimating noise caused by frequency selection characteristic when signal propagating in the indoor environments without any narrowband or wideband interference. The initial CSI database could provide a foundation to accurately fetch the pure subcarrier-level interference in normal communication links. The theory and procedures are detailed in Section 4 (2).

(3) Interfered range of target subcarriers determination module (module 3). This module is used to determine the range of interfered subcarriers according to the features of wireless measurements. The threshold-based methods were proposed to evaluate the interfered subcarriers more accurately. The procedures of this module are detailed in Section 4 (3).

3.1. Channel State Information Collection and Analysis Module. Channel state information collection and analysis module is to analyze the signal-to-noise ratios (SNRs) of the received signal for each subcarrier. The method is to

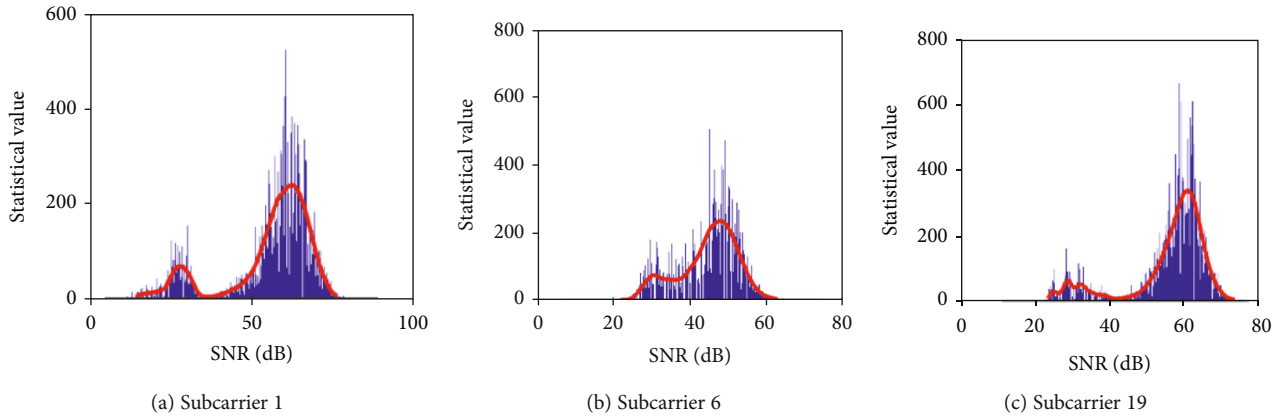


FIGURE 6: The initial state variations belong to different subcarriers: (a) the variations of subcarrier 1; (b) the variations of subcarrier 6; (c) the variations of subcarrier 19.

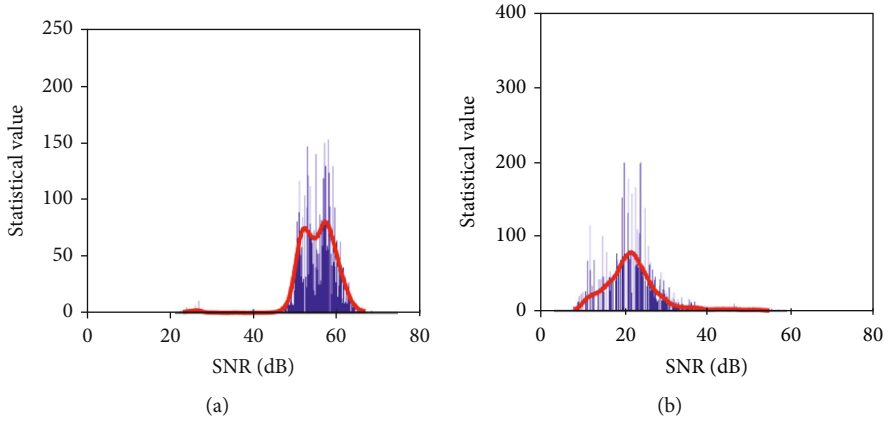


FIGURE 7: The initial state variations collected from different scenarios: (a) data from subcarrier 19 in indoor deployment and (b) data from subcarrier 19 in the outdoor deployment.

collect the CSI measurements to get the detailed information of each subcarrier, and the data structure of CSI received at a time t is shown in Equation (1) [14].

$$\text{CSI matrix}(t) = \begin{bmatrix} \text{csi}_{1,1}^t & \text{csi}_{1,2}^t & \cdots & \text{csi}_{1,m}^t \\ \text{csi}_{2,1}^t & \text{csi}_{2,2}^t & \cdots & \text{csi}_{2,m}^t \\ \text{csi}_{3,1}^t & \text{csi}_{3,2}^t & \cdots & \text{csi}_{3,m}^t \end{bmatrix}, \quad (1)$$

where each element in the matrix represents the j^{th} subcarrier from i^{th} communication link and each element in the CSI matrix is a complex number shown in Equation (2) [14].

$$\text{csi}_{i,j}^t = a + jb. \quad (2)$$

Therefore, at the time t, the signal-to-noise ratio (SNR) of j^{th}, the subcarrier of i^{th} communication link, can be calculated by Equation (3) [14].

$$\text{SNR}_{i,j}^t = \sqrt{a^2 + b^2}. \quad (3)$$

In order to verify that the CSI measurement could detect subcarrier-level SNRs, this paper uses Intel 5300 CSI acquisition tool to get CSI samples from wideband technology with interference from narrowband technology (ZigBee), and the experimental results are shown in Figure 3. The results show that the SNRs of the interfered subcarriers are different from those without any interference. Therefore, we propose to use CSI features to locate the interfered subcarriers when multiple devices content the same channel. The experimental results are analyzed from data via https://github.com/p01aris/csi_subcarrier_sniffer.

To observe the differences between interfered subcarriers and noninterfered clearly, Figure 4 is the two-dimension colormap of Figures 3(c) and 3(d). The interfered subcarriers are indicated by red rectangles showed in Figure 4, in which Figure 4(a) shows the data collected from antenna 1 while Figure 4(b) shows the data from antenna 2.

3.2. CSI Initial States Establish and Denoising Module. The experimental results in Figure 3 also show that the signal-to-noise ratios (SNRs) of different subcarriers also vary greatly without interference. The variations are different with different frequencies of each subcarrier. Figure 5(a) shows the variation of SNR series in subcarrier 1, while Figure 5(b)

shows the variation of SNR series in subcarrier 6, and Figure 5(c) shows the variation of SNRs series in subcarrier 19.

In summary, Figure 4 shows that the variations of SNRs of different subcarriers are various since the frequency selection features. Therefore, it is important to perceive the initial state of each subcarrier. The initial subcarrier states are also affected by the propagation environments.

Figure 5 shows the different variations without any interference from subcarrier 1, subcarrier 6, and subcarrier 19, respectively. The experimental results show that the variations belong to different subcarriers are various.

Figure 6 shows the statics value from subcarrier 1, subcarrier 6, and subcarrier 19, respectively, and the results further proved that the variations from different subcarriers are various without interference.

Figure 7 shows the statistical results of subcarrier 19 in indoor and outdoor deployment scenarios, respectively.

In summary, sensing the real-time initial state of each subcarrier is essential, and the specific procedures to fetch and store the initial CSI status are shown in Figure 2. If the RSS is in a normal state, recording the status of CSI as the initial state, the main reason is RSS seeks the average value of the received energy of the entire channel which could represent whether there is existing interference in the whole channel. Normally, when the jitter range of the RSS does not exceed 3 dB, it can be judged that there is no interference in the current time.

3.3. Interfered Range of Target Subcarrier Determination Module. Due to the existence of harmonic energy, the interference reflected from wideband technologies is wider than the narrowband interference. Figure 8 shows the practical interference caused by narrowband technology.

Figure 9 also shows the narrowband interference which is overlapped on the wideband signals. Therefore, the interference range determination process is significant to accurately locate the subcarrier-level interference when multiple heterogeneous wireless technologies coexist in the same deployment scenarios.

Figures 8 and 9 show that the energy generated by the harmonic portion also affects the wideband signal. Therefore, the subcarriers with harmonics also need to be considered as interfered subcarriers. To range the interfered subcarriers, the jitter value due to the interference should be calculated firstly by Equation (4).

$$\Delta \mathrm{snr}_{i,j}^{\tau} = \mathrm{SNR}_{i,j}^{\tau} - \mathrm{SNR}_{i,j}^{h}. \tag{4}$$

Suppose $\mathrm{flag}_{i,j}^{\tau}$ indicates whether subcarrier j on communication link i interferes at τ the time, $\mathrm{flag}_{i,j}^{\tau} = 0$ indicates that the subcarrier does not interfere, and $\mathrm{flag}_{i,j}^{\tau} = 1$ indicates that the subcarrier i interferes. The calculation method $\mathrm{flag}_{i,j}^{\tau}$ is as shown in Equation (5).

$$\mathrm{flag}_{i,j}^{\tau} = \begin{cases} 0\left(\left|\Delta \mathrm{snr}_{i,j}^{\tau}\right| < \varepsilon\right), \\ 1\left(\left|\Delta \mathrm{snr}_{i,j}^{\tau}\right| < \varepsilon\right), \end{cases} \tag{5}$$

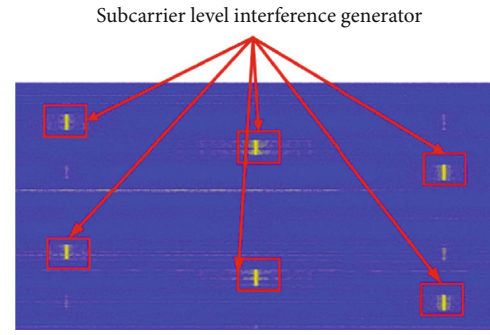


Figure 8: Interference of energy generated by harmonics on other subcarriers.

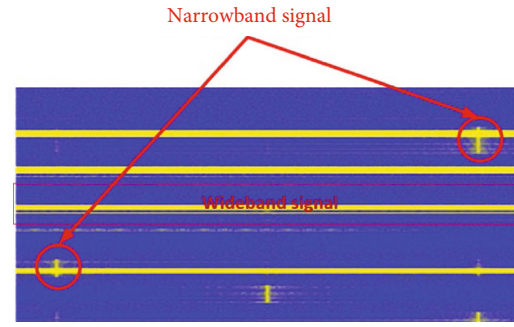


Figure 9: Interference caused by narrowband signals.

where ε is a threshold for determining whether subcarrier j on communication link i interfered. This value can be derived from the historical data of the communication error rate based on different values of ε. The experience value is around 6-10 dB.

4. Implementation and Experimental Results

This section implements the *Subcarrier-Sniffer* system and verified the experimental results according to the model in Section 4, we evaluated the performance of *Subcarrier-Sniffer* under different settings.

4.1. Implementation. The *Subcarrier-Sniffer* system contains three parts: (i) the wideband system transmitting and receiving end, the narrowband interfered signal generators, and the going-on signal monitor end. The configurations of the transmitting and receiving end of the wideband signals are shown in Table 1. The narrowband interfered signals are generated by USRP B200min. The going-on signal monitor end is implemented by USRP B200 and GNU radio platforms, and the structure of the monitor end is shown in Figure 10. The monitor end is to evaluate the performance of our proposed method, and the receiving end also could perform the monitoring function by using CSI value to judge the subcarrier-level interference in practical applications.

4.2. Experimental Results. To verify the performance of *Subcarrier-Sniffer*, this section evaluates the accuracy of the subcarrier-level spectrum sensing algorithm with the

Table 1: Wideband signal transmission and receiver configurations.

	Computer model	Network card model	Operating system	CSI tool	Other instructions
Sending end	DELL latitude E6440	AR9580	Ubuntu 14.04	Atheros CSI tool	Turn off the NIC back-off function
Receiving end	ThinkPad E40	AR9462	Ubuntu 14.04	Atheros CSI tool	Turn off the NIC back-off function

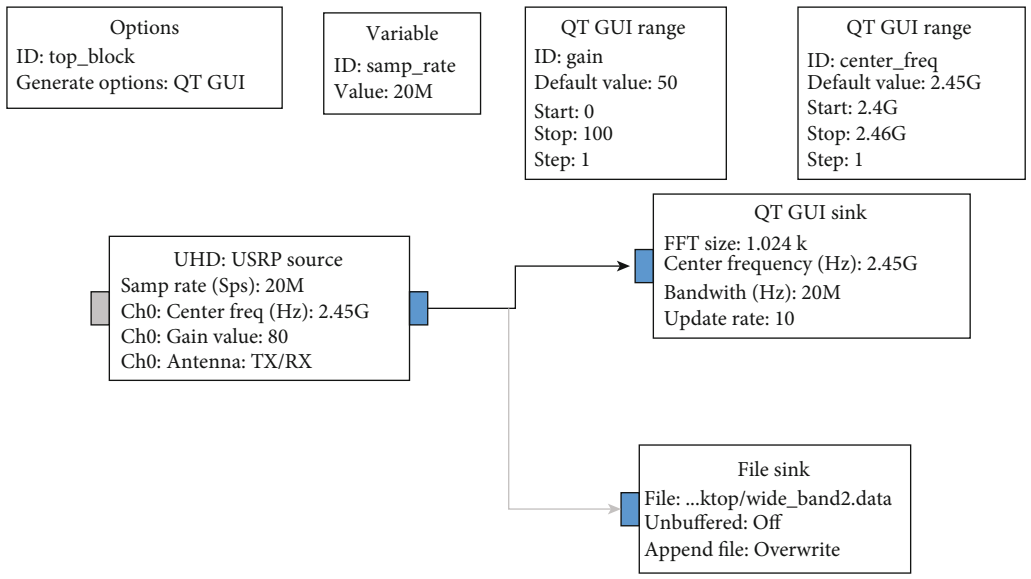


Figure 10: Implementation going-on signal monitor based on GNU radio.

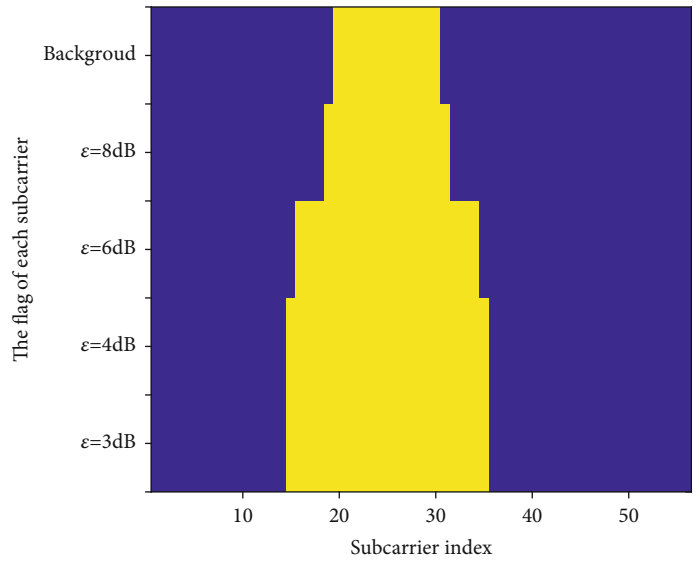


Figure 11: Subcarrier-level interference sensing results with the changes of ε.

changing of parameter ε, and the settings in this experiment are as follows: $dw = 7$m (the distance between wideband sender and wideband receiver) and di = 7m (the distance between narrowband interference generator and the wideband receiver). The sensing accuracies are shown in Figure 11, the bandwidth of the narrowband interference generator is 10 subcarriers, and the background is detailed in Table 2. The interference range determined by Subcarrier-Sniffer increases when ε is getting small, and trends are getting stable when $\varepsilon \leq 4$dB.

In order to further evaluate the performance of *Subcarrier-Sniffer*, we evaluate the sensing accuracy with the parameter di (the distance between narrowband interference generator and the wideband receiver); in this experiment, we assume $\varepsilon = 6$dB, and the results are shown in Table 3.

Table 2: Subcarrier interference sensing results in changes with ε.

Scenarios with different parameters	Interfered subcarriers
Background	[-5 -4 -3 -2 -1 1 2 3 4 5]
Subcarrier-Sniffer (ε = 8dB)	[-6 -5 -4 -3 -2 -1 1 2 3 4 5 6]
Subcarrier-Sniffer (ε = 6dB)	[-8 -7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7 8]
Subcarrier-Sniffer (ε = 4dB)	[-9 -8 -7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7 8 9]
Subcarrier-Sniffer (ε = 3dB)	[-9 -8 -7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7 8 9]

Table 3: The situation of subcarrier interference sensing result changes with di.

Scenarios	Interfered subcarrier
Background (ε = 6dB)	[-5 -4 -3 -2 -1 1 2 3 4 5]
Subcarrier-Sniffer (di = 0.5m)	[-12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7 8 9 10 11 12]
Subcarrier-Sniffer (di = 2.5m)	[-10 -9 -8 -7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7 8 9 10]
Subcarrier-Sniffer (di = 5m)	[-8 -7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7 8]
Subcarrier-Sniffer (di = 10m)	[-5 -4 -3 -2 -1 1 2 3 4 5]
Subcarrier-Sniffer (di = 12.5m)	[null]

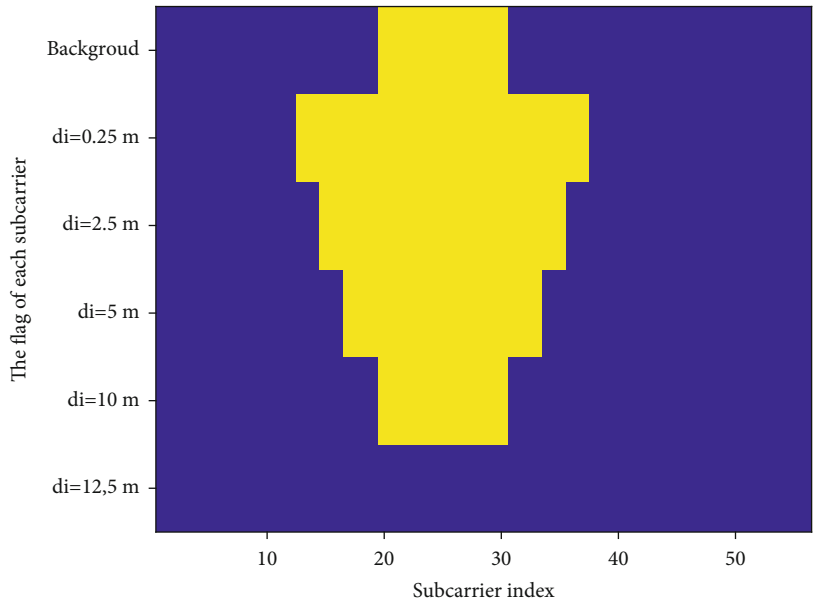


Figure 12: Subcarrier-level interference sensing results with the changes of di.

The results in Table 3 show that the interference determination range is larger when di is smaller (detailed in Figure 12); the reason is when di is smaller, the harmonic energy generated by the narrowband signal is relatively higher. When di ≥ 12.5m, the number of interfered subcarriers sensed by *Subcarrier-Sniffer* is 0. The main reason is the interference generator is far from the signal receiver, and it could not interfere with the received signals.

5. Conclusion

To further optimize the spectrum utilization of the ISM band, this paper proposes *Subcarrier-Sniffer*, a subcarrier-level spectrum sensing algorithm. *Subcarrier-Sniffer* contains three parts: (i) channel state information collection and analysis module, (ii) CSI initial states establish and denoising module, and (iii) interfered range of target subcarrier determination module. Compared with traditional channel-

based spectrum sensing algorithms, subcarrier-level spectrum sensing algorithms could further improve the ISM spectrum utilization. The experimental results show that the sensing accuracies of *Subcarrier-Sniffer* are changing with ε and di, where ε is the threshold to decide when the subcarrier is interfered with by other technologies and di is the distance between narrowband interference generator and the wideband receiver.

Subcarrier-Sniffer could be widely used in cross-technology coexisting environment, and we will introduce *Subcarrier-Sniffer* to WiFi and ZigBee and WiFi and LTE coexisting wireless networks to increase the spectrum utilization in our future works; WiFi and ZigBee coexisting work should be an example for any wideband and narrowband coexisting scenarios, while WiFi and LTE coexisting work should be an example solution for wideband and wideband coexisting scenarios; more precise fine-grained spectrum sensing and sharing methods are also our future concerns to further optimize the spectrum utilization in ISM bands.

Data Availability

The data that support the findings of this study are openly available in csi_subcarrier_sniffer at https://github.com/p01aris/, reference number Bsg-c000083.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This project is supported by the CERNET Innovation Project (NGII20180315), Undergraduate Training Programs for Innovation and Entrepreneurship Project of Jilin Province (2020JLSFDX-JSJ03 and 2020JLSFDX-JSJ05), and Science and Technology Project of Jilin Provincial Department of Education (JJKH20210457KJ). We also thank anonymous reviewers and their valuable comments.

References

[1] T. Barnett, S. Jain, U. Andra, and T. Khurana, *Cisco visual networking index (VNI) global mobile data traffic, 2017-2022*, Americas/EMEAR Cisco Knowledge Network (CKN) Presentation, 2018.

[2] K. Davaslioglu, S. Soltani, T. Erpek, and Y. E. Sagduyu, "DeepWiFi: cognitive WiFi with deep learning," *IEEE Transactions on Mobile Computing*, vol. 20, no. 2, pp. 429–444, 2021.

[3] Y. Zhu, L. Liu, J. Li, J. Yu, and C. Long, "OutSense: out-of-band sensing with ZigBee sensors for channel adaptation in wireless LANs," in *Proceedings of the International Conference on Parallel and Distributed Systems*, pp. 92–99, Melbourne, VIC, Australia, 2015.

[4] X. Zhang, Y. Zhang, Y. Ma, and Y. Gao, "RealSense: real-time compressive spectrum sensing testbed over TV white space," in *IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 1–3, Macau, 2017.

[5] M. Hirzallah, W. Afifi, and M. Krunz, "Full-duplex spectrum sensing and fairness mechanisms for Wi-Fi/LTE-U coexistence," in *IEEE Global Communications Conference*, pp. 1–6, Washington, DC, 2016.

[6] E. Pei, X. Lu, B. Deng, J. Pei, and Z. Zhang, "The impact of imperfect Spectrum sensing on the performance of LTE licensed assisted access scheme," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1966–1978, 2020.

[7] C. Cano, D. J. Leith, A. Garcia-Saavedra, and P. Serrano, "Fair coexistence of scheduled and random access wireless networks: unlicensed LTE/WiFi," *IEEE/ACM Transactions on Networking*, vol. 25, no. 6, pp. 3267–3281, 2017.

[8] P. Gawlowicz, A. Zubow, and A. Wolisz, "Enabling cross-technology communication between LTE unlicensed and WiFi," in *IEEE Conference on Computer Communications*, pp. 144–152, Honolulu, HI, 2018.

[9] D. Saliba, R. Imad, S. Houcke, and B. E. Hassan, "WiFi dimensioning to offload LTE in 5G networks," in *IEEE 9th Annual Computing and Communication Workshop and Conference*, pp. 0521–0526, Las Vegas, NV, USA, 2019.

[10] Y. Gao, Z. Qin, Z. Feng, Q. Zhang, O. Holland, and M. Dohler, "Scalable and reliable IoT enabled by dynamic spectrum management for M2M in LTE-A," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1135–1145, 2016.

[11] H. Sun, Z. Fang, Q. Liu, Z. Lu, and T. Zhu, "Enabling LTE and WiFi coexist in 5GHz for efficient spectrum utilization," *Journal of Computer Networks and Communications*, vol. 2017, Article ID 5156164, 17 pages, 2017.

[12] F. Li, J. Luo, G. Shi, and Y. He, "ART: adaptive frequency-temporal co-existing of ZigBee and WiFi," *IEEE Transactions on Mobile Computing*, vol. 16, no. 3, pp. 662–674, 2017.

[13] F. Chiti, R. Fantacci, F. Nizzi, L. Pierucci, and T. Pecorella, "A cooperative spectrum sensing protocol for IEEE 802.15.4m wide-area WSNs," in *IEEE International Conference on Communications*, pp. 1–6, Paris, 2017.

[14] Z. Chi, Y. Li, H. Sun, Y. Yao, Z. Lu, and T. Zhu, "B2W2: N-way concurrent communication for IoT devices," in *14th ACM Conference on Embedded Network Sensor Systems*, pp. 245–258, New York, NY, USA, 2016.

[15] Z. Chi, Y. Li, Z. Huang, H. Sun, and T. Zhu, "Simultaneous bidirectional communications and data forwarding using a single ZigBee data stream," in *IEEE Conference on Computer Communications*, pp. 577–585, Paris, France, 2019.

[16] S. Pan, X. Zhao, and Y. C. Liang, "Robust power allocation for OFDM-based cognitive radio networks: a switched affine based control approach," *IEEE Access*, vol. 5, pp. 18778–18792, 2017.

[17] H. Zhang, L. L. Yang, and L. Hanzo, "Compressed sensing improves the performance of subcarrier index-modulation-assisted OFDM," *IEEE Access*, vol. 4, pp. 7859–7873, 2016.

[18] T. Agrawal, V. Lakkundi, A. Griffin, and P. Tsakalides, "Compressed sensing for OFDM UWB systems," in *IEEE Radio and Wireless Symposium*, pp. 190–193, Phoenix, AZ, USA, 2011.

[19] A. Razavi, M. Valkama, and D. Cabric, "Compressive identification of active OFDM subcarriers in presence of timing offset," in *IEEE Global Communications Conference*, pp. 1–6, San Diego, CA, USA, 2015.

[20] Y. Zeng, S. W. Oh, and R. Mo, "Subcarrier sensing for distributed OFDMA in powerline communication," in *IEEE International Conference on Communications*, pp. 1–5, Dresden, Germany, 2009.

[21] W. Xu, W. Xiang, M. Elkashlan, and H. Mehrpouyan, "Spectrum sensing of OFDM signals in the presence of carrier frequency offset," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6798–6803, 2016.

[22] H. Sun, Z. Fang, X. Liu, and Z. Lu, "Insightful causes and potential resolutions for spectrum crisis in ISM bands: a survey," *The Journal of Communication*, vol. 11, no. 9, pp. 862–870, 2016.

[23] J. Fang, K. Tan, Y. Zhang et al., "Fine-grained channel access in wireless LAN," *IEEE/ACM Transactions on Networking*, vol. 21, no. 3, pp. 772–787, 2013.

[24] S. Han, X. Zhang, and K. G. Shin, "Fair and efficient coexistence of heterogeneous channel widths in next-generation wireless LANs," *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2749–2761, 2016.

[25] S. Yun, D. Kim, and L. Qiu, "Fine-grained spectrum adaptation in WiFi networks," in *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*, pp. 327–338, New York, NY, USA, 2013.

[26] Y. Xiao, "IEEE 802.11n: enhancements for higher throughput in wireless LANs," *IEEE Wireless Communications*, vol. 12, no. 6, pp. 82–91, 2005.

[27] IEEE, "802.11ac," https://standards.ieee.org/standard/802_11ac-2013.html.

[28] S. Rayanchu, V. Shrivastava, S. Banerjee, and R. Chandra, "Fluid: improving throughputs in enterprise wireless LANs through flexible channelization," *IEEE Transactions on Mobile Computing*, vol. 11, no. 9, pp. 1455–1469, 2012.

[29] F. Wu, N. Singh, N. Vaidya, and G. Chen, "On adaptive-width channel allocation in non-cooperative, multi-radio wireless networks," in *Proceedings IEEE INFOCOM*, pp. 2804–2812, Shanghai, China, 2011.

[30] R. Chandra, R. Mahajan, T. Moscibroda, R. Raghavendra, and P. Bahl, "A case for adapting channel width in wireless networks," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, pp. 135–146, 2008.

[31] A. Hussain and N. A. Saqib, "Effects of adaptable channelization in Wi-Fi networks," in *11th International Conference on Innovations in Information Technology*, pp. 166–171, Dubai, United Arab Emirates, 2015.

[32] T. Moscibroda, R. Chandra, Y. Wu, S. Sengupta, P. Bahl, and Y. Yuan, "Load-aware spectrum distribution in wireless LANs," in *IEEE International Conference on Network Protocols*, pp. 137–146, Orlando, FL, USA, 2018.

[33] H. Rahul, N. Kushman, D. Katabi, C. Sodini, and F. Edalat, "Learning to share: narrowband-friendly wideband networks," in *Proceedings of the ACM SIGCOMM 2008 conference on Data communication*, pp. 147–158, New York, NY, USA.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6699010, 11 pages
https://doi.org/10.1155/2021/6699010

Research Article

Construct a Teaching System Combining Image Linguistics and Multimedia Technology

Yan Yuejue,[1] **Sun Xinze**,[2] **Li Bingyue**,[1] **and Wang Xiaoyuan**[1]

[1]*School of Media Art, Jilin Business and Technology College, Changchun 130507, China*
[2]*School of Journalism, Chongqing University, Chongqing 401331, China*

Correspondence should be addressed to Yan Yuejue; yyj001@cumt.edu.cn

Received 30 December 2020; Revised 8 April 2021; Accepted 13 April 2021; Published 27 April 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Yan Yuejue 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 research on the theoretical system of multimedia image linguistics in my country is very limited. In order to further improve and develop the theoretical system of multimedia picture linguistics, this article extracts effective knowledge from the research results of educational technology, linguistics, psychology, and semiotics and establishes a teaching system based on image linguistics. Multimedia picture linguistics consists of three parts: structure, semantics, and pragmatics, corresponding to the artistic laws of multimedia pictures, the cognitive laws of multimedia pictures, and the humanization and naturalization of picture design. Therefore, the research in this article is an important part of establishing multimedia image linguistics. Starting from the research methods of multimedia technology and picture linguistics, this article introduces the theoretical framework and the relationship between the multimedia learning cognitive theory and Mayer's multimedia learning cognitive theory and conducts in-depth discussion and demonstration. The thesis constructs a teaching system that combines picture linguistics and multimedia technology and makes it a language tool for developing and designing multimedia learning materials. Experiments have proved that the teaching system combining picture linguistics and multimedia technology proposed in this paper can provide a framework, ideas, and methods for the related research of multimedia picture linguistics and multimedia technology and make multimedia picture language a multimedia learning material design and development The language tools that can be followed can effectively improve the teaching effect under the background of information technology.

1. Introduction

"Multimedia Picture Linguistics" is a new subject category that has emerged recently. The direct reason for its emergence is the need to carry out the construction of multimedia teaching materials widely at home and abroad at present, and the deeper reason is that in the information age. It is urgent to recognize and standardize the popular language of this era, the multimedia picture language [1]. Multimedia picture linguistics is an innovative theory born and grown up in China. It is a new design category formed in the information age. It has important application value and is an important research field of educational technology [2]. At present, the research of the theoretical system is still in the primary stage, and it is necessary to further explore and improve the theoretical framework, research content, research methods, and theoretical basis and so on. It can be used to provide a reference framework, ideas, and methods for the follow-up related research.

"Symbol" is a meaningful medium used to convey information in the process of communication. Multimedia pictures are considered to be a language because it conforms to the definition of language; that is, it can be regarded as a symbol system and can be used to transmit knowledge information to view aesthetic hearing [3]. This definition shows that language has two basic properties: language is a "symbol" system, and language is used for communication and thinking. This paper is divided into six chapters. The first section is the introduction part, which mainly leads to the design method of this paper. The related work is given in the second section. The third section introduces the theory and core concepts of the teaching system which combines picture linguistics and multimedia technology, and it can give readers a more intuitive feeling. Section 4 describes in detail

how to construct the teaching system of the combination of picture linguistics and multimedia technology [4]. The fifth section mainly from the experimental point of view proves that the combination of the picture linguistics and multimedia technology teaching system can better improve the teaching efficiency. Finally, Section 6 gives a summary of the article.

The research contributions of the thesis include the following:

(1) The thesis establishes a teaching system based on image linguistics

(2) This article starts with the research methods of multimedia technology and picture linguistics, introduces the theoretical framework and relationship between multimedia learning cognitive theory and Mayer multimedia learning cognitive theory, and conducts in-depth discussion and demonstration

(3) This paper constructs a teaching system that combines picture linguistics and multimedia technology and makes it a language tool for developing and designing multimedia learning materials

2. Related Work

In 2002, Professor you Sewing, a famous educational technology expert in China, put forward the concept of "multimedia picture language". In recent years, under the guidance of professor yet, the research team of our country has been guided by professor yet. On the basis of "multimedia picture language" as the starting point, this paper makes a reasonable creation of the multimedia picture art theory and designs some rules to be followed in view of the multimedia learning material interface. In recent years, the Chinese research team has put forward the theory system of "multimedia picture linguistics", which has further expanded its research to the level of linguistics [5]. The main purpose of the research on the theory of "multimedia picture linguistics" is to improve the teaching effect in the context of information technology and to make the application, development, and design of various multimedia language materials follow certain rules. It is an innovative theory growing up in China, and the value of its application is very high [6]. We need to continuously explore and improve it in order to provide better methods, frameworks, and ideas for subsequent research [7]. It plays an important role in the field of Chinese educational technology research. However, in view of the current research status, the research on the theoretical system of multimedia picture linguistics is still in its infancy.

From the perspective of general linguistics, this paper reveals the essential attributes and linguistic characteristics of the multimedia picture language. This paper discusses the theoretical basis for the division of multimedia picture linguistics into three parts: construction, semantics, and pragmatics, in order to verify the reasonableness of multimedia picture linguistics [8]. The ideas and conclusions of studying multimedia picture linguistics can be found in general linguistics. General linguistics defines language as the symbolic system of human communication and thinking. Multimedia infrastructure is shown in Figure 1.

Computer network platform can provide a good environment for students, so that students can receive information both visually and audibly, thus improving their learning efficiency. We made statistics on a large number of data: visual can get 82% of the information, and auditory can get 11% of the information [9]. Therefore, the human sense of hearing and vision is the most accepting information. We compared the same material, using hearing alone and hearing at the same time to get 70% of the information. Two days later, the amount of information held down to 10%, and visual can get 83% of the information. Then, two days later, the information kept down to 20%. If both vision and hearing receive information at the same time, the amount of information held can still reach 65% two days later. Therefore, it is very important to build a computer network platform, which can provide students with a good learning environment and improve their English ability [10].

3. The Theory and Core Concepts of the Teaching System Based on the Combination of Picture Linguistics and Multimedia Technology

3.1. Multimedia Graphics. Multimedia picture is a new kind of comprehensive information picture that appears with the appearance of multimedia. First of all, the multimedia learning materials and the knowledge information exchange between people, in the multimedia picture. It also has the following several important characteristics: (1) based on the basis of plane display, multimedia images can only be based on electronic plane display, such as mobile phones and computers. In order to complete the input and output of various media information in the multimedia interactive environment, the information it needs to present and the related interactive operations also need to be carried out with the help of the display screen. (2) The picture is interactive and dynamic. With the change of time, multimedia picture will also change, and multimedia picture is not a single picture but a "picture group." In this way, the information interaction between human and machine can be realized on the basis of controlling picture group connection and change. (3) The audio and visual media can be combined effectively to form a multimedia picture, in which the sound of the screen is synchronized with the change of the picture. Although its performance has some differences, it is an essential part of multimedia pictures. (4) There are many types of multimedia pictures. Multimedia learning materials include network course, teaching course, electronic teaching plan, teaching animation, and multimedia courseware. From the point of view of presentation form, multimedia images can be divided into tablet screen, mobile phone screen, and computer screen. Two languages meet the conditions of "language" that is shown in Table 1.

The multimedia picture language is a symbol system. The basic elements of traditional art, such as painting, photography, animation, film, and television, have already formed the

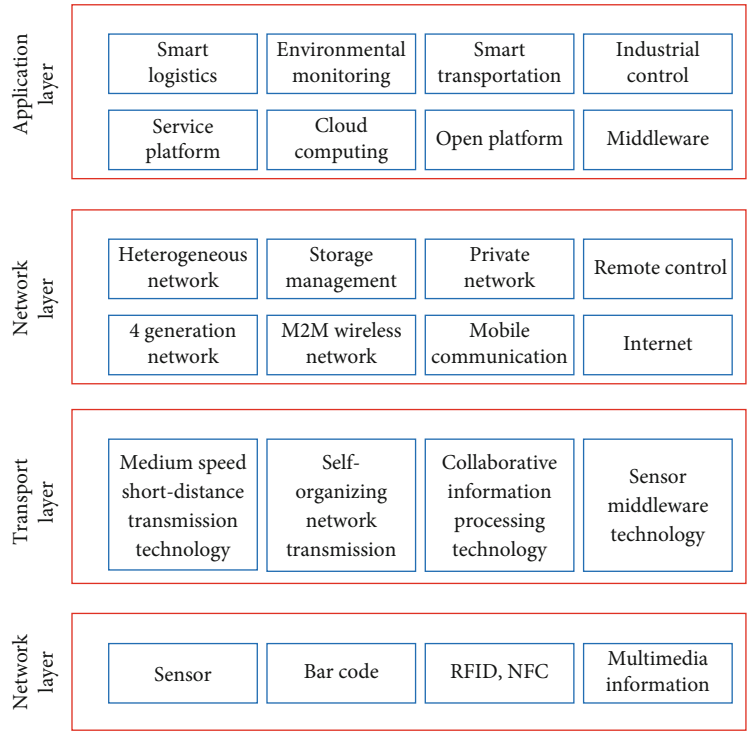


Figure 1: Multimedia infrastructure.

Table 1: Two languages meet the conditions of "language".

Conditions satisfied by written language	Requirements for multimedia picture language
A sufficient number and type of vocabulary and vocabulary have formed a complete set of rules (such as grammar and syntax) On the occasion of language communication, consensus has been formed.	The basic elements and elements of a picture consisting of sufficient numbers and types A set of rules (such as picture grammar rules) has been formed initially. Consensus has been formed in information-based teaching and art appreciation occasions.

symbol system in each art category. The basic elements of multimedia picture are composed of the basic elements of the above traditional art, which are divided into four media types: picture, text, sound, and image [2]. And the use of multimedia picture art rules (that is, the grammatical rules of picture language) is to regulate the relationship between the four media symbols, so that the symbol system formed a set of rigorous picture language structure as shown in the following expression.

$$F(x) = \sum_{i=1,2,..,3} \lambda_i f_i(x) - \sum_{i=4,5} \lambda_i f_i(x),$$
$$s.t. \quad \sum_{i=1}^{5} \lambda_i = 1, \qquad (1)$$
$$\sum_{j=1}^{m} x_{ij} = 1.$$

The multimedia picture language can be used to convey knowledge or visual sense of hearing. The multimedia teaching material is written in the multimedia picture language, because it contains all kinds of media such as picture, text and audio, and video, which can make full use of the visual, auditory, and visual and conceptual channels of learners. So, the effect of transmitting knowledge and aesthetic sense is better than that of written language and traditional art. It must be pointed out that the use of multimedia screen language symbols for communication is necessary to pay attention which is different from the two characteristics of the character symbol, namely, "in the form of meaning" and "time-sharing presentation" as shown in the following expression.

$$\text{Lifetime}[i] = \text{MinLT} + \alpha \frac{\text{fitness}[i]\text{-WorstFit}}{\text{AvgFit-WorstFit}}, \qquad (2)$$
$$\text{if fitness}[i] \le \text{AvgFit}.$$

The multimedia picture language can also be used for thinking. Multimedia picture language introduces information content into the human brain through visual and

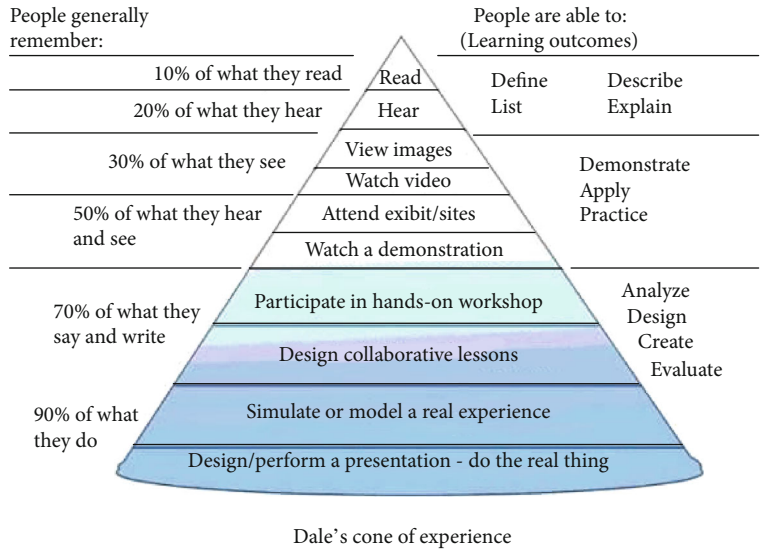


Figure 2: Experience of the tower.

auditory visual channels and conceptual channels. These information content can be processed in the form of representation and concept in the human mind [6]. Therefore, it can be considered that this is the form of thinking using the multimedia picture language. Experience of the towers is given in Figure 2.

$$\text{Lifetime}[i] = \frac{1}{2}(\text{MinLT} + \text{MaxLT}) + \alpha \frac{\text{fitness}[i]\text{-WorstFit}}{\text{BestFit-WorstFit}}, \text{if fitness}[i] > \text{AvgFit}. \tag{3}$$

3.2. Multimedia Pictures Have Three Features of General Language. In linguistics, language should generally have three characteristics: symbolism, systematizes, and venerability [2]. Multimedia pictures also have these characteristics of linguistic description language. Symbolic: linguistics holds that language symbols have hierarchical structures: phrases are composed of morpheme and words, and sentences are composed of phrases [11]. Multimedia picture language symbols also have a similar hierarchy [3]. For example, the hierarchical structure of graphic language symbols is composed of points, lines and surfaces, color, texture, and shadow tone, by form, attribute, sound, motion, and so on constitute the moving picture (namely the television picture, the computer picture, the multimedia picture).

$$M_p, t_s = \{NB, NM.NS, ZP, PS, \text{PM}, \text{PB}\}. \tag{4}$$

Systematic: linguistics holds that language is a system of linguistic symbols maintained by grammatical rules. The multimedia picture language is a language symbol system, its language symbols are maintained by multimedia picture art rules (grammar rules), and this maintenance should pay attention to two main points: (1) the rules of multimedia picture art regulate the symbol of visual (auditory) perception, that is, the evolution of the symbol of the basic element of the norm in the picture [6]. This point is emphasized because it clarifies the thinking of the basic theory of the language (multimedia picture art theory), and grammar rules need to be applied according to this idea [12]. (2) On the scope, the multimedia picture art rules are standardized from eight aspects. The basic rules of "prominent theme (or subject)," "Media matching," and "orderly change" are used to cover the static presentation and dynamic presentation of "Graph." The text and interpretation of the two aspects of the art rules on the form of the text language sound in the presentation of the screen to regulate [13]. The emphasis on this point is to illustrate that the rules are inextricably linked to the basic theory of the language; that is, the grammatical rules are extracted from the theory as shown in the following expression [9].

$$\begin{cases} K_p = K_p' + \left\{ M_{p(\text{t})}, t_{s(\text{t})} \right\}_p, \\ K_i = K_i' + \left\{ M_{p(\text{t})}, t_{s(\text{t})} \right\}_i, \\ K_d = K_d' + \left\{ M_{p(\text{t})}, t_{s(\text{t})} \right\}_d. \end{cases} \tag{5}$$

Generative: linguistics holds that although the system of linguistic symbols is made up of limited linguistic symbols and grammatical rules, it can produce an infinite variety of sentences [14]. The multimedia picture language is a language symbol system composed of limited language symbols (basic elements, visual elements) and grammar rules (multimedia picture art rules) [10]. It is precisely because of this limitation that the multimedia picture language can be made to be learnable and can be used to develop infinite kinds of multimedia teaching materials [15]. Therefore, the multimedia picture can be regarded as the same as the text, the two belong to two different types of artificial language [16].

3.3. Theoretical Framework and Research Content of Multimedia Picture Linguistics

3.3.1. Determination of Theoretical Framework Based on Semiotics Theory. Charles William Morris, a famous American pragmatism philosopher, divides semiotics into three parts based on logic and law, empiricism semantics, and pragmatism semiotics that are linguistics, semantics, and pragmatics [4]. Morphology mainly studies the relationship between charms, namely, the structure of symbols, which is also called grammar, construction, or syntax. Semantics mainly studies the interrelation between the symbol form and the object represented by the symbol, that is, the relationship between the information content of the symbol object expressed and transmitted by the symbol form and the symbol [17, 18]. Pragmatics mainly studies the relationship among the patterns of symbols, objects, and the specific situations in which symbols are applied [5]. From the perspective of the three components of semiotics, semiotics studies the three relationships between symbols and their related factors. As a symbolic tool for interpersonal communication, language also belongs to the category of semiotics [6].

If the multimedia picture language is compared to the symbolic language, the following viewpoints can be obtained: (1) the basic symbols used in the multimedia picture are all kinds of media; (2) the content of the information expressed and transmitted through the various media symbols is the teaching content; (3) the concrete situation of all kinds of media symbols in multimedia picture is the real information teaching environment. Based on the theoretical framework of semiotics and on the basis of the above viewpoints, the theoretical framework of multimedia picture linguistics can be preliminarily formed. Multimedia picture linguistics consists of three parts: picture construction, picture semantics, and picture pragmatics. Picture construction studies the structure and relationship between various media, while screen semantics studies the relationship between various media and the information they express or disseminate. Based on the theoretical framework of semiotics, and based on the above viewpoints, the theoretical framework of multimedia picture linguistics can be initially formed [19, 20]. Picture pragmatics mainly studies the relationship between various media and information teaching environment. The theoretical framework is shown in Figure 3 on the page below. The theoretical framework of multimedia picture linguistics is given in Figure 3

$$E = E\left(y_1^l, \cdots, y_n^l\right) = \sum_j^h \left(y_j^l - t_j\right). \tag{6}$$

4. The Basic Theory and Application of Constructing Picture Linguistics and Multimedia Technology

4.1. The Concepts of "Language" and "Speech" Exist in Linguistics. After entering the 20th century, the Swiss linguist Saussure first puts forward the concepts of "language" and "speech" in linguistics, which he called "the first bifurcation

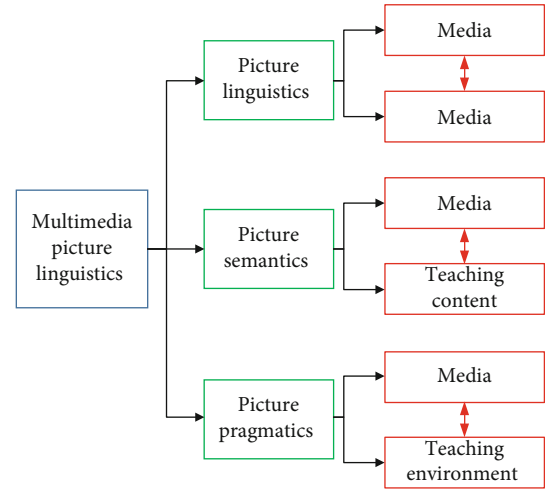


Figure 3: The theoretical framework of multimedia picture linguistics.

of the theory of speech activity." That is to say, for linguistics, it is necessary to distinguish language from speech [4]. Saussure's "language" actually refers to the language structure regulated by grammatical rules, while "speech" refers to the process and result of expressing with grammar [21, 22]. If "language" is regarded as a tool, then "speech" can be regarded as the use of tools. In 1957, Chomsky (American linguist) based on the theory of cognitive psychology created "transformational generative linguistics", which proposed the concepts of "language competence" and "language use." It corresponds to Saussure's concepts of "language" and "speech." Functional linguistics, which emerged in the late 1970s, is characterized by its emphasis on the function of language, which holds that grammar alone is not enough, and needs to be explained by function and cognition to explain syntax and other linguistic phenomena [23].

$$w_{im}^{l-1} = w_{im}^{l-1} - \eta \times \frac{\partial E}{\partial w_{im}^{l-1}}. \tag{7}$$

In the study of multimedia picture textualism, the research path of "constitution" can be adopted, and the constituent elements are analyzed in terms of the design rules of the elements and the relationship between the elements, which generalize the design object into a visual and auditory object with certain characteristics. This particular object is analyzed as a constituent element, and then the design rules of the constituent elements and the relationship between the elements are summarized as the design rules [7]. Therefore, it can be determined that there are three aspects in the study of screen construction, namely, the analysis of the constituent elements of multimedia pictures, the study of the design rules of each element, and the study of the relationship between the elements [24, 25]. The analysis of the constituent elements of multimedia picture is the basis of the other two parts of the research content in the construction of picture, and the study of the design rules of each element and the relationship between the elements are actually the grammatical

Table 2: The basic attributes of the elements in the Multimedia Picture.

Multimedia screen elements	Basic attribute
Text	Color, font, size, word spacing, line spacing, layout, location, etc.
Image	Point, line, face, composition, light, color, size, position, resolution, picture type, position
Sound	Tone, intonation, speed, rhythm, rhythm, melody, strength, length, strength, height, sound type, etc.
Animation, video	Color, light, composition, shooting, editing, length, size, resolution, type, position, etc.
Interaction	Input mode, output mode, application, interactive device features, etc.

rules of the multimedia picture language. The basic attributes of the elements in the multimedia picture are given in Table 2.

It can be seen from the development of linguistics that although the study of linguistics is deepening, Saussure's original idea of distinguishing language from speech has continued to this day. Based on the existing research results, the idea can be described as follows. Language is a relatively stable lexical and grammatical system which has formed a consensus. It can help people communicate with the outside world and can help people to think internally. One refers to speech (including writing), which is an act of communication, and the other is that the words spoken (including the words written) need to take into account the results of the communication [26]. The expression is given as follows.

$$o\text{Map}N = \left(\frac{i\text{Map}N - C\text{Window}}{C\text{Interval}} + 1\right). \tag{8}$$

4.2. Analysis of the Elements of Multimedia Picture. The constituent elements are the basic unit of the picture composition. The constituent elements of multimedia picture can be summarized as four kinds of media symbols, such as picture, text, sound and image, and the interactive function of the picture, that is, graph picture, including drawing figure and still image. The intersection-interactive function is essentially a multimedia picture of the group function. The presentation of the interactive output result and the trigger of the interactive input operation are realized by the multimedia screen which contains the elements of the interactive function [8]. The interactive form of multimedia picture not only depends on the input and output of these two kinds of picture, but also depends on the characteristics of the different multimedia picture presentation terminal interactive information input technology. Teaching content classification diagram is given in Figure 4.

Second, this paper analyses the new visual and auditory effects derived from the coordination of various attributes in the picture elements, studies its rules, and then comes up with the corresponding picture grammar rules in turn. Although the visual and auditory effects are also related to the basic attribute components of the elements of the picture, they are not equal to the sum of these elements, but they are a new whole completely independent of these elements [4]. This whole has an independent "new quality" that transcends all parts, which is called the "Gestalt quality", a process in which implicit stimuli are generated from explicit stimuli. Taking text as an example, new visual effects can be derived

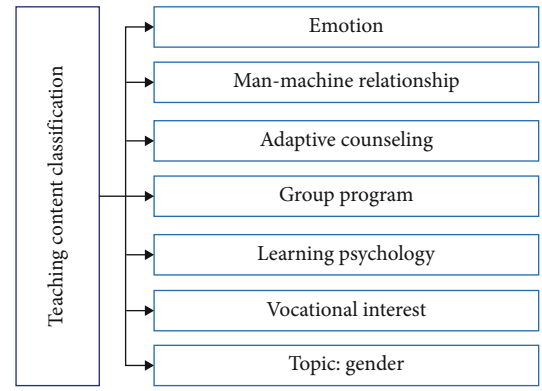


Figure 4: Teaching content classification diagram.

from the changes, layout, combination, and collocation of the basic properties of the text, such as font, font size, color, spacing, location, and layout. At the same time, it presents a comprehensive style of text, which generates "new quality" in the process of learners' perception. If this comprehensive style meets the aesthetic needs of learners, it will produce harmonious visual effects; otherwise, it will become a failure in multimedia pictures, which will have a negative impact on learners' learning. Investigation on the situation creation of mental health education in primary and secondary school is given in Table 3. The first 2 columns show the results, and the back shows the number of votes.

From the perspective of the social function, language is nothing more than a tool. The relationship between language and speech is like the relationship between tools and the use of tools. In general linguistics, sentences are the dividing line between "language" and "speech." The static units below sentences, including words, sounds, and meanings, belong to the study category of language, which is the internal structural model of language as a relatively stable communication tool [5]. The dynamic units above sentences, including the flow of language, intonation, sentence group, paragraph, paragraph group, and text, are completed in a certain context by people using the language system, which belongs to the category of speech research. Moreover, the model of sentence belongs to the category of language, and the use of specific sentence belongs to the category of speech.

$$C\text{Params} = (i\text{Map} \times C\text{Window} + 1) \times o\text{Map}. \tag{9}$$

4.3. The Distinction between Language and Speech in Multimedia Picture Linguistics. In multimedia picture linguistics, according to Saussure's idea of distinguishing language from speech, the language of multimedia picture

Table 3: Investigation on the situation creation of mental health education in primary and secondary school.

Content	Totally agree	More agree	General	Disagree	Strongly disagree
You prefer the form of teaching in a multimedia environment	26	23	13	2	1
You are easy to adapt to multimedia and interactive teaching	23	24	14	6	1
More concerned about the teaching of multimedia and interactive materials	11	16	19	9	2
Your focus is more focused with the multimedia and interactive teaching,	12	18	26	8	3
The use of multimedia materials makes it easier for you to enter the appropriate situation	18	21	24	15	4
After using the multimedia presentation context, your understanding of the content is easier	20	26	13	3	1
The use of multimedia context makes you more memorable to the content	16	17	18	4	1

is divided into two aspects: basic theory and theoretical application. In the view of multimedia picture linguistics, the teaching content in multimedia teaching material is expressed through multimedia picture, or in a broad sense, it is expressed by moving picture (including TV picture, animation picture, and multimedia picture). This shows that the moving picture should be equivalent to the "sentence" in general linguistics, which is the dividing line between language and speech in multimedia picture linguistics. The expressions are given as follows.

$$t_{nl}^{\text{out}} = f_{sub}\left(t_{nq}^{\text{in}}, t_{n(q+1)}^{\text{in}}\right), \tag{10}$$

$$\text{DoMap}N = \left(\frac{o\text{Map}N}{D\text{Window}}\right). \tag{11}$$

The art of multimedia picture is divided into the art of media presentation and the art of picture group, which embodies the idea of dividing the picture according to the multimedia picture. The former regards the basic elements of all kinds of media on the screen as "words" in the multimedia picture language (symbol) system. The rules of multimedia picture art are regarded as "grammar" to regulate the evolution of these words (the basic elements after evolution are called visual elements). Therefore, the multimedia picture is through these "grammar" standard "vocabulary" and has formed each kind of performance teaching content structure pattern, namely, "sentence." Obviously, these belong to the theory of multimedia picture art, corresponding to the category of "language" in linguistics. Obviously, these contents belong to the application of basic theories and correspond to the study of "phonetics" in linguistics. The latter uses editing functions or interactive functions to group each picture according to the requirements of "disciplines", according to the requirements of different disciplines and environments. The purpose is to accurately express the content of each discipline and achieve good teaching results.

4.4. Interpretation of Three Parts: Construction, Semantics, and Pragmatics. Modern linguistics holds that linguistics is a discipline that studies language tools (i.e., symbolic systems), tool internal structure systems, and the processes and results of people using language system tools in different contexts [6]. Linguistics is divided into "linguistic linguistics" and "linguistic linguistics."

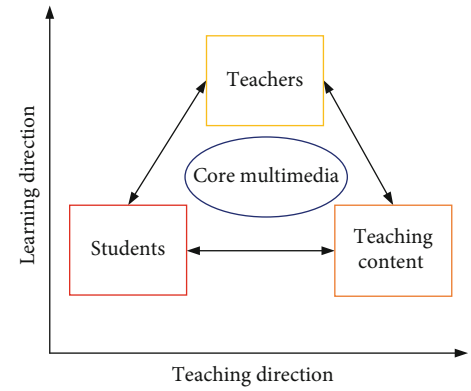


Figure 5: Four elements of the teaching system.

Language linguistics is a kind of linguistics which takes "language" as its only object. "language" refers to the lexical and grammatical system of the combination of sound and meaning. Linguistic study of speech: the process and results of using the linguistic system tools in different contexts [7]. American philosopher Miles further divides semiotics into three parts: morphology semantics and pragmatics. The study of the relationship between symbolism morphology is also called grammar or syntax, and semantics studies the relationship between symbols and symbolic objects, that is, the information about symbolic objects transmitted by symbols through symbols. Pragmatics studies the relationships among charms, objects, and symbolic situations. Symbolic context refers to the specific environment in which symbols are used between users to convey their thoughts and feelings [8].

Multimedia picture linguistics draws lessons from the views of modern linguistics and semiotics, and it is divided into three parts: construction, semantics, and pragmatics [9]. Four elements of the teaching system are given in Figure 5.

In this teaching system, the relationship between teaching content and all kinds of media belongs to the above multimedia picture semantic research content. The medium in the teaching system refers to the physical carrier of storing and transmitting information, such as information equipment and network. The relationship and law between

Table 4: Evaluation in the PE learning content.

Evaluation	Physical fitness	Motor skill	Theoretical knowledge	Learning attitude	Spirit of cooperation
Number	35	39	35	34	21
Percentage	87.5	97.5	87.5	85	52.5

the media and various kinds of media as the form of information presentation in multimedia pictures belong to the research content of picture pragmatics. It focuses on the different characteristics of different media carriers such as computers and mobile devices and the matching rules of various media symbols in multimedia images [6].

Among them, the study on the structure of multimedia picture belongs to the category of "linguistic linguistics", which consists of linguistic symbols and grammatical rules. In addition, according to modern linguistics, the context in linguistics of speech can be divided into small context and large context. Multimedia picture pragmatics uses the "big background" as the classroom teaching environment to study different teaching forms of picture language to meet the needs of classroom teaching strategies. According to the needs of compiling multimedia teaching materials, multimedia picture semantics extends the concept of "small situation" to the teaching content of different subjects and studies different design formats of screen languages to meet the needs of these subjects. The small context refers to the context before and after oral or written language, while the large context refers to the identity, situation, emotion, and social and cultural environment of both parties. Therefore, multimedia picture semantics and pragmatics can be regarded as the study category of "linguistics of speech."

5. Experimental Verification Analysis

Picture semantics studies the relationship between all kinds of media and the teaching content it expresses or conveys, from which to sum up some understanding of the regularity of different types of teaching content and form the "picture semantic rules". According to pedagogy, the teaching content refers to all kinds of materials and information dynamically generated by teachers and students in teaching communication activities to serve the purpose of teaching. The teaching content includes not only the contents of teaching materials [9]. It also includes the role of motivation, guidance, methodology guidance, normative concepts, and value judgment. It is the sum total of the knowledge and skills, the indoctrination of ideas and views, and the behaviors and habits cultivated by schools and teachers. Although the teaching material content is not the only component of the teaching content, it is the most important component. As a form of teaching material, multimedia learning material is responsible for presenting or carrying all kinds of knowledge information in the teaching content. Therefore, more accurately, picture semantics is actually to study the relationship between all kinds of media and the expressed or transmitted knowledge information and to excavate the matching rule between various media and different types of knowledge information—"semantic rules of picture." Evaluation in the PE learning content is given in Table 4.

Table 5: Student achievement statistics.

S1	S2	S3	S4
100	80	90	92
95	85	91	93
93	89	94	96
98	93	98	99

As a result of the above analysis, it can be confirmed that the research content of picture semantics is to study the picture, text, sound, image, interactive form, and the factual knowledge, conceptual knowledge expressed, or transmitted in the multimedia picture. The relationship and matching rules between procedural knowledge and metacognitive knowledge and the semantic rules suitable for different knowledge types are summarized. Through the specification of semantic rules, the design of multimedia picture can present and express all kinds of knowledge more accurately as follow. Student achievement statistics is shown in Table 5.

$$\frac{f(X/w_i)f(w_i)}{f(X)} = \max_j \left(\frac{f(X/w_j)f(w_j)}{f(X)} \right). \tag{12}$$

It can be seen from Table 5 that there is a direct proportional relationship between the student's performance and the student's cognitive level. Different theories of teaching elements reflect different perspectives and emphases of scholars, but in essence, there is no fundamental opposition and conflict between different theories of teaching elements. Teaching is a communication activity between teachers and students as a way of cultural inheritance. Therefore, teaching is essentially a cognitive activity, a form of existence, and a social phenomenon, which includes both the subject and the object of action [7]. Teaching is by teachers and students as the main body, culture as the content, the media as a means of cultural exchanges, and inheritance of an activity. Therefore, the basic elements of the teaching system should include teachers, students, teaching content, and media. The four elements are not independent of each other, but an organic whole that interacts and interrelates, thus forming a relatively stable teaching system, as shown in Figure 6.

As can be seen from the Figure 6, with the continuous development of picture linguistics and multimedia technology, the system performance of the teaching system is constantly improving. This is because the teaching system based on the combination of picture linguistics and multimedia technology can better combine the advantages of picture linguistics and multimedia technology. Fully applying the

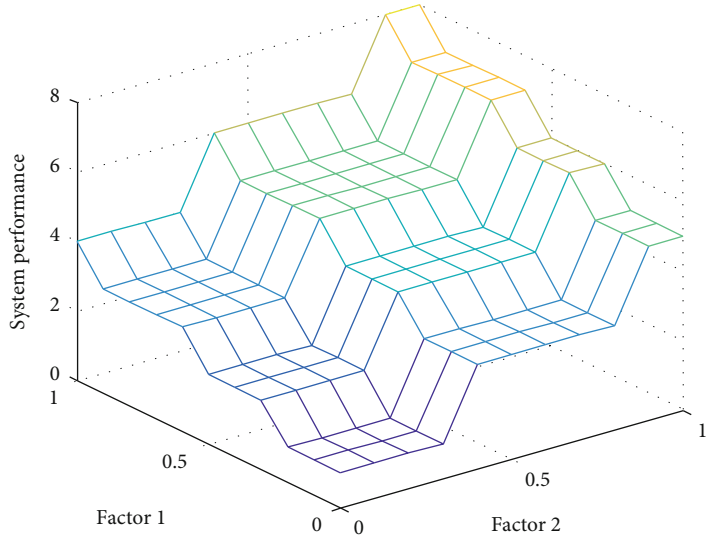


Figure 6: Performance of the teaching system based on image linguistics.

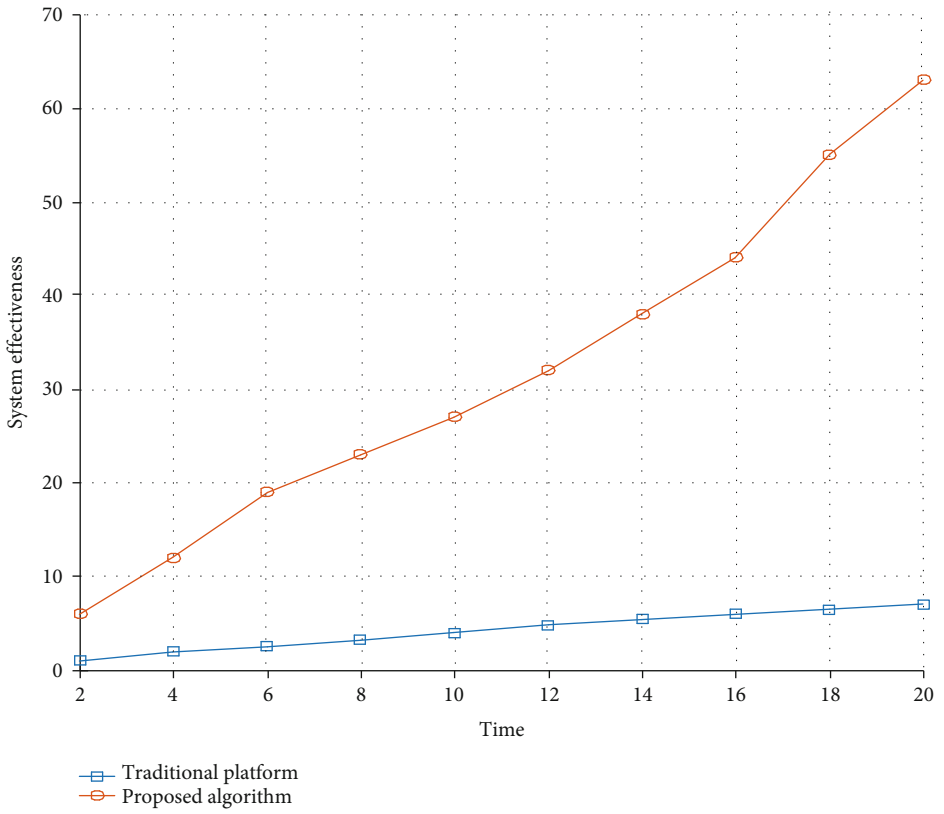


Figure 7: The relationship between the image linguistic teaching system and time.

advantages of the two to the teaching system can better adapt to the development of the teaching model.

It can be seen from the Figure 7, compared with the traditional teaching system, that this paper fully combines picture linguistics with multimedia technology, so that it can better cope with the teaching changes in multimedia technology. As a result, the teaching system has been in a steady growth of efficiency situation. Figure 8 shows the comparison between model prediction accuracy and other methods.

We compare the model established in this article with the methods in the latest international references, such as support vector regression machine (SVR) and extreme learning machine (ELM). The comparison uses model-based accuracy as an indicator. Figure 8 shows the comparison between model prediction accuracy and other methods. As shown in Figure 8, in the prediction comparison, the model established in this paper has better accuracy and higher prediction accuracy under the same test condition.

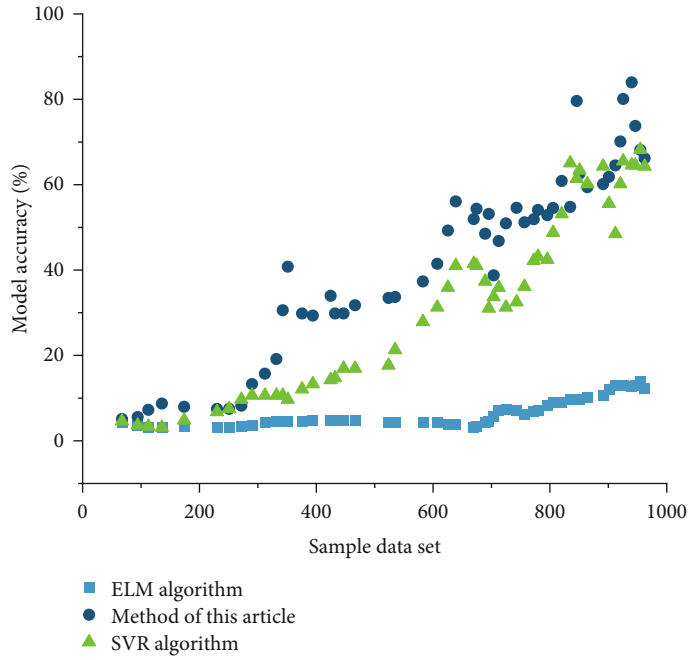


Figure 8: Comparison of the accuracy of the model established in this article and other methods.

6. Conclusion

Multimedia picture linguistics is composed of three parts: construction, semantics, and pragmatics, which correspond to the art law of multimedia picture, the cognitive law of multimedia picture, and the humanization and naturalization of picture design. Therefore, the study of this paper is an important part of creating multimedia picture linguistics. At present, the research on the theoretical system of multimedia picture linguistics in China is very limited. In order to further improve and develop the theoretical system of multimedia picture linguistics, this paper draws lessons from the research results of educational technology, linguistics, psychology, and semiotics. Then, starting from the research methods and theoretical framework of multimedia technology and picture linguistics, the relationship between multimedia learning cognitive theory and Mayer's multimedia learning cognitive theory is deeply explored and demonstrated. Experiments have proved that the teaching system that combines picture linguistics and multimedia technology proposed in this paper can provide a framework, ideas, and methods for the research of multimedia picture linguistics and multimedia technology. It is an information teaching application. The method proposed in this paper can provide more frameworks, methods, and ideas to construct a teaching system that combines picture linguistics and multimedia technology and make it a language tool for developing and designing multimedia learning materials. In the future, we will further devote ourselves to related research to provide reference and reference for the development of the education industry.

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(s) declare(s) that they have no conflicts of interest.

References

[1] Z. Zheng, H. Y. Jeong, T. Huang, and J. Shu, "KDE based outlier detection on distributed data streams in multimedia network," *Multimedia Tools and Applications*, vol. 76, no. 17, pp. 18027–18045, 2017.

[2] X. Li, Z. Lv, Z. Zheng, C. Zhong, I. H. Hijazi, and S. Cheng, "Assessment of lively street network based on geographic information system and space syntax," *Multimedia Tools and Applications*, vol. 76, no. 17, pp. 17801–17819, 2017.

[3] S. Pan, W. Sun, and Z. Zheng, "Video segmentation algorithm based on superpixel link weight model," *Multimedia Tools and Applications*, vol. 76, no. 19, pp. 19741–19760, 2017.

[4] S. Shivani, S. Tiwari, K. K. Mishra, Z. Zheng, and A. K. Sangaiah, "Providing security and privacy to huge and vulnerable songs repository using visual cryptography," *Multimedia Tools and Applications*, vol. 77, no. 9, pp. 11101–11120, 2018.

[5] B. D. Nye, "Intelligent tutoring systems by and for the developing world: a review of trends and approaches for educational technology in a global Context," *International Journal of Artificial Intelligence in Education*, vol. 25, no. 2, pp. 177–203, 2015.

[6] A. A. Genlott and Å. Grönlund, "Closing the gaps – Improving literacy and mathematics by ict-enhanced collaboration," *Computers & Education*, vol. 99, pp. 68–80, 2016.

[7] J. I. Liontas, "Reading and multimedia applications: going beyond bells and whistles, hotlinks and pop-up windows," *IALLT Journal of Language Learning Technologies*, vol. 33, no. 1, 2017.

[8] A. Öman and S. S. Hashemi, "Design and redesign of a multimodal classroom task – implications for teaching and

learning," *Journal of Information Technology Education: Research*, vol. 14, no. 1, pp. 139–159, 2015.

[9] S. Fraihat and Q. Shambour, "A framework of semantic recommender system for e-learning," *Journal of Software*, vol. 10, no. 3, pp. 317–330, 2015.

[10] K. Kozan, G. Erçetin, and J. C. Richardson, "Input modality and working memory: effects on second language text comprehension in a multimedia learning environment," *System*, vol. 55, pp. 63–73, 2015.

[11] C.-H. Chen, "A cell probe-based method for vehicle speed estimation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E103.A, no. 1, pp. 265–267, 2020.

[12] C. F. Nicholson, M. I. Gómez, and O. H. Gao, "The costs of increased localization for a multiple-product food supply chain: dairy in the United States," *Food Policy*, vol. 36, no. 2, pp. 300–310, 2011.

[13] C. Sel, B. Bilgen, J. M. Bloemhof-Ruwaard, and J. G. van der Vorst, "Multi-bucket optimization for integrated planning and scheduling in the perishable dairy supply chain," *Computers & Chemical Engineering*, vol. 77, no. 9, pp. 59–73, 2015.

[14] Y. L. XuJiuping and Z. Xiaodan, "A multi-objec-tivechance-constrained network optimal model with random fuzzy coefficients and its application tologistics distribution center location problem," *Fuzzy Optimization and Decision Making*, vol. 10, no. 3, pp. 255–285, 2011.

[15] E. M. Golonka, A. R. Bowles, V. M. Frank, D. L. Richardson, and S. Freynik, "Technologies for foreign language learning: a review of technology types and their effectiveness," *Computer Assisted Language Learning*, vol. 27, no. 1, pp. 70–105, 2014.

[16] C. Xin, P. Zhan, and L. Pan, "Coordinating inventory control and pricing strategies for perishable products," *Operations Research*, vol. 62, no. 2, pp. 284–300, 2014.

[17] J. Bensmail, "A 1-2-3-4 result for the 1-2-3 conjecture in 5-regular graphs," *Discrete Applied Mathematics*, vol. 257, pp. 31–39, 2019.

[18] X. Xianhao, C. Chengyuan, and G. Shen, "Short lifecycle product inventory model based on controllable lead time and delayed delivery," *Chinese Management Science*, vol. 18, no. 2, pp. 42–47, 2010.

[19] C. Fang, S. Fang, and C. Guo, "Research on network design and cost optimization of dairy product supply chain under matching theory," *Chinese Management Science*, vol. 23, no. S1, pp. 610–616, 2015.

[20] Z. Ayag, F. Samanlioglu, and G. Buyukozkan, "A fuzzy QFD approach to determine supply chain management strategies in the dairy industry," *Journal of Intelligent Manufacturing*, vol. 24, no. 6, pp. 1111–1122, 2013.

[21] C. Tang, H. Yang, E. Cao, and K. K. Lai, "Channel competition and coordination of a dual-channel supply chain with demand and cost disruptions," *Applied Economics*, vol. 50, no. 46, pp. 4999–5016, 2018.

[22] S. Heidarzadeh, A. Doniavi, and M. Solimanpur, "Development of supply chain strategy in the Iranian automotive industry based on system dynamics and game theory. ScientiaIranica. Transaction E," *Industrial Engineering*, vol. 24, no. 6, pp. 3345–3354, 2017.

[23] M. R. Shaharudin, K. Govindan, S. Zailani, K. C. Tan, and M. Iranmanesh, "Product return management: linking product returns, closed-loop supply chain activities and the effectiveness of the reverse supply chains," *Journal of Cleaner Production*, vol. 149, pp. 1144–1156, 2017.

[24] D. Ivanov, "OR/MS methods for structural dynamics in supply chain risk management," in *Structural Dynamics and Resilience in Supply Chain Risk Management*, pp. 115–159, Springer, Cham, 2018.

[25] H. Jafari, A. Nyberg, and P. Hilletofth, "Postponement and logistics flexibility in retailing: a multiple case study from Sweden," *Industrial Management & Data Systems*, vol. 116, no. 3, pp. 445–465, 2016.

[26] D. Aravind, "End to end supply chain planning for a fashion retailer in India," in *Handbook of Research on Strategic Supply Chain Management in the Retail Industry*, pp. 489–505, IGI Global, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5591543, 15 pages
https://doi.org/10.1155/2021/5591543

Research Article

Quasi-Affine Transformation Evolutionary with Double Excellent Guidance

Tien-Wen Sung, Baohua Zhao, and Xin Zhang

Fujian Provincial Key Laboratory of Big Data Mining and Applications, College of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, China

Correspondence should be addressed to Tien-Wen Sung; tienwen.sung@gmail.com

Received 5 February 2021; Revised 29 March 2021; Accepted 1 April 2021; Published 19 April 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Tien-Wen Sung 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 Quasi-Affine Transformation Evolutionary (QUATRE) algorithm is a swarm-based collaborative optimization algorithm, which has drawn attention from researchers due to its simple structure, easy implementation, and powerful performance. However, it needs to be improved regarding the exploration, especially in the late stage of evolution, and the problem of easy falling into a local optimal solution. This paper proposes an improved algorithm named Quasi-Affine Transformation Evolutionary with double excellent guidance (QUATRE-DEG). The algorithm uses not only the global optimal solution but also the global suboptimal solution to guide the individual evolution. We establish a model to determine the guiding force by the distance between the global optimal position and the suboptimal position and propose a new mutation strategy through the double population structure. The optimization of population structure and the improvement of operation mechanisms bring more exploration for the algorithm. To optimize the algorithm, the experiments on parameter settings were made to determine the size of the subpopulation and to achieve a balance between exploration and development. The performance of QUATRE-DEG algorithm is evaluated under CEC2013 and CEC2014 test suites. Through comparison and analysis with some ABC variants known for their strong exploration ability and advanced QUATRE variants, the competitiveness of the proposed QUATRE-DEG algorithm is validated.

1. Introduction

Optimization problems have been a hot research field since optimization solutions are crucial to various engineering practices. Scientific theories have indicated that appropriate optimization methods applied in different fields [1–7] will bring significant benefits. Accordingly, powerful and efficient algorithms have drawn the attention of researchers in different application fields.

Inspired by the laws of the biological community and natural world, various intelligent optimization algorithms have been proposed. At the early stages, evolutionary algorithms were first developed, such as the Genetic Algorithm (GA) [8] and Differential Evolution (DE) algorithm [9]. Genetic Algorithm is a global optimization algorithm developed by imitating the genetic and evolutionary characteristics of organisms. The core idea of the algorithm is "survival of the fittest in natural selection." The other algorithm, Differential Evolution, is also an efficient evolutionary algorithm. In its process of searching, the better individuals are continuously selected with the cooperation and competition among particles in the population to find the optimal solution. At the later stages, numerous swarm intelligence algorithms inspired by biology emerged. Particle Swarm Optimization (PSO) [10] is an optimization algorithm proposed by observing the living habits of birds. The algorithm imitates the behavior of continuous migration and clustering in the process of searching foods. It has the characteristics of simple structure and easy implementation. The artificial bee colony (ABC) algorithm [11] is a swarm intelligence algorithm firstly proposed by Karabora in 2005, inspired by the bee colony behavior of honey source searching. Compared with the premature convergence of PSO, the ABC algorithm is relatively mature in the application of swarm intelligence because of its strong searching ability. In recent years, with the development of computational intelligence, many novel

global optimization algorithms have been proposed, such as Multiverse Optimizer (MVO) [12] algorithm, grasshopper optimization algorithm (GOA) [13], and Quasi-Affine Transformation Evolutionary (QUATRE) algorithm [14]. These algorithms also have a simple structure and excellent performance.

It is well known that particle exploitability and exploratory are considered of equivalent importance in swarm intelligence algorithms, and they should be treated equally in the process of updating particle position. QUATRE algorithm is evolved from DE algorithm. It solves the position deviation problem of DE algorithm and has a stronger development ability. But it is slightly insufficient in search performance. In this paper, a new QUATRE-DEG algorithm is proposed, which divides the population into two subpopulations to increase the diversity of the population and uses the positions of the global optimal individual and the global suboptimal individual to establish the guiding force model, which improves the original mutation strategy and enhances the searching ability of the algorithm. The performance of the proposed QUATRE-DEG algorithm is verified by the experiments on CEC2013 benchmark set [15] and CEC2014 benchmark set [16]. The specific contributions of this paper are as follows:

(1) The two-population structure of mutation in two central regions is proposed, which maintains the diversity of particles, so that more benchmark environments can be explored

(2) The gravity model based on the relationship between the global optimal solution and the global suboptimal solution is established and applied to the new operation strategy, so as to enhance the information interaction between the excellent individuals

(3) In the benchmark set of CEC2013 and CEC2014, some famous ABC variants and advanced QUATRE variants are compared and analyzed with our QUATRE-DEG algorithm. The results show that the performance of the new QUATRE-DEG algorithm is competitive

The next few sections are arranged as follows. The second section reviews the development of QUATRE and ABC algorithms. The third section mainly introduces the QUATRE-DEG algorithm proposed in this paper. In the fourth section, the performance of the QUATRE-DEG algorithm is evaluated by using CEC2013 and CEC2014 benchmarks, respectively. The final section provides the summary.

2. Related Works

QUATRE is a relatively new global optimization algorithm first proposed in 2016. The term QUATRE is the abbreviation of Quasi-Affine Transformation Evolution. The algorithm is based on the DE algorithm and uses particle cooperative evolution in the update process. It replaces the parameter CR in the DE algorithm with a new collaborative search matrix M, which solves the position deviation problem in the DE algorithm. It is a powerful and simple optimization algorithm. Its operation mechanism is shown in the following formula:

$$\begin{cases} B = X_{g\text{best},G} + F \bullet (X_{r1,G} - X_{r2,G}) \\ X \longleftarrow \bar{M} \otimes X + M \otimes B, \end{cases} \tag{1}$$

$$\begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ & & \cdots & \\ 1 & 1 & \cdots & 1 \end{bmatrix} \sim \begin{bmatrix} & & 1 & \\ & 1 & & 1 \\ & & \cdots & \\ 1 & 1 & \cdots & 1 \end{bmatrix} \sim \begin{bmatrix} & 1 & & \\ & \cdots & & \\ 1 & 1 & \cdots & 1 \\ & 1 & 1 & \end{bmatrix} = M, \tag{2}$$

where F is an adjustable mutation size factor. In the original QUATRE algorithm, its optimal value is 0.7. $X_{g\text{best},G}$ is a global optimal matrix, and all row vectors consist of global optimal individuals. $X_{r1,G}$ and $X_{r2,G}$ are all randomly generated matrices. They are generated by randomly replacing the row vector of X, but the total elements remain unchanged. The M is a cooperative search matrix, which is the key to the evolutionary performance of the algorithm. The required M matrix can be generated in three steps. Firstly, the lower triangular matrix is initialized, and the elements in the matrix are 0 or 1. Then, all the elements in the row vector of the matrix are randomly displaced. Finally, the M matrix can be obtained by randomly displacing the row vector in the matrix. $\bar{M}$ is the binary inverse operation of all elements in M, and $\otimes$ represents the component multiplication.

The QUATRE algorithm has the characteristics of fewer control parameters, simple to use, and powerful performance, which attracts more and more scholars to study it. Several successful variants have been proposed [17–21], including the competition-based QUATRE algorithm (C-QUATRE) [22] which employs a pairing competition. The winning individuals are directly retained to the next generation, while the losers need to evolve like the original QUATRE algorithm. It effectively improves the performance of the algorithm. Like the DE algorithm, the QUATRE has the disadvantages of premature convergence and easy to fall into a local optimal solution. The diversity of particles can effectively reduce premature convergence, but the convergence of particles is also important, so finding the balance between particle convergence and diversity is the key. A new sorting-based QUATRE algorithm (S-QUATRE) [23] was proposed, which not only keeps the convergence of particles but also enhances the diversity of particles. Its main improvement scheme is to divide the population into better and worse groups by sorting strategy and to evolve the individuals in the worse group, which improves the search performance of the algorithm. In addition, the performance of the algorithm is greatly affected by the disturbance strategy and the parameters of the disturbance. For example, in QUATRE, a larger disturbance factor F will enhance the algorithm in search, while a smaller one will pay more attention to development. In [24], the BP-QUATRE algorithm is proposed and the original population is divided into different subpopulations by a new sorting

Initialization: preset maximum number of iterations G, the number of individuals SN, dimension D, and parameters a_1, a_2, z.
Generate the initial population X, to calculate the fitness of all individuals and get the optimal individual $x_{g\text{best}}$ and the suboptimal individual $x_{s\text{best}}$.
According to Eq. (3), individuals are divided into two fixed subpopulations X_1 and X_2 and generate the corresponding global optimal matrix $X_{g\text{best},G}$ and global suboptimal matrix $X_{s\text{best},G}$.
for $g = 1$ to G
  Randomly generate the matrix M according to Eq. (2), then generate the $\bar{M}$ matrix.
  Calculate the Euclidean distance d_{gs} between the global optimal individual and the global suboptimal individual.
  Calculate the guiding force $\mathrm{AF}_{gs,G}$ through Eq. (4).
  Generate the matrix B by Eq. (5).
  Evolve individuals in each group using Eq. (6).
  Calculate the fitness value of each individual.
  for $i = 1$ to SN
    if $f(xi) < = f(x_{p\text{best}_i})$
      $x_{p\text{best}_i} = x_i$
    end if
  end for
  Update the current optimal individual $x_{g\text{best}}$ and the current suboptimal individual $x_{s\text{best}}$.
  Generate the new global optimal matrix $X_{g\text{best},G}$ and global suboptimal matrix $X_{s\text{best},G}$.
end for
Output: Individual of global optimal solution $x_{g\text{best}}$ and global optimal value $f(x_{g\text{best}})$.

Algorithm 1: QUATRE-DEG.

strategy. Different subpopulations use different disturbance modes, and the linear descent parameter F is used to enhance the global optimization ability of the algorithm.

3. Proposed Approach

The original QUATRE algorithm has an inherent shortcoming of paying too much attention to development, which makes it easy to fall into the local optimal solution. The diversity of the population can make up for this shortcoming and enhance the search algorithm. In the iterative process, the global optimal solution of each iteration is widely used in the operation mechanism of a large number of algorithms. It can greatly enhance the convergence of the algorithm, make it easier to find the real global optimal, and also can improve the operation speed of the algorithm. And the excellent individuals in each iteration are also worthy of reasonable utilization, such as the global suboptimal solution. In this study, we propose a new QUATRE algorithm based on double excellent guidance (QUATRE-DEG) which divides the population into two subgroups, including the optimal guided group and the suboptimal guided group. It makes the possible to use the global suboptimal solution to enhance the performance of the algorithm. The details of the new QUATRE algorithm are as follows.

3.1. Initialization. Generate a new population and obtain the current optimal solution and suboptimal solution, as well as set a maximum number of iterations.

3.2. Double Population Structure. Firstly, all individuals are divided into two fixed subpopulations, including the optimal solution population X_1 and suboptimal solution population X_2. The number of individuals of optimal solution population and suboptimal solution population is determined by Eq. (3). In the population of optimal solution, B in Eq. (1) still takes the location of the optimal solution $x_{g\text{best}}$ as the center for mutation strategy. But in the population of suboptimal solution, B is the mutation centered on the position of suboptimal solution $x_{s\text{best}}$. In this way, the individuals in the population will search in two central regions in each iteration, so as to explore more environment regions.

$$\begin{cases} \mathrm{SN}_g = (1 - z) * \mathrm{SN}, \\ \mathrm{SN}_s = z * \mathrm{SN}, \end{cases} \tag{3}$$

where SN is the total population and z is the parameter which are used to determine the individual number SN_g of the optimal solution population and SN_s of the suboptimal solution population.

3.3. Relation between Optimal Solution and Suboptimal Solution. For the global optimization algorithm, the task is to find the global optimal solution. In the process of updating and iteration, excellent individuals can play an important role. It should be concerned that the neighborhood of the current global optimal solution may not be better than the one of the global suboptimal solution, which means that both the global suboptimal solution and the global optimal solution contain extremely valuable information. If the information can be shared between them, it can enhance the performance of the algorithm, since there is a strong information interaction between individuals in the grasshopper optimization algorithm (GOA). Inspired by the GOA, this study proposes a model to determine the guiding force

Table 1: Comparison on 10D optimization under CEC2013 test suite with different values of parameter z.

QUATRE-DEG	$z = 0.1$	$z = 0.2$	$z = 0.3$	$z = 0.4$	$z = 0.5$	$z = 0.6$
f	Mean/std	Mean/std	Mean/std	Mean/std	Mean/std	Mean/std
f_{a1}	5.80E-04/8.5E-05	5.90E-04/1.2E-04	5.97E-04/7.9E-05	6.20E-04/1.2E-04	6.21E-04/6.9E-05	6.73E-04/1.0E-04
f_{a2}	2.68E+05/1.1E+05	3.40E+05/1.1E+05	4.23E+05/1.9E+05	2.63E+05/1.6E+05	3.29E+05/1.1E+05	3.55E+05/1.4E+05
f_{a3}	9.18E+06/2.6E+07	5.44E+05/1.0E+06	6.73E+05/1.0E+06	2.89E+05/5.4E+05	3.20E+05/4.9E+05	2.02E+06/3.2E+06
f_{a4}	2.01E+01/1.5E+01	2.48E+01/1.8E+01	1.79E+01/1.1E+01	2.77E+01/1.7E+01	2.12E+01/1.5E+01	3.45E+01/2.8E+01
f_{a5}	2.82E-04/6.7E-05	2.93E-04/3.8E-05	2.99E-04/4.7E-05	2.80E-04/3.5E-05	3.30E-04/6.2E-05	3.72E-04/7.1E-05
f_{a6}	6.98E+00/1.0E+01	9.43E+00/1.1E+01	3.35E+00/1.4E+00	4.92E+00/7.5E+00	4.47E+00/7.7E+00	4.97E+00/7.6E+00
f_{a7}	1.67E+01/8.5E+00	1.80E+01/1.4E+01	1.86E+01/1.0E+01	2.19E+01/1.5E+01	1.89E+01/1.2E+01	1.30E+01/1.2E+01
f_{a8}	2.10E+01/5.3E-02	2.09E+01/7.2E-02	2.10E+01/6.0E-02	2.10E+01/5.1E-02	2.09E+01/5.9E-02	2.10E+01/3.1E-02
f_{a9}	1.58E+01/2.9E+00	1.41E+01/5.5E+00	1.67E+01/5.9E+00	1.71E+01/4.6E+00	1.79E+01/3.9E+00	1.42E+01/5.2E+00
f_{a10}	6.09E-02/1.4E-02	5.72E-02/1.4E-02	5.72E-02/1.4E-02	5.68E-02/1.6E-02	6.47E-02/1.4E-02	5.88E-02/1.0E-02
f_{a11}	2.23E+01/5.0E+00	2.62E+01/8.8E+00	2.39E+01/7.5E+00	2.03E+01/5.9E+00	1.87E+01/3.8E+00	2.05E+01/6.7E+00
f_{a12}	6.10E+01/9.1E+00	5.16E+01/1.5E+01	6.41E+01/1.5E+01	5.51E+01/1.6E+01	5.64E+01/2.0E+01	5.35E+01/1.5E+01
f_{a13}	1.22E+02/3.1E+01	1.16E+02/3.1E+01	9.83E+01/3.3E+01	1.13E+02/2.9E+01	1.21E+02/3.4E+01	1.03E+02/3.5E+01
f_{a14}	7.67E+02/1.9E+02	6.38E+02/2.2E+02	6.50E+02/1.7E+02	7.10E+02/1.4E+02	7.66E+02/2.5E+02	7.07E+02/2.5E+02
f_{a15}	3.57E+03/5.9E+02	3.85E+03/6.2E+02	3.46E+03/6.8E+02	4.02E+03/5.2E+02	3.75E+03/7.2E+02	3.92E+03/5.5E+02
f_{a16}	3.11E-01/2.0E-01	3.34E-01/1.1E-01	3.59E-01/1.2E-01	3.40E-01/1.0E-01	3.20E-01/1.5E-01	2.42E-01/1.1E-01
f_{a17}	5.41E+01/4.8E+00	5.15E+01/6.2E+00	4.89E+01/3.5E+00	5.02E+01/6.1E+00	4.59E+01/4.6E+00	4.63E+01/5.3E+00
f_{a18}	8.44E+01/2.0E+01	8.87E+01/1.7E+01	8.49E+01/1.0E+01	7.62E+01/1.8E+01	8.12E+01/1.1E+01	7.65E+01/1.0E+01
f_{a19}	3.04E+00/7.1E-01	2.49E+00/2.9E-01	3.05E+00/5.8E-01	2.19E+00/5.3E-01	2.49E+00/4.6E-01	2.41E+00/5.0E-01
f_{a20}	1.16E+01/5.5E-01	1.17E+01/4.9E-01	1.17E+01/3.4E-01	1.12E+01/6.4E-01	1.13E+01/4.8E-01	1.14E+01/6.0E-01
f_{a21}	2.75E+02/7.7E+01	3.18E+02/1.1E+02	3.33E+02/8.2E+01	3.33E+02/8.2E+01	3.28E+02/1.0E+02	2.65E+02/8.0E+01
f_{a22}	1.09E+03/2.5E+02	6.64E+02/2.7E+02	6.45E+02/1.3E+02	7.37E+02/2.1E+02	6.98E+02/1.6E+02	5.98E+02/2.5E+02
f_{a23}	3.86E+03/4.1E+02	4.10E+03/5.3E+02	3.96E+03/2.9E+02	3.90E+03/8.2E+02	3.93E+03/4.9E+02	4.08E+03/8.0E+02
f_{a24}	2.33E+02/1.2E+01	2.30E+02/1.2E+01	2.37E+02/8.2E+00	2.24E+02/1.4E+01	2.29E+02/1.6E+01	2.25E+02/1.3E+01
f_{a25}	2.60E+02/9.5E+00	2.54E+02/7.2E+00	2.55E+02/3.5E+00	2.54E+02/8.0E+00	2.50E+02/7.8E+00	2.52E+02/9.3E+00
f_{a26}	2.69E+02/7.2E+01	2.38E+02/6.1E+01	2.13E+02/4.0E+01	2.51E+02/6.5E+01	2.23E+02/4.7E+01	2.27E+02/5.6E+01
f_{a27}	6.49E+02/1.0E+02	6.39E+02/7.1E+01	6.61E+02/1.3E+02	6.10E+02/5.8E+01	6.11E+02/7.3E+01	6.50E+02/7.7E+01
f_{a28}	3.97E+02/3.0E+02	3.01E+02/5.1E-02	3.01E+02/6.9E-02	3.01E+02/9.1E-02	3.01E+02/5.2E-02	3.01E+02/1.1E-01
Best quantity	2	4	5	9	4	4

between the global optimal solution and the global suboptimal solution by the distance between them, as shown in Eq. (4).

$$\begin{cases} f(r) = a_1 * e^{-r/a_2} - e^{-r}, \\ r = \dfrac{d_{gs}}{D+2}, \\ \mathrm{AF}_{gs} = \dfrac{1}{2} * (\mathrm{ub} - \mathrm{lb}) * f(r) * \dfrac{x_{g\mathrm{best}} - x_{s\mathrm{best}}}{d_{gs}}, \end{cases} \tag{4}$$

where d_{gs} is the Euclidean distance between the global optimal solution and the global suboptimal solution and D represents the dimension. In the function of $f(r)$, the parameter values of a_1 and a_2 are recommended to be set to 0.5 and 1.5, respectively. The AF_{sg} represents the attraction of the global optimal solution to the global suboptimal solution. The ub and lb represent the upper and lower bounds of the search space, respectively. In the $f(r)$ function, when r is equal to 2.079, the function value is 0, and when r value increases, the function value will first increase to the maximum value of 0.02 and then gradually decrease to 0. A distance-based attractiveness model is constructed by using the characteristics of $f(r)$. When the distance between the global optimal solution and the global suboptimal solution is close, the attractiveness AF_{gs} of the global optimal solution and the global suboptimal solution is positively correlated with the distance d_{gs}. However, when the distance between the global optimal solution and the global suboptimal solution is far, the attraction AF_{gs} is negatively correlated with the distance d_{gs}.

Table 2: Comparison on 10D optimization under CEC2013 test suite with different algorithms.

Algorithms f	QUATRE-DEG Mean/std	C-QUATRE Mean/std	QUATRE Mean/std	MABC Mean/std	IABC Mean/std
f_{a1}	1.32E-04/5.1E-05	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)
f_{a2}	8.42E-01/7.4E-01	1.98E-04/1.4E-03(>)	5.29E-03/1.8E-02(>)	1.48E+06/5.5E+05(<)	1.51E+06/6.6E+05(<)
f_{a3}	7.64E+01/8.7E+01	5.60E-01/1.7E+00(>)	1.63E-01/8.9E-01(>)	7.18E+06/6.5E+06(<)	5.54E+07/3.6E+07(<)
f_{a4}	1.82E-02/8.3E-03	6.92E-10/1.4E-09(>)	4.53E-07/8.7E-07(>)	1.27E+04/3.3E+03(<)	5.63E+03/1.7E+03(<)
f_{a5}	7.75E-05/3.5E-05	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)
f_{a6}	4.90E+00/4.9E+00	3.17E+00/4.6E+00(>)	5.26E+00/1.0E+01(<)	4.87E+00/3.6E+00(>)	3.22E+00/4.2E+00(>)
f_{a7}	6.14E-01/2.1E+00	2.37E+00/9.4E+00(<)	7.23E-01/2.2E+00(<)	1.54E+01/3.3E+00(<)	1.89E+01/4.8E+00(<)
f_{a8}	2.03E+01/9.1E-02	2.05E+01/9.7E-02(<)	2.05E+01/8.8E-02(<)	2.04E+01/8.2E-02(<)	2.04E+01/7.7E-02(<)
f_{a9}	1.71E+00/1.2E+00	1.74E+00/1.3E+00(<)	1.98E+00/1.3E+00(<)	5.05E+00/5.8E-01(<)	4.15E+00/7.8E-01(<)
f_{a10}	1.73E-01/7.6E-02	1.75E-01/1.1E-01(<)	1.82E-01/1.2E-01(<)	7.61E-01/1.6E-01(<)	1.03E+00/2.2E-01(<)
f_{a11}	2.56E+00/1.7E+00	2.98E+00/1.6E+00(<)	2.75E+00/1.7E+00(<)	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)
f_{a12}	1.23E+01/5.5E+00	1.47E+01/7.9E+00(<)	1.38E+01/5.9E+00(<)	1.77E+01/4.2E+00(<)	1.57E+01/4.6E+00(<)
f_{a13}	1.75E+01/8.1E+00	2.16E+01/7.3E+00(<)	2.21E+01/8.5E+00(<)	1.90E+01/4.5E+00(<)	1.77E+01/5.1E+00(<)
f_{a14}	6.05E+01/7.3E+01	9.90E+01/9.7E+01(<)	9.80E+01/8.3E+01(<)	3.89E+01/2.7E+01(>)	6.99E+00/4.2E+00(<)
f_{a15}	8.24E+02/2.7E+02	1.08E+03/3.0E+02(<)	9.81E+02/3.0E+02(<)	1.12E+03/1.6E+02(<)	7.15E+02/1.8E+02(>)
f_{a16}	1.15E-01/1.0E-01	1.27E+00/3.4E-01(<)	1.19E+00/3.7E-01(<)	1.04E+00/1.6E-01(<)	1.01E+00/1.6E-01(<)
f_{a17}	1.03E+01/3.1E+00	1.01E+01/4.0E+00(>)	1.12E+01/2.5E+00(<)	1.22E+01/5.7E-01(<)	1.01E+01/1.3E-02(>)
f_{a18}	2.25E+01/5.7E+00	3.27E+01/9.0E+00(<)	3.10E+01/7.0E+00(<)	3.70E+01/3.3E+00(<)	3.42E+01/5.8E+00(<)
f_{a19}	5.23E-01/1.9E-01	6.87E-01/2.0E-01(<)	6.64E-01/2.7E-01(<)	5.24E-01/1.9E-01(<)	5.78E-01/1.8E-01(<)
f_{a20}	2.78E+00/5.7E-01	3.04E+00/5.4E-01(<)	2.92E+00/6.0E-01(<)	3.21E+00/2.3E-01(<)	2.79E+00/3.1E-01(<)
f_{a21}	3.88E+02/4.8E+01	3.57E+02/9.0E+01(>)	3.79E+02/6.1E+01(>)	2.41E+02/9.6E+01(>)	3.68E+02/8.6E+01(>)
f_{a22}	1.40E+02/1.0E+02	1.74E+02/9.3E+01(<)	1.96E+02/1.4E+02(<)	1.21E+02/6.5E+01(>)	3.80E+01/5.9E+01(>)
f_{a23}	8.82E+02/3.3E+02	1.05E+03/3.5E+02(<)	9.66E+02/3.8E+02(<)	1.26E+03/1.5E+02(<)	9.46E+02/2.1E+02(<)
f_{a24}	2.04E+02/5.2E+00	2.05E+02/4.7E+00(<)	2.06E+02/5.2E+00(<)	1.49E+02/9.1E+00(>)	1.38E+02/2.8E+01(>)
f_{a25}	2.04E+02/3.8E+00	2.05E+02/4.3E+00(<)	2.00E+02/1.9E+01(>)	1.84E+02/1.4E+01(>)	1.91E+02/2.3E+01(>)
f_{a26}	1.62E+02/4.4E+01	1.72E+02/5.9E+01(<)	1.75E+02/5.4E+01(<)	1.39E+02/7.7E+00(>)	1.32E+02/2.7E+01(>)
f_{a27}	3.31E+02/2.9E+01	3.55E+02/9.0E+01(<)	3.37E+02/7.8E+01(<)	3.33E+02/3.2E+01(<)	3.45E+02 /3.8E+0(<)
f_{a28}	2.89E+02/4.8E+01	2.84E+02/5.4E+01(>)	2.92E+02/5.2E+01(<)	2.43E+02/6.3E+01(>)	2.93E+02/1.6E+02(<)
$w/t/l$		19/0/9	21/0/7	17/0/11	16/0/12

3.4. Arithmetic Strategy. The model of $\mathrm{AF}_{gs,G}$ establishes the relationship between the global optimal solution and the global suboptimal solution so that the algorithm can effectively search the neighborhood between the global optimal solution and the global suboptimal solution. The characteristics of the function $f(r)$ can help us balance the search and development of the algorithm. The new arithmetic strategy is shown in Eq. (5) and Eq. (6).

$$\begin{cases} B_g = X_{g\mathrm{best},G} + F\bullet(X_{r1,G} - X_{r2,G}) - \mathrm{AF}_{gs,G}, \\ B_s = X_{s\mathrm{best},G} + F\bullet(X_{r1,G} - X_{r2,G}) + \mathrm{AF}_{gs,G}, \\ B = [B_g ; B_s], \end{cases} \tag{5}$$

$$\begin{cases} X = [X_1 ; X_2], \\ X \longleftarrow \bar{M} \otimes X + M \otimes B, \end{cases} \tag{6}$$

where $X_{g\mathrm{best},G}$ and $X_{s\mathrm{best},G}$ are the matrices generated, respectively, by taking the position coordinates of the global optimal solution and the global suboptimal solution as the row vector elements, and the number of columns of the matrix is determined by the number of individuals in the population. The $\mathrm{AF}_{gs,G}$ matrix can be obtained by the same principle. As shown in Eq. (6), the matrix M and the crossover strategy remain unchanged. The algorithm assigns individuals to two fixed subpopulations to increase the diversity of the population. Each subpopulation has different operation

Table 3: Comparison on 30D optimization under CEC2013 test suite with different algorithms.

Algorithms f	QUATRE-DEG Mean/std	C- QUATRE Mean/std	QUATRE Mean/std	MABC Mean/std	IABC Mean/std
f_{a1}	6.17*E*-04/1.1*E*-04	0.00*E*+00/0.0*E*+0(>)	2.23*E*-14/6.8*E*-1(>)	0.00*E*+00/0.0*E*+0(>)	0.00*E*+00/0.0*E*+0(>)
f_{a2}	3.16*E*+05/1.5*E*+05	5.42*E*+05/2.8*E*+0(<)	2.60*E*+05/1.0*E*+0(>)	6.89*E*+07/1.3*E*+0(<)	2.13*E*+07/3.9*E*+0(<)
f_{a3}	1.62*E*+06/4.2*E*+06	4.55*E*+05/1.8*E*+0(>)	2.41*E*+06/4.8*E*+0(<)	3.35*E*+09/8.6*E*+0(<)	4.53*E*+09/1.2*E*+0(<)
f_{a4}	2.39*E*+01/1.6*E*+01	8.33*E*+01/4.9*E*+0(<)	1.59*E*+01/1.0E+0(>)	8.06*E*+04/1.0*E*+0(<)	4.69*E*+04/7.1*E*+0(<)
f_{a5}	3.08*E*-04/5.0*E*-05	9.59*E*-14/4.2*E*-1(>)	1.11*E*-13/1.6*E*-1(>)	0.00*E*+00/0.0*E*+0(>)	0.00*E*+00/0.0*E*+0(>)
f_{a6}	5.01*E*+00/8.0*E*+00	1.05*E*+01/1.3*E*+0(<)	6.51*E*+00/1.1*E*+0(<)	1.76*E*+01/8.9*E*+0(<)	3.59*E*+01/2.4*E*+0(<)
f_{a7}	1.39*E*+01/1.0*E*+01	8.53*E*+00/7.8*E*+0(>)	2.17*E*+01/1.5*E*+0(<)	9.79*E*+01/8.1*E*+0(<)	9.64*E*+01/8.2*E*+0(<)
f_{a8}	2.09*E*+01/5.6*E*-02	2.10*E*+01/5.9*E*-0(<)	2.10*E*+01/5.0*E*-0(<)	2.09*E*+01/5.3*E*-0(=)	2.09*E*+01/6.1*E*-0(=)
f_{a9}	1.65*E*+01/5.3*E*+00	1.51*E*+01/5.6*E*+0(>)	1.69*E*+01/6.3*E*+0(<)	3.43*E*+01/1.2*E*+0(<)	2.92*E*+01/1.9*E*+0(<)
f_{a10}	5.98*E*-02/1.4*E*-02	2.69*E*-02/1.7*E*-0(<)	2.48*E*-02/1.6*E*-0(>)	6.22*E*-01/5.0*E*-0(<)	2.25*E*+00/7.4*E*-0(<)
f_{a11}	2.25*E*+01/5.6*E*+00	1.91*E*+01/5.1*E*+0(>)	2.69*E*+01/7.2*E*+0(<)	9.60*E*+01/8.0*E*+0(<)	0.00*E*+00/0.0*E*+0(>)
f_{a12}	6.19*E*+01/1.9*E*+01	8.75*E*+01/2.4*E*+0(<)	7.87*E*+01/2.5*E*+0(<)	2.28*E*+02/1.2*E*+0(<)	2.84*E*+02/1.2*E*+0(<)
f_{a13}	1.12*E*+02/2.8*E*+01	1.13*E*+02/3.0*E*+0(<)	1.25*E*+02/3.0*E*+0(<)	2.25*E*+02/1.4*E*+0(<)	2.83*E*+02/2.4*E*+0(<)
f_{a14}	6.88*E*+02/2.7*E*+02	6.48*E*+02/2.8*E*+0(>)	8.74*E*+02/3.2*E*+0(<)	4.14*E*+03/2.8*E*+0(<)	9.08*E*+01/3.2*E*+0(>)
f_{a15}	3.77*E*+03/6.1*E*+02	5.59*E*+03/8.8*E*+0(<)	5.23*E*+03/6.2*E*+0(<)	7.37*E*+03/2.9*E*+0(<)	5.00*E*+03/4.2*E*+0(<)
f_{a16}	3.01*E*-01/1.6*E*-01	2.43*E*+00/3.6*E*-0(<)	2.51*E*+00/4.2*E*-0(<)	2.43*E*+00/2.7*E*-0(<)	2.20*E*+00/2.5*E*-0(<)
f_{a17}	4.71*E*+01/4.9*E*+00	4.94*E*+01/5.5*E*+0(<)	5.46*E*+01/6.8*E*+0(<)	1.43*E*+02/8.9*E*+0(<)	3.08*E*+01/1.8*E*-0(>)
f_{a18}	8.61*E*+01/2.5*E*+01	1.74*E*+02/2.3*E*+0(<)	1.69*E*+02/2.4*E*+0(<)	2.52*E*+02/9.2*E*+0(<)	3.30*E*+02/2.8*E*+0(<)
f_{a19}	2.53*E*+00/5.5*E*-01	3.23*E*+00/7.4*E*-0(<)	3.47*E*+00/7.4*E*-0(<)	1.33*E*+01/1.0*E*+0(<)	2.75*E*+00/5.1*E*-0(<)
f_{a20}	1.14*E*+01/6.0*E*-01	1.19*E*+01/5.6*E*-0(<)	1.18*E*+01/5.8*E*-0(<)	1.38*E*+01/2.7*E*-0(<)	1.30*E*+01/3.0*E*-0(<)
f_{a21}	2.93*E*+02/7.5*E*+01	2.93*E*+02/8.8*E*+0(=)	2.84*E*+02/6.5*E*+0(>)	2.61*E*+02/4.9*E*+0(>)	3.28*E*+02/5.8*E*+0(<)
f_{a22}	6.09*E*+02/2.6*E*+02	6.47*E*+02/2.4*E*+0(<)	8.62*E*+02/3.0*E*+0(<)	4.68*E*+03/2.2*E*+0(<)	1.23*E*+02/3.0*E*+0(>)
f_{a23}	3.92*E*+03/7.6*E*+02	5.58*E*+03/6.3*E*+0(<)	5.37*E*+03/8.0*E*+0(<)	7.86*E*+03/3.0*E*+0(<)	5.82*E*+03/5.1*E*+0(<)
f_{a24}	2.31*E*+02/1.1*E*+01	2.32*E*+02/1.4*E*+0(<)	2.37*E*+02/1.2*E*+0(<)	2.87*E*+02/3.4*E*+0(<)	2.53*E*+02/5.9*E*+0(<)
f_{a25}	2.51*E*+02/7.4*E*+00	2.53*E*+02/7.0*E*+0(<)	2.59*E*+02/7.9*E*+0(<)	2.94*E*+02/4.2*E*+0(<)	2.63*E*+02/2.6*E*+0(<)
f_{a26}	2.47*E*+02/6.5*E*+01	2.56*E*+02/6.5*E*+0(<)	2.58*E*+02/6.8*E*+0(<)	2.08*E*+02/1.5*E*+0(>)	2.00*E*+02/5.4*E*-0(>)
f_{a27}	6.43*E*+02/1.1*E*+02	6.34*E*+02/1.0*E*+0(>)	6.51*E*+02/1.0*E*+0(<)	1.12*E*+03/1.5*E*+0(<)	8.71*E*+02/1.0*E*+0(<)
f_{a28}	3.22*E*+02/1.5*E*+02	3.83*E*+02/2.9*E*+0(<)	3.16*E*+02/1.4*E*+0(>)	3.00*E*+02/3.4*E*-1(>)	2.80*E*+02/6.0*E*+0(>)
$w/t/l$		18/1/9	21/0/7	22/1/5	19/1/8

strategies and can share information in the process of searching. When searching with the global optimization as the center, it will be guided by the global suboptimal force; similarly, when searching with the global suboptimal as the center, it will be guided by the global optimal force. This expands the algorithm's search for excellent regions and avoids falling into the local optimum. The double optimal guidance mechanism not only ensures the convergence of both global optimal position and global suboptimal position, but also brings more exploration to benchmark environments. Accordingly, the proposed algorithm can be an effective approach to enhance the search performance of the algorithm.

3.5. Algorithm Process. The process of QUATRE-DEG algorithm is given as follows.

4. Experiment and Analysis

In order to verify the performance of the algorithm, we carried out experiments with 58 real-parameter single objective optimizations in the CEC2013 and CEC2014 test suites. All benchmark functions are considered as black-box problems, and the search range is set to [-100, 100]. In CEC2013, there are 28 test functions, including unimodal functions (fa1-fa5), basic multimodal functions (fa6-f20), and composition functions (fa21-fa28). In CEC2014, there are 30 test functions including four different categories: unimodal function (fb1-fb3), simple multimodal function (fb4-fb16), hybrid function 1 (fb17-fb22), and composition function (fb23-fb30).

We have learned that the QUATRE algorithm focuses on the development, but the search is slightly inadequate. Our

Table 4: Comparison on 50D optimization under CEC2013 test suite with different algorithms.

Algorithms f	QUATRE-DEG Mean/std	C-QUATRE Mean/std	QUATRE Mean/std	MABC Mean/std	IABC Mean/std
f_{a1}	1.31*E*-03/1.4*E*-04	1.25*E*-13/1.1*E*-13(>)	2.27*E*-13/0.0*E*+00(>)	0.00*E*+00/0.0*E*+00(>)	0.00*E*+00/0.0*E*+00(>)
f_{a2}	1.06*E*+06/3.1*E*+05	8.70*E*+05/2.8*E*+05(>)	1.01*E*+06/4.3E+05(>)	2.99*E*+08/3.3*E*+07(<)	3.66*E*+07/5.6*E*+06(<)
f_{a3}	1.76*E*+07/1.7*E*+07	3.27*E*+07/4.3*E*+07(<)	2.46*E*+07/3.1*E*+07(<)	5.73*E*+10/6.6*E*+09(<)	1.48*E*+10/2.3*E*+09(<)
f_{a4}	1.45*E*+02/6.1*E*+01	8.71*E*+01/6.1*E*+01(>)	1.36*E*+02/8.2*E*+01(>)	1.54*E*+05/1.1*E*+04(<)	8.10*E*+04/8.9*E*+03(<)
f_{a5}	6.06*E*-04/8.1*E*-05	1.18*E*-13/2.2*E*-14(>)	1.78*E*-13/6.5*E*-14(>)	0.00*E*+00/0.0*E*+00(>)	0.00*E*+00/0.0*E*+00(>)
f_{a6}	4.36*E*+01/8.1*E*-01	4.39*E*+01/1.5*E*+00(<)	4.37*E*+01/8.0*E*-01(<)	4.45*E*+01/9.7*E*-01(<)	7.77*E*+01/2.2*E*+01(<)
f_{a7}	4.42*E*+01/1.6*E*+01	4.32*E*+01/1.7*E*+01(>)	6.56*E*+01/2.6*E*+01(<)	1.79*E*+02/8.7*E*+00(<)	1.13*E*+02/7.7*E*+00(<)
f_{a8}	2.11*E*+01/3.3*E*-02	2.12*E*+01/3.7*E*-02(<)	2.12*E*+01/3.1*E*-02(<)	2.11*E*+01/3.5*E*-02(=)	2.11*E*+01/5.0*E*-02(=)
f_{a9}	3.58*E*+01/1.0*E*+01	3.72*E*+01/1.1*E*+01(<)	3.58*E*+01/1.1*E*+01(=)	6.89*E*+01/1.3*E*+00(<)	5.66*E*+01/2.4*E*+00(<)
f_{a10}	1.08*E*-01/2.1*E*-02	6.63*E*-02/3.2*E*-02(>)	4.57*E*-02/2.8*E*-02(>)	3.30*E*+02/4.1*E*+01(<)	3.25*E*+00/6.0*E*-01(<)
f_{a11}	6.59*E*+01/1.4*E*+01	7.60*E*+01/1.9*E*+01(<)	8.56*E*+01/1.6*E*+01(<)	2.91*E*+02/1.6*E*+01(<)	1.73*E*-01/6.1*E*-01(>)
f_{a12}	1.06*E*+02/2.6*E*+01	1.84*E*+02/4.0*E*+01(<)	1.80*E*+02/4.7*E*+01(<)	4.83*E*+02/1.6*E*+01(<)	4.52*E*+02/3.4*E*+01(<)
f_{a13}	2.24*E*+02/4.5*E*+01	2.58*E*+02/5.4*E*+01(<)	2.62*E*+02/4.2*E*+01(<)	4.84*E*+02/1.8*E*+01(<)	5.70*E*+02/4.0*E*+01(<)
f_{a14}	1.67*E*+03/4.6*E*+02	2.29*E*+03/4.6*E*+02(<)	2.32*E*+03/4.8*E*+02(<)	9.50*E*+03/3.2*E*+02(<)	1.44*E*+02/5.8*E*+01(>)
f_{a15}	7.13*E*+03/8.1*E*+02	1.14*E*+04/1.3*E*+03(<)	1.11*E*+04/1.0*E*+03(<)	1.42*E*+04/3.4*E*+02(<)	1.06*E*+04/6.4*E*+02(<)
f_{a16}	5.13*E*-01/1.7*E*-01	3.34*E*+00/5.1*E*-01(<)	3.20*E*+00/5.3*E*-01(<)	3.38*E*+00/2.7*E*-01(<)	2.85*E*+00/3.2*E*-01(<)
f_{a17}	1.12*E*+02/1.4*E*+01	1.39*E*+02/1.5*E*+01(<)	1.40*E*+02/1.6*E*+01(<)	3.61*E*+02/1.5*E*+01(<)	5.24*E*+01/7.5*E*-01(>)
f_{a18}	1.47*E*+02/2.2*E*+01	3.58*E*+02/4.6*E*+01(<)	3.37*E*+02/3.9*E*+01(<)	5.30*E*+02/1.4*E*+01(<)	6.98*E*+02/5.0*E*+01(<)
f_{a19}	6.17*E*+00/1.6*E*+00	8.69*E*+00/1.9*E*+00(<)	8.79*E*+00/1.9*E*+00(<)	3.63*E*+01/1.7*E*+00(<)	6.89*E*+00/7.3*E*-01(<)
f_{a20}	2.10*E*+01/6.3*E*-01	2.18*E*+01/5.6*E*-01(<)	2.16*E*+01/7.2*E*-01(<)	2.38*E*+01/2.3*E*-01(<)	2.21*E*+01/4.3*E*-01(<)
f_{a21}	7.23*E*+02/4.2*E*+02	6.85*E*+02/4.4*E*+02(>)	7.39*E*+02/4.3*E*+02(<)	9.43*E*+02/3.3*E*+02(<)	8.98*E*+02/1.2*E*+02(<)
f_{a22}	1.95*E*+03/5.0*E*+02	2.27*E*+03/4.8*E*+02(<)	2.46*E*+03/5.0*E*+02(<)	9.99*E*+03/2.9*E*+02(<)	7.92*E*+01/5.7*E*+01(>)
f_{a23}	7.60*E*+03/9.7*E*+02	1.13*E*+04/1.1*E*+03(<)	1.14*E*+04/1.2*E*+03(<)	1.49*E*+04/3.9*E*+02(<)	1.18*E*+04/7.0*E*+02(<)
f_{a24}	2.72*E*+02/1.5*E*+01	2.73*E*+02/1.3*E*+01(<)	2.80*E*+02/1.1*E*+01(<)	3.72*E*+02/4.0*E*+00(<)	3.33*E*+02/1.0*E*+01(<)
f_{a25}	3.06*E*+02/1.2*E*+01	3.13*E*+02/1.5*E*+01(<)	3.13*E*+02/1.0*E*+01(<)	3.92*E*+02/4.4*E*+00(<)	4.20*E*+02/1.2*E*+01(<)
f_{a26}	3.63*E*+02/5.1*E*+01	3.65*E*+02/3.7*E*+01(<)	3.58*E*+02/6.0*E*+01(>)	2.35*E*+02/5.4*E*+00(>)	2.01*E*+02/1.5*E*-01(>)
f_{a27}	1.04*E*+03/1.6*E*+02	1.08*E*+03/1.7*E*+02(<)	1.13*E*+03/1.5*E*+02(<)	2.03*E*+03/3.4*E*+01(<)	1.70*E*+03/7.6*E*+01(<)
f_{a28}	8.13*E*+02/1.0*E*+03	9.90*E*+02/1.2*E*+03(<)	8.73*E*+02/1.1*E*+03(<)	4.23*E*+02/1.6*E*+02(>)	4.04*E*+02/2.8*E*+01(>)
$w/t/l$		21/0/7	21/1/6	23/1/4	19/1/8

QUATRE-DEG algorithm focuses on improving the search ability of the algorithm to improve the overall performance of the algorithm. In the experiments, to verify the overall performance of the algorithm, it is not only compared with the advanced QUATRE variant but also with the famous variants of the ABC algorithm, which is famous for its strong search ability. The new QUATRE-DEG algorithm is a further improvement of QUATRE algorithm, and the search strategy of ABC algorithm determines its strong exploration performance. In QUATRE-DEG, the performance of the algorithm is mainly improved by enhancing the exploration ability. Therefore, we use some famous ABC variants and advanced QUATRE variants to compare with our proposed algorithm to verify that the performance of QUATRE-DEG is reliable. The comparison includes the original QUATRE [14], C-QUATRE [22], IABC [25], MABC [26], and QUATRE-DEG algorithms. In all QUATRE-based algorithms, the population size is set to 100, while the population size of all ABC algorithms is set to 50, and the parameter limit is set to D^* SN/2, where D is the dimension size and SN is the population size. In addition, in the QUATRE, C-QUATRE, and QUATRE-DEG algorithms, the value of parameter F is set to the recommended value of 0.7. In the IABC algorithm, MR is set to 0.3.

The experiments were complete on Windows 10 Professional system with 8 GB RAM and Intel i5-4210 m processor. The evaluation tool is MATLAB 2019a and the *maxFES* is set to $10e4^*D$.

4.1. Parameter Setting. There is an important parameter z in the QUATRE-DEG algorithm, which determines the size of the population allocated to the global optimal and global

Table 5: Comparison on 10D optimization under CEC2014 test suite with different algorithms.

Algorithms f	QUATRE-DEG Mean/std	C-QUATRE Mean/std	QUATRE Mean/std	MABC Mean/std	IABC Mean/std
f_{b1}	2.84E-01/2.0E-01	1.20E-10/1.8E-10(>)	6.59E-08/7.5E-08(>)	1.34E+05/6.7E+04(<)	3.14E+05/1.3E+05(<)
f_{b2}	5.82E+00/2.9E+01	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)	1.08E+02/1.5E+02(<)	4.47E+02/5.1E+02(<)
f_{b3}	1.14E-03/3.8E-04	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)	1.50E+02/2.5E+02(<)	1.44E+02/8.5E+01(<)
f_{b4}	2.36E+01/1.6E+01	2.10E+01/1.7E+01(>)	2.17E+01/1.6E+01(>)	1.16E+00/1.4E+00(>)	1.12E+00/1.7E+00(>)
f_{b5}	1.96E+01/2.8E+00	2.02E+01/5.5E-02(<)	1.90E+01/4.8E+00(>)	1.99E+01/1.2E+00(<)	1.40E+01/8.8E+00(>)
f_{b6}	2.33E-01/3.6E-01	6.54E-01/7.8E-01(<)	4.09E-01/6.2E-01(<)	6.54E-01/4.2E-01(<)	3.61E-01/2.9E-01(<)
f_{b7}	9.98E-02/4.8E-02	8.97E-02/6.8E-02(>)	1.03E-01/6.1E-02(<)	1.36E-02/1.5E-02(>)	1.36E-02/1.3E-02(>)
f_{b8}	2.69E+00/1.5E+00	3.02E+00/1.4E+00(<)	3.06E+00/1.7E+00(<)	0.00E+00/0.0E+00(>)	0.00E+00/0.0E+00(>)
f_{b9}	7.92E+00/4.1E+00	9.30E+00/4.0E+00(<)	9.35E+00/4.4E+00(<)	9.07E+00/2.0E+00(<)	5.25E+00/1.5E+00(>)
f_{b10}	4.23E+01/6.4E+01	7.09E+01/7.5E+01(<)	6.07E+01/6.3E+01(<)	1.88E+00/2.9E+00(>)	1.63E-01/7.2E-02(>)
f_{b11}	3.48E+02/2.1E+02	3.89E+02/1.9E+02(<)	3.32E+02/2.2E+02(>)	5.28E+02/1.1E+02(<)	1.50E+02/8.6E+01(>)
f_{b12}	2.30E-02/2.6E-02	3.16E-01/1.8E-01(<)	2.75E-01/1.9E-01(<)	4.50E-01/1.1E-01(<)	3.56E-01/8.0E-02(<)
f_{b13}	1.17E-01/3.2E-02	1.48E-01/7.8E-02(<)	1.48E-01/6.1E-02(<)	1.83E-01/3.1E-02(<)	1.71E-01/2.6E-02(<)
f_{b14}	1.07E-01/4.2E-02	1.52E-01/5.9E-02(<)	1.58E-01/5.2E-02(<)	1.86E-01/3.1E-02(<)	1.44E-01/2.6E-02(<)
f_{b15}	8.24E-01/2.7E-01	1.30E+00/4.5E-01(<)	1.06E+00/4.2E-01(<)	1.33E+00/2.2E-01(<)	1.00E+00/1.9E-01(<)
f_{b16}	1.79E+00/5.6E-01	1.93E+00/5.5E-01(<)	1.94E+00/5.5E-01(<)	2.72E+00/1.6E-01(<)	1.80E+00/3.2E-01(<)
f_{b17}	8.07E+01/7.4E+01	8.26E+01/8.0E+01(<)	9.87E+01/6.7E+01(<)	9.45E+03/6.1E+03(<)	1.82E+04/1.7E+04(<)
f_{b18}	2.74E+00/2.9E+00	2.39E+00/2.0E+00(<)	3.87E+00/4.4E+00(<)	3.67E+02/2.9E+02(<)	2.45E+03/1.4E+03(<)
f_{b19}	3.94E-01/4.6E-01	8.26E-01/9.7E-01(<)	7.46E-01/7.6E-01(<)	8.59E-01/2.4E-01(<)	3.95E-01/3.1E-01(<)
f_{b20}	1.09E+00/8.2E-01	1.15E+00/7.1E-01(<)	1.92E+00/2.5E+00(<)	1.14E+02/1.0E+02(<)	3.72E+02/3.6E+02(<)
f_{b21}	3.13E+01/4.7E+01	5.02E+01/6.0E+01(<)	5.19E+01/6.2E+01(<)	1.02E+03/9.4E+02(<)	2.82E+03/1.9E+03(<)
f_{b22}	1.15E+01/1.1E+01	8.59E+00/1.1E+01(>)	1.60E+01/3.0E+01(<)	5.96E-01/7.3E-01(>)	5.77E+00/1.8E+01(>)
f_{b23}	3.29E+02/8.1E-07	3.29E+02/2.9E-13(=)	3.29E+02/2.9E-13(=)	3.18E+02/4.8E+01(>)	3.19E+02/6.5E+01(>)
f_{b24}	1.19E+02/4.9E+00	1.18E+02/6.1E+00(>)	1.19E+02/5.8E+00(<)	1.16E+02/3.4E+00(>)	1.09E+02/3.6E+00(>)
f_{b25}	1.73E+02/3.8E+01	1.85E+02/3.3E+01(<)	1.76E+02/3.7E+01(<)	1.52E+02/7.3E+00(>)	1.34E+02/6.3E+00(>)
f_{b26}	1.00E+02/4.1E-02	1.00E+02/7.4E-02(=)	1.00E+02/7.0E-02(=)	1.00E+02/4.8E-02(=)	1.00E+02/2.4E-02(=)
f_{b27}	2.25E+02/1.8E+02	2.28E+02/1.9E+02(<)	2.30E+02/1.7E+02(<)	3.38E+01/7.2E+01(>)	4.69E+01/1.2E+02(>)
f_{b28}	4.06E+02/5.4E+01	4.24E+02/7.6E+01(<)	4.32E+02/6.0E+01(<)	3.62E+02/4.9E+00(>)	3.69E+02/8.9E+01(>)
f_{b29}	4.20E+04/3.0E+05	1.17E+05/4.7E+05(<)	1.22E+05/4.9E+05(<)	3.56E+02/5.4E+01(>)	4.04E+02/7.0E+01(>)
f_{b30}	5.18E+02/7.5E+01	5.56E+02/1.9E+02(<)	5.46E+02/1.3E+02(<)	5.23E+02/3.2E+01(<)	5.46E+02/4.6E+01(<)
$w/t/l$		20/2/8	22/2/6	18/1/11	15/1/14

suboptimal subpopulations. The size of the population will affect the performance of the algorithm, especially when the two populations affect each other. Accordingly, it is necessary to determine the size of the two subpopulations first. Ideologically, it usually does not allow the size of the global optimal subpopulation to be lower than that of the global suboptimal. The population should be dominated by the global optimal subpopulation, which conforms to the basic logic of finding the global optimal solution. In our setting, when z value is larger, the more individuals are allocated to the global optimal subpopulation and more development will be brought to the global optimal position; on the contrary, when the z value is smaller, the less individuals are allocated to the global optimal subpopulation and the less the development will be. Therefore, in this experiment, the parameter z is set to 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6, respectively, and the dimension is set to 30D. The algorithms run 10 times for each different z value. The results are shown in Table 1. The evaluation compares the average and variance of the running results. It is obvious that when z is set to 0.4, the algorithm QUATRE-DEG is the most competitive. It can also be analyzed that when z is larger, the algorithm will pay more attention to exploration, and when z is smaller, the algorithm will have a better development. It also shows that when $z = 0.4$ the exploration and development of the algorithm will tend to a balance; it is undoubtedly a good choice of the z value.

Table 6: Comparison on 30D optimization under CEC2014 test suite with different algorithms.

Algorithms f	QUATRE-DEG Mean/std	C-QUATRE Mean/std	QUATRE Mean/std	MABC Mean/std	IABC Mean/std
f_{b1}	1.15E+05/8.1E+04	2.10E+05/1.7E+05(<)	1.37E+05/1.0E+05(<)	6.65E+07/1.1E+07(<)	2.07E+07/3.9E+06(<)
f_{b2}	1.86E+01/5.1E+01	5.57E-16/4.0E-15(>)	1.39E-14/1.4E-14(>)	2.79E+01/4.3E+01(<)	7.21E+03/3.9E+03(<)
f_{b3}	3.84E-03/5.8E-04	1.00E-14/2.2E-14(>)	3.34E-14/2.8E-14(>)	4.69E+02/3.0E+02(<)	7.78E+02/3.2E+02(<)
f_{b4}	3.78E+00/1.5E+01	6.42E+00/1.9E+01(<)	2.53E+00/1.2E+01(>)	7.65E+01/1.5E+01(<)	7.86E+01/1.3E+01(<)
f_{b5}	2.00E+01/4.1E-02	2.05E+01/8.3E-02(<)	2.05E+01/9.4E-02(<)	2.09E+01/4.7E-02(<)	2.05E+01/4.9E-02(<)
f_{b6}	2.60E+00/2.0E+00	2.08E+00/1.8E+00(>)	3.69E+00/2.3E+00(<)	2.65E+01/1.1E+00(<)	9.34E+00/1.9E+00(<)
f_{b7}	5.90E-03/6.3E-03	4.16E-03/7.6E-03(>)	7.77E-03/1.2E-02(<)	0.00E+00/0.0E+00(>)	1.20E-04/2.1E-04(>)
f_{b8}	2.20E+01/5.4E+00	1.89E+01/4.8E+00(>)	2.85E+01/7.1E+00(<)	8.68E+01/8.0E+00(<)	8.31E-02/2.7E-01(>)
f_{b9}	5.53E+01/1.4E+01	7.11E+01/2.3E+01(<)	7.07E+01/2.0E+01(<)	1.84E+02/1.2E+01(<)	6.98E+01/9.2E+00(<)
f_{b10}	3.86E+02/1.8E+02	3.15E+02/1.7E+02(>)	4.99E+02/2.5E+02(<)	3.26E+03/1.8E+02(<)	2.45E+00/1.3E+00(>)
f_{b11}	2.41E+03/6.5E+02	3.30E+03/7.2E+02(<)	3.35E+03/6.9E+02(<)	6.15E+03/2.4E+02(<)	2.88E+03/3.7E+02(<)
f_{b12}	8.04E-02/3.7E-02	5.04E-01/2.2E-01(<)	4.43E-01/2.3E-01(<)	1.75E+00/1.5E-01(<)	6.95E-01/7.6E-02(<)
f_{b13}	2.82E-01/7.2E-02	3.39E-01/7.1E-02(<)	3.10E-01/8.4E-02(<)	4.26E-01/4.1E-02(<)	2.58E-01/3.0E-02(<)
f_{b14}	2.80E-01/1.1E-01	2.70E-01/6.7E-02(>)	3.38E-01/1.3E-01(<)	2.47E-01/2.9E-02(>)	1.36E-01/1.8E-02(>)
f_{b15}	3.97E+00/1.0E+00	8.37E+00/2.2E+00(<)	7.41E+00/2.1E+00(<)	1.75E+01/1.2E+00(<)	1.28E+01/1.4E+00(<)
f_{b16}	1.03E+01/6.3E-01	1.04E+01/7.6E-01(<)	1.03E+01/6.8E-01(=)	1.26E+01/1.9E-01(<)	1.05E+01/3.5E-01(<)
f_{b17}	1.68E+03/8.1E+02	2.66E+03/2.7E+03(<)	2.31E+03/2.6E+03(<)	1.69E+06/4.9E+05(<)	1.41E+06/5.6E+05(<)
f_{b18}	6.05E+01/3.4E+01	3.85E+01/2.1E+01(>)	8.77E+01/4.6E+01(<)	4.06E+02/1.4E+02(<)	1.90E+04/2.2E+04(<)
f_{b19}	4.11E+00/1.5E+00	4.46E+00/1.7E+00(<)	4.75E+00/1.4E+00(<)	9.25E+00/2.8E-01(<)	8.57E+00/1.5E+00(<)
f_{b20}	2.57E+01/1.6E+01	2.59E+01/1.2E+01(<)	3.89E+01/3.7E+01(<)	4.63E+03/1.5E+03(<)	2.70E+03/1.4E+03(<)
f_{b21}	4.60E+02/2.0E+02	5.74E+02/5.5E+02(<)	4.94E+02/2.7E+02(<)	2.22E+05/7.6E+04(<)	2.06E+05/7.0E+04(<)
f_{b22}	2.91E+02/1.9E+02	3.27E+02/1.6E+02(<)	3.14E+02/1.5E+02(<)	1.77E+02/5.1E+01(>)	2.27E+02/6.9E+01(>)
f_{b23}	3.15E+02/4.4E-07	3.15E+02/4.0E-13(=)	3.15E+02/4.0E-13(=)	3.15E+02/4.4E-13(=)	3.15E+02/3.6E-01(=)
f_{b24}	2.24E+02/3.9E+00	2.25E+02/3.2E+00(<)	2.28E+02/5.8E+00(<)	2.24E+02/1.5E+00(=)	2.00E+02/3.1E-03(>)
f_{b25}	2.03E+02/4.7E-01	2.03E+02/6.5E-01(=)	2.03E+02/8.4E-01(=)	2.15E+02/2.0E+00(<)	2.00E+02/0.0E+00(>)
f_{b26}	1.00E+02/2.0E+01	1.00E+02/8.3E-02(=)	1.00E+02/1.0E-01(=)	1.00E+02/5.2E-02(=)	1.00E+02/3.7E-02(=)
f_{b27}	3.86E+02/5.3E+01	3.93E+02/4.5E+01(<)	4.17E+02/7.1E+01(<)	5.00E+02/2.1E+01(<)	4.05E+02/1.7E+00(<)
f_{b28}	8.51E+02/7.6E+01	8.73E+02/1.2E+02(<)	9.20E+02/1.5E+02(<)	9.44E+02/2.0E+01(<)	8.93E+02/4.7E+01(<)
f_{b29}	1.70E+05/1.2E+06	1.77E+05/1.3E+06(<)	1.81E+05/1.3E+06(<)	2.37E+03/6.4E+02(>)	3.56E+03/1.7E+03(>)
f_{b30}	1.53E+03/7.5E+02	1.58E+03/6.4E+02(<)	1.80E+03/7.5E+02(<)	4.52E+03/6.2E+02(<)	5.40E+03/9.4E+02(<)
$w/t/l$		19/3/8	24/3/3	24/2/4	20/2/8

4.2. Comparison of Optimization Accuracy. For the data reliability, 51 independent experiments were carried out with the comparison of the five algorithms. The mean and variance of the outcome of each algorithm are also compared and analyzed. Firstly, the dimensions were set to 10D, 30D, and 50D for the CEC2013 test suite, and then, 28 benchmark functions were carried out. The comparative experimental results are shown in Tables 2–4 for the dimensions of 10D, 30D, and 50D, respectively. Table 2 shows that the proposed QUATRE-DEG obtains better performance in 19 functions compared with the C-QUATRE algorithm, better performance in 21 functions compared with the QUATRE algorithm, and better performance in 17 functions compared with both of the MABC and IABC algorithms. Table 3 shows the QUATRE-DEG obtains better performance in 22 functions as well as similar performance in 1 function compared with the MABC, better performance in 21 functions compared with the QUATRE, better performance in 19 functions as well as similar performance in 1 function compared with the IABC algorithms, and better performance in 18 functions as well as similar performance in 1 function compared with the C-QUATRE. Table 4 shows the QUATRE-DEG obtains better performance in 23 functions as well as similar performance in 1 function compared with the MABC, better performance in 21 functions as well as similar performance in 1 function compared with the QUATRE, better performance

Table 7: Comparison on 50D optimization under CEC2014 test suite with different algorithms.

Algorithms f	QUATRE-DEG Mean/std	C-QUATRE Mean/std	QUATRE Mean/std	MABC Mean/std	IABC Mean/std
f_{b1}	8.15*E*+05/3.3*E*+05	6.41*E*+05/2.5*E*+05(>)	6.64*E*+05/3.1*E*+05(>)	2.55*E*+08/2.9*E*+07(<)	1.84*E*+07/3.4*E*+06(<)
f_{b2}	6.95*E*+03/6.5*E*+03	1.37*E*-03/3.4*E*-03(>)	1.46*E*-04/3.4*E*-04(>)	3.78*E*+03/3.5*E*+03(>)	1.93*E*+04/1.4*E*+04(<)
f_{b3}	2.73*E*-01/2.0*E*-01	1.50*E*-03/4.7*E*-03(>)	3.16*E*-04/9.1*E*-04(>)	9.86*E*+04/8.6*E*+03(<)	2.87*E*+03/8.4*E*+02(<)
f_{b4}	4.96*E*+01/4.1*E*+01	4.93*E*+01/3.7*E*+01(>)	5.05*E*+01/4.2*E*+01(<)	9.64*E*+01/2.1*E*+00(<)	1.33*E*+02/3.3*E*+01(<)
f_{b5}	2.00*E*+01/9.3*E*-02	2.07*E*+01/9.0*E*-02(<)	2.07*E*+01/7.6*E*-02(<)	2.11*E*+01/3.6*E*-02(<)	2.06*E*+01/4.6*E*-02(<)
f_{b6}	8.63*E*+00/4.6*E*+00	9.40*E*+00/3.6*E*+00(<)	1.13*E*+01/4.3*E*+00(<)	5.74*E*+01/1.3*E*+00(<)	1.94*E*+01/2.4*E*+00(<)
f_{b7}	6.60*E*-03/5.5*E*-03	2.08*E*-03/4.2*E*-03(>)	2.27*E*-03/5.2*E*-03(>)	0.00*E*+00/0.0*E*+00(>)	8.38*E*-03/1.2*E*-02(<)
f_{b8}	5.91*E*+01/1.4*E*+01	6.87*E*+01/1.6*E*+01(<)	7.09*E*+01/1.4*E*+01(<)	2.56*E*+02/1.2*E*+01(<)	1.44*E*+00/2.7*E*+00(>)
f_{b9}	1.12*E*+02/2.3*E*+01	1.40*E*+02/3.5*E*+01(<)	1.59*E*+02/3.9*E*+01(<)	4.21*E*+02/1.4*E*+01(<)	1.63*E*+02/1.5*E*+01(<)
f_{b10}	1.36*E*+03/3.5*E*+02	1.67*E*+03/5.4*E*+02(<)	1.69*E*+03/5.2*E*+02(<)	8.74*E*+03/3.3*E*+02(<)	6.20*E*+00/1.9*E*+00(>)
f_{b11}	5.07*E*+03/8.8*E*+02	7.34*E*+03/1.3*E*+03(<)	7.14*E*+03/1.0*E*+03(<)	1.26*E*+04/3.6*E*+02(<)	6.60*E*+03/6.2*E*+02(<)
f_{b12}	1.62*E*-01/6.2*E*-02	4.81*E*-01/2.5*E*-01(<)	4.72*E*-01/2.5*E*-01(<)	2.85*E*+00/2.2*E*-01(<)	8.41*E*-01/1.2*E*-01(<)
f_{b13}	3.94*E*-01/5.7*E*-02	4.50*E*-01/9.5*E*-02(<)	4.64*E*-01/1.0*E*-01(<)	5.86*E*-01/3.6*E*-02(<)	4.43*E*-01/3.2*E*-02(<)
f_{b14}	4.34*E*-01/2.2*E*-01	4.40*E*-01/2.2*E*-01(<)	4.39*E*-01/2.1*E*-01(<)	3.00*E*-01/2.7*E*-02(>)	2.18*E*-01/1.7*E*-02(>)
f_{b15}	8.42*E*+00/2.2*E*+00	1.67*E*+01/3.7*E*+00(<)	1.69*E*+01/4.3*E*+00(<)	4.08*E*+01/1.7*E*+00(<)	3.09*E*+01/2.4*E*+00(<)
f_{b16}	1.94*E*+01/9.8*E*-01	1.96*E*+01/8.6*E*-01(<)	1.99*E*+01/8.2*E*-01(<)	2.24*E*+01/1.8*E*-01(<)	1.91*E*+01/4.4*E*-01(>)
f_{b17}	2.35*E*+04/1.4*E*+04	2.65*E*+04/2.1*E*+04(<)	2.75*E*+04/1.4*E*+04(<)	1.35*E*+07/1.7*E*+06(<)	5.19*E*+06/1.1*E*+06(<)
f_{b18}	8.88*E*+02/9.2*E*+02	2.84*E*+02/5.8*E*+02(>)	2.42*E*+02/1.9*E*+02(>)	1.37*E*+03/8.9*E*+02(<)	4.43*E*+04/8.5*E*+04(<)
f_{b19}	8.65*E*+00/2.0*E*+00	8.20*E*+00/2.4*E*+00(>)	8.88*E*+00/2.4*E*+00(<)	3.19*E*+01/7.5*E*+00(<)	1.49*E*+01/6.3*E*+00(<)
f_{b20}	1.03*E*+02/6.3*E*+01	1.14*E*+02/6.3*E*+01(<)	1.32*E*+02/6.8*E*+01(<)	4.84*E*+04/8.5*E*+03(<)	6.29*E*+03/2.2*E*+03(<)
f_{b21}	5.63*E*+03/4.5*E*+03	1.55*E*+04/2.3*E*+04(<)	9.09*E*+03/7.4*E*+03(<)	4.20*E*+06/1.1*E*+06(<)	1.56*E*+06/4.6*E*+05(<)
f_{b22}	8.76*E*+02/2.4*E*+02	1.11*E*+03/2.9*E*+02(<)	1.07*E*+03/2.9*E*+02(<)	1.00*E*+03/1.3*E*+02(<)	7.93*E*+02/1.4*E*+02(>)
f_{b23}	3.44*E*+02/2.2*E*-07	3.44*E*+02/5.3*E*-13(=)	3.44*E*+02/4.4*E*-13(=)	3.44*E*+02/4.6*E*-13(=)	2.68*E*+02/7.3*E*+01(>)
f_{b24}	2.74*E*+02/2.9*E*+00	2.76*E*+02/3.2*E*+00(<)	2.76*E*+02/3.5*E*+00(<)	2.58*E*+02/1.1*E*+00(>)	2.00*E*+02/2.0*E*-03(>)
f_{b25}	2.07*E*+02/1.4*E*+00	2.09*E*+02/2.1*E*+00(<)	2.07*E*+02/1.7*E*+00(=)	2.53*E*+02/5.0*E*+00(<)	2.00*E*+02/0.0*E*+00(>)
f_{b26}	1.25*E*+02/5.4*E*+01	1.44*E*+02/7.9*E*+01(<)	1.28*E*+02/6.5*E*+01(<)	1.01*E*+02/4.6*E*-02(>)	1.49*E*+02/5.0*E*+01(<)
f_{b27}	5.69*E*+02/8.9*E*+01	6.08*E*+02/9.0*E*+01(<)	6.58*E*+02/9.0*E*+01(<)	1.57*E*+03/7.7*E*+01(<)	7.53*E*+02/9.6*E*+01(<)
f_{b28}	1.33*E*+03/2.8*E*+02	1.42*E*+03/3.7*E*+02(<)	1.44*E*+03/2.7*E*+02(<)	1.49*E*+03/3.7*E*+01(<)	1.46*E*+03/1.1*E*+02(<)
f_{b29}	7.21*E*+05/5.1*E*+06	1.77*E*+06/8.8*E*+06(<)	1.97*E*+06/9.9*E*+06(<)	3.22*E*+04/1.5*E*+04(>)	3.44*E*+03/9.0*E*+02(>)
f_{b30}	9.54*E*+03/9.8*E*+02	9.92*E*+03/9.4*E*+02(<)	9.94*E*+03/8.8*E*+02(<)	1.63*E*+04/1.7*E*+03(<)	1.26*E*+04/1.3*E*+03(<)
$w/t/l$		22/1/7	23/2/5	23/1/6	21/0/9

in 21 functions compared with the C-QUATRE, and better performance in 19 functions as well as similar performance in 1 function compared with the IABC. Based on the above results obtained from the CEC2013 test suite, it is obvious the proposed QUATRE-DEG performs better than the other algorithms.

The CEC2014 test suite also was used in the experiment for the performance evaluation of the proposed QUATRE-DEG. The results of 30 benchmark functions of CEC2014 are shown in Tables 5–7 for the dimensions of 10D, 30D, and 50D, respectively. Table 5 shows that the proposed QUATRE-DEG obtains better performance in 21 functions and similar performance in 3 function compared with the QUATRE algorithm, better performance in 20 functions and similar performance in 2 function compared with the C-QUATRE algorithm, better performance in 18 functions and similar performance in 1 function compared with the MABC algorithm, and better performance in 15 functions and similar performance in 1 function compared with the IABC algorithm. Table 6 shows the QUATRE-DEG obtains better performance in 23 functions and similar performance in 4 functions compared with the QUATRE, better performance in 23 functions and similar performance in 3 functions compared with the MABC, better performance in 19 functions and similar performance in 2 functions compared with the IABC, and better performance in 19 functions and

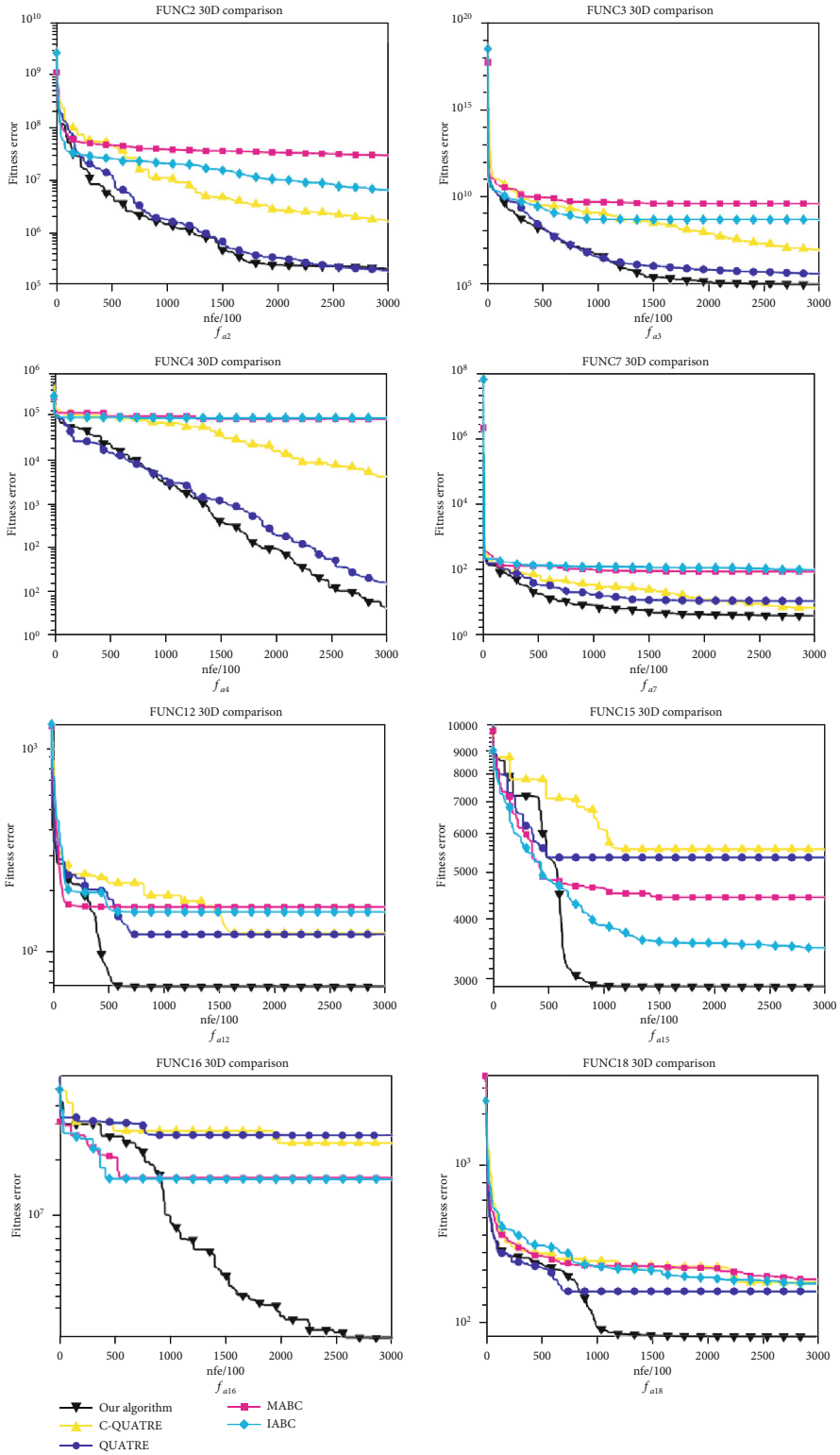


Figure 1: Comparison of convergence curves on 30D optimization under CEC2013 with different algorithms.

similar performance in 3 functions compared with the C-QUATRE. Table 7 shows the QUATRE-DEG obtains better performance in 23 functions and similar performance in 2 functions compared with the QUATRE, better performance in 23 functions and similar performance in 1 function compared with the MABC, better performance in 22 functions and similar performance in 1 function compared with the C-QUATRE, and better performance in 21 functions

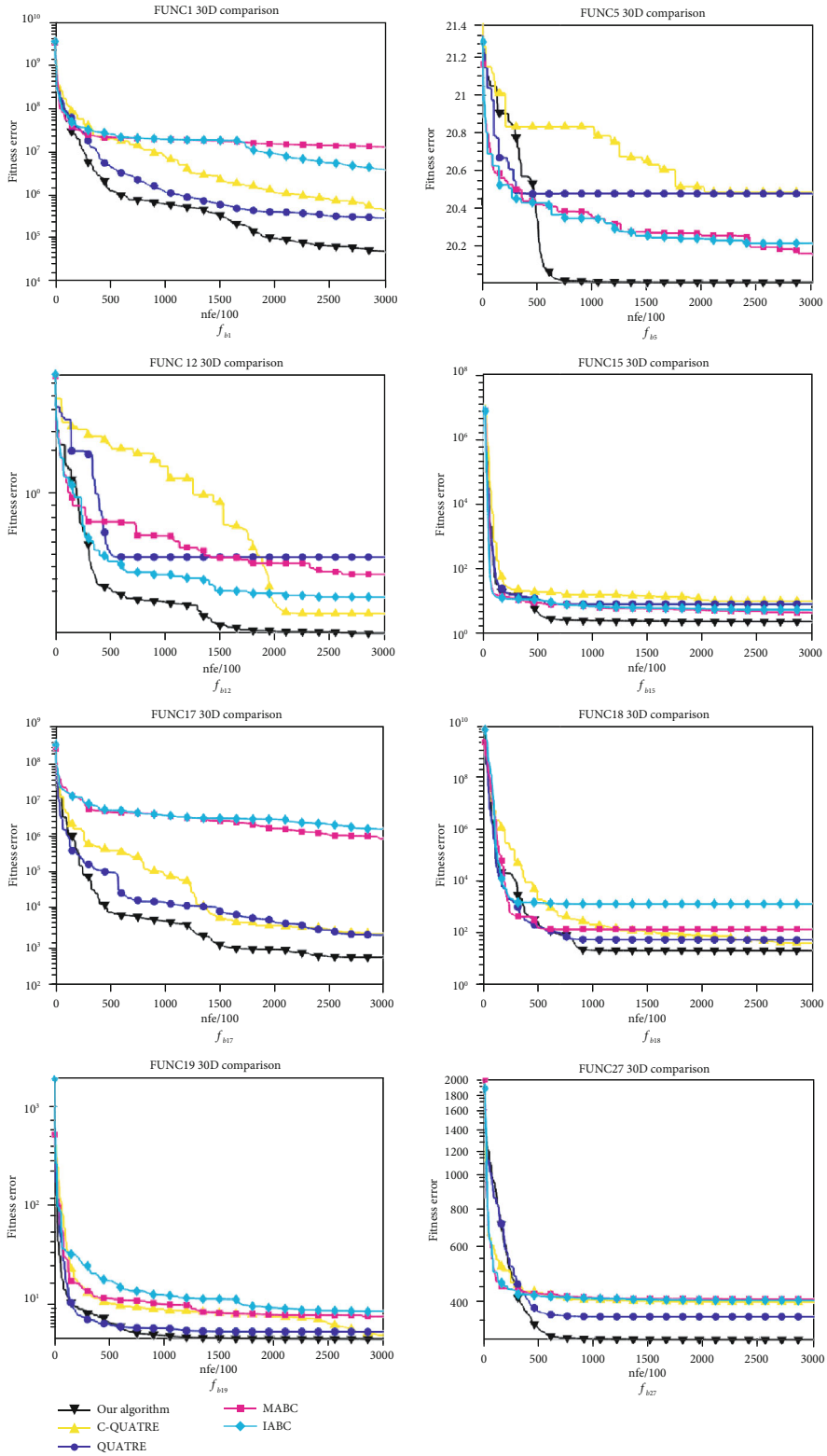


Figure 2: Comparison of convergence curves on 30D optimization under CEC2014 with different algorithms.

compared with the IABC. Based on the above results obtained from the CEC2014 test suite, it is also obvious the proposed QUATRE-DEG performs better than the other algorithms.

4.3. Analysis of Convergence Rate. The QUATRE-DEG algorithm was also evaluated from the perspective of convergence speed. Figures 1 and 2 show the convergence curves obtained by the representative benchmark functions in the CEC2013

Table 8: Wilcoxon's test among C-QUATRE, QUATRE, MABC, IABC, and QUATRE-DEG algorithms over all dimensions under CEC2013 test suite.

D	Algorithms	R^+	R^-	p value	+	=	-	Dec.
10	QUATRE-DEG vs. C-QUATRE	298	108	0.02941	19	0	9	+
	QUATRE-DEG vs. QUATRE	321	85	0.00606	21	0	7	+
	QUATRE-DEG vs. MABC	247	159	0.32721	17	0	11	=
	QUATRE-DEG vs. IABC	226	180	0.61364	16	0	12	=
30	QUATRE-DEG vs. C-QUATRE	273	105	0.04291	18	1	9	+
	QUATRE-DEG vs. QUATRE	327.5	78.5	0.0035	21	0	7	+
	QUATRE-DEG vs. MABC	345	33	4.74602E-5	22	1	5	+
	QUATRE-DEG vs. IABC	283	95	0.02294	19	1	8	+
50	QUATRE-DEG vs. C-QUATRE	329.5	76.5	0.00296	21	0	7	+
	QUATRE-DEG vs. QUATRE	325	53	5.9849E-4	21	1	6	+
	QUATRE-DEG vs. MABC	346	32	4.09335E-5	23	1	4	+
	QUATRE-DEG vs. IABC	280	98	0.02802	19	1	8	+

Table 9: Wilcoxon's test among C-QUATRE, QUATRE, MABC, IABC, and QUATRE-DEG algorithms over all dimensions under CEC2014 test suite.

D	Algorithms	R^+	R^-	p value	+	=	-	Dec.
10	QUATRE-DEG vs. C-QUATRE	313	93	0.01102	20	2	8	+
	QUATRE-DEG vs. QUATRE	299	79	0.007	21	3	6	+
	QUATRE-DEG vs. MABC	258	177	0.39265	18	1	11	=
	QUATRE-DEG vs. IABC	221	214	0.94907	15	1	14	=
30	QUATRE-DEG vs. C-QUATRE	291.5	86.5	0.01224	19	3	8	+
	QUATRE-DEG vs. QUATRE	325	26	3.18587E-5	23	4	3	+
	QUATRE-DEG vs. MABC	335.5	42.5	1.64852E-4	23	3	4	+
	QUATRE-DEG vs. IABC	307	99	0.01673	19	2	9	+
50	QUATRE-DEG vs. C-QUATRE	338	97	0.00801	22	1	7	+
	QUATRE-DEG vs. QUATRE	325	81	0.00441	23	2	5	+
	QUATRE-DEG vs. MABC	370	65	5.50576E-4	23	1	6	+
	QUATRE-DEG vs. IABC	345	120	0.01966	21	0	9	+

and CEC2014 test suites, respectively. Figure 1 shows the proposed QUATRE-DEG performs much better than the other algorithms in functions fa12, fa15, fa16, and fa18 and performs similar to the other algorithms in functions fa2, fa3, fa4, and fa7. Figure 2 shows the convergence speed of the proposed QUATRE-DEG is significantly better than the other algorithms in functions fb1, fb5, fb12, and fb17 and is slightly better than the other algorithms in fb15, fb18, fb19, and fb27 functions. The ABC variant is easy to converge prematurely and fall into local optimal solution, such as the performance of fa3, fa4, fa7, fa12, fb17, and fb27 benchmark functions. However, the performance of QUATRE variants in the functions of fa2, fa3, fa4, fa7, fb18, and fb19 is only slightly insufficient. In summary, the proposed QUATRE-DEG can perform with a very competitive convergence speed compared with the QUATRE variants and ABC variants.

4.4. The Wilcoxon Signed Rank Test. Wilcoxon signed rank test is needed when two results present two different populations [27]. Here, we also give the Wilcoxon signed rank test results. Tables 8 and 9 give the Wilcoxon signed rank test results of CEC2013 test suites and CEC2014 test suites, respectively. The results are calculated by using Origin 2019b statistical analysis packages, and our proposed QUATRE-DEG algorithm is verified to be well performed.

The statistical analysis results of applying Wilcoxon's test between QUATRE-DEG and other comparison algorithms using CEC2013 for all dimensions are summarized in Table 8. As the table states, QUATRE-DEG shows a significant improvement over all comparison algorithms, with a level of significance $\alpha = 0.05$ in 30 and 50 dimensions, over QUATRE and C-QUATRE, with a level of significance $\alpha = 0.05$ in 10 dimensions. On the other hand, in all comparisons, QUATRE-DEG gets higher R^+ values than R^- in all dimensions. Therefore, the performance of QUATRE-DEG algorithm is proved to be better than the other algorithms in this benchmark.

The statistical analysis results of applying Wilcoxon's test between QUATRE-DEG and other comparison algorithms using CEC2014 for all dimensions are summarized in

Table 10: Time complexity comparison on 30D optimization under CEC2013 benchmark 14.

Algorithms	T_0	T_1	$\widehat{T_2}$	$\frac{\widehat{T_2}-T_1}{T_0}$
QUATRE-DEG	0.0906	0.7336	2.7126	21.8433
C-QUATRE	0.0906	0.7336	2.2314	16.5320
QUATRE	0.0906	0.7336	2.0011	13.9901
MABC	0.0906	0.7336	5.7920	55.8322
IABC	0.0906	0.7336	4.7662	44.5099

Table 9. As the table states, QUATRE-DEG shows a significant improvement over all comparison algorithms, with a level of significance $\alpha = 0.05$ in 30 and 50 dimensions, over QUATRE and C-QUATRE, with a level of significance $\alpha = 0.05$ in 10 dimensions. On the other hand, in all comparisons, QUATRE-DEG gets higher R^+ values than R^- in all dimensions. Therefore, the performance of QUATRE-DEG algorithm is proved to be better than the other algorithms in this benchmark.

4.5. Comparison of Complexity. The evaluation method of algorithm complexity adopted by the CEC2013 competition was used to evaluate the proposed QUATRE-DEG algorithm. Let T_0 be the standard computation time for the reference benchmark given by CEC2013 and T_1 be the standard computation time of the specified function fa14 for 200000 evaluations with a dimension of 30D. Let T_2 be the computation time of function fa14 using a specified algorithm for 200000 evaluations with the same dimension of 30D, and $\widehat{T_2}$ is the average of T_2 values acquired by 51 runs. Finally, the algorithm complexity is determined by $(\widehat{T_2} - T_1)/T_0$. Table 10 shows the result of comparing the proposed QUATRE-DEG with the algorithms. It indicates that the QUATRE-DEG can obtain lower complexity compared with either the MABC or IABC. However, the QUATRE-DEG takes more computation time compared with the QUATRE and C-QUATRE. It is reasonable due to the additional process in the proposed algorithm.

5. Conclusion

QUATRE is a relatively new algorithm proposed in recent years. Although it is powerful, it still can be improved in some concerns. For example, in the later stage of evolution, it is easier to fall into local optimal solutions and could not explore more. This study proposes a new QUATRE-DEG algorithm to solve this problem. In the proposed algorithm, excellent individuals in the current iteration process are fully utilized for the optimization. Not only the global optimal solution but also the global suboptimal solution is used to guide the evolution of individuals. According to the distance between the global optimal individuals and the global suboptimal individuals, a new mutation strategy is proposed to expand the current search environment. In addition, the population is divided into two subpopulations, including the subpopulation with the global optimal location as the search center and the subpopulation with the global suboptimal location as the search center. The expansion of search coverage and the increase of population diversity provide better conditions for more exploration, so as to avoid falling into the local optimal solution. Although QUATRE-DEG has stronger search performance and can explore more benchmark environments, it also reduces its search speed and sometimes makes it not easy to completely converge to the global optimal solution, which indicates the possible limitation of the approach. In the experiments, we use CEC2013 and CEC2014 to verify the performance of the proposed QUATRE-DEG algorithm. The results show that the optimization accuracy and convergence speed are excellent, and the time complexity is tolerable. In future works, the proposed algorithm will be applied to several practical engineering problems such as UAV path planning [28, 29]. Moreover, CEC2017 benchmark suite will be used to evaluate the further works.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the Fujian University of Technology under Grant GY-Z20016 and GY-Z18183, National Natural Science Foundation of China under Grant 61773415, and Fujian Provincial Natural Science Foundation under Grant 2017J01730.

References

[1] X. Wang, J.-S. Pan, and S.-C. Chu, "A parallel multi-verse optimizer for application in multilevel image segmentation," *IEEE Access*, vol. 8, pp. 32018–32030, 2020.

[2] N. Liu, J. S. Pan, C. Sun, and S. C. Chu, "An efficient surrogate-assisted quasi-affine transformation evolutionary algorithm for expensive optimization problems," *Knowledge Based Systems*, vol. 209, article 106418, 2020.

[3] S. Zhang and S. Liu, "A discrete improved artificial bee colony algorithm for 0–1 knapsack problem," *IEEE Access*, vol. 7, pp. 104982–104991, 2019.

[4] X. Pan, Q. Zhang, and H. Pan, "Improved artificial bee colony algorithm and its application to fundus retinal blood vessel image binarization," *IEEE Access*, vol. 8, pp. 123726–123734, 2020.

[5] Z.-G. Du, J.-S. Pan, S.-C. Chu, H.-J. Luo, and P. Hu, "Quasi-affine transformation evolutionary algorithm with communication schemes for application of RSSI in wireless sensor networks," *IEEE Access*, vol. 8, pp. 8583–8594, 2020.

[6] N. Liu, J.-S. Pan, J. Wang, and T.-T. Nguyen, "An adaptation multi-group quasi-affine transformation evolutionary algorithm for global optimization and its application in node localization in wireless sensor networks," *Sensors*, vol. 19, no. 19, article 4112, 2019.

[7] J. S. Pan, N. Liu, S. C. Chu, and T. Lai, "An efficient surrogate-assisted hybrid optimization algorithm for expensive optimization problems," *Information Sciences*, vol. 561, pp. 304–325, 2021.

[8] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[9] R. Storn and K. Price, "Differential Evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces," Technical Report TR-95-012, International Computer Science Institute, Vacaville, CA, USA, 1995.

[10] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Neural Networks, 1995. Proceedings., IEEE International Conference on*, vol. 4, pp. 1942–1948, Perth, WA, Australia, 2002.

[11] 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.

[12] M. Seyedali, M. Mohammad, and H. Abdolreza, "Multi-verse optimizer," *Neural Computing and Applications*, vol. 27, pp. 495–513, 2016.

[13] S. Saremi, S. Mirjalili, and A. Lewis, "Grasshopper optimisation algorithm: Theory and application," *Advances in Engineering Software*, vol. 105, pp. 30–47, 2017.

[14] Z. Meng, J.-S. Pan, and H. Xu, "QUasi-Affine TRansformation Evolutionary (QUATRE) algorithm: a cooperative swarm based algorithm for global optimization," *Knowledge Based Systems*, vol. 109, pp. 104–121, 2016.

[15] J. J. Liang, B. Y. Qu, P. N. Suganthan, and A. G. Hernández-Díaz, "Problem definitions and evaluation criteria for the CEC 2013 special session on real-parameter optimization," Technical Report 201212, Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou, China, 2013.

[16] J. J. Liang, B. Y. Qu, and P. N. Suganthan, "Problem definitions and evaluation criteria for the CEC 2014 special session and competition on single objective real-parameter numerical optimization," Technical Report 201311, Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou, China, 2013.

[17] Z. Meng, Y. Chen, X. Li, C. Yang, and Y. Zhong, "Enhancing QUasi-Affine TRansformation Evolution (QUATRE) with adaptation scheme on numerical optimization," *Knowledge Based Systems*, vol. 197, article 105908, 2020.

[18] S.-C. Chu, Y. Chen, F. Meng, C. Yang, J.-S. Pan, and Z. Meng, "Internal search of the evolution matrix in QUasi-Affine TRansformation Evolution (QUATRE) algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 5, pp. 5673–5684, 2020.

[19] B.-Q. Jiang and J.-S. Pan, "A Parallel Quasi-Affine Transformation Evolution algorithm for global optimization," *J. Netw. Intell.*, vol. 4, pp. 30–46, 2019.

[20] Z. Meng and J.-S. Pan, "QUasi-Affine TRansformation Evolution with External ARchive (QUATRE-EAR): an enhanced structure for Differential Evolution," *Knowledge Based Systems*, vol. 155, pp. 35–53, 2018.

[21] B. Zhao, T. W. Sung, and X. Zhang, "A quasi-affine transformation artificial bee colony algorithm for global optimization," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 3, pp. 5527–5544, 2021.

[22] Z. Meng and J.-S. Pan, "A Competitive QUasi-Affine TRansformation Evolutionary (C-QUATRE) Algorithm for global optimization," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 1644–1649, Budapest, Hungary, 2016.

[23] J.-S. Pan, Z. Meng, S.-C. Chu, and J. F. Roddick, "QUATRE Algorithm with Sort Strategy for Global Optimization in Comparison with DE and PSO Variants," in *The Euro-China Conference on Intelligent Data Analysis and Applications*, pp. 314–323, Springer, 2017.

[24] N. Liu, J.-S. Pan, and T.-T. Nguyen, "A bi-population QUasi-Affine TRansformation Evolution algorithm for global optimization and its application to dynamic deployment in wireless sensor networks," *Eurasip Journal on Wireless Communications and Networking*, vol. 2019, no. 1, 2019.

[25] Y. Cao, Y. Lu, X. Pan, and N. Sun, "An improved global best guided artificial bee colony algorithm for continuous optimization problems," *Cluster Computing*, vol. 22, no. S2, pp. 3011–3019, 2019.

[26] W. Gao and S. Liu, "A modified artificial bee colony algorithm," *Computers & Operations Research*, vol. 39, no. 3, pp. 687–697, 2012.

[27] J. Derrac, S. García, D. Molina, and F. Herrera, "A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms," *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 3–18, 2011.

[28] T.-W. Sung, L. Sun, and K.-C. Chang, "Multi-hop Route Planning Based on Environment Information for Path-Following UAVs," in *The International Conference on Artificial Intelligence and Computer Vision*, pp. 831–839, Cairo, Egypt, 2020.

[29] T.-W. Sung, B. Zhao, and K.-C. Chang, "Experimental Comparison of Different Feature Detection Algorithms for UAV Obstacle Avoidance," in *The International Conference on Artificial Intelligence and Computer Vision*, pp. 840–849, Cairo, Egypt, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 8851121, 8 pages
https://doi.org/10.1155/2021/8851121

Research Article

Computational Resource Allocation Strategy in a Public Blockchain Supported by Edge Computing

Sujie Shao, Weichao Gong, Shaoyong Guo, and Xuesong Qiu

Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Weichao Gong; gongweichao@bupt.edu.cn

Received 15 July 2020; Revised 21 March 2021; Accepted 1 April 2021; Published 15 April 2021

Academic Editor: Feng-Jang Hwang

Copyright © 2021 Sujie Shao 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.

Blockchain, as an emerging distributed data management technology, has attracted extensive attention in recent years. In particular, a public blockchain network can ensure data security by addressing computationally intensive cryptographic tasks. Therefore, for node devices, sufficient computing power is required. However, mobile devices with limited computing power do not meet the conditions required by public blockchain network applications (OZEX, CoininAsia, BitRewards, etc.). To copy with the mentioned problems, nodes can offload computing tasks to edge computing services with low latency. This paper mainly focuses on the trade between edge computing providers (ECP) and nodes. We build a computational resource market model based on auction. Meanwhile, we propose two strategies to deal with two methods of offloading to achieve higher system profit. We also prove that the proposed strategy has individual rationality, authenticity under resource constraints. The simulation results have significance for administrators of a public blockchain network to improve the efficiency of computing resource allocation.

1. Introduction

Since blockchain was first proposed in 2008, it has gradually expanded from cryptographic digital currencies to innovative applications in all walks of life [1]. Blockchain implements a decentralized system that can reach a consensus of trust in the digital space. By reaching a consensus among peers in the peer-to-peer (P2P) network, nodes jointly maintain a data chain that connects data blocks in a chain way in a chronological order [2]. Each block on the data chain is protected from tampering and forgery by digital signature and other data verification methods [3].

Proof of work (PoW) is an important consensus mechanism in the public blockchain, which is used to prove that a task is completed and the calculation is paid for [4]. The task is to compute a cryptographic security-related nonce value that leads the hash of the nonce and block data to be less than a given upper limit (depending on the mining difficulty). This calculation process is called mining [5], and the nodes that complete tasks to reach consensus are called miners. Miners broadcast transactions on the network after calculating the nonce value. Other nodes immediately verify the consensus after receiving the block containing transactions. After verification, the block will be added to their local node in the chain of block maintenance [4]. Only data which is packaged by the node that calculates the nonce value fastest can be added to the blockchain. In return, the miners can receive a reward as an incentive of mining. The reward consists of a fixed bonus and a variable transaction fee [6].

Unfortunately, solving the PoW puzzle in public blockchain requires sustained high computing power that mobile devices cannot afford [7]. In order to alleviate the problem of insufficient computing power, consensus nodes can offload their mining tasks to edge computing services, supporting applications based on the public blockchain [8]. Edge computing closer to data source is a comprehensive platform that provides network, computing, and storage. It enables data storage and computing power to be deployed closer to the edge of the user, thereby reducing network latency. Edge computing services can support more consensus nodes to mine tasks, which will significantly improve the robustness and security of the public blockchain network, further attracting more developers to join to promote the construction of public blockchain.

Further, edge computing matches the distributed model of blockchain on networking, computing, and storage. On the one hand, when blockchain nodes are deployed in edge node devices, edge computing services provide computing resources and storage capacity to a large number of scattered nodes in the blockchain. To copy with the problem of delay transmission in the case of coexistence of a large number of nodes, edge computing with low latency is a good response to the application demands of blockchain on the edge side. On the other hand, blockchain provides a trusted and secure environment for edge services, which ensure the integrity and authenticity of storage, achieving the collaborative management of resources among multiple subjects.

In this paper, we divide the nodes into full nodes and light nodes in a public blockchain network. Full nodal participation in the consensus process, i.e., miners, light nodes do not participate in the consensus process, in spite of can be rewarded by performing a portion of the computing tasks for full nodes for collaborative mining [9]. We also call light nodes "sharer." We consider the trading between the edge compute provider (ECP) and the full node as well as the trading between the full node and the light node. From a system perspective, our goal is to maximize the social welfare of the public blockchain networks. Social welfare can be understood as system efficiency in [10]. In order to achieve the goal, we propose an edge computing resource market model based on auction for a public blockchain network. In addition, we designed an individual and truthful computing resource allocation mechanism.

The specific contributions of this paper are summarized as follows:

(i) In the proposed edge computing resource market model, we consider the competitive relationship between miners in the public blockchain and maximize social welfare under appropriate network effect parameters

(ii) Two offloading modes are proposed in a public blockchain network based on edge computing. Miners can offload their mining tasks to ECP or idle nodes, which takes advantage of idle resources in a network and reduces load on an edge server

(iii) Two auction algorithms were used to improve network efficiency. The trade between ECP and miners was conducted through the Vickrey-Clarke-Groves (VCG) auction. The trade between miners and sharers was realized through Continuous Double Auction (CDA). Algorithms are designed to meet resource constraints, individual rationality, and authenticity

2. Related Work

There have been many excellent works studying resource allocation problems among the edge providers and miners in a blockchain network. In [11], authors study the interactions among the cloud/edge providers and miners in blockchain using a multileader multifollower game-theoretic approach, in order to support proof-of-work-based blockchain application. In [12], the authors formulate resource allocation between ECPs and computationally constrained devices as a double auction game. Meanwhile, the experience-weighted attraction (EWA) algorithm is introduced to reach Nash equilibrium. The authors in [13] proposed a novel edge computing offloading framework, which applies decentralized resource management strategy based on auction to perform resource allocation to ECPs. In order to intelligently manage decentralized applications in a blockchain network, Edgence (EDGe + intelligENCE) is proposed by [14] to realize self-governing of the edge devices. In [15], dynamic resource allocation by combinatorial auction is mainly concentrated and the CA-PROVISION algorithm is used for providing pricing strategies. In [16], authors present a two-layer computation offloading paradigm that includes an edge computing provider (ECP) and a cloud computing provider (CCP) to jointly maximize the profits of each service provider and the payoffs of individual miners. In [17], a new model-free deep reinforcement learning-based online computation offloading approach was proposed for blockchain-empowered mobile edge computing in which both mining tasks and data processing tasks are considered. In [18], the authors introduced a broker to manage and adjust the trading market. Then, an iterative double-sided auction scheme was proposed for computing resource trading, where the broker solves an allocation problem to achieve the maximum social welfare meanwhile protecting the privacies of the buyers and the sellers. However, the above papers do not make good use of idle computing resources in the network. In this article, we consider the existence of competition among miners in the computing resource market model and propose two offloading models. Our model supports miners to offload tasks to idle nodes, which improves the utilization of computing resources in the network and reduces the load on edge servers. Further, on the basis of the above-mentioned papers, this paper realizes the optimal resource allocation of the public blockchain network through auction competition between miners and ECP.

3. System Model

3.1. Edge Computing Resources Trade Market Model. As shown in Figure 1, we propose a resources trade market model. There are two types of nodes: full nodes $N = \{1, 2, \cdots, n\}$ and light nodes $M = \{1, 2, \cdots, m\}$ in the public blockchain network with PoW as the consensus strategy. We consider an offload model where the full node could choose to offload the mining calculation task to ECP or to a light node for collaborative mining. In the proposed market model, auction strategy is used to guide trading. The miners submit their resource demands $d = \{d_1, \cdots, d_i, \cdots, d_n\}$ and corresponding bids $b = \{b_1, b_2, \cdots, b_n\}$. After receiving miner's demands and bids, ECP uses different strategies to determine win miners $X = \{x_1, x_2, \cdots, x_n\}$, $x_i \in \{0, 1\}$, and service prices $P = \{p_1, p_2, \cdots, p_n\}$ according to different ways of offloading. In the case of task offloaded to edge sever, ECP uses the VCG auction strategy; in the case of task offloaded to light nodes, ECP uses the CDA strategy. This will be covered in detail in Section 3. To

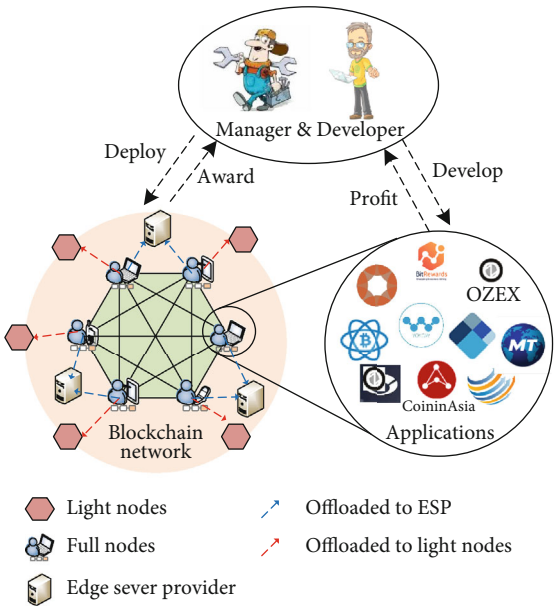


FIGURE 1: The network model of the public blockchain network based on edge computing.

be clear, $x_i = 1$ means that miner i wins the competition and can be allocated to resources at price p_i, while $x_i = 0$ means that no resources are allocated and $p_i = 0$.

3.2. Mining Mechanism Supported by Edge Computing. The efficiency of blockchain is largely determined by network computing power, so how to ensure the availability of computing resources and the participation of devices will be the driving factors. Chang et al. [19] have discussed the incentive mechanism for edge computing-based blockchain in detail, so the focus of this chapter is on the mining mechanism. In order to balance market prices, optimize resource allocation, and minimize marginal loss, auction theory is applied to the computational resource allocation of the edge blockchain.

In the article, we assume that both the miner and ESP are rational and will bid the computational resources according to their actual needs. Device i's hash capability γ_i is represented as

$$\gamma_i(d_i, x_i) = \frac{d_i^{\alpha} x_i}{\sum_{j \in N} d_j^{\alpha} x_j}, \quad i \in N, \tag{1}$$

where $\sum_{i \in N} \gamma_i = 1$, d_i^{α} is the computational resource unit required by device i, and α is an exponential parameter of hash power function [20].

In mining competition, miners need to be the first to find the correct nonce to solve the PoW problem and broadcast the block to reach consensus. The process of generating the new block can be described as a Poisson process with intensity $1/\lambda$ [21]. Firstly, miners collect unconfirmed transactions represent by $S = \{s_1, s_2, \cdots, s_n\}$ into their blocks. Secondly, miner i broadcast its block to a blockchain network to reach consensus. The broadcast delay is affected by the size of transactions s_i. Miner i's expected reward R_i consists of a fixed bonus T and a variable transaction fee affected by s_i. R_i can be expressed as equation (2), where $P_i(\gamma_i, s_i)$ is the probability that miner i reaches consensus and gets rewards.

$$R_i = (T + rs_i) P_i(\gamma_i, s_i). \tag{2}$$

Through the above mining competition, miners will only be rewarded if they succeed in creating blocks and reaching consensus. However, in the block broadcast process, the block that encapsulates more transactions has a higher latency. So, it may not reach a consensus first because of the high broadcast latency. It has a higher latency that the block encapsulates more transactions. This type of block is called an orphaned block. The miner i's block broadcast latency τ_i is linear to the size of the encapsulated transactions, i.e., $\tau_i = \xi s_i$, (ξ is a constant) [22]. As mentioned earlier, the generation process of a block is a Poisson process with intensity $1/\lambda$, so miner i's orphaning probability can be described as $P_i^o = 1 - e^{-(1/\lambda)\tau_i}$. Furthermore, the probability that miners will be rewarded P_i can be calculated as follows:

$$P_i = P_i^m (1 - P_i^o) = \gamma_i e^{-(1/\lambda)\xi\tau_i}. \tag{3}$$

4. Resource Allocation Based on Auction

In this part, we mainly introduce the auction strategy that the system executes when the computing task is offloaded to ECP or light nodes. Regarding how to decide which task is offloaded to ECP and which task is offloaded to light nodes, we have discussed in more detail in the previous study [23]. To rephrase briefly, Lyapunov theory and proposed deviation update decision algorithm (DUDA) are used to solve computation offloading decision-making and offloading update order-making.

4.1. Task Are Offloaded to ECP. When the miners' tasks were offloaded to ECP, the VCG auction strategy was executed to determine winners and trade prices. Assume that c_E is the unit cost of computing services provided by ECP. Moreover, in this offloading scenario, the social welfare of a network is the sum of all the miners' rewards minus all the ECP's costs. Therefore, maximizing social welfare can be described as the following optimization problem:

$$\begin{aligned} &\text{Max} \sum_{i \in N} \left\{ (T + rs_i) \frac{d_i^{\alpha} x_i}{\sum_{j \in N} d_j^{\alpha} x_j} e^{-1/\lambda \xi \tau_i} - c_E d_i \right\} i \in N \\ &\text{s.t.} \sum_{i \in N} d_i x_i \le C, \quad x_i \in \{0, 1\}, \ \forall i \in N, \end{aligned} \tag{4}$$

where $\sum_{i \in N} d_i x_i \le C$ is the constraint that needs to be met, and C is the most computing resource unit that ECP can afford.

VCG auction is a method of distribution of goods with the optimization goal of maximizing social welfare. It uses the marginal damage caused by bidders to others as the bidding pricing rule. Therefore, VCG ensures that the best strategy for each bidder is to bid according to actual needs and value.

Input: $d = \{d_1, \cdots, d_i, \cdots, d_n\}$; $b = \{b_1, b_2, \cdots, b_n\}$
Output: $P = \{p_1, p_2, \cdots, p_n\}$;$X = \{x_1, x_2, \cdots, x_n\}$

For each $i \in N$
$x_i = 0$;$p_i = 0$;
End for
$W = \varnothing$;$W_t = \varnothing$;$S = 0$;$S_t = 0$;
While $S \le S' \& W \ne N \& |W| \le C$;
$j = \arg\ \max_{j \in N \setminus W}(b_j/d_j)$;
$W = W_t$;$W_t = W \cup \{j\}$;$S = S_t$;
$s_t = \sum_{i \in N}(T + rs_i)(d_i^{\alpha} x_i / \sum_{j \in N} d_j^{\alpha} x_j) e^{-(1/\lambda)\xi\tau_i} - \sum_{i \in N} c_E d_i$
End while
For each $j \in W$;
$x_i = 1$;$N_{-j} = N \setminus \{j\}$;$W_{-j} = W \setminus \{j\}$;
$W' = \varnothing$;$W'_t = \varnothing$;$S' = 0$;$S'_t = 0$;
While $S' \le S'_t \& W'_t \ne N_{-j} \& |W'| \le C$;
$k = \arg\ \max_{k \in N_{-j} \setminus W'}(b_k/d_k)$;
$W' = W'_t$;$W'_t = W' \cup \{k\}$;$S' = S'_t$;
$s'_t = \sum_{i \in w'_t}(T + rs_i)(d_i^{\alpha} x_i / \sum_{j \in N} d_j^{\alpha} x_j) e^{-(1/\lambda)\xi\tau_i} - \sum_{i \in w'_t} c_E d_i$
End while
$p_i = \sum_{j \in W'_t} S_t - \sum_{j \in W_t \setminus i} S'_t$
End for

Algorithm 1: VCG auction algorithm.

First, consider ECP auctioning a set of computing resources $d = \{d_1, \cdots, d_i, \cdots, d_n\}$. Buyers can declare a set of bids $b = \{b_1, b_2, \cdots, b_n\}$, and bidders are sealed (only visible to the auction system), so they cannot see other people's bids at any time. After all bids are completed, the auction will close.

Then, the auction system considers all possible bid combinations and reserves the bid combination that maximizes the total amount of bids. It should be noted that the price they pay $P = \{p_1, p_2, \cdots, p_n\}$ is not the price they originally bid, but the marginal damage $p_i = \sum_{j \in W'_t} S_t - \sum_{j \in W_t \setminus i} S'_t$ they cause to other bidders (at most as high as the original bid).

At the end of the auction, if the agent is completely rational and there is no collusion, we can assume that the willingness to pay has been truthfully reported, because each participant will only be accused of marginal damage to other bidders.

We solved this integer programming problem through a VCG auction using the following algorithm: first, determine the list of winning miners and then determine the final trade price based on marginal damage [24].

4.2. Task Are Offloaded to Light Nodes. In our scenario, there are some light nodes (sharers) $M = \{1, 2, \cdots, m\}$ in the blockchain network that have computing resources $cr = \{cr_1, cr_2, \cdots, cr_m\}$ but do not participate in mining; they only help the full nodes (miners) $N = \{1, 2, \cdots, n\}$ perform computing tasks. When the miners' tasks were offloaded to light nodes, the Continuous Double Auction (CDA) strategy was executed to determine winners and trade price [25].

Input: $\{a_1, \cdots a_j, \cdots a_m\}$;$\{b_1, \cdots b_i, \cdots b_n\}$;$\{d_1, \cdots d_i, \cdots d_n\}$;$\{cr_1, \cdots cr_j, \cdots cr_m\}$
Output: matching relationship $i \longleftrightarrow j$; trading price p_i
1. Sort bids of buyers a_j with decreasing value.
2. Sort bids of sellers b_i with increasing value.
3. For each $i \in N$;
4. If $\exists i : a_j \le b_i$ then $p_i = b_i + a_j/2, i \in N, j \in M$;
5. Output $j \longleftrightarrow i$, p_i;
6. End for

Algorithm 2: CDA Algorithm.

In the CDA strategy, the miner's (buyer) resource demand is $d = \{d_1, \cdots, d_i, \cdots, d_n\}$, and corresponding bids are $b = \{b_1, b_2, \cdots, b_n\}$, assuming the trade price for the buyer is $P = \{p_1, p_2, \cdots, p_n\}$. The sharers (sellers) will also bid for the computing resource as$a = \{a_1, a_2, \cdots, a_m\}$. Let us define the seller's trade price as $Q = \{q_1, q_2, \cdots, q_m\}$. ECP acts as an auctioneer, and his income φ is that the seller's trade price subtracted from the buyer's.

$$\varphi = \sum_{i \in N} p_i - \sum_{j \in M} q_j. \tag{5}$$

Then, we discuss the validity of the CDA algorithm. Similar to Zhang et al. [26], we try to prove the validity from the following:

(1) Individual rationality: both sellers and buyers need to make a profit

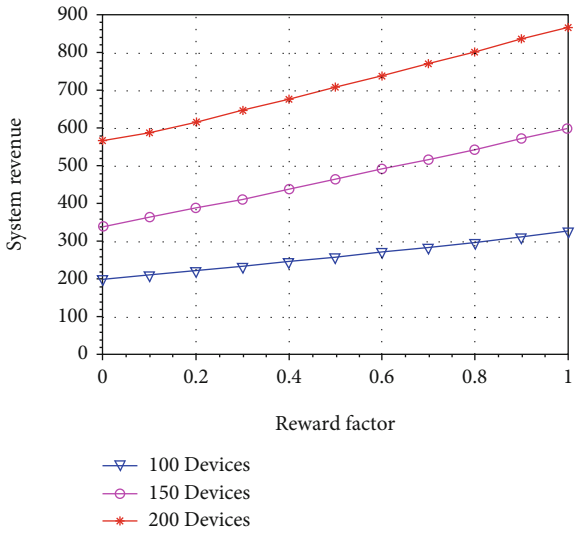


Figure 2: The influence of reward factors on system profits under different equipment quantity.

$$\begin{cases} q_j \geq a_j, & j \in M, \\ p_i \leq b_i, & i \in N \end{cases} \tag{6}$$

(2) Resource constraints: if both seller j and buyer i agree to trade, equations (7) and (8) need to be met

$$d_i \leq cr_j, \quad i \in N, j \in M, \tag{7}$$

$$\varphi = \left(\sum_{i \in N} p_i - \sum_{j \in M} q_j \right) \geq 0 \tag{8}$$

(3) Truthful: the seller j's profit is defined as χ_j^M, and the buyer i's profit is defined as χ_i^N. In the CDA process, the highest profits occur when participants compete based on the true price of the resource. In other words, the constraints of equations (11) and (12) have to be satisfied:

$$\chi_j^M = \begin{cases} q_j - a_j, & j \text{ succeeded in the CDA auction,} \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

$$\chi_i^N = \begin{cases} a_i - p_i, & i \text{ succeeded in the CDA auction,} \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

$$\chi_j^M (a_j, M_{-j}) \geq \chi_j^M \left(a_j', M_{-j} \right), \quad \forall a_j' \neq a_j, j \in M, \tag{11}$$

$$\chi_i^N (b_i, N_{-i}) \geq \chi_i^N \left(b_i', N_{-i} \right), \quad \forall b_i' \neq b_i, i \in N \tag{12}$$

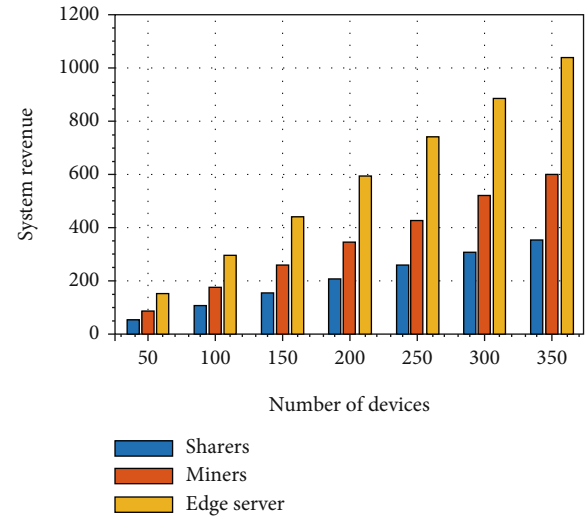


Figure 3: Comparison of profits for different device roles.

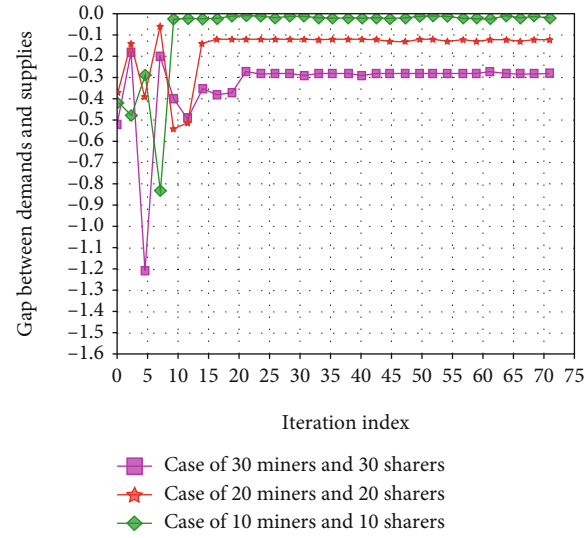


Figure 4: Gap between demands and supplies in different cases.

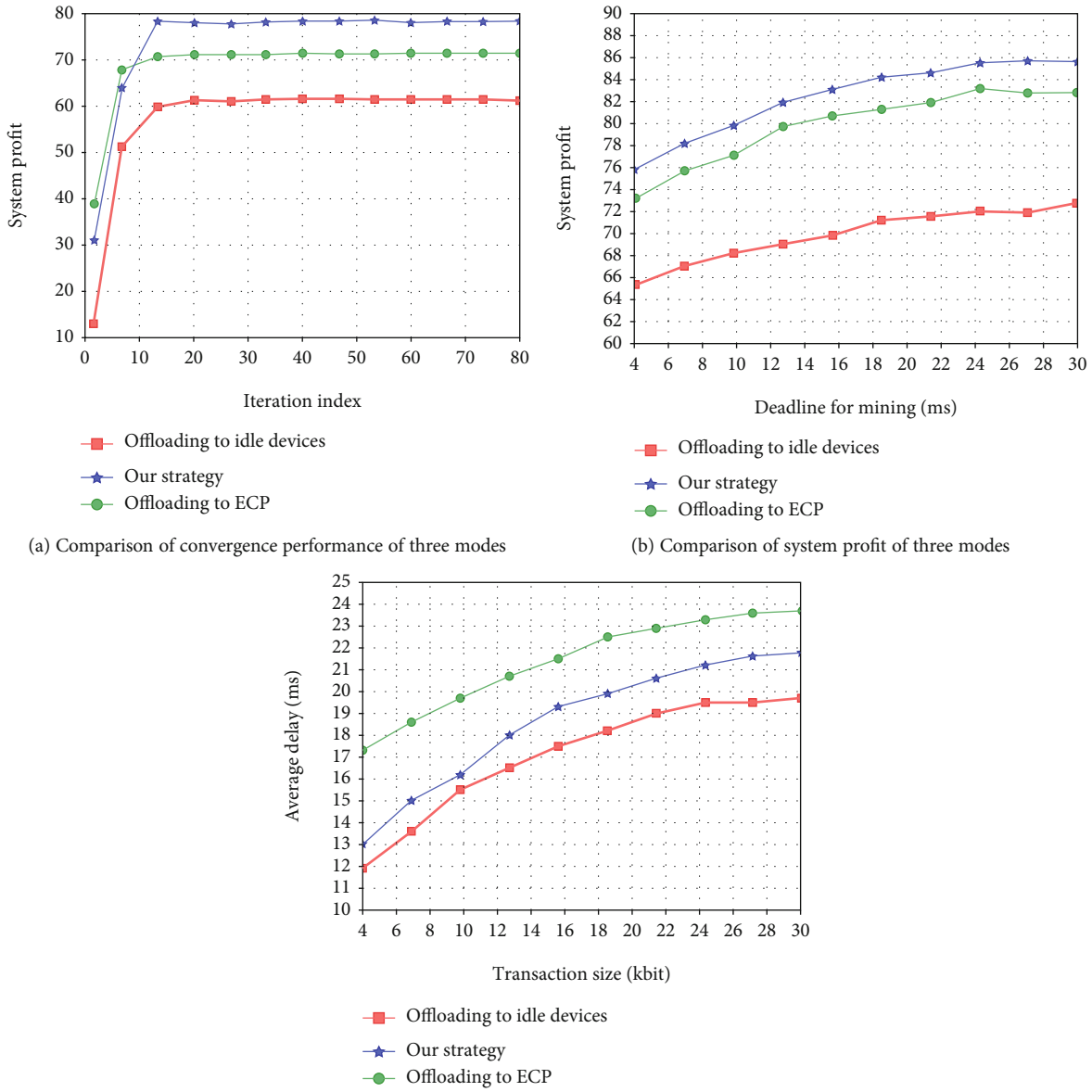
In the process of CDA in this paper, we set $\varphi = 0$ and used equation (13) to determine the trade price, with $\delta = 0.5$.

$$p_i = q_j = \delta b_i + (1 - \delta) a_j, \quad i \in N, j \in M. \tag{13}$$

5. Experiment Results and Analysis

In this chapter, the auction strategy we proposed is simulated to verify the effectiveness and rationality. Simulation result proves that the strategy proposed has better system profits.

According to Figure 2, the value of reward factor also affects the system's profits. It can be clearly seen that, with increasing of the reward factor in the number of different devices, the system's profits have increased significantly. This is because the grown of reward factor incentivizes devices to



(a) Comparison of convergence performance of three modes

(b) Comparison of system profit of three modes

(c) Comparison of average delay of three modes

Figure 5: Comparison of performance of the three modes.

rent computing resources from edge servers at a higher cost, leading to better expected profits.

The profits of different network roles are shown in Figure 3. Under the same circumstances, device-sharing computing resources have lower profits. Devices acting as a miner can get higher profits. Meanwhile, edge server provider can also get higher profits. It should be noted that as the number of devices increases, the profits to all roles in the system increased.

In Figure 4, we analyze the gap between demands and supplies in different cases. It is obvious that as the number of iterations increases, the gap gradually approaches zero. This shows that the strategy we proposed can meet the demand dynamically.

We compared the proposed strategy with mode A (offloading all tasks to ECP) and mode B (offloading all tasks to adjacent idle devices) in Figure 5. In (a), the convergence performance of the three models is compared. The system profit increases sharply in the first 20 iterations and then reaches a stable state, indicating that our mechanism has a faster convergence than the other two modes. This is because our model can offload mining tasks to idle devices, improving the utilization of computing resources. Moreover, because there are offloading decisions to be made, the plateau is reached relatively late. In (b), we compare the system profit of the three modes. With the increase of time, the system profit of all three modes will increase, but our strategy has higher returns. Offloading all mining tasks to idle devices causes longer delays, so mode B (offloading all tasks to adjacent idle devices) has the lowest return. In (c), we compare the average delay of the three modes. Since the miners who choose mode

B are closer to the provider, the average delay is lower. Taken as a whole, our strategy performs better.

6. Conclusions

In this paper, we offload computing tasks to ECP or light nodes to alleviate the PoW problem in a public blockchain network. When miners offload tasks to the ECP, the VCG auction strategy is adopted to maximize the social welfare of resource allocation; when tasks are offloaded to light nodes, the CDA strategy is adopted to realize efficient allocation of computing resources. The authenticity, individual rationality, and effectiveness of the strategy are also proved in this paper. The simulation results prove that the proposed strategy can improve the profits of the system, as well as providing reference significance for the administrators of the public blockchain network.

Data Availability

Data are available upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work is supported by the National Key R&D Program of China (2018YFB1402704).

References

[1] G. B. Mermer, E. Zeydan, and S. S. Arslan, "An overview of blockchain technologies: principles, opportunities and challenges," in *2018 26th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, Izmir, 2018.

[2] Z. Guo, Z. Gao, H. Mei, M. Zhao, and J. Yang, "Design and optimization for storage mechanism of the public blockchain based on redundant residual number system," *IEEE Access*, vol. 7, pp. 98546–98554, 2019.

[3] D. Ding, K. Li, L. Jia, Z. Li, J. Li, and Y. Sun, "Privacy protection for blockchains with account and multi-asset model," *China Communications*, vol. 16, no. 6, pp. 69–79, 2019.

[4] T. P. Keenan, "Alice in blockchains: surprising security pitfalls in PoW and PoS blockchain systems," in *2017 15th Annual Conference on Privacy, Security and Trust (PST)*, pp. 400–4002, Calgary, AB, 2017.

[5] A. Kiayias, E. Koutsoupias, M. Kyropoulou, and Y. Tselekounis, "Blockchain mining games," in *Proceedings of the 2016 ACM Conference on Economics and Computation*, pp. 365–382, 2016.

[6] Z. Yu, X. Liu, and G. Wang, "A survey of consensus and incentive mechanism in blockchain derived from P2P," in *2018 IEEE 24th international conference on parallel and distributed systems (ICPADS)*, pp. 1010–1015, Singapore, Singapore, 2018.

[7] K. Suankaewmanee, D. T. Hoang, D. Niyato, S. Sawadsitang, P. Wang, and Z. Han, "Performance analysis and application of mobile blockchain," in *2018 International Conference on Computing, Networking and Communications (ICNC)*, pp. 642–646, 2018.

[8] Z. Xiong, Y. Zhang, D. Niyato, P. Wang, and Z. Han, "When mobile blockchain meets edge computing," *IEEE Communications Magazine*, vol. 56, no. 8, pp. 33–39, 2018.

[9] A. Palai, M. Vora, and A. Shah, "Empowering light nodes in blockchains with block summarization," in *2018 9th IFIP international conference on new technologies, mobility and security (NTMS)*, pp. 1–5, Paris, 2018.

[10] X. Zhang, Z. Huang, C. Wu, Z. Li, and F. C. M. Lau, "Online auctions in IaaS clouds: welfare and profit maximization with server costs," *IEEE/ACM Transactions on Networking*, vol. 25, no. 2, pp. 1034–1047, 2017.

[11] Z. Xiong, J. Kang, D. Niyato, P. Wang, and V. Poor, "Cloud/edge computing service management in blockchain networks: multi-leader multi-follower game-based ADMM for pricing," *IEEE Transactions on Services Computing*, vol. 13, no. 2, pp. 356–367, 2020.

[12] Q. Li, H. Yao, T. Mai, C. Jiang, and Y. Zhang, "Reinforcement-learning- and belief-learning-based double auction mechanism for edge computing resource allocation," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 5976–5985, 2020.

[13] C. Avasalcai, C. Tsigkanos, and S. Dustdar, "Decentralized resource auctioning for latency-sensitive edge computing," in *2019 IEEE international conference on EDGE computing (EDGE)*, pp. 72–76, Milan, Italy, 2019.

[14] J. Xu, S. Wang, A. Zhou, and F. Yang, "Edgence: a blockchain-enabled edge-computing platform for intelligent IoT-based dApps," *China Communications*, vol. 17, no. 4, pp. 78–87, 2020.

[15] E. I. Nehru, J. I. S. Shyni, and R. Balakrishnan, "Auction based dynamic resource allocation in cloud," in *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, pp. 1–4, Nagercoil, 2016.

[16] S. Jiang, X. Li, and J. Wu, "Hierarchical edge-cloud computing for mobile blockchain mining game," in *2019 IEEE 39th international conference on distributed computing systems (ICDCS)*, pp. 1327–1336, Dallas, TX, USA, 2019.

[17] X. Qiu, L. Liu, W. Chen, Z. Hong, and Z. Zheng, "Online deep reinforcement learning for computation offloading in blockchain-empowered mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8050–8062, 2019.

[18] Z. Li, Z. Yang, and S. Xie, "Computing resource trading for edge-cloud-assisted internet of things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3661–3669, 2019.

[19] Z. Chang, W. Guo, X. Guo, Z. Zhou, and T. Ristaniemi, "Incentive mechanism for edge-computing-based blockchain," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 11, pp. 7105–7114, 2020.

[20] Y. Jiao, P. Wang, D. Niyato, and Z. Xiong, "Social welfare maximization auction in edge computing resource allocation for mobile blockchain," in *2018 IEEE international conference on communications (ICC)*, pp. 1–6, Kansas City, MO, 2018.

[21] D. Kraft, "Difficulty control for blockchain-based consensus systems," *Peer-to-Peer Networking and Applications*, vol. 9, no. 2, pp. 397–413, 2016.

[22] A. Narayanan, J. Bonneau, E. Felten, A. Miller, and S. Goldfeder, *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*, Princeton Univ. Press, Princeton, NJ, USA, 2016.

[23] L. Ruan, Z. Liu, X. Qiu, Z. Wang, S. Guo, and F. Qi, "Resource allocation and distributed uplink offloading mechanism in fog environment," *Journal of Communications and Networks*, vol. 20, no. 3, pp. 247–256, 2018.

[24] S. Yang and B. Hajek, "VCG-Kelly mechanisms for allocation of divisible goods: adapting VCG mechanisms to one-dimensional signals," in *2006 40th annual conference on information sciences and systems*, pp. 1391–1396, Princeton, NJ, 2006.

[25] J. Wang, Q. Wang, and N. Zhou, "A decentralized electricity transaction mode of microgrid based on blockchain and continuous double auction," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, Portland, OR, 2018.

[26] D. Zhang, Z. Chang, T. Hämäläinen, and F. R. Yu, "Double auction based multi-flow transmission in software-defined and virtualized wireless networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8390–8404, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6610273, 11 pages
https://doi.org/10.1155/2021/6610273

Research Article

Appling an Improved Method Based on ARIMA Model to Predict the Short-Term Electricity Consumption Transmitted by the Internet of Things (IoT)

Ni Guo[1], **Wei Chen**[1,2], **Manli Wang**[3], **Zijian Tian**[1] and **Haoyue Jin**[1]

[1]*School of Mechanical Electronic and Information Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China*
[2]*School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China*
[3]*School of Physics &Electronic Information Engineering, HeNan Polytechnic University, Jiaozuo 454000, China*

Correspondence should be addressed to Wei Chen; chenwdavior@163.com

Received 24 December 2020; Revised 1 March 2021; Accepted 29 March 2021; Published 12 April 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Ni Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rapid development of the Internet of Things (IoT) has brought a data explosion and a new set of challenges. It has been an emergency to construct a more robust and precise model to predict the electricity consumption data collected from the Internet of Things (IoT). Accurately forecasting the electricity consumption is a crucial technology for the planning of the energy resource which could lead to remarkable conservation of the building electricity consumption. This paper is focused on the electricity consumption forecasting of an office building with a small-scale dataset, and 117 daily electricity consumption of the building are involved in the dataset, among which 89 values are selected as the training dataset and the remaining 28 values as the testing dataset. The hybrid model ARIMA (autoregression integrated moving average)-SVR (support vector regression) is proposed to predict the electricity consumption with different prediction horizons ranging from 1 day to 28 days. The model performances are assessed by three evaluation indicators, respectively, are the mean squared error (MSE), the root mean square error (RMSE), and the mean absolute percentage error (MAPE). The proposed model ARIMA-SVR is compared with the other four models, respectively, are the ARIMA, ARIMA-GBR (gradient boosting regression), LSTM (long short-term memory), and GRU (gated recurrent unit) models. The experiment result shows that the ARIMA-SVR model has lower prediction errors when the prediction horizon is within 20 days, and the ARIMA model is better when the prediction horizon is in the interval of 20 to 28 days. The provided method ARIMA-SVR has higher flexibility, and it is a great choice for electricity consumption prediction with more accurate results.

1. Introduction

Nowadays, with the continuous increase of the electricity consumption in buildings, the problem of excessive waste of resources has occurred. Internet of things (IoT) has been popularly applied to smart city controls for collecting electricity consumption; specifically, the real-time electricity consumption data is transmitted to the electricity consumption monitoring system by the distributed wireless sensor network (WSN). The wireless sensor is mainly composed of many intelligent distributed wireless sensor nodes, and each of which has the function of sending a message. Forecasting electricity demand in advance and scheduling accordingly is an essential measure to achieve energy conservation, and it could also assist policymakers and energy managers to make reasonable strategies to promote environmental protection and reduce carbon emission.

Electricity consumption is influenced by many factors such as weather conditions, occupant behavior, and the physical parameters of buildings; this could be verified by many relevant publications. Besides, these impact factors are also considered in the experimental analysis as input vectors by researchers. Yannan et al. proposed a framework for data-driven occupant-behavior analytics in Ref. [1] which will

help build an analytics feedback loop from behavior impact to incentive design for energy saving: the approach is designed based on machine learning techniques such as k-means and kernel ridge regression. Climate factors as a type of important factor are analyzed frequently by many researchers in the forecasting field. Caro et al. have studied the temperature's influence and the effect of public holidays on short-term electric load forecasting for Spanish insular electric systems in Ref. [2]; a mathematical 24 h Reg-ARIMA model as a proposed algorithm is utilized to predict the electric load demand of the next day for ten insular systems located in the two Spanish archipelagos. Although it could improve the precision of the electricity forecast when the impact factors are involved into the model, there are still many limitations when collecting data on influencing factors. Therefore, it is an efficient measure to address the challenge of electricity consumption forecasting by extracting data features from the historical time series.

For univariate time series forecasting, the typical ARIMA algorithm is commonly utilized in many works as a result of requiring very few assumptions in model training and the flexibility of application [3–5]. As mentioned by Jamil [6], the hydroelectricity consumption of Pakistan is predicted up to the year 2030 by using the ARIMA model, and exhaustive statistical analysis and validations have been performed in this research; moreover, a sensitivity analysis is also conducted to study the relation of hydroelectricity consumption to the annual population and GDP growth rate of the country. Li et al. have applied four models to forecast the carbon emission intensity in 2030; the experiment result shows that the ARIMA model is the best-fitted model compared with the other three models [7].

Following many studies in the forecasting field, it can be indicated that satisfactory results could not be attained just with the individual models for all situations. For instance, the characteristics of the time series data may not be captured adequately by individual ARIMA model due to the fact that both the linear and nonlinear features existed in historical data, while the ARIMA model just experts in learning the linear trend of the sample data. Artificial neural networks (ANN) perform well only with sufficient information relying on a large number of historical data. Abundant studies have indicated that the combined methodologies are an advantage of solving the complicated problems concerning time series forecasting owing to the method could benefit from each composition algorithm. In general, the hybrid model has always contained a linear model and a nonlinear model component, and the typical ARIMA is frequently used as a linear model component owing to its advantage of capturing the line characters existing in the data. With regard to the nonlinear model component in the hybrid model, there are many choices available such as machine learning algorithms and statistical methodologies. Machine learning methods that support vector regression (SVR) and gradient boosting regression (GBR) are selected to combine the ARIMA model due to these models having shown great potential in dealing with nonlinear patterns using a small dataset.

The main contributions of this study are demonstrated as follows:

(1) The proposed model ARIMA-SVR can be developed using a small training set while maintaining high accuracy, and few studies are found in electricity consumption prediction using the ARIMA-SVR hybrid model

(2) The proposed model does not need any additional variables just based on a value of its historic observation, and the model is very flexible and explanatory; very few parameters need to be tuned and the model is easy to be implemented

(3) The proposed model combines the advantages of the ARIMA and SVR models, and the nonlinear and linear characters could be well extracted by this hybrid model

This paper is organized as follows: Section 2 has exhibited a comprehensive overview of some notable findings related to this work recently, and the literature reviews are illustrated in different aspects. Section 3 is devoted to introducing the related methodologies which appeared in this paper. Section 4 has presented the procedure of data preprocessing and the construction procedure of prediction models. Section 5 has assessed the simulation results utilizing three evaluation indicators and discussed the performances of prediction models. Finally, the conclusion of the paper is provided in the last section.

2. Literature Review

The improvement of forecasting technology has become a focus in researches towards obtaining a higher accuracy of electricity consumption prediction. This section has presented a comprehensive overview of some superior combined methodologies, and the summary is demonstrated from different aspects, such as the perspective of long term and short term. For instance, many authors have demonstrated the prediction situations using specific cases in different forecast horizons. Kaytez has come up with a hybrid model based on the ARIMA model and LSSVR (least-square support vector machine) and applied it to conduct long-term forecasting of net electricity consumption for Turkey until 2022; the final results demonstrate that the hybrid model ARIMA-LSSVM can generate more realistic and reliable forecasts [8]. With the aim to implement the 24 h-ahead forecasting of the district heat demand of buildings, Eseye and Lehtonen has tested several ML (machine learning) approaches and the excellent result is obtained by the integrated model, namely, EMD-ICA-SVM; it has achieved outperformed forecasting accuracy enhancement compared to the other nine evaluated models [9]. The paper [10] has introduced two novels deep supervised machine learning models including RFEM-GKR (Gaussian Kernel regression model with random feature expansion) model and NPK-NNM (nonparametric based k-NN) model for large-scale utilities and buildings' short- and medium-term load requirement forecasts; the hybrid method RFEM-GKR has remarkable predictor improvements and is proved superior with its high accuracy and stability; the proposed model can be taken as a successful tool to predict energy consumption. The study [11] is focused on the

ultrashort-term (15-minute) predictions of residential electricity of consumption by developing a hybrid model which is based on the Holt-Winters (HW) method and extreme learning machine (ELM) network; the single-model Holt-Winters (HW), extreme learning machine (ELM) network, and long short-term memory network were also established in this research; the experiment result has shown that the proposed HW-ELM model offers more outstanding performance compared with other relevant models.

In addition, some other up-to-date publications about the combined models are also exhibited here. Tascikaraoglu and Uzunoglu have contributed a comprehensive review about wind power forecasting which has outlined various combined forecasting approaches and an up-to-date annotated bibliography of the wind forecasting literature [12]. In Ref. [13], the authors propose a hybrid method based on the combination of autoregressive integrated moving average (ARIMA), artificial neural network (ANN), and the proposed support vector regression (SVR) technique to forecast the yearly peak load and total energy demand of Iran National Electric Energy System; the parameters of the SVR technique are optimized using a particle swarm optimization (PSO) method. Nepal et al. have combined the clustering and the ARIMA model toward electricity load forecasting,; the result has proved that the proposed approach has provided improved accuracy as well as superior performances than that using the ARIMA model alone [14]. The combined method which consists of the ARIMA and NGM methods, namely, the NGM-ARIMA model has been put forward by Ma et al. aimed at accurately predicting South Africa's energy consumption in 2017-2030 [15]; the highest prediction accuracy was achieved by the NGM-ARIMA model, and the prediction result is more close to the actual energy consumption compared to the single ARIMA and NGM model. Gulay and Duru have combined three different models: ARDL (autoregressive distributed lag model), EMD (empirical mode decomposition), and ANN (artificial neural network) for the predictive analytics of energy systems and prices; the proposed hybrid forecasting algorithms provided better results by improving the forecasting accuracy [16]. The publication [17] has displayed a novel hybrid model ANFIS which consolidates both ANN and fuzzy frameworks for prediction future power utilization; the result has proved that this hybridizing approach has the potential of improving prediction performance since it has more significant accuracy and leads to smaller errors contrasted with other models. Likewise, the advantages of the hybrid approach were also verified by many studies [18–20]。.

Different hybrid methodologies are used in different kinds of literature, and the desired results are obtained in various forecasting fields. Generally speaking, it can be concluded that regarding methodologies, the hybrid models or the combined models are composed of the linear and nonlinear models, and each of which carries out a part in the process of the prediction. In addition, after reviewing many related publications in this field, it can be indicated that the hybrid model has a significant advantage of capturing the hidden linear and nonlinear components which are embedded in the original dataset.

3. Methodology

3.1. Autoregressive Integrated Moving Average (ARIMA) Algorithm. The typical ARIMA (autoregressive integrated moving average) algorithm has been proved to be an efficient and reliable method for dealing with the univariable time series. The emphasized advantage is that the ARIMA algorithm does not need any additional variables just based on the values of its historic observations. And the required conditions previous to conduct the ARIMA model process should be satisfied with two conditions; one is that it should be a stationary time series, and the other is the recommended minimum amount of the sample data is at least 50 [21].

Actually, the ARIMA algorithm was integrated by autoregression (AR) and moving average (MA) method with an addition of integrative module; it is characterized by three terms, respectively, p, d, and q, and the general format of the model is ARIMA(p, d, q). Here, p is the order of the AR term, q is the order of the MA term, and d is the number of differencing required for obtaining a stationary time series.

The forecasting equation of the ARIMA (p, d, q) can be expressed as follows:

$$y_t = c + \sum_{i=1}^{p} \varnothing_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t. \tag{1}$$

In the above equation, c is the constant representing the intercept, $\varnothing_i$ and y_{t-i}, respectively, are the parameters and regressors for the AR part of the model, while θ_j and ε_{t-j}, respectively, represent the parameters and regressors of the MA part of the model, whereas ε_t is the white noise error term of the model.

3.2. Support Vector Regression (SVR) Algorithm. SVR (support vector regression) is a good choice to characterize the nonlinear statistical features which existed in the small-scale dataset. This algorithm was firstly proposed by Vapnik et al. in literature [22] and was frequently applied by many researchers in recent years [23, 24]. The fundamental principle of the model is mapping the input data into a high-dimensional space to explore the nonlinear relationship between the input data and output variables; the input dataset is assumed as $\{((x_1, y_1) \cdots (x_n, y_n))\}$, and the optimization is described by the following formula:

$$\min \frac{1}{2} w^T w + C \frac{1}{n} \sum_{i=1}^{n} \left(\xi_i + \xi_i^*\right), \tag{2}$$

$$w^T \varnothing(x_i) + b - y_i \le \varepsilon + \xi_i, \tag{3}$$

$$y_i - w^T \varnothing(x_i) - b \le \varepsilon + \xi_i^*, \quad \xi_i, \xi_i^* \ge 0, i = 1, \cdots.., n, \tag{4}$$

where w, b, ξ_i, and ξ_i^* are the decision variable parameters of the optimization problem; $w^T w$ is a regularized term, and ξ_i and ξ_i^* are the slack variables; C is the penalty parameter, ε is the insensitive loss coefficient, it represents a ε tube, if the predicted value is within the tube, the loss is zero, while if it

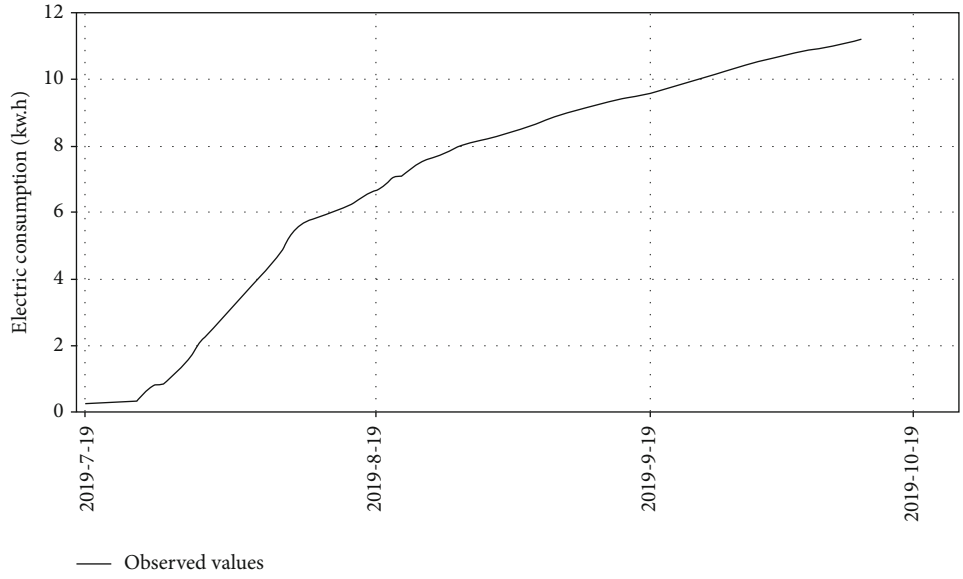


Figure 1: The curve of the original time series.

is outside the tube, the loss is the magnitude of the difference between the predicted value and the radius ε of the tube. What is more, the term $\varnothing(x_i)$ is the feature map and the radial basis function(RBF) is used as the kernel function in this paper; the expression was written as $K(x_i, x_j) = \varnothing(x_i)^T \varnothing(x_i)$.

3.3. Gradient Boosting Regression (GBR) Algorithm. GBR (gradient boosting regression) algorithm is proven as an efficient forecasting technique to capture the nonlinear relationship between the input and output datasets in previous studies. It is one of the boosting regression which has combined a bunch of weak learners [25], and the weak learner would be increasing to minimize the forecasting errors by iteratively training. The main mathematical principles of the gradient boosting regression (GBR) are detailed in the literature [26].

4. Predict Models

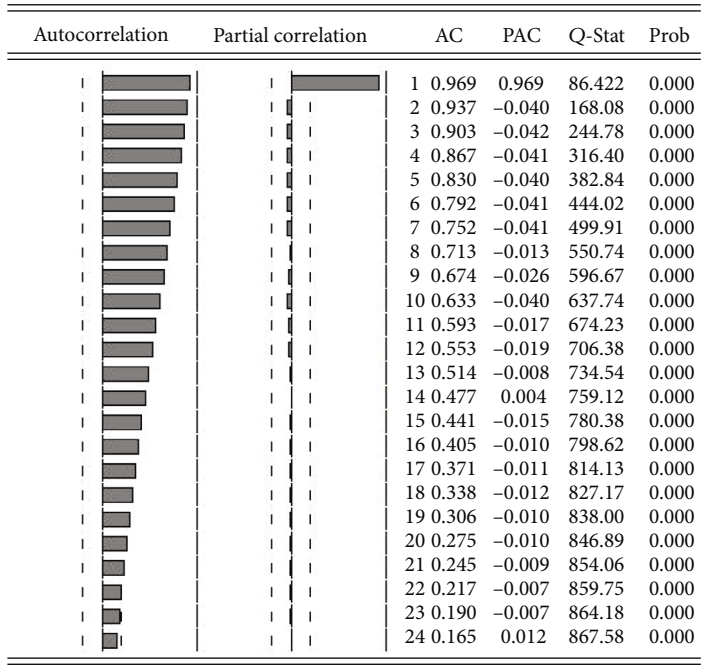
4.1. Data Preprocessing. In many studies, it can be observed that better results are usually obtained when with data preprocessing, and this procedure could make a considerable contribution to the prediction performance of the model. There are two main kinds of processing for the original dataset as the following illustration.

4.1.1. Fill the Missing Data and Normalization. The data is collected from the energy consumption monitoring system by the Internet of Things (IoT), and the monitoring electricity consumption belongs to an office building which is located in Changyang Peninsula in Fangshan District of Beijing, China. The dataset involved 117 daily electricity consumption of the building, and among which, 89 values are selected as the training dataset and the remaining 28 values as the testing dataset. The collected dataset was a series of successive and univariate time series over the period of 19 July 2019 to 12 November 2019. The training dataset was ranged from 19 July 2019 to 15 October 2019, and the testing dataset was the following 28 days. It is an essential step to complete the information of the samples by means of filling the missing data, and the lack information of the data in this paper is filled by calculating the median value of the former and the latter in samples which would be considered to replace the missing data so as to improve the efficiency of the forecasting model. Additionally, data normalization could help reduce the influence of different magnitude levels on the predicted results. The normalization is completed rely on the mapping function which is described as follows:

$$x' = \frac{x - \min}{\max - \min}, \tag{5}$$

where x is the original value of the data and x' is the normalized data through the transform function above, and min and max which appeared in the formula, respectively, represent the maximum and minimum values of the located column in the data. The data could be ultimately mapped to the interval which ranges from 0 to 1.

4.1.2. Differencing. Since the statistical modeling methods assume or require the time series to be stationary, which makes the model easier and more effective, it is necessary to check the stationarity of the time series before it is put into the ARIMA model. The initial time series fluctuation is exhibited in Figure 1; it is clear from the plot that the time series is not stationary which means that it contains the trend and seasonal components. In order to verify the stationarity of the initial data, the detection measures were using the autocorrelation function (ACF) and partial autocorrelation (PACF), and the graph was presented in Figure 2. By observing the values of the autocorrelation in Figure 2, it decays very slowly to zero and it indicates that the initial time series data is not stationary. The procedure



	AC	PAC	Q-Stat	Prob
1	0.969	0.969	86.422	0.000
2	0.937	−0.040	168.08	0.000
3	0.903	−0.042	244.78	0.000
4	0.867	−0.041	316.40	0.000
5	0.830	−0.040	382.84	0.000
6	0.792	−0.041	444.02	0.000
7	0.752	−0.041	499.91	0.000
8	0.713	−0.013	550.74	0.000
9	0.674	−0.026	596.67	0.000
10	0.633	−0.040	637.74	0.000
11	0.593	−0.017	674.23	0.000
12	0.553	−0.019	706.38	0.000
13	0.514	−0.008	734.54	0.000
14	0.477	0.004	759.12	0.000
15	0.441	−0.015	780.38	0.000
16	0.405	−0.010	798.62	0.000
17	0.371	−0.011	814.13	0.000
18	0.338	−0.012	827.17	0.000
19	0.306	−0.010	838.00	0.000
20	0.275	−0.010	846.89	0.000
21	0.245	−0.009	854.06	0.000
22	0.217	−0.007	859.75	0.000
23	0.190	−0.007	864.18	0.000
24	0.165	0.012	867.58	0.000

Figure 2: The ACF and PACF graph of the original time series.

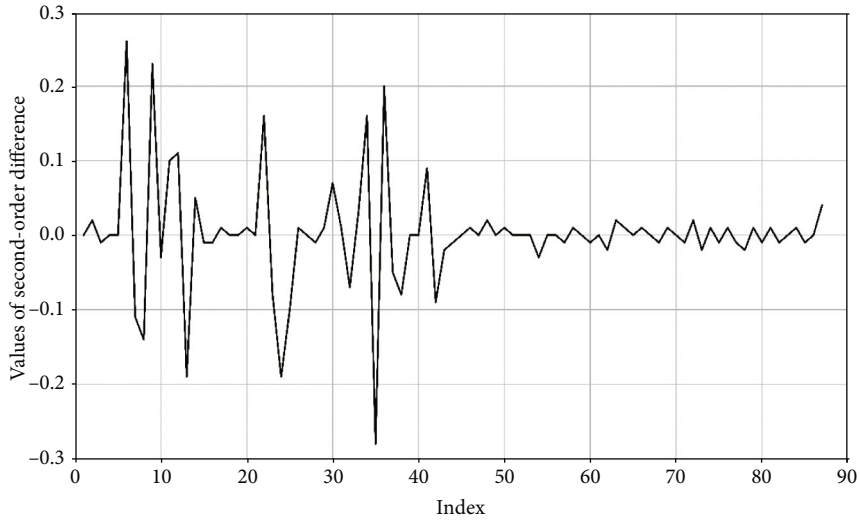
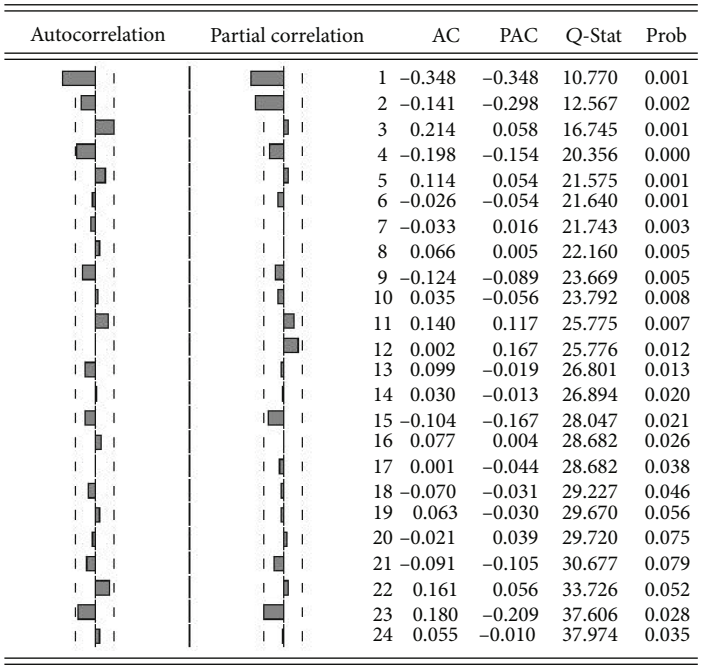


Figure 3: The curve of the second difference time series.

of differencing is an essential step to make time series stationary. The number of differencing required to make sequence stationary in this paper is gained by ndiffs() function of R programming language, and the term d of the ARIMA model is ultimately determined to be 2. Figure 3 is the second-difference graph of the original data; it can be seen intuitively that the curve fluctuates around 0 and the trend is stationary. The ACF and PACF graph of the second-difference of original data is displayed in Figure 4; as can be observed in the figure, the values of the function are sharply dropped down to 0 and fluctuated in the confidence interval. In addition, the measure of Augmented Dickey-Fuller test (ADF) is applied to identify the stationary or nonstationary processes of the second difference time series, the results are presented in Table 1, the significance level is set as 1%, 5%, and 10%, and the small p value suggests that the second difference time series is stationary.

4.2. ARIMA Model Construction. The autoregressive integrated moving average (ARIMA) model is based on the classical Box-Jenkins methodology for forecasting time series data. The precondition of establishing the ARIMA model is obtaining a stationary time series, and it has been fulfilled by the aforementioned differencing procedure. In this research, the terms of the ARIMA model are determined automatically by auto.arima function of R programming



Autocorrelation	Partial correlation		AC	PAC	Q-Stat	Prob
		1	−0.348	−0.348	10.770	0.001
		2	−0.141	−0.298	12.567	0.002
		3	0.214	0.058	16.745	0.001
		4	−0.198	−0.154	20.356	0.000
		5	0.114	0.054	21.575	0.001
		6	−0.026	−0.054	21.640	0.001
		7	−0.033	0.016	21.743	0.003
		8	0.066	0.005	22.160	0.005
		9	−0.124	−0.089	23.669	0.005
		10	0.035	−0.056	23.792	0.008
		11	0.140	0.117	25.775	0.007
		12	0.002	0.167	25.776	0.012
		13	0.099	−0.019	26.801	0.013
		14	0.030	−0.013	26.894	0.020
		15	−0.104	−0.167	28.047	0.021
		16	0.077	0.004	28.682	0.026
		17	0.001	−0.044	28.682	0.038
		18	−0.070	−0.031	29.227	0.046
		19	0.063	−0.030	29.670	0.056
		20	−0.021	0.039	29.720	0.075
		21	−0.091	−0.105	30.677	0.079
		22	0.161	0.056	33.726	0.052
		23	0.180	−0.209	37.606	0.028
		24	0.055	−0.010	37.974	0.035

Figure 4: The ACF and PACF graph of the second difference time series.

Table 1: The statistical results of the ADF test for the second difference time series.

		t-statistic	*p*-value
Augmented Dickey-Fuller test statistic		-10.05383	≤0.001
Test critical values:	1% level	-3.510259	
	5% level	-2.896346	
	10% level	-2.585396	

language, and the ARIMA model which with the minimum AIC (Akaike Information Criterion) score will be selected for future forecasts. In this context, the final model is determined as ARIMA(0, 2, 1) with a minimum AIC score -2.499856. The next step is to validate the applicability of the model using Box-Ljung statistics, if the calculated residuals are white noise, then the model will be determined as the definitive model for further forecasting; otherwise, a more suitable model needs to be found. The Box-Ljung test for the calculated residuals of the model is achieved using the Box.test function of R programming language; the statistical p value of the Box-Ljung test is 0.8993, and it indicates that this ARIMA model is suitable for further forecasting due to the p value of the Box-Ljung test being greater than 0.05. The final ARIMA model is utilized to conduct the further forecasting, and the prediction results are shown in Figure 5; the dashed line represents the prediction value; it can be seen that the fluctuation of the electricity consumption is on the rise in the next 28 days; the specific predicted values of the ARIMA model are presented in Table 2.

4.3. The Construction of Hybrid Model. It is a popular trend to combine the statistical model with machine learning methods to establish the electricity consumption prediction method. In the hybrid model, incorporating the residual error values into the predicted series deduces the better precision of the predictions [12]. The general construction process of the hybrid models is depicted in Figure 6. As presented in the following pipeline, the construction procedure of the hybrid model generally consists of two main steps; firstly, the ARIMA model is applied in this hybrid model as a well-known linear model; secondly, the residual error prediction is also considered in the hybrid model for the purpose of extracting the nonlinear features of the input data. The residual error is calculated from the ARIMA predictions and furthermore applied to the nonlinear model as input data. What is more, the procedures of normalization and renormalization are also contained in the design process with the aim of removing the influence of the magnitude levels. The obtained corrected predictions are taken as the ultimate results of the hybrid model and applied as the final predicted value.

In general, it is more effective to combine individual models for forecasting energy consumption. The hybrid method which combines with machine learning and univariate ARIMA method has been used more frequently because the hybrid method could benefit from both of them. Just as its name implies, the ARIMA-SVR algorithm is a combination of the ARIMA model and SVR model. Similarly, the ARIMA-GBR model was also established in the same way, and the performance comparison of the five models is presented in Table 2.

Figure 7 has depicted the forecasting results of the proposed model and relevant models with different predicted days. Within a prediction interval of 16 days, the output of the ARIMA-SVR model more closely approximates the

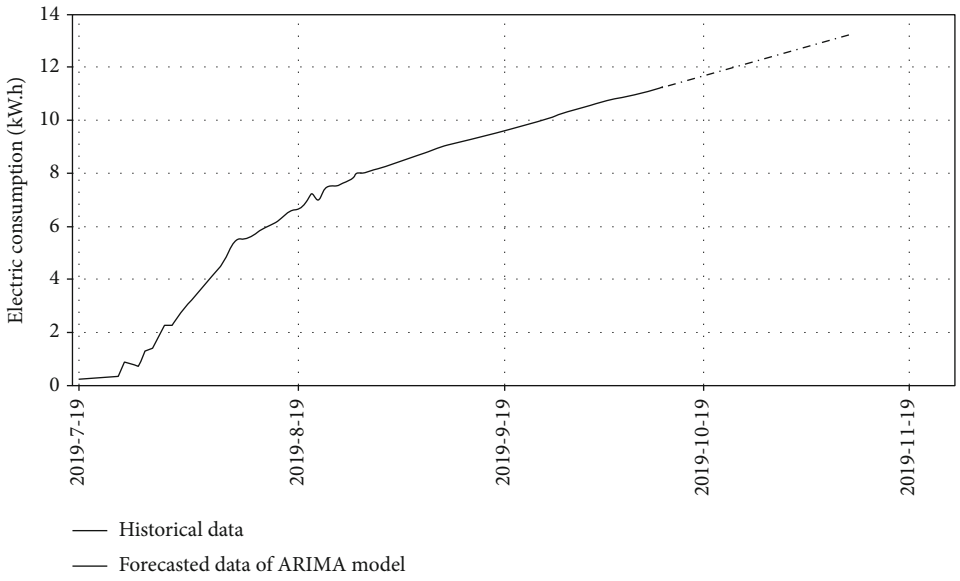


Figure 5: The prediction curve of the ARIMA model (dash line represents the prediction value).

Table 2: Forecasting results of the proposed model and relevant models.

Date	Actual values (kw/h)	Predicted values (kw/h)				
		ARIMA model	ARIMA-SVR model	ARIMA-GBR model	LSTM model	GRU model
2019-10-16	11.31	11.2915	11.3282	11.2899	11.2491	11.2279
2019-10-17	11.39	11.3629	11.4004	11.3614	11.3041	11.3220
2019-10-18	11.48	11.4344	11.4726	11.3554	11.5189	11.4989
2019-10-19	11.56	11.5059	11.5448	11.4269	11.4577	11.4398
2019-10-20	11.65	11.5773	11.617	11.4983	11.6105	11.5731
2019-10-21	11.71	11.6488	11.6892	11.6213	11.6519	11.6113
2019-10-22	11.77	11.7203	11.7614	11.7061	11.5522	11.601
2019-10-23	11.85	11.7917	11.8336	11.7775	11.6989	11.7306
2019-10-24	11.90	11.8632	11.9058	11.8490	11.8505	11.7955
2019-10-25	11.96	11.9347	11.9779	11.9205	11.8838	11.8988
2019-10-26	12.04	12.0061	12.0501	11.9919	11.8921	11.9203
2019-10-27	12.12	12.0776	12.1223	12.0634	11.8078	11.8065
2019-10-28	12.19	12.1491	12.1945	12.1349	11.9552	11.8758
2019-10-29	12.27	12.2205	12.2667	12.2446	11.8943	11.9488
2019-10-30	12.35	12.2920	12.3389	12.3160	12.041	12.0671
2019-10-31	12.40	12.3635	12.4111	12.3852	11.9898	12.0154
2019-11-1	12.44	12.4349	12.4833	12.4567	12.2121	12.6195
2019-11-2	12.48	12.5064	12.5555	12.5095	12.1576	12.5813
2019-11-3	12.53	12.5779	12.6276	12.5809	12.3366	12.553
2019-11-4	12.57	12.6494	12.6998	12.6524	12.4153	12.6426
2019-11-5	12.62	12.7208	12.7720	12.7239	12.5056	12.8841
2019-11-6	12.66	12.7923	12.8442	12.7953	12.6973	13.0932
2019-11-7	12.71	12.8638	12.9164	12.8668	12.6343	12.8843
2019-11-8	12.80	12.9352	12.9886	12.9383	12.7165	13.0761
2019-11-9	12.85	13.0067	13.0608	13.0097	12.5879	12.4027
2019-11-10	12.89	13.0781	13.133	13.0812	12.6287	12.5881
2019-11-11	12.94	13.1496	13.2052	13.1527	12.7034	12.7740
2019-11-12	12.99	13.2211	13.2774	13.2241	12.9237	13.1884

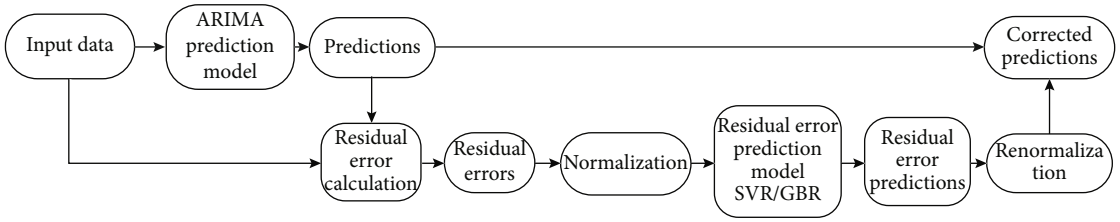


Figure 6: The flow chart of the hybrid model construction.

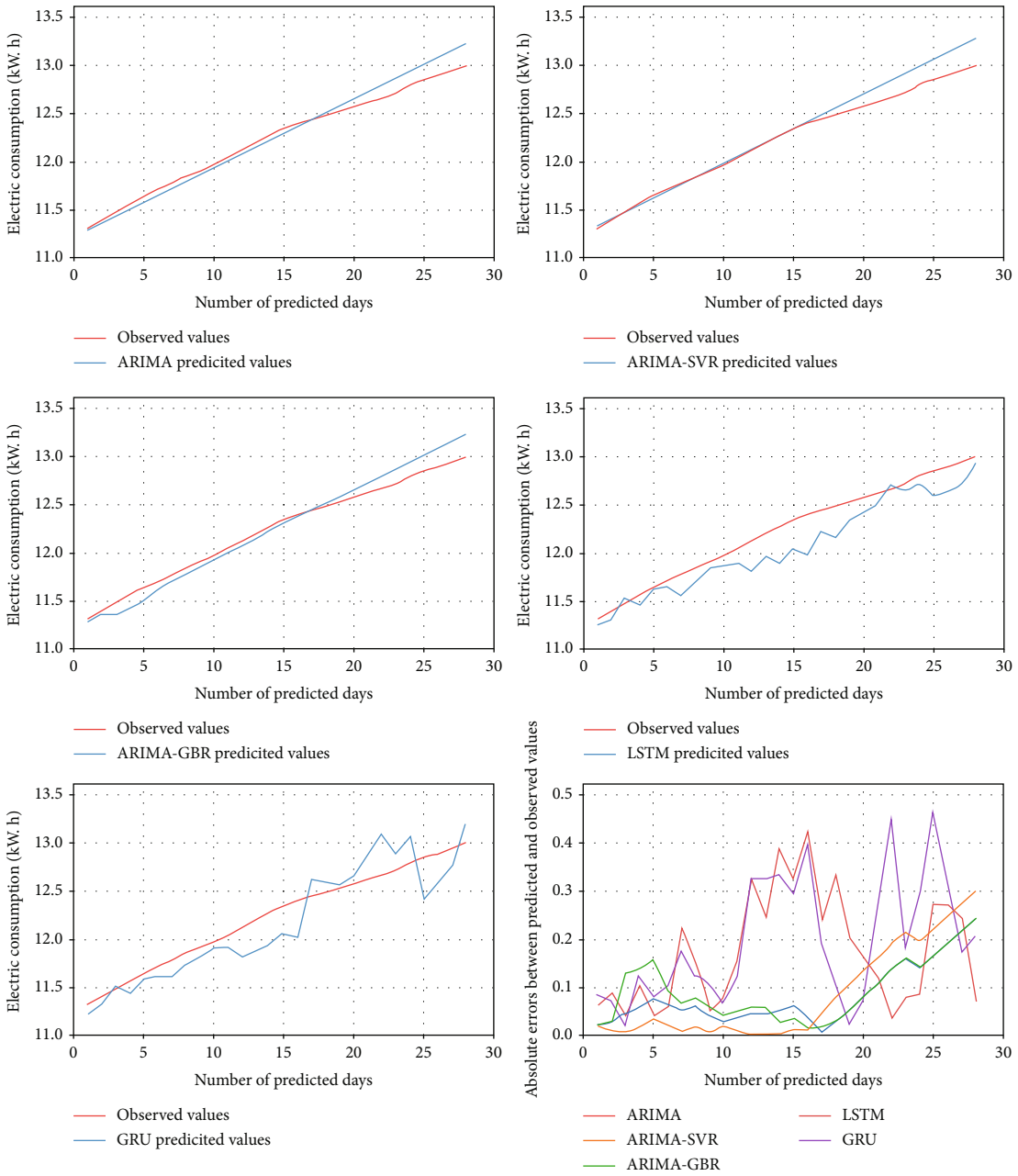


Figure 7: Forecasting results of the proposed model and relevant models.

measured electricity consumption than that other relevant models. The last subgraph above is the absolute error curve between the observed and predicted values; the yellow line represents the absolute error between the predicted values of the ARIMA-SVR model and the observed data; it fluctuates lower than the other curves stability until the predicted

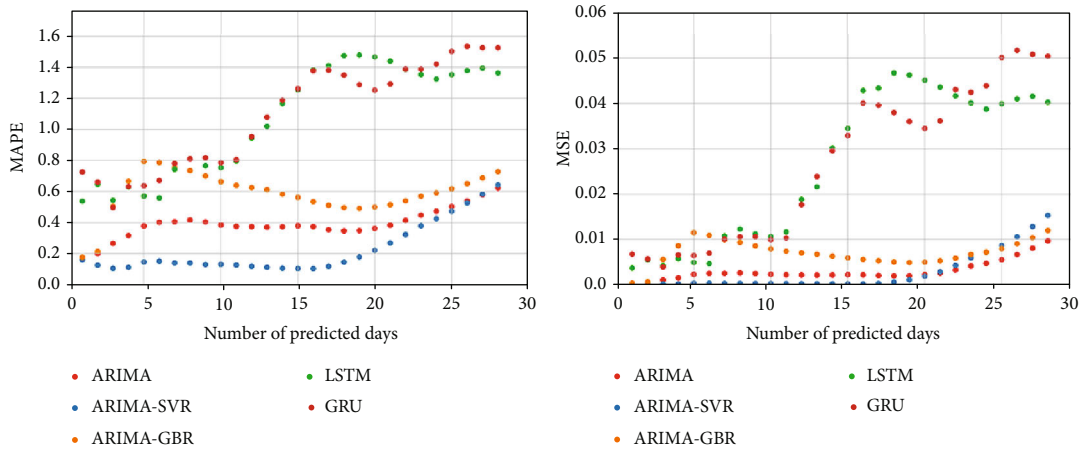


Figure 8: The MAPE and MSE of the five forecasting models.

point reached at 16 days; the absolute error curve of the single ARIMA model is under than that of ARIMA-SVR model when the predicted point beyond 16 days, and it almost overlapped with the error curve of the ARIMA-GBR hybrid model. The results come out to be that the hybrid model ARIMA-SVR could improve the prediction performance of the individual ARIMA model to a certain degree, and the drawback of the single linear model could be overcome when it is in conjunction with some nonlinear models which have the capability of capturing nonlinear features in the dataset. This is because the time series is often neither purely linear nor purely nonlinear. Thus using an individual model alone could not capture the data characteristics adequately. Besides, it is worth noting that the hybrid model could improve the accuracy of forecasting efficiently only in certain situations or in a certain prediction horizon. The hybrid models are more and more used by many researchers due to their excellent predictive performance.

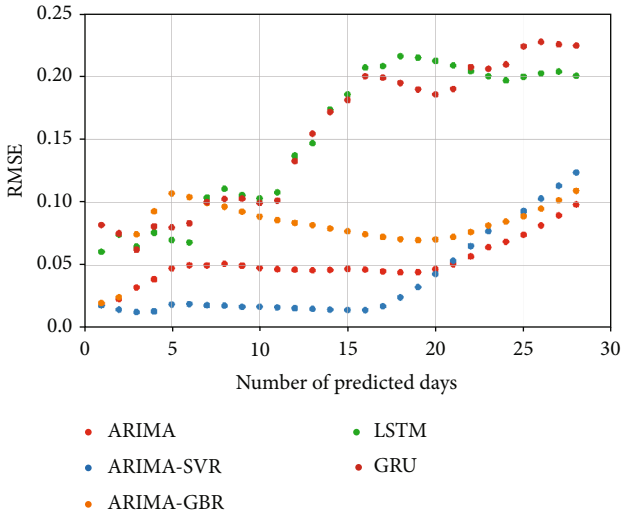


Figure 9: The RMSE of the five forecasting models.

5. Simulation Results and Discussion

The paper uses three evaluation criteria to assess the performance of the five models, respectively, are the mean squared error (MSE), the root mean square error (RMSE), and the mean absolute percentage error (MAPE); the formulations are detailed as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - y\wedge_i|^2, \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - y_i\wedge)^2}, \tag{7}$$

$$\mathrm{MAPE} = \sum_{i=1}^{n} \left|\frac{y_i - \widehat{y}_i}{y_i}\right| \times \frac{100}{n}. \tag{8}$$

Here, y_i is the observed data, $\widehat{y}_i$ is the predicted value of the forecast model, and n is the number of the observed dataset.

Figures 8 and 9 have depicted the prediction errors MAPE, MSE, and RMSE of five forecasting models; it can be seen that the MAPE values of the ARIMA-SVR model are the lowest values among the other relevant models within the prediction horizon of 26 days, and the MSE and RMSE values of the ARIMA-SVR model are the lowest values within 20 predicted days. Also, the MSE and RMSE values of the ARIMA-SVR model maintain the lowest within 20 predicted days whereas the lowest region would be taken place by the errors of the single ARIMA model when prediction horizon beyond 20 days. Table 3 has presented the RMSE values of the five predictive models. The predictive performance of the LSTM and GRU models is not as good as the ARIMA-SVR model; the main reason is probably that the dataset is small in this paper. In addition, it is notable that the prediction performance of the ARIMA-SVR model is more stable compared to the other models, and the prediction results showed that the ARIMA-SVR model is a good choice when the electricity consumption is predicted within 20 days, and the ARIMA model is suitable for the prediction over 20 days.

Table 3: The root mean square error (RMSE) of the five predictive models.

Predicted days	Root mean square error (RMSE)				
	ARIMA model	ARIMA-SVR model	ARIMA-GBR model	LSTM model	GRU model
1	0.0185	0.0182	0.0201	0.0609	0.0821
2	0.0232	0.0148	0.0247	0.0745	0.0754
3	0.0324	0.0128	0.0747	0.0648	0.0625
4	0.039	0.0135	0.0928	0.0759	0.0809
5	0.0477	0.019	0.1072	0.0702	0.0801
6	0.0502	0.0194	0.1044	0.0683	0.0835
7	0.0501	0.0182	0.0996	0.1038	0.1003
8	0.0512	0.018	0.0966	0.1108	0.1029
9	0.0498	0.0171	0.0927	0.1058	0.103
10	0.0479	0.0172	0.0888	0.1032	0.0997
11	0.0468	0.0167	0.0859	0.108	0.1016
12	0.0465	0.016	0.0838	0.1372	0.1329
13	0.0461	0.0154	0.082	0.147	0.1546
14	0.0463	0.0148	0.0793	0.1736	0.1719
15	0.0472	0.0146	0.0771	0.1858	0.1814
16	0.0466	0.0144	0.0748	0.2071	0.2003
17	0.0452	0.0175	0.0726	0.2083	0.1991
18	0.0444	0.0246	0.0709	0.2163	0.195
19	0.0446	0.0328	0.07	0.2151	0.1898
20	0.0469	0.0432	0.0707	0.2125	0.1857
21	0.0508	0.0536	0.0726	0.2089	0.1902
22	0.057	0.0655	0.0766	0.2042	0.2075
23	0.0644	0.0772	0.0817	0.2004	0.2062
24	0.0688	0.0848	0.0848	0.1969	0.2096
25	0.0743	0.0932	0.0891	0.1999	0.224
26	0.0817	0.103	0.095	0.2026	0.2275
27	0.0898	0.1132	0.1018	0.204	0.2255
28	0.0983	0.1238	0.1094	0.2007	0.2246

The experiment is conducted on a PC with Intel Core i5-8300H CPU @2.30GHz, 8.00 GB RAM, and 64-bit operating systems. Exhibited graphs and hybrid models are implemented in Pycharm using Python language.

6. Conclusions

The results have indicated that the hybrid method ARIMA-SVR has great capability for forecasting the electricity consumption of the buildings; it is efficient for enhancing the accuracy of the electricity consumption prediction in certain conditions. In the prediction horizon of 20 days, the hybrid model ARIMA-SVR has significant superiority than the other four models while the ARIMA model is a better choice when the prediction horizon exceeds 20 days.

The commended hybrid ARIMA-SVR model in this paper has provided a new perspective of understanding a complex energy data structure. The general process of the hybrid model building is always decomposing the original time series into a stationary linear component and a fluctuant nonlinear residual. The limitations of the proposed model are that it highly dependents on time series and requires the univariate time series to be stationary or to be stationary after differencing; data preprocessing is an essential step before building the model. As we all know, the overall electricity consumption of buildings was also related to energy-using behaviors and climatic factors. In further works, we plan to consider some relevant eigenvalues such as temperature and humidity into model construction and focus on combining some popular algorithms testing on different size datasets; on the other hand, we will focus on extending the effective scope of the proposed model in future studies.

Forecasting electricity consumption in advance is of great importance in achieving energy conservation, and it could provide data support for policymakers and energy managers to make reasonable strategies to promote environmental protection.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under Grant 52074305, in part by the National Natural Science Foundation of China under Grant 51874300 and 51874299, in part by the National Natural Science Foundation of China and Shanxi Provincial People's Government Jointly Funded Project of China for Coal Base and Low Carbon under Grant U1510115, and in part by the Open Research Fund of Key Laboratory of Wireless Sensor Network and Communication, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, under Grant 20190902.

References

[1] Y. Sun, W. Hao, Y. Chen, and B. Liu, "Data-driven occupant-behavior analytics for residential buildings," *Energy*, vol. 206, p. 118100, 2020.

[2] E. Caro and J. Juan, "Short-term load forecasting for Spanish insular electric systems," *Energies*, vol. 13, no. 14, article 3645, 2020.

[3] S. L. Ho and M. Xie, "The use of ARIMA models for reliability forecasting and analysis," *Computers & Industrial Engineering*, vol. 35, no. 1-2, pp. 213–216, 1998.

[4] H. Tandon, P. Ranjan, T. Chakraborty, and V. Suhag, "Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future," 2020, https://arxiv.org/abs/2004.07859.

[5] A. Barman, "Time series analysis and forecasting of COVID-19 cases using LSTM and ARIMA models," 2020, https://arxiv.org/abs/2006.13852.

[6] R. Jamil, "Hydroelectricity consumption forecast for Pakistan using ARIMA modeling and supply-demand analysis for the year 2030," *Renewable Energy*, vol. 154, pp. 1–10, 2020.

[7] Y. Li, Y. Wei, and Z. Dong, "Will China achieve its ambitious goal?—forecasting the CO_2 emission intensity of china towards 2030," *Energies*, vol. 13, no. 11, article 2924, 2020.

[8] F. Kaytez, "A hybrid approach based on autoregressive integrated moving average and least- square support vector machine for long-term forecasting of net electricity consumption," *Energy*, vol. 197, article 117200, 2020.

[9] A. T. Eseye and M. Lehtonen, "Short-term forecasting of heat demand of buildings for efficient and optimal energy management based on integrated machine learning models," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7743–7755, 2020.

[10] T. Ahmad and H. Zhang, "Novel deep supervised ML models with feature selection approach for large- scale utilities and buildings short and medium-term load requirement forecasts," *Energy*, vol. 209, article 118477, 2020.

[11] C. Liu, B. Sun, C. Zhang, and F. Li, "A hybrid prediction model for residential electricity consumption using holt- winters and extreme learning machine," *Applied Energy*, vol. 275, article 115383, 2020.

[12] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 243–254, 2014.

[13] M. R. Kazemzadeh, A. Amjadian, and T. Amraee, "A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting," *Energy*, vol. 204, article 117948, 2020.

[14] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji, "Electricity load forecasting using clustering and ARIMA model for energy management in buildings," *Japan Architectural Review*, vol. 3, no. 1, pp. 62–76, 2020.

[15] M. Ma and Z. Wang, "Prediction of the energy consumption variation trend in South Africa based on ARIMA, NGM and NGM-ARIMA Models," *Energies*, vol. 13, no. 1, p. 10, 2020.

[16] E. Gulay and O. Duru, "Hybrid modeling in the predictive analytics of energy systems and prices," *Applied Energy*, vol. 268, article 114985, 2020.

[17] K. Balachander and D. Paulraj, "ANN and fuzzy based household energy consumption prediction with high accuracy," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, pp. 1–15, 2020.

[18] T. Silveira Gontijo and M. Azevedo Costa, "Forecasting hierarchical time series in power generation," *Energies*, vol. 13, no. 14, article 3722, 2020.

[19] A. J. del Real, F. Dorado, and J. Durán, "Energy demand forecasting using deep learning: applications for the French grid," *Energies*, vol. 13, no. 9, article 2242, 2020.

[20] F. Prado, M. C. Minutolo, and W. Kristjanpoller, "Forecasting based on an ensemble autoregressive moving average - adaptive neuro - fuzzy inference system - neural network - genetic algorithm framework," *Energy*, vol. 197, article 117159, 2020.

[21] G. E. Box and G. C. Tiao, "Intervention analysis with applications to economic and environmental problems," *Journal of the American Statistical Association*, vol. 70, no. 349, pp. 70–79, 1975.

[22] V. Vapnik, S. E. Golowich, and A. J. Smola, "Support vector method for function approximation, regression estimation and signal processing," *Advances in Neural Information Processing Systems*, vol. 9, pp. 281–287, 2008.

[23] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.

[24] M. Awad, R. Khanna, M. Awad, and R. Khanna, "Support vector regression," in *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, pp. 67–80, Apress, Berkeley, CA, 2015.

[25] A. Chokor and M. El Asmar, "Data-driven approach to investigate the energy consumption of LEED-certified research buildings in climate zone 2B," *Journal of Energy Engineering*, vol. 143, no. 2, article 05016006, 2017.

[26] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5574255, 14 pages
https://doi.org/10.1155/2021/5574255

Research Article

A New Method for Reconstructing Data on a Single Failure Node in the Distributed Storage System Based on the MSR Code

Miao Ye,[1,2] **Ruoyu Wei**,[1] **Wei Guo**,[3] **Qiuxiang Jiang**,[4] **Hongbing Qiu**,[1,2] **and Yong Wang**[3,5]

[1]*School of Information and Communications, Guilin University of Electronic Technology, Guilin Guangxi 541004, China*
[2]*Key Laboratory of Cognitive Radio and Information Processing, Guilin University of Electronic Technology, Guilin 541004, China*
[3]*School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin Guangxi 541004, China*
[4]*School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin Guangxi 541004, China*
[5]*Guangxi Education Collaborative Innovation Center for Big Data and Network Security, Guilin University of Electronic Technology, Guilin Guangxi 541004, China*

Correspondence should be addressed to Hongbing Qiu; qiuhb@guet.edu.cn

Received 1 February 2021; Revised 22 February 2021; Accepted 19 March 2021; Published 1 April 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Miao Ye et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a storage method for a distributed storage system, an erasure code can save storage space and repair the data of failed nodes. However, most studies that discuss the repair of fault nodes in the erasure code mode only focus on the condition that the bandwidth of heterogeneous links restricts the repair rate but ignore the condition that the storage node is heterogeneous, the cost of repair traffic in the repair process, and the influence of the failure of secondary nodes on the repair process. An optimal repair strategy based on the minimum storage regenerative (MSR) code and a hybrid genetic algorithm is proposed for single-node fault scenarios to solve the above problems. In this work, the single-node data repair problem is modeled as an optimization problem of an optimal *Steiner tree* with constraints considering heterogeneous link bandwidth and heterogeneous node processing capacity and takes repair traffic and repair delay as optimization objectives. After that, a hybrid genetic algorithm is designed to solve the problem. The experimental results show that under the same scales used in the MSR code cases, our approach has good robustness and its repair delay decreases by 10% and 55% compared with the conventional tree repair topology and star repair topology, respectively; the repair flow increases by 10% compared with the star topology, and the flow rate of the conventional tree repair topology decreases by 40%.

1. Introduction

Distributed storage first refers to a storage system built on large-scale and low-cost commercial hardware (network-attached storage (NAS)). It uses a hardware and software cooperation mechanism to encapsulate the interface to provide external programmable calls and command access [1]. Compared with the traditional centralized storage system, the design framework of a distributed storage system brings many advantages. For example, the storage capacity of a cluster can be expanded by constantly adding new underlying hardware devices or upgrading old hardware devices, and the adoption of large-scale and inexpensive commercial hardware reduces the cost of cluster construction. The advantages of a distributed storage system enable it to cope with the various demands of mass data storage. Therefore, distributed storage systems are widely favored worldwide.

With the increase in the number of storage nodes and the diversity and complexity of storage nodes in distributed storage systems, nodes often fail [2]. There are two ways to improve the reliability of a storage system: the copy mode [3] and erasure code [4]. The copy method stores the data in multiple copies, and when the storage node where one copy is located fails, the storage system can switch the service to the remaining copies. This method does not involve special coding and reconstruction algorithms; it has good fault-tolerant performance but very low storage utilization [5]. The erasure code method uses the data encoding method to store the coded blocks obtained after data encoding. When a storage node where a coded block is located fails, the

storage system can calculate the coding block and recover the fault node. Erasure code technology has the advantage of low storage costs; however, its repair cost is high, and it generates considerable repair traffic that occupies considerable bandwidth.

Although the copy strategy has advantages over the erasure correction code in terms of the speed of reading surviving data and recovering lost data, its high storage space cost is contrary to the original intention of distributed storage systems. The erasure code can provide almost the same reliability as the multireplica strategy while effectively improving the utilization of storage space and providing higher data writing efficiency, making it widely used in distributed storage systems.

Traditional erasure codes, such as the *Vandermonde* and Reed-Solomon (RS) codes [6] $(k+r, k)$, divide the original data of size m into k data blocks with the size of M/k. By applying the *Vandermonde* matrix in the *Galois* domain $\mathrm{GW}(2^w)$, the loss within any r block in $k+r$ blocks can be obtained by obtaining all the data from any k nodes from the remaining living nodes through the replacement node, and the codec operation can be performed to obtain the original data and repair the lost data. Traditional coding usually needs to obtain the data with the size of M from k nodes to recover the failed nodes and requires other nodes to transmit much repair traffic; this is the main reason for a large amount of repair traffic in the correction code method. The minimum storage regenerative (MSR) code can recover the fault node by performing more complex encoding and decoding operations on the original data and can recover the fault node when the total amount of data transmitted by the node providing the data is M/k, which greatly reduces the repair flow and repair delay so that the stored data can be repaired faster.

When the erasure code is applied to repair a failed node, the storage system provides a new node in the cluster topology to replace the failed node. This node is called the *newcomer*. The *newcomer* downloads data from several surviving nodes and performs the data recovery operation. The nodes that provide data to the *newcomer* are called the *provider* nodes. In the traditional data repair process, the repair topology in which *provider* nodes transmit the repair data to a *newcomer* is of the star type. The repair only starts when all *provider* nodes transmit their data to the *newcomer*, so the delay in repair depends on the bottleneck bandwidth. To reduce the repair delay, researchers proposed a tree repair topology based on the star repair topology that creates a tree repair topology with the *newcomer* as the root and then selects links with higher available bandwidth to join the topology to reduce the repair delay.

Compared with the star repair topology, the traditional tree repair topology does improve the data repair rate, but it brings more traffic overhead to the storage cluster network. In the construction of the repair tree, the optimal repair topology with the best repair delay can be obtained. This also causes more node selection problems, while the traditional greedy algorithm causes more computing overhead and increases the delay in repairing failed nodes due to the increase in nodes. Additionally, compared with the conventional star repair topology, the robustness of the conventional tree repair topology is very low when the repair operation of a failed node in the storage cluster is in progress. If another node fails in the repair topology, the traditional star topology just picks a live node to continue the repair. Nevertheless, when node losses occur in a traditional tree repair topology, all nodes in a subtree with a lost node as the root are impacted and the repair topology needs to be reconstructed, which results in a reduction in the repair rate.

In recent years, some improvements of the data construction for a single failure node have emerged, but the performances of the time delay, the transmitting traffic, and the robustness of the repair topology still need to be improved significantly. As the *Steiner tree* with constrained conditions for the data construction is a kind of NP-hard problem, it is hard to find the global optimal solution. For the above problem, in this paper, some nodes with excellent bandwidth properties in the storage cluster are added to the repair node set to participate in the construction of the repair tree as alternative intermediate nodes, and based on the characteristics of the MSR coded intermediate node computation, a *Steiner tree* model with constraints is established to optimize the repair time delay and the traffic flow in the repair topology. Then, by designing a hybrid genetic algorithm to calculate the approximate global optimal solution, finally, an optimal repair topology construction method based on the MSR code and a hybrid genetic algorithm is proposed. The simulation results show that under the same size of MSR codes, the repair delay is 70%-90% of that of the traditional tree repair topology and only 35%-45% of that of the traditional star repair topology. Although the repair traffic in the topology is slightly improved compared to the traditional star topology, the repair traffic is only 45%~60% of that of the traditional tree repair topology, and the constructed repair topology has higher robustness.

In summary, the innovations of this paper are as follows:

(1) We consider the heterogeneity of the nodes, and the node set participating in the repair joins the nodes with better bandwidth properties in the cluster and participates in the construction of the repair topology as alternative intermediate nodes to obtain the repair tree with the optimal repair characteristics and to enhance the robustness of the repair topology simultaneously. As far as we know, there is no work that considers time delay and robustness at the same time in the available literature we can find

(2) The MSR code with better storage characteristics is used to construct an optimal repair tree with constraints based on the calculation of the intermediate nodes, which reduces the repair traffic cost of erasure code tree repair

(3) To solve the NP-hard problem of the optimal repair tree, a hybrid genetic algorithm is designed according to the background characteristics of the problem to obtain the optimal repair *Steiner tree* based on the repair node set

The rest of this article is organized as follows. Section 2 introduces the research status of the distributed storage

system and erasure code. Section 3 analyzes the characteristics and topology structure of the MSR code, the influence of the repair flow, and the node performance during the repair process to deduce the definition of the repair topology problem that needs to be constructed in this paper, and a mathematical model is established. Section 4 describes in detail the design and algorithm flow of each operator of the genetic algorithm proposed in this paper. In Section 5, experiments are designed to compare and analyze the performance differences between the proposed algorithm and the traditional star and tree repair algorithms and to analyze the influence of the parameters in the algorithm. Section 6 summarizes the whole paper and proposes future research directions.

2. Related Work

There are two ways to guarantee the reliability of nodes. One is multireplica, and the other is the erasure code [7]. Erasure codes originated in the field of communication and were originally used to solve the problems of error detection and error correction in communication transmission. Later, erasure codes were gradually applied to data error detection and error correction in storage systems to improve the reliability of storage systems and were gradually improved and popularized with the characteristics of storage system applications [8].

With the increasing scale of storage systems, the multireplica method has difficulty meeting the requirements of mass storage systems for redundant backup mechanisms in disk utilization and fault tolerance [9]. The advantages of the erasure code method with respect to its low storage cost become increasingly obvious, and research on erasure code technology has become a research hotspot in the industry. At present, the research literature on the disaster recovery strategy of erasure code methods mainly focuses on improving the coding mechanism and the construction of topology in the repair process.

In terms of the improved coding mechanisms, Wang [10] fixed defects with excessive local characteristics in the traditional minimum bandwidth regenerated (MBR) code and variable fractional repetition (VFR) code according to the idea of different data heat and different repeatability. Lin et al. [11] proposed a packet repair code (PRC) to overcome the high cost of multimode failure repair. Teng et al. [12] proposed sparse random erasure codes based on the high probability of the full rank property of a random matrix with equal row weights and sparsity. Based on the RS code, the Microsoft team designed a local reconstruction code (LRC) [13]. Dimakis et al. [14] obtained a regenerated code by improving the basic maximum distance separable (MDS) erasure code. Compared with the traditional erasure code, the regenerated code reduced the repair flow rate. Jiekak et al. [15] proved that the regenerated code repair traffic in the process of repairing the data transmitted by a failed node is approximately equal to 16% of the RS code at the same level.

In terms of repair topology construction, Huang [16] proposed a pipeline repair scheme, but it was essentially a serial transmission mode and could not reach the optimal repair rate. Jia et al. [17] considered the heterogeneity of bandwidth in a storage cluster, constructed a mathematical model based on this, and solved the routing algorithm for repairing data provided by a data-providing node to a newborn node. Gong et al. [18] proposed selection criteria for nodes and tree repair topologies provided by data in the repair process for storage clusters with regenerative code mechanisms and analyzed their advantages. Wan [19] proposed a signal-to-noise ratio (STNR) algorithm for repairing traffic integration at intermediate nodes, but it did not attempt to further optimize the coding mechanism. Zhang et al. [20] considered the actual network topology. They used a field-programmable gate array (FPGA) to replace the switch for traffic integration, but due to the high price of FPGA, it is difficult to apply it in an actual environment.

In this paper, we consider constructing an optimal *Steiner tree* topology with constraints based on the MSR code with better storage characteristics and a hybrid genetic algorithm through a parallel repair network, which not only reduces the bandwidth and computational cost but also increases the robustness of the topology and effectively solves the problem of the low efficiency of a serial repair.

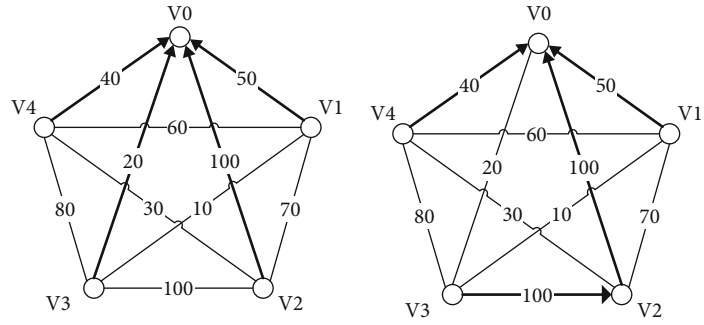
3. Modeling and Formulation of the Single-Node Failure Repair Problem

Given the problems mentioned above in the research of the erasure code repair process, this section analyzes the influence on the single-node fault repair process from three aspects: repair topology, topology repair traffic, and node processing capacity heterogeneity.

3.1. MSR Code. To further reduce the traffic generated by repair operations in clustered networks, Wu et al. proposed the minimal storage regenerating code based on regenerative code [21]. A more complex codec operation of the original data is used to reduce the size of the data block sent by each *provider* in the process of repair and further reduce the repair traffic in the overall topology.

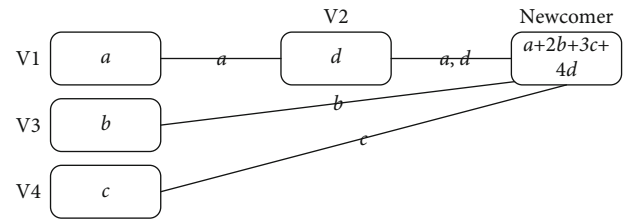
The MSR code with (n, k, d) means the original data block with a size of M is transformed into $k * (d - k + 1)$ encoded blocks which are stored in n storage nodes. Every storage node stores the encoded blocks with the same size of M/k. In general, when a storage node is selected as the provider node, the encoded block with a size of M/k will be transformed into a data block with a size of $M/k(d - k + 1)$ which is sent to the newcomer. When the *newcomer* receives $d * M/(k(d - k + 1))$ data blocks from d providers, the decoding operation can be performed to restore the data of size M/k stored on the failed node. As the data is transmitted by each *provider*, $\beta = M/(k(d - k + 1))$, under the MSR code disaster tolerance mechanism, the *newcomer* also needs to directly access the data from the $d - k + 1$ node during the repair process. Therefore, any repair topology that uses the MSR code must satisfy [22]

$$\text{degree}_{\text{newcomer}} \geq d - k + 1. \quad (1)$$

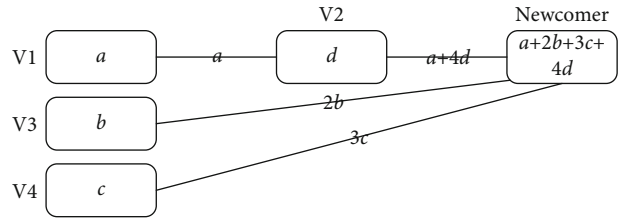


(a) Traditional star repair topology (b) Traditional tree repair topology

Figure 1: Traditional star repair topology and tree repair topology (unit Mbps).



(a) The repair traffic of each link in the case of the traditional tree topology



(b) The repair traffic of each link in the case of the tree topology with an intermediate node aggregation mechanism

Figure 2: Repair flow with the traditional tree repair topology and tree repair topology adopted an intermediate node aggregation.

3.2. Optimal Bottleneck Bandwidth Repair Tree. To demonstrate the optimization of repair delay between the tree repair topology and the star repair topology under the same environment, a network topology with an erasable code disaster recovery cluster is assumed to have five nodes, V0, V1, V2, V3, and V4. The available bandwidth of the link between them is also provided (unit Mbps), and the data of $M = 600\,\text{MB}$ are stored in the nodes. In the case of (5, 3, and 4) MSR codes and the traditional star repair topology, V0 is adopted as a *newcomer*, as shown in Figure 1(a). Each node of the topology needs to pass $\beta = M/(k(d-k+1)) = 100\,\text{MB}$ during the repair process, and the repair delay is $t = \beta/20\,\text{Mbps} = 5\,\text{s}$. The tree repair topology is shown in Figure 1(b), the repair delay is $t = \text{Max}(\beta/40\,\text{Mbps}, 2\beta/100\,\text{Mbps}) = 2.5\,\text{s}$, and the bottleneck link changes from (V0, V3) to (V0, V4). The tree repair topology is more diversified in the choice of the optimal repair path than the star topology and can make full use of some relatively idle links in the cluster network to improve the overall repair efficiency.

Therefore, the problem of finding the optimal repair delay topology can be transformed into a classical minimum spanning tree problem. A tree repair topology T is found in the whole network topology, and the bottleneck repair delay is required to be the shortest which satisfies

$$T = \min(\max\{t\}), \tag{2}$$

where T represents the transmission delay on each link in the current repair topology.

3.3. Repair Traffic Overhead. Although the tree topology promotes the efficient repair of failure nodes, compared with a star topology, when the overall repair flow is large and when the scale of rectifying the deleted code increases, the topological structure of the repair becomes more complicated. Correcting the gap flow is also increasingly clear because the traffic of each link is the number of *providers* in the subtree times β. The links that are close to the *newcomer* experience more burden than the links close to the *provider*; hence, more network problems, such as congestion, occur.

Given the problem that the total amount of repair traffic in the tree repair topology is too large, considering the

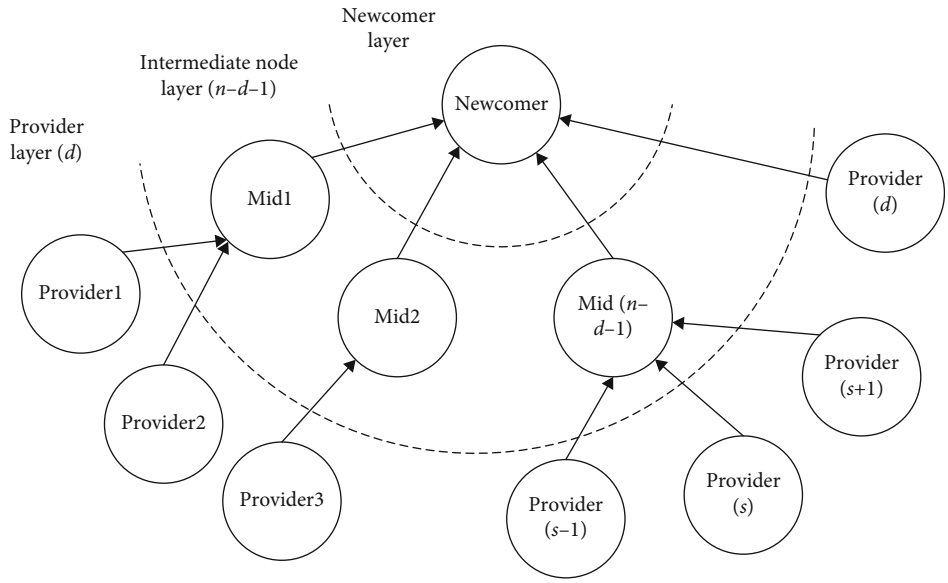


Figure 3: The three-tier repair structure of the topology.

encoding and decoding nature of the erasure code itself, the computing operation finally performed at the *newcomers* can be divided into several aggregation operations and then implemented at the intermediate nodes.

Figure 2(a) shows the repair traffic on each link in the topology when the traditional tree repair is performed. When the data of $a + 2b + 3c + 4d$ need to be repaired, the traffic in the overall repair topology is $a + b + c + d + a + b$, which is 6β in total. In Figure 2(b), the data are preliminarily calculated and processed at each node; then, the repair flow of the subtrees is consolidated at the middle node M1, and the complete decoding operation is finally completed at the *newcomer* node. At this time, the flow in the repair topology is $a + 2b + 3c + 4d + (a + 2b)$, which is 5β in total. Compared with the traditional tree repair topology, the repair flow is reduced by 1 unit.

After the traffic aggregation operation of the intermediate nodes is performed, the traffic saved by each intermediate node is 2 less than its degree in the repair topology. The repair traffic saved by the overall topology is shown in

$$\text{traffic}_{\text{save}} = \sum \text{degree}_{\text{mid}} - 2. \tag{3}$$

Therefore, the traffic saved by the topology is directly proportional to the number of intermediate nodes and nodes in the subtree of the intermediate nodes. However, too many intermediate nodes can generate much extra transmission overhead in the topology, which is contrary to the research goal of the erasure code repair process. Therefore, a balance must be struck between the topology repair rate and the traffic savings. The minimum spanning tree problem mentioned in the previous section is transformed into a *Steiner tree* problem with constraints when a mechanism for the traffic aggregation operation of the intermediate nodes is added to the traditional tree repair topology. In the network topology, a *newcomer* becomes the root of tree T to fix the topology,

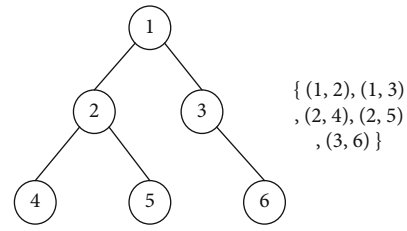


Figure 4: An example of tree coding.

and the repair time and the flow topology are minimized to satisfy

$$T = \min \left\{ c_{\text{time}} \max \left(\frac{\beta}{\omega} \right) + \frac{c_{\text{traffic}}}{\sum \text{degree}_{\text{mid}} - 2} \right\}, \tag{4}$$

where $c_{\text{time}} + c_{\text{traffic}} = 1$, c_{time} represents the repair delay weight coefficient, and c_{traffic} represents the repair traffic weight coefficient. Determining the constructed repair topology takes precedence over the repair delay or repair traffic and represents the available bandwidth of the link.

3.4. Heterogeneous Node Processing Capabilities. Although most research studies on the repair process of erasure codes ignore the processing delay of each node for data by default, the heterogeneity of nodes caused by the nature of the distributed storage system makes the processing speed of each node in the cluster to repair data vary. Compared with most erasure codes, MSR codes cost less in data repair operation at the encoding and decoding operation. In addition, the intermediate node processing mechanism in Section 3.3 will make massive data that are gathered in the middle of provider nodes for processing; so for this repair scenario, we cannot ignore the node processing data delay effect on the efficiency of the whole repair process.

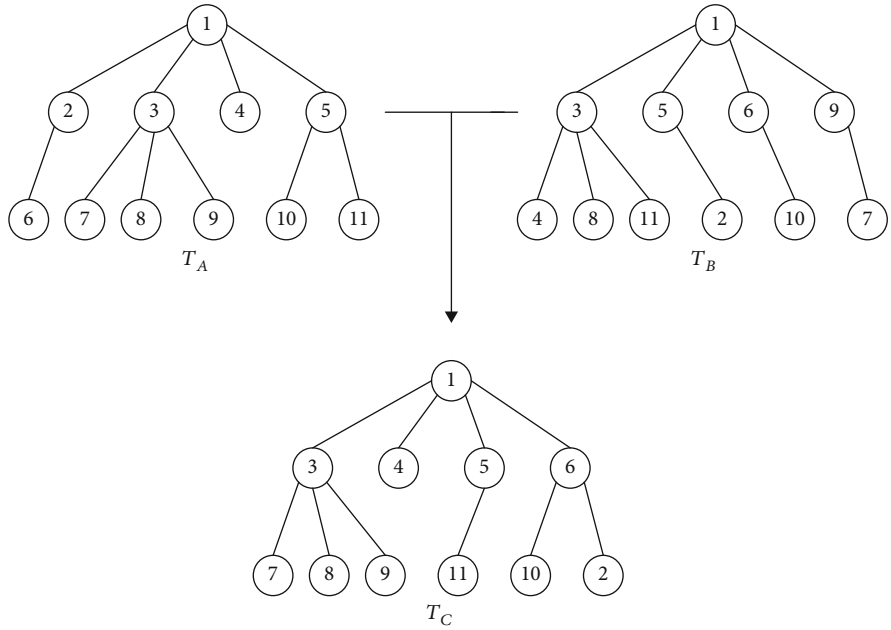


Figure 5: Example of a crossover operator.

This section adopts the definition of node processing capacity in the repair topology proposed by Qi et al. [23]; uses the sequence $x_1, x_2, \cdots, x_k, \cdots, x_m$ to represent the factors that affect the processing capacity of a node, such as a disk I/O, CPU cores, and main frequency; and assigns the corresponding weight factors $\gamma_1, \gamma_2, \cdots, \gamma_k, \cdots, \gamma_m$, where M represents the corresponding serial number. Therefore, the processing capacity of each storage node in the repair topology can be expressed as

$$\text{process}_i = \sum_{k=1}^{m} x_k r_k. \tag{5}$$

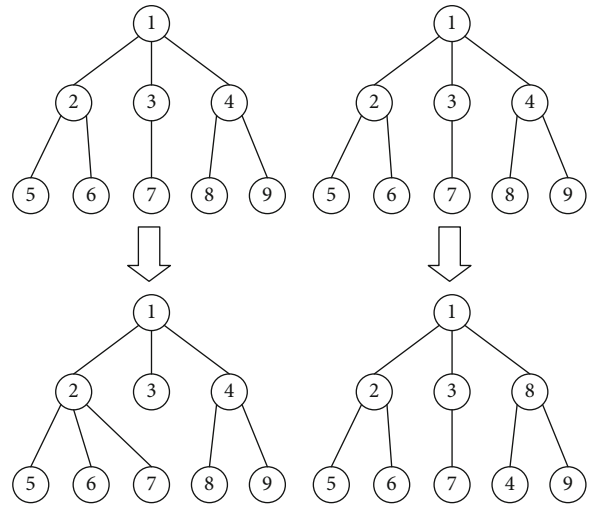
It is assumed that the repair traffic on node V_i is represented by D_i, and the processing time of node V_i for reading, coding, and forwarding the repair traffic is expressed as

$$t_{V_i} = \alpha \frac{D_i}{\text{process}_i}, \tag{6}$$

where α is the capacity conversion coefficient. When considering the processing delay of the intermediate nodes, the repair delay $t(V_i, V_j)$ of the link (V_i, V_j) is converted from β/ω to

$$t_{(V_i,V_j)} = \alpha \frac{D_i}{\text{process}_i} + \frac{\beta}{\omega(V_i, V_j)} + \alpha \frac{D_j}{\text{process}_j}. \tag{7}$$

3.5. The Proposed Multiobject Optimal Model of the Single Failure Node. Based on the above analysis of the MSR code coding mechanism and the three main problems in the repair process of the tree repair topology, the optimal repair topology model based on the MSR code and a genetic algorithm is obtained in this paper.

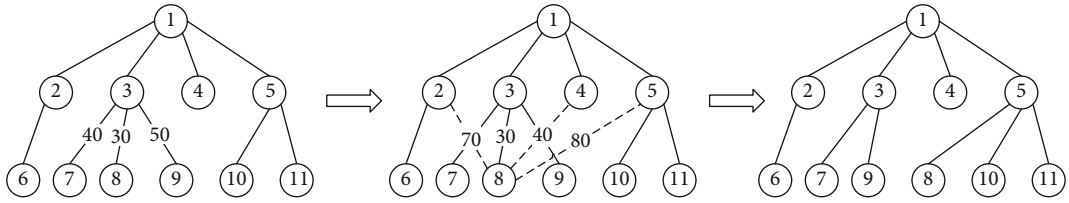


(a) Mutation operator 1 (b) Mutation operator 2

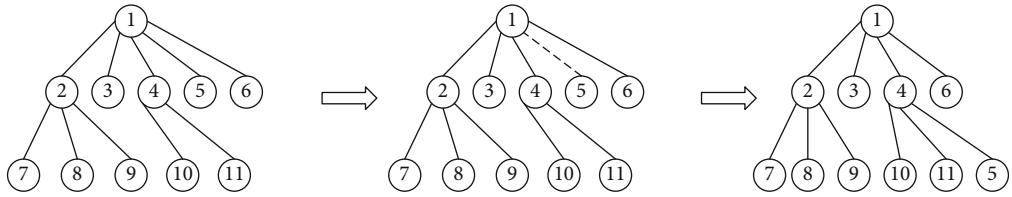
Figure 6: Example of the mutation operator.

This section simplifies the storage cluster topology based on the MSR code disaster recovery mechanism represented by (n, k, d) and uses a completely undirected connectivity weight graph $G = (V, E, W)$ to represent the topology structure of the distributed storage system, where $V = \{v_1, v_2, \cdots, v_n\}$ represents the collection of n nodes in the distributed storage system; $E = \{e_1, e_2, \cdots, e_m\}$, where $e_s(1 \le s \le m) = (v_i, v_j)$ represents the links that exist between nodes v_i and v_j in the topology; and $W = \{\omega(v_i, v_j) \mid (v_i, v_j \in E)\}$ represents the available bandwidth on the link and the processing capacity of the v_i node itself when $v_i = v_j$.

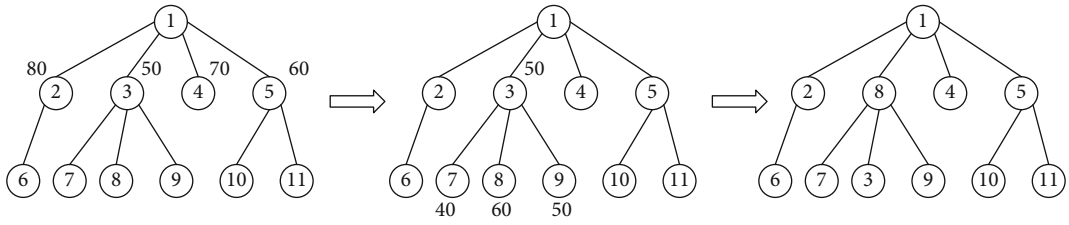
For the single-node failure scenario, to ensure the minimum repair delay and minimum repair flow at the same



(a) Local search for bandwidth



(b) Local search for traffic



(c) Local search for node processing power

Figure 7: Examples of the proposed local search operators.

time, a *Steiner tree* with constraints is constructed and satisfies

$$\min \quad \left\{ c_{\text{time}} \max \left[\alpha \frac{D_i}{\text{process}_i} + \frac{\beta}{\omega(V_i, V_j)} + \alpha * \frac{D_j}{\text{process}_j} \middle| (V_i, V_j) \in E \right] + \frac{c_{\text{traffic}}}{\sum_{\text{mid}-1}^{n-d-1} \text{degree}_{V_{\text{mid}}} - 2} \right\}$$

$$\text{s.t.} \quad \text{degree}_{\text{newcomer}} \geq d - k - 1. \tag{8}$$

The *Steiner tree* topology has a three-tier structure, as shown in Figure 3. At the top, the *newcomer* node is taken as the root, and the constraint in formula (1) is met to ensure that the amount of data required to restore the data block of size M/k is available. The intermediate layer consists of $n - d - 1$ intermediate nodes, which are responsible for the aggregation of the repair flow. The bottom layer contains d *provider* nodes connected to the intermediate node and the root node and provides the data needed to repair the failed node. The advantage of the three-layer structure is that it can avoid some weak links in the cluster topology by leveraging the advantage of the tree repair topology to reduce the repair delay. Second, the aggregation mechanism of the intermediate nodes can reduce the repair flow in the topology. In addition, the three-layer topology also limits the extra transfer overhead imposed by the intermediate nodes to the greatest extent so that the repair traffic in the whole topology can be reduced as much as possible. Finally, compared with the traditional tree topology, which cannot tolerate a secondary node failure, this paper designed a three-layer structure. It can avoid reconstructing the topology when the number of nonintermediate failure nodes is less than $n - d - 1$. The intermediate node can also be used as a substitute when a *provider* node fails; this allows the repair flow to complete the repair process. In exchange for a more robust topology, there is a small increase in the flow repair cost.

4. The Proposed Hybrid Genetic Algorithm

As shown in the model in Section 3.5, the goal of this paper is to improve the efficiency of repairing failed nodes in a distributed storage cluster topology with heterogeneous available bandwidth to increase the node processing performance and to reduce the repair traffic generated during the repair process as much as possible. After adding intermediate nodes to realize the aggregation mechanism of repair traffic, the topological problem of constructing a tree repair structure has been transformed from the traditional minimum spanning tree problem to the *Steiner tree* problem. However, under the repair scenario in this paper, the problem differs from the traditional *Steiner tree* problem in the following aspects. First, due to the nature of the MSR code, the degree constraint exists in the root node of the tree. Second, the root node of the tree has a degree constraint. Third, the *d provider* nodes used as leaves and the $n - d - 1$ nodes used as intermediate nodes are not fixed, and the nodes themselves have heterogeneous processing capabilities. These differences prevent us from using the traditional *Steiner tree* algorithm to solve this problem. In this paper, we design an optimal repair topology construction method based on a hybrid genetic algorithm.

In contrast to the classical mathematical approaches, stochastic optimization methods (such as an evolutionary

Encoding and decoding mode and population initialization
1: Assuming that the fault node number is r, add r to A and add else $n-1$ node to B, $s=0$
2: Choose $b \in B$, connect r and b as an edge of tree T, $A = A \cup \{b\}$ and $B = B - \{b\}$, $s = s+1$;
3: **if** $s < d-k+1$ **then** go to step 2
4: **else** go to step 6
5: **end if**
6: Take any element from set A and random element b from set B, join a and b as an edge and add it to tree T, and $B = B - \{b\}$, $k = k+1$, $s = s+1$;
7: **if** $s = n-1$ **then** the construction of individual T is completed;
8: **else** go to step 6.
9: **end if**

Algorithm 1

Crossover operator
1: Merge the intermediate node set $m\text{Node}_A(2,3,4,5)$ and $m\text{Node}_B(3,5,6,9)$ of T_A and T_B, obtain the intermediate node candidate $c\text{Node}(2,3,4,5,6,9)$, randomly select $d-k+1$ intermediate node for the intermediate node set $m\text{Node}_C(3,5,6,9)$ of the crossover descendant.
2: Execute Algorithm 1 based on the intermediate node set $m\text{Node}_C(3,4,5,6)$ to obtain the crossover subgeneration T_C.

Algorithm 2

Framework of ensemble learning for our system
1: p = random [0, 1].
2: **if** $p < 0.5$ **then** go to step 5
3: **else** go to step 7
4: **end if**
5: Function (mutation operator 1)
6: Select any leaf node and add it to another intermediate node or root node at random to obtain the mutation subtree, go to step 11
7: Function (mutation operator 2)
8: Select any leaf node and an intermediate node.
9: **if** links exist between the leaf node and all leaf nodes of the intermediate node **then** swap the positions of the two nodes in the repair topology to obtain the mutation subtree.
10: **end if**
11: EndFunction

Algorithm 3

algorithm) only utilize the objective function value with lower requirements for the continuity and differentiability of the objective function, which are suitable for the most complex engineering optimization problems [24–26]. As one of the well-developed stochastic optimization methods, a genetic algorithm [27–29] is adopted to design the solving algorithm to resolve our problem defined in Equation (8).

4.1. Encoding and Decoding Mode, Population Initialization, and Fitness Function. Since the subsequent crossover and mutation operations of a genetic algorithm are carried out based on population coding, the advantages and disadvantages of the population coding mode directly affect the search efficiency of the genetic algorithm. This section uses edge coding [30] and takes the edge set of the tree repair topology itself as the code, as shown in Figure 4. In a tree topology scene with a large number of constraints for the algorithm designed in this paper, edge coding can avoid a large number of infeasible solutions in this scene by adopting the crossover and mutation design, which is based on the characteristics of the tree topology and makes the execution of genetic operators more efficient.

According to the following, the tree repair topology constructed by the algorithm proposed in this paper should satisfy multiple constraints, including the constraint that all *provider* nodes are leaf nodes. The number of intermediate nodes is $n-d-1$, and the degree constraint condition is given in formula (1). Since the generation of an appropriate initial population is crucial to the genetic algorithm, this paper designs an initial population generation algorithm based on random methods and avoids infeasible solutions to ensure population diversity. Its pseudocode is as follows.

The fitness function is the main index for evaluating and describing the quality of an individual population, and it is

Local search operator
1: Function (bandwidth local search)
2: An intermediate node and its subtree are randomly selected from the parental topology to find the leaf node connected by the link with the least available bandwidth.
3: **if** there is more available bandwidth between the node and the other intermediate nodes **then** the node is connected to that intermediate node.
4: **end if**
5: EndFunction
6: Function (traffic local search).
7: **if** the parental repair topology meets the $\text{degree}_{\text{newcomer}} > d - k + 1$**then** randomly select a node from the leaf node directly connected to the root node and connect it to a random intermediate node $(d - k + 1 = 4)$;
8: EndFunction
9: Function (local search of node processing capacity)
10: A node with stronger processing capacity is randomly selected from the leaf node of the subtree to replace the worst processing capacity intermediate node, the positions of these two nodes in the topology are exchanged.
11: EndFunction

Algorithm 4

Local search operator
1: Randomly select individuals from the current population to join the next-generation population;
2: **if** the population of the next generation reaches pop $*$ 0.9**then** go to step 5
3: **else** go to step 1
4: **end if**
5: Arrange the current population individuals in ascending order according to fitness value, and select the former pop $*$ 0.1 individuals to inherit directly to the next-generation population.

Algorithm 5

also an important reference factor for the genetic algorithm to obtain high-quality solutions. The objective of the algorithm optimization problem in this paper is to improve the repair efficiency of the online repair process of the failure nodes while reducing the repair flow as much as possible. Therefore, in combination with the above mathematical model and formula (8), the fitness function is formulated as follows:

$$f = c_{\text{time}} \max \left[\alpha \frac{D_i}{\text{process}_i} + \frac{\beta}{\omega(V_i, V_j)} + \alpha * \frac{D_j}{\text{process}_j} \middle| (V_i, V_j) \in E \right] + \frac{c_{\text{traffic}}}{\sum_{\text{mid}-1}^{n-d-1} \text{degree}_{V_{\text{mid}}} - 2}. \tag{9}$$

The smaller the individual fitness function value is, the better the tree repair topology performance is.

4.2. Crossover Operator. The crossover operator ensures the diversity of the population based on inheriting the parental gene fragment to improve the searching ability of the algorithm. For the two trees T_A and T_B under the three-layer topology in this paper, the crossover operation is shown in Figure 5. The scheme is as follows.

4.3. Mutation Operator. The mutation operator maintains the search range of the algorithm in the individual neighborhood of the population. To ensure that every solution in the variation parental neighborhood is reachable, two different variation operators are designed and executed with the same probability.

Mutation operator 1 makes it possible for every leaf node to be connected with any intermediate node or root node, while mutation operator 2 makes it possible for every leaf node to be an intermediate node. Here, in combination with Figure 6, the mutation operator scheme is as follows.

4.4. Local Search Operator. Due to the large-scale random search mechanism of the genetic algorithm, the convergence speed of the algorithm is too slow. The algorithm in this paper designed some operators to perform a local search for three targets in the fitness function: bottleneck bandwidth, topology saving traffic, and node processing capacity, as shown in Figure 7. These three operators are executed in turn, and the pseudocode is as follows.

4.5. Select Operator. This paper adopts the method of random selection combined with elite retention for selection. The random selection operator allows every individual in the current population to inherit the next generation through a random selection of the current population, maintains the diversity of the population, and avoids premature convergence to the local optimal algorithm. Elite retention selects some of the best individuals in the current population and passes them directly to the next generation. Although some population diversity is lost, elite retention can guarantee the convergence of the whole algorithm. The combined selection

The optional topology construction method based on a hybrid genetic algorithm
1: Function (parameter initialization)
2: Initialize the population size parameter pop, maximum evolutionary generation G, and cross variation probability P_C, P_M;
3: Function (population initialization)
4: Repeat Algorithm 1 pop times to get the initial population $P(0)$ with the number of pop. Calculate the fitness value of each individual according to formula (9) and set $t = 0$.
5: Function (crossover operation)
6: Randomly select two individuals from the current population $P(t)$
7: **if** probability $P_C < 0.5$**then** perform Algorithm 2 to generate crossover descendants
8: **else** parental individuals directly join the crossover descendants
9: **end if**
10: Repeat the crossover descendants until the number of crossover descendants is equal to the pop, and the individual set of all these crossed individuals is denoted as a O_1;
11: Function (mutation operation)
12: From a random individual in the current population $P(t)$,
13: **if** probability $P_M < 0.5$**then** Algorithm 3 is executed to generate mutation progeny
14: **else** parental individuals are directly added to cross progeny
15: **end if**
16: Repeat the mutation operation until the number of descendants of the mutation is equal pop, denoted as O_2 in the topology.
17: Function (local search)
18: Perform Algorithm 4 successively for each individual in set $P(t) \cup O_1 \cup O_2$, perform a local search, and get three equally large subpopulations of L_1, L_2, and L_3.
19: Function (select operation)
20: Perform Algorithm 5 and set $P(t) \cup O_1 \cup O_2 \cup L_1 \cup L_2 \cup L_3$. Select individuals of pop as the next-generation population $P(t+1)$, and let $t = t + 1$;
21: EndFunction
22: **if** the generations = G**then** the individual with the lowest fitness function value is the solution
23: **else** go to step 3
24: **end if**

Algorithm 6

operator inherits the advantages of the two selection methods so that the algorithm does not fall into local optima based on good convergence. Assuming that the population size is pop, the specific steps are as follows.

4.6. Optimal Topology Construction Method Based on a Hybrid Genetic Algorithm. Based on the operators designed above, the whole genetic algorithm proposed is as follows.

5. Experiment and Evaluation

5.1. Experimental Environment. We used PyCharm 2019.1.3 and Python 3.7 to code the genetic algorithm simulation program on an Intel® Core™ i7-9750H @ 2.60 GHz CPU, 16 GB of memory, and 512 GB SSD hardware to conduct the contrast experiment. The experiment simulated the performance of the storage nodes in the cluster topology parameter selection for the I/O, CPU, memory, chips; the weighting of the corresponding parameters of 40%, 30%, 20%, and 10%; and the corresponding values range for [21,265], [1, 90], [0.3,76], and [0.5,23] [31]. The heterogeneity of the available bandwidth of the links in the cluster is governed by the parameters in the literature. Its distribution is obtained in PlanetLab [32], which is shown in Table 1. In the experiment, the size of the original data M was 1 GB, and the crossover probability and mutation probability of the algorithm in this paper were set as 0.9 and 0.1, respectively, after repeated experiments.

Table 1: Bandwidth distribution of the virtual storage nodes (conversion results and statistics).

Bandwidth C (KB/s)	The percentage (%)
$C < 30$	2.7
$30 \le C < 200$	11.3
$200 \le C < 500$	26.2
$500 \le C < 1000$	56.1
$1000 \le C$	3.7

5.2. Repair Efficiency. The optimal repair topology construction method based on the MSR code and the genetic algorithm is proposed in this paper.

The repair delay of all links traversing the *newcomer* direct connection is sorted, and the traditional star repair topology of the optimal solution is selected. The prim algorithm based on the minimum spanning tree algorithm PrimMST is used to construct the traditional tree repair topology with four MSR codes of different sizes, $(10, 6, 8)$, $(20,12,16)$, $(40,24,32)$, and $(80,48,64)$, considering the storage cluster scenario with heterogeneous node processing performance and link available bandwidth. The optimal repair delays of the three topologies were compared. The results are shown in Figure 8.

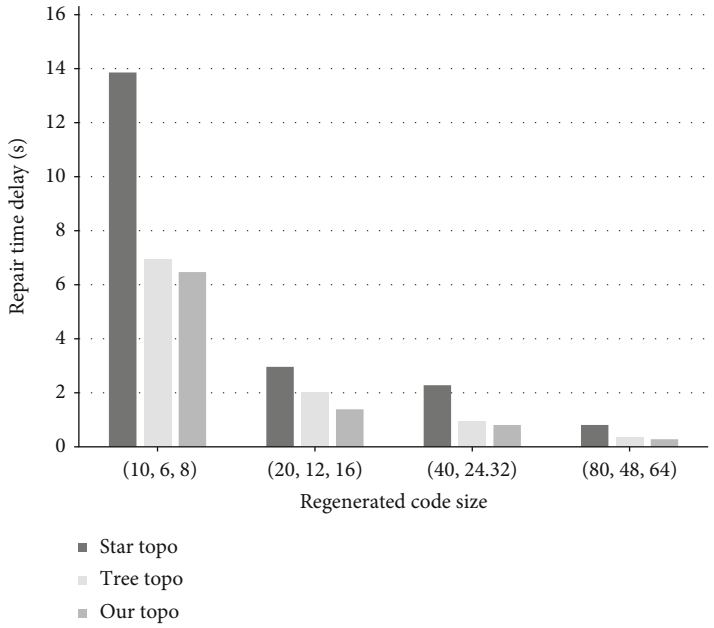


Figure 8: Optimal topology repair delay of each algorithm.

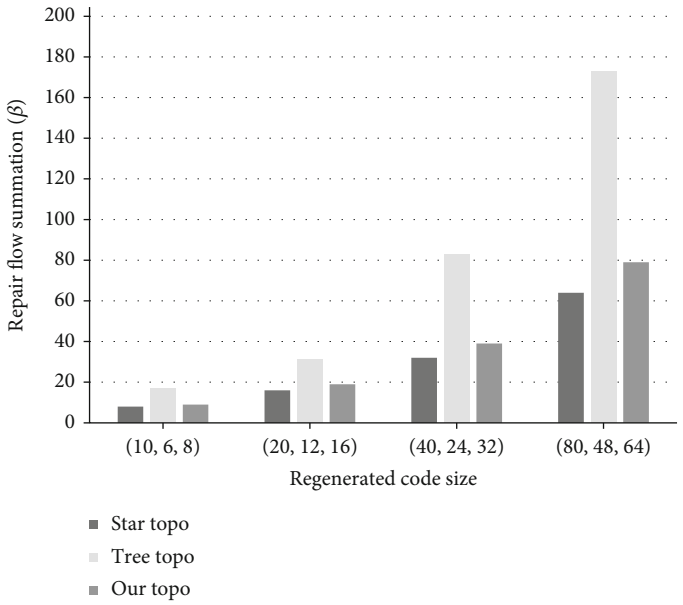


Figure 9: Optimal topology repair flow of each algorithm.

It can be seen that the star topology cannot avoid some links with poor available bandwidth in the selection of topology links, leading to a long repair delay. However, the PrimMST algorithm based on a greedy mechanism can no longer obtain the global optimal solution after the introduction of the processing capability of node heterogeneity. Under the same storage scale, the repair delay of the approximate global optimal topology constructed by the algorithm proposed in this paper is only 70%~90% and 35%~45% of the traditional tree and star repair topologies.

5.3. Repair Flow. As mentioned in Section 3.3 of this paper, the size of the repair traffic in the repair topology is also an

Table 2: Influence of four $(c_{\text{time}}, c_{\text{traffic}})$ values on MSR code repair delays of different sizes.

$(c_{\text{time}}, c_{\text{traffic}})$ Size of MSR	(0.25,0.75)	(0.5,0.5)	(0.75,0.25)	(1, 0)
(10, 6, 8)	6.47	6.47	6.47	6.47
(20,12,16)	1.39	1.39	1.39	1.39
(40,24,32)	0.82	0.82	*0.81*	*0.80*
(80,48,64)	*0.29*	*0.28*	*0.27*	*0.26*

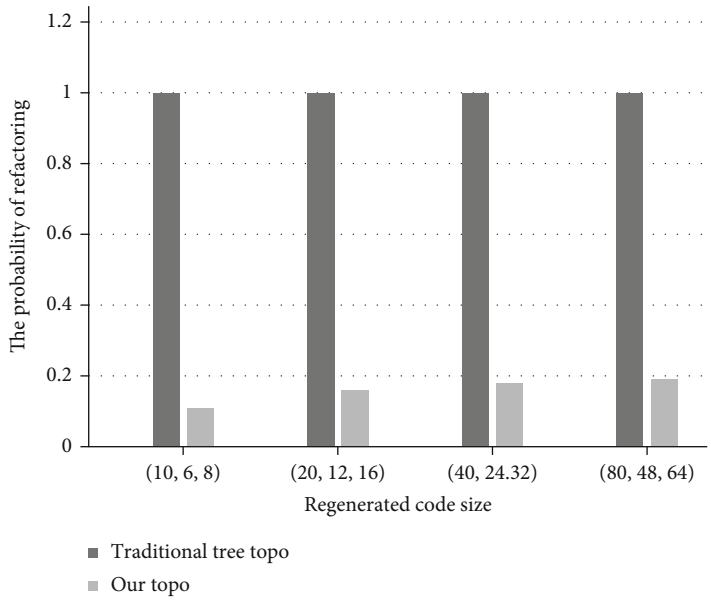


Figure 10: Probability of topology reconstruction of the single-node failure in the repair topology under four MSR codes.

important indicator for judging the merits of the repair topology. Taking the β-size traffic sent by each *provider* in the topology as a unit, we calculated the sum of the repair traffic passed on each link in each topology in the same experimental environment as in the previous section. The results are shown in Figure 9.

Although star topologies have the lowest repair traffic, the previous section has shown that their repair efficiency is too low for practical use. Combined with the experimental results in the previous section, it can be found that compared with the traditional tree repair topology construction method, the proposed method has a higher repair efficiency and the traffic in the topology is only 50% to 60% of that in the traditional tree topology. Although the extra path brought by the intermediate nodes increases the traffic overhead slightly compared with the star topology, the algorithm in this paper can use links with larger available bandwidth in the topology to transmit the repair traffic, so the repair efficiency is far higher than that of the star topology. Considering the two factors of repair efficiency and repair flow, the method in this paper surpasses the two traditional topological algorithms in terms of comprehensive performance, which is consistent with the original design goal.

5.4. Parameter Analysis. We adjust c_{time} and c_{traffic} to test and analyze the relationship between the repair rate and the repair flow in the topology during the overall repair process. The larger the c_{time} is, the greater the algorithm focuses on the repair rate, and the larger the value is, the less the repair delay of the algorithm construction topology is. When $c_{\text{time}} = 1$, the topology reduces the repair delay as much as possible at the cost of improving the repair traffic, and the topology repair delay is the shortest. However, the effect of adjusting the weight parameters on the whole topology structure can be almost ignored because of the small amount of traffic saved when the size of the MSR code is small. Table 2 shows how the repair rate of the topology constructed by the algorithm in this paper changes with the repair delay weight coefficient c_{time} and the repair flow weight coefficient c_{traffic} under different sizes of the regenerated code. When $(c_{\text{time}}, c_{\text{traffic}})$ is $(0.25,0.75)$ and $(0.5,0.5)$, the repair delays of the corresponding topologies in the scales of $(40,24,32)$, $(20,12,16)$, and $(10, 6, 8)$ do not change. When $(c_{\text{time}}, c_{\text{traffic}})$ is $(0.75,0.25)$ and $(1, 0)$, although the repair delays corresponding to the scales $(20,12,16)$ and $(10, 6, 8)$ remain unchanged, the repair delays of the topologies $(80,48,64)$ and $(40,24,32)$ are reduced to a certain extent, which is consistent with our analysis.

5.5. Robust Analysis of the Repair Topology. To demonstrate the robustness of the repair topology constructed by the proposed method, we also analyzed the probability that the repair topology in the MSR code storage cluster with scales of $(10, 6, 8)$, $(20,12,16)$, $(40,24,32)$, and $(80,48,64)$ is reconstructed when nodes in the topology fail during the repair process. The results of the proposed method and the traditional tree repair topology are compared. The results are shown in Figure 10.

Due to the nature of the MSR code itself, when a node loss occurs during repair, the repair traffic must be supplied from the nonprovider node, but the traditional tree topology only has two types of nodes, *provider* and *newcomer*; even with the spare nodes in the structural topology, the newly built repair topology does not have global optimality and there is no guarantee that a failed node can be repaired. Therefore, the reconstruction probability of the proposed method is always 100% regardless of the storage size. The intermediate nodes of the repair topology constructed by this method are also selected from n coding nodes, so they are qualified to be *providers*. When one of the *provider* nodes fails in the repair topology, it is only necessary to specify an intermediate node to make it a *provider*. Therefore, unless the intermediate node fails, there is no need to reconstruct

the repair topology, and the repair topology is more robust than the traditional tree repair topology.

6. Summary and Prospects

In the practical application of a distributed storage system, the heterogeneity of the processing capacity of each node and the influence of the repair traffic on the state of the cluster cannot be ignored in the process of repairing the data on the failed nodes. Aiming at reconstructing data in a single failure node scenario, this paper establishes a repair model and proposes an optimal repair topology construction method based on a genetic algorithm to simultaneously optimize the repair rate and flow of the repair topology. Experiments show that the proposed repair topology can improve the efficiency of node repair and can optimize the flow of repair topology, and the proposed topology is more robust than the traditional tree repair topology. The method proposed in this paper is only applicable to the repair scenario of single-node failure in a distributed storage system. The multimode scenario that considers optimization objectives is worth further research.

Data Availability

The heterogeneity of the available bandwidth of the links in the cluster is governed within the article (Ref. [27]). Its distribution is obtained in PlanetLab (Ref. [28]) which is shown in our paper (Table 1) to support the findings of our study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (Nos. 61662018 and 61861013), the Science and Technology Major Project of Guangxi (No. AA18118031), the Guangxi Natural Science Foundation of China (No. 2018GXNSFAA050028), the Doctoral Research Foundation of Guilin University of Electronic Science and Technology (No. UF19033Y), and the Director Fund Project of Key Laboratory of Cognitive Radio and Information Processing of Ministry of Education (No. CRKL190102).

References

[1] Wikipedia, "Distributed file system for cloud[EB/OL]," 2019, https://en.wikipedia.org/wiki/Distributed_file_system_for_cloud.

[2] N. Lakshmi, G. R. Bairavasundaram, S. Pasupathy, and J. Schindler, "An analysis of latent sector errors in disk drives," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 35no. 1, pp. 289–300, 2007.

[3] Y. Wang and S. Li, "Research and performance evaluation of data replication technology in distributed storage systems," *Computers & Mathematics with Applications*, vol. 51, no. 11, pp. 1625–1632, 2006.

[4] J. S. Plank, "Erasure codes for storage applications," in *Tutorial Slides Presented at FAST-2005: The 4th USENIX Conference on File and Storage Technologies*, San Francisco, USA, 2005.

[5] Z. Fengyan, W. Yan, and L. Nianshuang, "Survey of heterogeneous-based data repair strategies for erasure codes," *Application Research of Computers*, vol. 36, no. 8, pp. 2241–2249+2255, 2019.

[6] L. Rizzo, "On the feasibility of software FEC," Univ di Pisa, Italy, 1997.

[7] Z. Yao, C. Jiajia, and W. Chuliang, "Survey on data updating in erasure-coded storage systems," *Journal of Computer Research and Development*, vol. 57, no. 11, pp. 2419–2431, 2020.

[8] Q. Zheng, "Research on erasure code for secure storage system," Shanghai Jiao Tong University, Shanghai, 2009.

[9] X. Luo and S. Jiwu, "Summary of research for erasure code in storage system," *Journal of Computer Research and Development*, vol. 49, no. 1, pp. 1–11, 2012.

[10] S. Wang, "Study on fast repair of failed nodes in distributed storage systems," Chang'an University, 2019.

[11] X. Lin, Y. Wang, P. Xiaoqiang, F. Xu, and F. Yongquan, "GRC: a high fault-tolerance and low recovery-overhead erasure code for multiple losses," *Journal of Computer Research and Development*, vol. 51, Supplement 2, pp. 172–181, 2014.

[12] P. Teng, L. Chen, D. Yuan, and X. Wang, "Sparse random erasure code: a fault tolerance scheme for large-scale storage system," *Journal of Xi'an Jiaotong University*, vol. 51, no. 5, pp. 48–53, 2017.

[13] D. S. Papailiopoulos and A. G. Dimakis, "Locally repairable codes," *IEEE Transactions on Information Theory*, vol. 12, no. 5, pp. 102–117, 2014.

[14] A. G. Dimakis, K. Ramchandran, Yunnan Wu, and Changho Suh, "A survey on network codes for distributed storage," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 476–489, 2011.

[15] S. Jiekak, A. M. Kermarrec, N. le Scouarnec, G. Straub, and A. van Kempen, "Regenerating codes," *ACM SIGOPS Operating Systems Review*, vol. 47, no. 2, pp. 23–32, 2013.

[16] H. Jianzhong, *Design and Optimization of Erasure-Coded Clustered Storage Systems*, Science Press, Beijing, 2016.

[17] C. Jia, J. Wang, Y. Zhu et al., "On the optimal provider selection for repair in distributed storage system with network coding," *International Conference on Algorithms and Architectures for Parallel Processing*, 2015, pp. 506–520, Springer, Guangzhou, China, 2015.

[18] Q. Gong, J. Wang, D. Wei, J. Wang, and X. Wang, "Optimal node selection for data regeneration in heterogeneous distributed storage systems," in *I ICPP '15 Proceedings of the 2015 44th International Conference on Parallel Processing (ICPP)*, pp. 390–399, Washington DC, USA, 2015.

[19] X. Wan, "The research on optimization of topology sensitive repair technology in distributed storage system," Nanjing University, 2015.

[20] H. Zhang, H. Li, and S. Y. R. Li, "Repair tree: fast repair for single failure in erasure-coded distributed storage systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 6, pp. 1728–1739, 2017.

[21] Y. Wu, R. Dimakis, and K. Ramchandran, "Deterministic regenerating codes for distributed storage," in *Allerton Conference on Control, Computing, and Communication*, IEEE Press, 2007.

[22] J. Li, S. Yang, X. Wang, and B. Li, "Tree-structured data regeneration in distributed storage systems with regenerating

codes," in *2010 Proceedings IEEE INFOCOM*, pp. 1–9, San Diego, CA, USA, March 2010.

[23] Q. Fenglin, G. Qingyuan, Z. Yangfan, and X. Wang, "Heterogeneity-aware node selection for data repair in distributed storage systems," *Journal of Computer Research and Development*, vol. 52, Supplement 2, pp. 68–74, 2015.

[24] H. Liu, Y. Wang, and N. Fan, "A hybrid deep grouping algorithm for large scale global optimization," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 6, pp. 1112–1124, 2020.

[25] X. Xue, "A compact firefly algorithm for matching biomedical ontologies," *Knowledge and Information Systems*, vol. 62, no. 7, pp. 2855–2871, 2020.

[26] C.-H. Chen, "A cell probe-based method for vehicle speed estimation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E103-A, no. 1, pp. 265–267, 2020.

[27] Y. Wang, H. Liu, F. Wei, T. Zong, and X. Li, "Cooperative coevolution with formula-based variable grouping for large-scale global optimization," *Evolutionary Computation*, vol. 26, no. 4, pp. 569–596, 2018.

[28] X. Xue and J. Chen, "Using Compact Evolutionary Tabu Search algorithm for matching sensor ontologies," *Swarm and Evolutionary Computation*, vol. 48, pp. 25–30, 2019.

[29] C.-H. Chen, "An arrival time prediction method for bus system," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4231-4232, 2018.

[30] G. R. Raidl and B. A. Julstrom, "Edge sets: an effective evolutionary coding of spanning trees," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 3, pp. 225–239, 2003.

[31] S. J. Lee, P. Sharma, S. Banerjee, S. Basu, and R. Fonseca, "Measuring bandwidth between PlanetLab nodes," *International Conference on Passive & Active Network Measurement*, 2005, pp. 292–305, Springer, Boston, USA, 2005.

[32] "Planet-Lab a platform for developing, deploying, and accessing planetary-scale services [EB/OL]," https://www.planet-lab.org/.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5547044, 10 pages
https://doi.org/10.1155/2021/5547044

Research Article

A Study on the Influencing Factors of Mobile Digital Music Selection in Subhealthy People Based on Perceived Value Theory

Ming Jiang,[1] **Liming Lin**,[1] **Weiguo Zhang**,[1] **Weiqing Zhuang**,[1] **Minghui Jin**,[1] **and Yi Lin**[2]

[1]*Fujian University of Technology, Fuzhou 350118, China*
[2]*Fujian Normal University, Fuzhou 350117, China*

Correspondence should be addressed to Ming Jiang; 19842158@fjut.edu.cn

Received 28 January 2021; Revised 24 February 2021; Accepted 18 March 2021; Published 30 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Ming Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. The mode of combining mobile terminals and mobile digital music has gradually entered a bottleneck period while promoting the development of the mobile terminal industry. Proposing personalized solutions for specific groups of people including subhealthy people has become the current direction for the further long-term development of this mode. Understanding the influencing factors that affect subhealthy people's selection of mobile digital music is an urgent problem that needs to be solved in the development of personalized solutions. *Objective*. This article analyzes the influencing factors of the purchase intent to buy mobile digital music for subhealthy people and provides suggestions for mobile terminal vendors on how to design personalized solutions for this group of people. *Methods*. In order to achieve the above goals, this article constructs an influencing factor model based on the theory of perceived value, collects data by means of questionnaires, and uses the method of structural equations to verify the proposed model. *Results*. The results show that perceived quality, perceived price, social value, and emotional value all have a significant effect on users' purchase intention. Meanwhile, the conditional value and epistemic value have a negative moderating effect on the relationship between the social value and subhealthy people's digital music purchase intention but strengthen the positive relationship between the emotional value and subhealthy people's digital music purchase intention. *Conclusions*. The analysis shows that mobile terminal vendors should have their own unique attitudes and feelings when designing personalized mobile digital music solutions for subhealthy people, clarify their position, find strategies to improve the experience of subhealthy people, and win the reputation of subhealthy people. Let subhealthy people have a sense of belongingness.

1. Introduction

Since the advent of the first Walkman in the late 1970s, mobile music terminals have become an indispensable product in people's lives, especially, after Apple Inc. launched the iPod and iTunes store, a combination of mobile terminals and mobile digital music. The market has undergone a classic transformation of creative destruction, but with the advent of the digital age, this model is gradually facing development bottlenecks, and the personalized crowd customization model has increasingly become the mainstream of current development. Subhealthy people account for 85% of the global population, and music therapy is widely regarded as the most acceptable and effective method for this group of people. It is a matter of course that it has become a customized group of personalized solutions that is widely concerned by mobile terminal vendors. The analysis of the influencing factors affecting the purchase intention of mobile digital music for this group of people is not only conducive to enriching the theoretical research in the field of information system management but also conducive to mobile terminal vendors to formulate appropriate personalized customized solutions on this basis.

The Yi Guan Research Center defines mobile digital music as the general term for users to obtain music and music-related services by using SMS and other access methods through mobile terminal devices like smartphones. At present, although scholars have carried out research on consumers' mobile digital music purchase intentions from the perspectives of consumer characteristics, personal personality

characteristics, network-related factors, and cultural-related factors, the overall theoretical results are still scarce, especially for subhealthy people. Research on specific populations is still limited. Mobile terminal vendors also need more reference materials to continuously inspire diversified and innovative thinking, so that the industry can develop more prosperously.

Perceived value theory was first proposed by Zeithaml, and it is widely used in related research on perceived value and purchase intention. A large number of studies have shown that perceived value can better predict consumer purchase intentions than satisfaction or quality. Perceived value is defined here as the overall subjective evaluation of the utility of mobile digital music after weighing the benefits that the subhealthy people can perceive and the cost it pays for it.

We build on these studies by developing and testing a model that proposes how the different dimensions of perceived value interact to influence subhealthy people's digital music purchase intention. Specifically, we first examine the effects of four dimensions of perceived value (i.e., perceived price, perceived quality, social value, and emotional value) on subhealthy people's digital music purchase intention; then, we evaluate the moderating effects of conditional value and epistemic value on the relationship between above four perceived value dimensions and subhealthy people's digital music purchase intention.

We test our model using primary data collected from subhealthy people in four Chinese hospitals. Empirical results show that the perceived quality, emotional value, and social value have positive effects on subhealthy people's digital music purchase intention. Meanwhile, perceived price is negatively related to subhealthy people's digital music purchase intention. Furthermore, both relationships between four perceived value dimensions and subhealthy people's digital music purchase intention were moderated by the conditional value and epistemic value.

Our research extends extant literature about perceived value and mobile digital music selection by systematically examining the main effects and the interaction effects of different dimensions of perceived value on subhealthy people's digital music purchase intention. Firstly, we apply perceived value theory to the adoption of digital music, and this extends the scope of perceived value theory. Secondly, we complement the perceived value theory by testing the interaction effects of different dimensions of perceived value on subhealthy people's digital music purchase intention.

This article is divided into four parts. The first part is a review of related research domestically and internationally. The second part proposes the basic assumptions and influencing factor models of the research. The third part uses the structural equation method to verify the proposed model. Based on the empirical results, the fourth part provides specific suggestions for mobile terminal vendors to develop personalized solutions for subhealthy people.

2. Related Work

Mobile digital music technology and mobile terminal technology have changed the production, storage, and dissemination of the music industry value chain, giving birth to a variety of business models. As the music content is no longer monopolized by the record company, a search engine, music studio, original individuals, music websites, and so on may become a music provider. Digital music, digital albums, online concerts, and other music products enrich the industry day by day. Currently, consumers, including subhealthy people, have access to mobile digital music by three ways: free ad, mobile terminal store, and pay per single songs or albums.

2.1. Literature Review on Consumer Behavior of Digital Music. With mobile wireless music as an example, Souza and Baldanza confirmed the moderating role of customer experience on customer perception and purchase intention, and the results show that the wireless music experience can significantly enhance the positive impact of perceived usefulness and perceived ease of use on perceived value and attitudes of use and the positive impact among perceived value, the attitudes of use, and purchase intention [1].

Lin et al. studied the downloading behavior of digital music in Taiwanese college students and found that awareness of digital music copyright would positively affect the paid download behavior but had no significant effect on perceived usefulness; perceived ease of use had positive effects on perceived usefulness and downloading behavior; perceived usefulness positively affected use intention [2].

Kim et al. based on the "perceived value-intention" model studied the purchase intention of online consumers in Taiwan and found that perceived value had a significant effect on purchase intention for online music, and the effect of perceived usefulness and perceived entertainment was more than that of perceived price [3].

Shaw and Sergueeva added the advantages of virtual products, inferior variables (convenience and perceived risk), and digital music characteristics (price, diversity, audition, and entertainment) into the rational behavioral model and held the ground that subjective norms, perceived risk, diversity, price, and perceived entertainment significantly affect consumer buying behavior [4].

In summary, scholars have conducted a preliminary exploration of the behavior of digital music consumers, and the main research content is divided into two aspects: influencing factors on the purchase intention of consumers from the physical music products to virtual digital music products and the factors driving consumers from free music to paid music in the case of the prevalence of pirated music. In respect of theory, the majority of studies have used the theory of consumer behavior as a research foundation.

In order to find out the driving factors that subhealthy people choose to purchase mobile digital music from mobile terminal and the influencing factor from free to payment, the research mainly takes the difference between mobile digital music and digital music for modeling.

2.2. Relevant Study on Perceived Value Theory. According to the "value-intention" model, perceived value is a stable component to predict consumers' purchase intention. A large number of empirical results also show that the more

Table 1: Definition of Operability of each dimension variable of the perceived value.

Variable	Definition
Perceived price	The perceived price of money to digital music
Social value	The extent to which consumers believe that their purchase of digital music can improve self-concept and get the certainty of others
Emotional value	The extent to which consumers believe that the consumption of digital music pleases them and makes them positive
Perceived quality	The extent to which the quality of product or service is perceived by the consumer
Conditional value	The extent to which consumers believe that in the time and space conditions, digital music can convey the value
Epistemic value	The extent to which consumers believe that they can experience the sense of innovation or knowledge from the consumption

consumers perceive the benefits of goods, the stronger their purchase intention [5–7].

With regard to the notion of perceived value, Itani et al. firstly portrayed the perceived value as a trade-off between consumer perception of perceived quality (intrinsic and extrinsic) and perceived price (money, time, energy, and search) from the perspective of economics [8]. Then, Sheth et al., Woodruff et al., and Holbrook et al. believed that the customer value can be divided into a product attribute layer, result layer, and target layer from the marketing, psychology, and social psychology, as the value judgment of consumers interacts in a specific context based on individual preferences. So the formation of perceived value not only is attributed to product or service but also is affected by emotional factors and situational factors [9]. Therefore, the measurement of perceived value needs to be carried out from multiple dimensions [10]. In addition to the functionality, material, and economics, objective rationality also includes the symbol, pleasure, and aesthetic and other emotional content [11].

Based on the above research on the concept of perceived value, Sweeney and Soutar developed a perceived value measurement scale [12] which includes 19 measures. On this basis, Pura developed a perceptual value model for studying mobile services by adding conditional value (time sufficiency, location proximity, equipment conditions, etc.) and epistemic value (curiosity and novelty) [13–15].

Based on the theory of consumers' perceived value, this paper divides the perceived value into profits and losses and then divides the two dimensions into external factors and related contents.

At present, there are few researches applying perceived value theory to study the driving factors of digital content (digital music, digital video, and digital news) on consumption intention. Therefore, this paper uses perceptual value as an independent variable to study the influencing factors of mobile digital music on subhealthy people's purchase intention. In the current market environment, exploring the driving factor of subhealthy people buying mobile digital music is the exploration of the application of perceived value theory in new fields.

3. Conceptual Model and Hypotheses

In this paper, our research context is digital music purchase intention. So we redefine each dimension of perceived value in Table 1.

3.1. Main Effect

3.1.1. Perceived Price. A negative correlation between perceived price and purchase intention has been proposed by many empirical studies. In the study of factors affecting consumer adoption of mobile banking, Luarn and Lin confirmed that the perceived economic costs will have a significant negative influence on user adoption [16]. Kuo and Yen confirmed the negative impact of perceived economic costs on purchase intention in its study of the factors influencing the use of 3G mobile value-added services by consumers [17]. Yingquan, in his study of paid online learning, also confirmed the negative correlation between perceived price and purchase intention.

In a study of why consumers are generally low on the purchase intention in digital music, Punj noted that the free ad-supported model gave most consumers a "$0" reference price, i.e., "free psychology." When digital music has a charge, a sense of loss was produced from a rational psychological point of view [18].

Based on the above analysis, this paper proposes the following hypothesis:

H1: perceived price has a negative impact on subhealthy people's digital music purchase intention.

3.1.2. Perceived Quality. Perceived quality is the most fundamental source of perceived value in practical value and is regarded as an antecedent variable of perceived value. Kim et al., in their study of consumer willingness to adopt mobile networks, proposed perceptual quality and perceived usefulness and such concepts. Perceived usefulness is an external manifestation of the perceived quality of product, and its research also confirms the positive impact of perceived quality on adoption [19].

In summary, product quality is consumer awareness and evaluation of product gains; in addition to products and services themselves, it is also based on the shopping environment and the experience process. For digital music, the perceptive quality measurement includes the objective quality of music content, richness and availability of music libraries, and consumer experience on music sites, and software shall also be included.

Based on the above analysis, this paper proposes the following hypothesis:

H2: perceived quality has a positive effect on subhealthy people's digital music purchase intention.

3.1.3. Social Value and Purchase Intention. Social value explores social impact from the individual to obtain social appreciation. Music is a personalized hedonic product, and its consumption is more reflected as a self-expression; in consideration of weak music copyright awareness and free enjoyment, paid music is a way of expressing individuality, status, and self-image in social situations [20], and in this way, consumers obtain social appreciation.

To sum up, this paper proposes the following hypothesis:

H3: social value has a positive effect on subhealthy people's digital music purchase intention.

3.1.4. Emotional Value. Emotional value is derived from the pleasure, enjoyment, and interesting experience that consumers obtain from the product or the service. The impulsive consumption behavior by emotion dominating the purchase shows the influence of emotional value on the purchase intention [21].

Cai and Xu's researches confirmed that the aesthetic qualities and layout of the site design played an important role in enhancing the consumer's shopping experience, influencing consumer perception of shopping and perceived value in the shopping process [22]. Childers et al. [23], Chen et al. [24], Van der Heijden [25] and other scholars found that, in the Internet environment, the emotional value of the purchase will have a positive effect, especially in the enjoyment of the Internet for the purpose of product and service consumption; the emotional value of the purchase-driven role is more significant.

Digital music content service also happens to be a highly enjoyable product wherein consumers get the spirit of relaxation, pleasure, and emotional pursuit, rather than out of rational practical needs.

Based on the above analysis, this paper proposes the following hypothesis:

H4: emotional value has a positive effect on subhealthy people's digital music purchase intention.

3.2. Moderating Effect

3.2.1. Conditional Value. The positive influence of consumption context on individual behavior decision is defined by Sheth et al. as conditional value, which depends on the shopping environment and shopping conditions.

Pura (2005) divides perceived price dimensions into context-dependent and content-related categories, where conditional values are environment-dependent variables that describe the context in which consumers are buying and their current situation and state of mind. Sheth et al. in the definition of perceived value pointed out that under different circumstances of choice, the perceived value and its relationship with the purchase intention are different. For example, the Christmas tree as a symbol of Christmas, compared with weekdays, is of very high value in Christmas.

Therefore, decision-making environments influence the decision-making by interfering with or enhancing the context-dependent perceived value and the path of decision-making behavior. The influence of perceived value on purchase intention is dependent on the specific environment. In analyzing the relationship between the consumption of digital music and the influence of perceived value and decision-making behavior on the decision-making path, it is necessary to consider the role of conditional value. For example, when consumers want to get a specific song and cannot find a free download path, due to the urgency of demand, its perception of price and risk will be relegated to the second place.

Since previous studies have shown that conditional value will not affect the perceived price and consumers' digital music purchase intention, as well as the perceived quality and consumers' digital music purchase intention, this time it is not included in this research.

Based on the above analysis, this paper proposes the following hypothesis:

H5a: conditional value strengthens the negative relationship between the social value and subhealthy people's digital music purchase intention.

H5b: conditional value weakens the positive relationship between the emotional value and subhealthy people's digital music purchase intention.

3.2.2. Epistemic Value. Epistemic value refers to the value of novelty and satisfaction of trying new things and learning new ways of doing things. In the case of digital music consumption, epistemic value includes the curiosity of consumers for paid music services and satisfaction by testing the new paid services. The curiosity of new products, innovation, and the pursuit of diversity are the most important triggers of consumer purchase motivation [26].

Duman and Matilla's research shows that when consumers are seeking hedonic value, the pursuit of novelty is more motivating than purchasing behavior for the purpose of seeking practical value [27]. In the same period, the results of domestic scholars show that consumer innovation has a significant moderating effect on the negative impact of functional risk on a persistent information search.

Digital music is usually used as a kind of hedonic relaxation product, and it is in the stage of transition from new things to daily necessities in our country. Because the pursuit of novelty is a one-time process and does not have repetitive characteristics, this paper holds that epistemic value moderates the relationship between perceived price, perceived risk, and purchase intention through higher pursuit of novelty or new knowledge [28, 29].

Since previous studies have shown that epistemic value will not affect the perceived price and consumers' digital music purchase intention, as well as the perceived quality and consumers' digital music purchase intention, this time it is not included in this research.

Based on the above analysis, this paper makes the following hypothesis:

H6a: epistemic value strengthens the positive relationship between the social price and subhealthy people's digital music purchase intention.

H6b: epistemic value weakens the positive relationship between the emotional value and subhealthy people's digital music purchase intention.

3.3. Research Models. Based on the above hypothesis, the research model is built and shown in Figure 1.

4. Methodology

Questionnaires were distributed online, and 346 questionnaires were retrieved, of which 92 were filled out. We finally got 254 valid respondents with a total valid recycling rate of 73.41%.

4.1. Reliability and Validity

4.1.1. Reliability Analysis. We used Cronbach's α to analyze the reliability of each construct in the research model, and the results are shown in Table 2.

From Table 2, we can see that Cronbach's alpha of all our research constructs is greater than 0.7; these indicate that all our constructs have good reliability.

4.1.2. Validity Analysis. All measurement scales of our research constructs are based on extant top journal literature and have been cited several times, which always show good content validity.

We use AMOS to do confirmatory factor analysis to test construct validity, and the results are shown in Table 3. The results of first-order confirmatory factor analysis show that the default model is well fitting the data ($\chi^2 = 498.430$, DF = 209, CFI = 0.954, TLI = 0.945, and RMSEA = 0.074). In addition, all the factor loadings of the seven constructs were higher than 0.6, all of the average variance extracted was greater than 0.5, and all of composite reliability was greater than 0.7; these indicate that the main constructs of this study have good convergent validity. In order to test the discriminant validity, the correlation coefficient between independent variables should be less than the square root of the AVE value of two variables at the same time [28]. As can be seen from Table 4, the max correlation coefficient between these seven constructs is less than the square root of their average variance extracted. Therefore, all constructs in our research have good discriminant validity. In summary, all seven constructs in this study have good construct validity.

4.2. Empirical Model. The hypotheses involved in this study are tested by the following models. The variables in the model are expressed as follows: Purchase_intention is a dependent variable; Perceived_price, Perceived_quality, Social_value, and Emotional_value are independent variables; Conditional_value and Epistemic_value are independent variables; control variables include gender, age, education level, and income level; and ε is a random disturbance term.

Model 1: Purchase_intention = $\beta_{10} + \beta_{11}$Perceived_ price + β_{12}Perceived_quality + β_{13}Social_value + β_{14}Emotional_ value + control + ε_1.

Model 2: Purchase_intention = $\beta_{20} + \beta_{21}$Perceived_price + β_{22}Perceived_quality + β_{23}Social_value + β_{24}Emotional_ value + β_{25}Conditional_value + control + ε_2.

Model 3: Purchase_intention = $\beta_{30} + \beta_{31}$Perceived_price + β_{32}Perceived_quality + β_{33}Social_value + β_{34}Emotional_ value + β_{35}Conditional_value + β_{36}Perceived_price × Conditional_value + β_{37}Perceived_quality × Conditional_ value + β_{38}Social_value × Conditional_value + β_{39}Emotional _value × Conditional_value + control + ε_3.

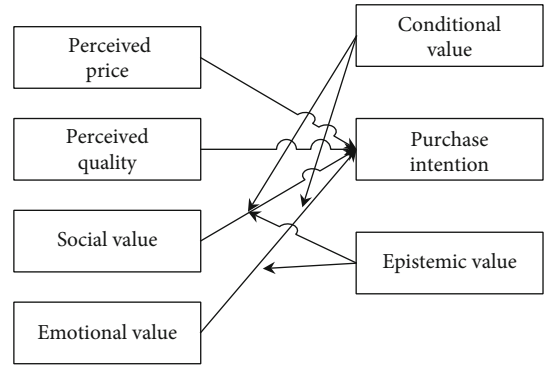


Figure 1: Conceptual model.

Model 4: Purchase_intention = $\beta_{40} + \beta_{41}$Perceived_price + β_{42}Perceived_quality + β_{43}Social_value + β_{44}Emotional_ value + β_{45}Epistemic_value + control + ε_4.

Model 5: Purchase_intention = $\beta_{50} + \beta_{51}$Perceived_price + β_{52}Perceived_quality + β_{53}Social_value + β_{54}Emotional_ value + β_{55}Epistemic_value + β_{56}Perceived_price × Epistemic _value + β_{57}Perceived_quality × Epistemic_value + β_{58}Social_ value × Epistemic_value + β_{59}Emotional_value × Epistemic_ value + control + ε_5.

Among all the models listed above, model 1 is used to test hypotheses H1, H2, H3, and H4. Models 2 and 3 are used to test hypotheses H5a, H5b, H5c, and H5d. Models 4 and 5 are used to test hypotheses H6a, H6b, H6c, and H6d.

4.3. Hypothesis Test. The regression results of models 1-5 are organized in Table 5.

4.3.1. Main Effect. In our study, hypothesis H1 suggests that perceived price has a positive impact on subhealthy people's digital music purchase intention. As we can see from the result of model 1 in Table 5, the regression coefficient of perceived price is $\beta_{11} = -0.052$ ($p > 0.10$). Thus, H1 has not received support.

Hypothesis H2 supposes that perceived quality is positively related to subhealthy people's digital music purchase intention. In model 1, the regression coefficient of perceived quality is $\beta_{12} = 0.153$ ($p < 0.01$). Therefore, H2 has received support.

Hypothesis H3 suggests that social value has a positive effect on subhealthy people's digital music purchase intention. In model 1, the regression coefficient of perceived quality is $\beta_{13} = 0.320$ ($p < 0.01$). Therefore, H3 is supported in this study.

Hypothesis H4 supposes that emotional value is positively related to subhealthy people's digital music purchase intention. In model 1, the regression coefficient of perceived quality is $\beta_{14} = 0.447$ ($p < 0.01$). Therefore, H4 has received support.

4.3.2. Moderating Effect. In this paper, the two moderator variables are conditional value and epistemic value. Both of

Table 2: Reliability analysis results.

	Perceived price	Perceived quality	Social value	Emotional value	Conditional value	Epistemic value	Purchase intention
Cronbach's α	0.852	0.925	0.968	0.943	0.860	0.920	0.933

Table 3: Validity analysis results.

	Perceived price	Perceived quality	Social value	Emotional value	Conditional value	Epistemic value	Purchase intention
Average variance extracted	0.663	0.757	0.908	0.850	0.689	0.739	0.820
Composite reliability	0.854	0.926	0.967	0.945	0.868	0.919	0.932

Table 4: Analysis of discriminant validity.

Variable	1	2	3	4	5	6	7
1. Perceived price	1						
2. Perceived quality	-0.044	1					
3. Social value	-0.047	0.646	1				
4. Emotional value	-0.113	0.697	0.916	1			
5. Conditional value	-0.056	0.751	0.736	0.852	1		
6. Epistemic value	-0.085	0.859	0.776	0.837	0.802	1	
7. Purchase intention	-0.132	0.702	0.879	0.907	0.849	0.821	1

Table 5: The results of linear regression analysis in this study.

Model name	Model 1	Model 2	Model 3	Model 4	Model 5
Gender	0.000 (0.105)	-0.004 (0.100)	0.030 (0.101)	-0.012 (0.104)	-0.018 (0.104)
Age	-0.085 (0.062)	-0.074 (0.059)	-0.064 (0.060)	-0.111 (0.063)	-0.099 (0.063)
Education	0.016 (0.058)	0.027 (0.056)	0.019 (0.055)	0.020 (0.058)	0.009 (0.057)
Income	0.055 (0.044)	0.047 (0.042)	0.028 (0.042)	0.070 (0.044)	0.052 (0.045)
Perceived price	-0.052 (0.037)	-0.065 (0.036)	-0.063 (0.036)	-0.050 (0.037)	-0.049 (0.037)
Perceived quality	0.153** (0.041)	0.059 (0.043)	0.043 (0.043)	0.087 (0.050)	0.108* (0.051)
Social value	0.320** (0.061)	0.336** (0.057)	0.438** (0.066)	0.301** (0.061)	0.351** (0.063)
Emotional value	0.447** (0.066)	0.300** (0.070)	0.221** (0.074)	0.402** (0.069)	0.360** (0.069)
Conditional value		0.272** (0.054)	0.279** (0.055)		
Perceived price × conditional value			-0.016 (0.023)		
Perceived quality × conditional value			0.017 (0.026)		
Social value × conditional value			-0.124** (0.041)		
Emotional value × conditional value			0.108* (0.046)		
Epistemic value				0.139* (0.061)	0.119 (0.061)
Perceived price × epistemic value					-0.011 (0.020)
Perceived quality × epistemic value					0.023 (0.025)
Social value × epistemic value					-0.113** (0.037)
Emotional value × epistemic value					0.090* (0.040)
N	254	254	254	254	254
F	103.78**	104.26**	75.21**	94.40**	67.87**
Adjusted R^2	0.765	0.784	0.773	0.769	0.775

Note: standard error of the regression coefficient in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$.

them are continuous variables, so the cross-product direct interaction method is adopted to test the interaction of moderating variables.

(1) Moderating Effect of Conditional Value. Hypothesis H5a proposes that conditional value weakens the positive relationship between the social value and subhealthy people's digital music purchase intention. As we can see from model 3 in Table 5, the regression coefficient of the interaction term between the conditional value and social value is $\beta_{38} = -0.124$ ($p < 0.01$). Thus, H5a has received support.

Hypothesis H5b suggests that conditional value strengths the positive relationship between the emotional value and subhealthy people's digital music purchase intention. In model 3, the regression coefficient of the interaction term between the conditional value and emotional value is $\beta_{39} = 0.108$ ($p < 0.05$). Therefore, H5b is supported in our study.

As we can see from Figure 2, when the conditional value is low, the relationship between the social value and subhealthy people's digital music purchase intention is positive, but when the conditional value is high, the relationship turns to be negative. Thus, the conditional value significantly weakens the relationship between the social value and subhealthy people's digital music purchase intention.

In Figure 3, when the conditional value is low, the relationship between the emotional value and subhealthy people's digital music purchase intention is positive but less; when the conditional value is high, the relationship turns to be more positive. Therefore, the conditional value significantly strengthens the relationship between the emotional value and subhealthy people's digital music purchase intention.

(2) Moderating Effect of Epistemic Value. Hypothesis H6a suggests that the epistemic value weakens the positive relationship between the social value and subhealthy people's digital music purchase intention. As we can see from model 5 in Table 5, the regression coefficient of the interaction term between the epistemic value and social value is $\beta_{58} = -0.113$ ($p < 0.01$). Therefore, H6a has received support.

Hypothesis H6b proposes that the epistemic value strengthens the positive relationship between the emotional value and subhealthy people's digital music purchase intention. In model 5, the regression coefficient of the interaction term between the epistemic value and emotional value is $\beta_{59} = 0.090$ ($p < 0.05$). Therefore, H5b is supported in this study.

As we can see from Figure 4, when the epistemic value is low, the relationship between the social value and subhealthy people's digital music purchase intention is positive, but when the epistemic value is high, such relationship turns to be negative. Thus, the epistemic value substantially weakens the relationship between the social value and subhealthy people's digital music purchase intention.

As shown in Figure 5, when the epistemic value is low, the relationship between the emotional value and subhealthy people's digital music purchase intention is positive but small; when the epistemic value is high, such positive relationship turns to be greater. Therefore, the epistemic value significantly strengthens the relationship between the emotional value and subhealthy people's digital music purchase intention.

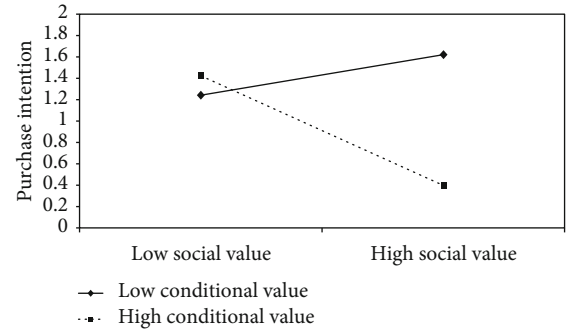


Figure 2: The interaction effect between the conditional value and social value.

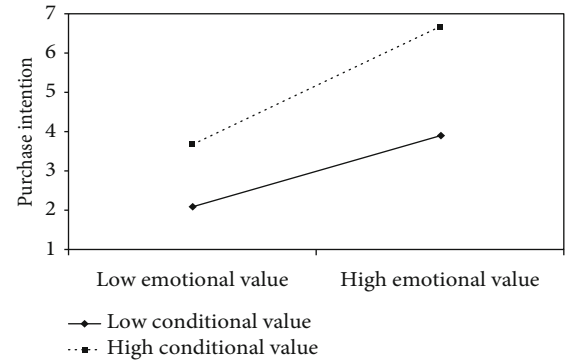


Figure 3: The interaction effect between the conditional value and emotional value.

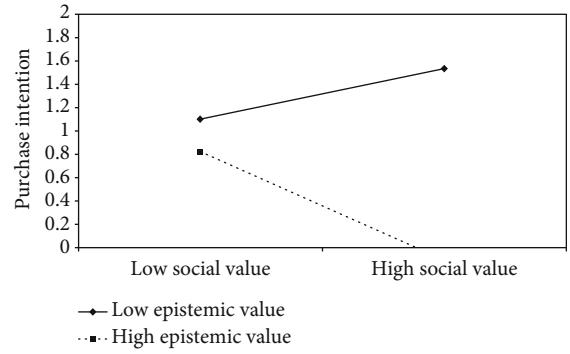


Figure 4: The interaction effect between the epistemic value and social value.

In summary, the moderating effects proposed in hypotheses H5a, H5b, H6a, and H6b all received support in our study. Specifically, both the conditional value and epistemic value moderate the relationship between the social value and subhealthy people's digital music purchase intention and the relationship between the emotional value and subhealthy people's digital music purchase intention.

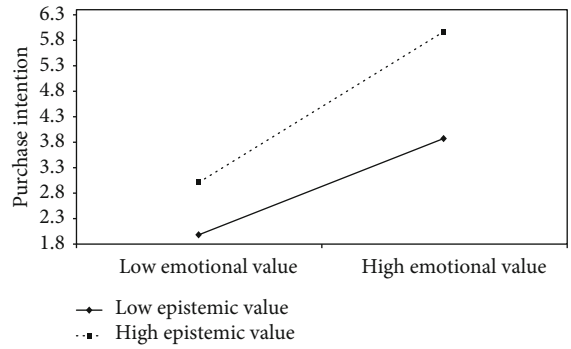


Figure 5: The interaction effect between the epistemic value and emotional value.

5. Discussion

5.1. Analysis Conclusion

(1) The effect of emotional value on purchase intention is the strongest among the perceived value dimensions. This result is consistent with the hedonic nature of mobile digital music; perceived quality has a significant positive effect on purchase intention, music quality, music library resources, mobile terminal design, and functional design, and other factors affect the user purchase intention. The quality of purchase intention also plays a significant positive effect. This conclusion is consistent with the made existing researches. Quality of digital music, music library resources, and interactive design of the mobile terminal and functional design are important factors affecting the subhealthy people's purchase intention

(2) Compared with the above significant conclusion, the effect of perceived price and purchase intention is not significant. The results are different from those of most scholars before, and the reason may be as follows: first, the reason that subhealthy people are reluctant to buy music is not caused by the absolute high price but based on the psychology of resisting pay for free product, or free music has met the demand; second, consumers believe that the price of mobile digital music is affordable compared to that of digital music; and third, paid mobile digital music is a virtual product of lower price, which does not help much for the promotion of social prestige and status, and no buying of mobile digital music can lead to ridicule. Revelation to the mobile terminal vendor is to improve consumer purchase intention which shall focus not only on lower prices but also more on improving quality strategies to enhance the user experience

(3) Social value has a significant impact on purchase intention. This article believes that the possible reasons are as follows: first, paid mobile digital music can show respect for others' labor, which is helpful to enhance social prestige and status and secondly, at present, Chinese people are concerned about digital music The gradual improvement of copyright knowledge does not mean that relevant laws are becoming more and more perfect. This behavior can be recognized by others

(4) Conditional value and epistemic value both have a negative moderating effect on the relationship between the social value and purchase intention. When subhealthy people perceive that paid music only produces a temporary sense of comfort in the context of its time and space, this value perception is not conducive to the formation of purchase intentions. If mobile terminal vendors can enable mobile digital music to bring more fresh experiences to subhealthy people and bring new knowledge or learning value, subhealthy people will be less sensitive to perceived loss. Mobile terminal vendors should create conditional value and increase the urgency of paid music, such as regularly launching prerelease versions of therapeutic music to increase the curiosity and entertainment of paid music services

(5) Conditional value and epistemic value have a positive moderating effect on the relationship between the emotional value and purchase intention. Subhealthy people can increase their satisfaction from music under the appropriate time and space background, which is conducive to the formation of purchase intention. If the mobile terminal vendor considers improving the details of the feelings of the subhealthy people when designing a personalized solution, this will help promote the purchase intention of the subhealthy people

5.2. Managerial Implication. Emotional value, social value, and perceived quality have a significant impact on purchase intention, and the mobile terminal vendor should be able to create emotional pleasure and good experience for subhealthy people.

(1) The mobile terminal vendor shall have unique attitude and feelings by clearing their own position, finding a strategy to improve the user experience to win reputation, and allowing users to feel a sense of belongingness, and in this way, they will willingly pay. The user interface shall be concise, clear, and easy to use, enhance the visual and layout logic, in line with the fragmentation characteristics of music products in modern times, and create use atmosphere. The use of comments is to enhance the user's emotional and adherence to content to improve listening experience; music library diversity and quality shall be enhanced by long-term cooperative relationship with record companies or musicians

(2) The perceived price is not a significant factor for the purchase intention; that is, consumers believe that the current pricing for digital music is not high, so the price reduction is not a good way to improve

purchase intention. However, the differentiation pricing strategy shows characteristics, such as an attempt to open up a pricing strategy and differential pricing; a combination of multiple payment models can improve user retention and loyalty and help attract potential customers

5.3. Limitations and Future Research. However, due to the objective conditions, there are some limitations and shortcomings in the study.

First of all, the geographical distribution of the questionnaire is relatively concentrated, and the respondents' academic record is higher than the overall sample, ignoring the possible impact of regional differences for an individual perceived value.

Secondly, most of the respondents in this study are subhealthy people who have already had mobile digital music purchase experience and cannot distinguish between prepurchase perception and postpurchase perception.

In the end, this paper does not explore the dimensions of antecedent variable of the perceived value. It only studies the influence of perceived quality on purchase intention and lack of variables which impact perceived quality, such as characteristics of the mobile terminal, smoothness of use, cross-platform service capability, and perceived ease of use. Therefore, the model is limited in practical application.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This article was supported by the Fujian Provincial Social Science Planning Project in 2020: Research on Capacity Building of Regional Medical Collaborative Innovation Community under the Background of "Internet + Medical" (No. FJ2020B038). The authors wish to thank Professor Peng LF and Ms. Wang Huan for their helpful comments on earlier drafts.

References

[1] D. Souza and R. Baldanza, "The e-consumer in light of the perceived value theory: a study on the acceptance of mobile commerce," *BASE-Revista de Administração e Contabilidade da Unisinos*, vol. 15, no. 3, pp. 238–253, 2018.

[2] K.-Y. Lin, Y.-T. Wang, and T. K. Huang, "Exploring the antecedents of mobile payment service usage," *Online Information Review*, vol. 44, no. 1, pp. 299–318, 2020.

[3] M. H. Kim, L. M. Marulis, J. K. Grammer, F. J. Morrison, and W. J. Gehring, "Motivational processes from expectancy-value theory are associated with variability in the error positivity in young children," *Journal of Experimental Child Psychology*, vol. 155, pp. 32–47, 2017.

[4] N. Shaw and K. Sergueeva, "The non-monetary benefits of mobile commerce: extending UTAUT2 with perceived value," *International Journal of Information Management*, vol. 45, pp. 44–55, 2018.

[5] Z. Zhang and H. Q. Wang, "How online social ties and product-related factors influence purchasing intention in mobile social commerce context," *International Journal of Mobile Communications*, vol. 17, no. 2, pp. 191–215, 2019.

[6] S. W. Joe, Y. H. Tsai, C. P. Lin, H. C. Ma, and C. K. Chiu, "Assessing perceived value: moderating effects of susceptibility to brand prestige and susceptibility to normative influence," *Review of Managerial Science*, vol. 11, no. 3, pp. 717–735, 2017.

[7] M. Choi, R. Law, and C. Y. Heo, "An investigation of the perceived value of shopping tourism," *Journal of Travel Research*, vol. 57, no. 7, pp. 962–980, 2018.

[8] O. S. Itani, A.-N. Kassar, and S. M. C. Loureiro, "Value get, value give: the relationships among perceived value, relationship quality, customer engagement, and value consciousness," *International Journal of Hospitality Management*, vol. 80, pp. 78–90, 2019.

[9] C. Lauckner, "The perceived value of online cancer resources among loved ones of people with cancer," *Oncology Nursing Forum*, vol. 45, no. 3, pp. 354–358, 2018.

[10] L. Aguiar, "Let the music play? Free streaming and its effects on digital music consumption," *Information Economics & Policy*, vol. 41, pp. 1–14, 2017.

[11] N. Xi, "Research on the design method of wireless music channel access network based on mobile Internet," *Revista De La Facultad De Ingenieria*, vol. 32, no. 3, pp. 740–748, 2017.

[12] F. J. F. Coelho, C. M. Bairrada, and A. F. M. Coelho, "Functional brand qualities and perceived value: the mediating role of brand experience and brand personality[J]," *Psychology Marketing*, vol. 4, pp. 41–55, 2020.

[13] X. Hu, "Evaluating mobile music services in China: an exploration in user experience," *Journal of Information Science*, vol. 45, no. 1, pp. 16–28, 2019.

[14] T. J. Gerpott and P. Meinert, "Not just every user of mobile music streaming shares the same characteristics: a classification analysis of mobile network operator subscribers in Germany," *Telematics Informatics*, vol. 41, pp. 19–33, 2019.

[15] S. McGrath and S. Love, "The user experience of mobile music making: an ethnographic exploration of music production and performance in practice," *Computers in Human Behavior*, vol. 72, pp. 233–245, 2017.

[16] M. Pura, "Linking perceived value and loyalty in location-based mobile services," *Managing Service Quality*, vol. 15, no. 15, pp. 509–538, 1991.

[17] Y. F. Kuo and S. N. Yen, "Towards an understanding of the behavioral intention to use 3G mobile value-added services," *Computers in Human Behavior*, vol. 25, no. 1, pp. 103–110, 2009.

[18] G. Punj, "The relationship between consumer characteristics and willingness to pay for general online content: implications for content providers considering subscription-based business models," *Marketing Letters*, vol. 26, no. 2, pp. 175–186, 2015.

[19] H. W. Kim, H. C. Chan, and S. Gupta, "Value-based adoption of mobile Internet: an empirical investigation," *Decision Support Systems*, vol. 43, no. 1, pp. 111–126, 2007.

[20] L. Leung and R. Wei, "More than just talk on the move: uses and gratifications of the cellular phone," *Journalism & Mass Communication Quarterly*, vol. 77, no. 2, pp. 308–320, 2000.

[21] M. P. Gardner, "Mood states and consumer behavior: a critical review," *Journal of Consumer Research*, vol. 12, no. 3, pp. 281–300, 1985.

[22] S. Cai and Y. Xu, "Designing not just for pleasure: effects of web site aesthetics on consumer shopping value," *International Journal of Electronic Commerce*, vol. 15, no. 4, pp. 159–188, 2014.

[23] T. L. Childers, C. L. Carr, J. Peck, and S. Carson, "Hedonic and utilitarian motivations for online retail shopping behavior," *Journal of Retailing*, vol. 77, no. 4, pp. 511–535, 2001.

[24] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Transactions on Information and Systems*, vol. E102.D, no. 7, pp. 1374–1383, 2019.

[25] H. Van der Heijden, "User acceptance of hedonic information systems," *MIS Quarterly*, vol. 28, no. 4, pp. 695–704, 2004.

[26] E. C. Hirschman, "Innovativeness, novelty seeking, and consumer creativity," *Journal of Consumer Research*, vol. 7, no. 3, pp. 283–295, 1980.

[27] T. Duman and A. S. Mattila, "The role of affective factors on perceived cruise vacation value," *Tourism Management*, vol. 26, no. 3, pp. 311–323, 2005.

[28] S. Onkar, "Digital rids a music academy from manual pain," *Dataquest*, vol. 35, no. 1, pp. 72–74, 2017.

[29] W. Dai and K. Yu, "Contestability in the digital music player market," *Journal of Industry Competition Trade*, vol. 19, no. 2, pp. 293–311, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5520009, 10 pages
https://doi.org/10.1155/2021/5520009

Research Article

Research on Music Multiterminal Audio Authentication Based on Wireless Network

Dongmei Li

Henan Polytechnic Institute, Nanyang, Henan 473000, China

Correspondence should be addressed to Dongmei Li; 2007003@hnpi.edu.cn

Received 27 January 2021; Revised 20 February 2021; Accepted 11 March 2021; Published 27 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Dongmei 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 current music multiterminal audio authentication algorithm does not consider the mutation of music signal, which leads to poor tamper detection ability and long time of audio authentication. By analyzing the characteristics and key technologies of wireless network, a wireless multiterminal audio system is established. The short-term energy calculation method is used to consider the sudden change of music signal. The music signal is divided into note segments, and chroma features of half order notes are extracted. The robust hash value is calculated by nonuniform quantization method. The dynamic time warping algorithm is used to align the notes, and the Hamming distance between the hash values of each two corresponding notes is calculated to obtain the measurement values of error series, statistical characteristics, and time distribution characteristics. According to the measurement value, the fuzzy classification method is applied to calculate the membership degree of the signals belonging to two different types of operation, and the authentication confidence degree is obtained. The tampered area of the music signal that has not passed the authentication is detected, and the music multiterminal audio authentication is realized. Experimental results show that the proposed algorithm has good tamper detection ability and can effectively shorten the audio authentication time.

1. Introduction

With the maturity of multimedia compression technology and the rapid popularization of Internet, the creation, storage, and transmission of multimedia digital works such as image, video, and audio have become extremely convenient [1]. The massive music information represented by MP3 has been widely spread on the Internet. With the wide use of modern audio processing tools, the processing of digital audio signal becomes very simple. However, it also means that the tampering of audio semantic information can be carried out at a lower cost [2]. For example, the semantics of audio signal may change fundamentally after a simple rearrangement or removal of a few small segments, which cannot be detected only by human auditory perception. Audio authentication technology is an effective technical means to protect the integrity and authenticity of music, voice, and other audio data [3]. It can ensure that the received audio data in the transmission process, without malicious editing and tampering by a third party, that is, in the sense of human perception system and the original audio, is exactly the same. The technology is widely used in many fields: government departments, national security, court defense, trade secrets, news, recorded speech, music recording and distribution, military, and so on.

In order to get secure multimedia applications, it is more important to protect and authenticate the authenticity and integrity of audio content. The ideal audio content authentication system should be able to accurately distinguish the content keeping operation and malicious tampering and can accurately locate the tampered area. At present, a large number of scholars have done research on a music multiterminal audio authentication algorithm and achieved some theoretical results. In reference [4], a dynamic authentication watermarking algorithm for IOT signals is proposed to detect network attacks. The watermark algorithm is based on the long-term and short-term memory structure of deep learning, which enables the Internet of things devices to extract a group of random features from the generated signals and embed these features into the signals dynamically. This

algorithm enables IOT gateway which collects signals from iotd to effectively verify the reliability of signals. In reference [5], a physical layer authentication scheme based on angle of arrival (AOA) estimation is proposed to cross verify the reported location information. Considering the multiantenna roadside unit under location deception attack, according to the Cramer Rao bound of AOA estimation and the existence of effective estimator, the basic limit of AOA estimation is given. The problem of determining whether the received signal is from the claimed GPS position is described as a two-sided hypothesis testing problem, and its solution is given by Wald test statics [6]. The closed form of correct decision probability (PD) and false alarm probability (PF) are given. The key to obtain reliable AOA measurement is the high accuracy of array response vector. But the above algorithm does not consider the signal mutation, resulting in poor tamper detection ability and long audio authentication time [7].

In order to solve the above problems, a music multiterminal audio authentication algorithm based on a wireless network is proposed. Using the characteristics and key technologies of wireless network, a wireless multiterminal audio system is established. Using the method of calculating short-term energy, considering the mutation of music signal, the music signal is segmented, chroma feature is extracted, and the robust hash value is calculated. The dynamic time warping algorithm is used to align the notes, the Hamming distance of the hash value is calculated, the error sequence is obtained, and the measurement index value is counted. The fuzzy classification method is applied to calculate the membership degree of the signal, and the authentication confidence is obtained. The tampered area is detected, and the music multiterminal audio authentication is realized. Experimental results show that the audio authentication time of the proposed algorithm is short, and it has good tamper detection ability [8].

The research contributions of the thesis include the following:

(1) By analyzing the characteristics and key technologies of wireless networks, a wireless multiterminal audio system is established

(2) The dynamic time warping algorithm is used to align the notes, and the Hamming distance between the hash values of each two corresponding notes is calculated to obtain the error sequence, the measurement value of the statistical feature, and the time distribution feature

(3) According to the measured value, the fuzzy classification method is used to calculate the membership degrees of the signals belonging to two different types of operations to obtain the authentication confidence. Detect the tampering area of the music signal that has not passed the authentication, and realize the music multiterminal audio authentication

The organization structure of the thesis is as follows. The second part discusses the related technologies of wireless network, the third part discusses the music multiterminal audio authentication algorithm, the fourth part conducts an experimental simulation, and the fifth part summarizes the paper.

2. Wireless Network Technology

2.1. Type of Wireless Network. The so-called wireless network refers to the network that can realize the interconnection of various communication devices without wiring. Wireless network technology covers a wide range, including global voice and data networks that allow users to establish long-distance wireless connections, as well as infrared and RF technologies that optimize short-range wireless connections [9]. According to the different network coverage, a wireless network can be divided into wireless wide area network (WWAN), wireless local area network (WLAN), wireless metropolitan area network (WMAN), wireless personal area network (WPAN), and wireless mesh network [10].

(1) WWAN: it mainly refers to the data communication through mobile communication, satellite, etc., with the largest coverage. Representative technologies include 3G, 4G, and 5G, and the general data transmission rate is above 2 Mb/s. As the standards of 3GPP and 3GPP2 are becoming more and more mature, some international standardization organizations are aiming at the next generation mobile communication system, which can provide higher wireless transmission rate and flexible and unified all IP network platform, generally known as 3G, enhanced IMT-2000, System Beyond IMT-2000, or 4G.

(2) WLAN: it is generally used for wireless communication between regions, and its coverage is small. The representative technology is IEEE802.11 series, which is also called WiFi network. The data transmission rate is 11~56 Mb/s, even higher.

(3) WMAN: it is a type of wireless network connecting several wireless LANs. Mobile data communication mainly through mobile phones or vehicle devices can cover most areas of the city [11]. The representative technology is IEEE802.20 series, which mainly studies the mobile broadband wireless access (MBWA) technology and the formulation of relevant standards. The standard emphasizes more on mobility, which is developed from the broadband wireless access (BBWA) of IEEE802.16.

(4) WPAN: at present, there are two wireless personal area network standards: wireless personal area network (WPAN, IEEE802.15.1) (Bluetooth) and low-speed wireless personal area network (LR-WPAN, IEEE802.15.4) (ZigBee). Bluetooth devices are generally battery devices with a coverage radius of 10 meters. ZigBee is more committed to ultralow power consumption networks. For example, for devices that can last about 10 years without changing the battery, ZigBee has a coverage radius of 50 meters [12].

(5) Wireless mesh network: wireless mesh network is a multihop ad hoc network, which is composed of mesh routers and mesh clients. Mesh routers constitute the backbone network and connect with a wired internet network, which is responsible for providing multihop wireless Internet connection for mesh clients. Wireless mesh network provides a wider network topology by storing and forwarding messages between AP. It can extend the existing wired or wireless network. Its biggest characteristic is that AP can not only act as an access point but also store and forward messages, playing the role of wireless router [13].

2.2. Characteristics of Wireless Network. Compared with a wired network, the main feature of wireless network is to completely eliminate the limitations of wired network and realize the wireless transmission of information. Specific wireless network features are as follows:

(1) Strong mobility: wireless network transmits network signals by transmitting radio waves. As long as it is within the range of transmission, it can use the corresponding receiving equipment to realize the connection to the corresponding network. This greatly gets rid of the limitation of space and time, which is beyond the traditional network.

(2) Strong expansibility: wireless network can be carried out by wireless signal at any time, and its network expansion performance is relatively strong, which can effectively realize the network expansion and configuration settings. Users will also become more efficient and convenient in accessing information. Wireless network not only expands the space range of people to use the network but also improves the use efficiency of the network [14].

(3) Low cost: generally speaking, the process of installing the wired network is more complicated. In addition to a large number of network cables and network cable connectors, the later maintenance cost of the wired network is very high [15]. The wireless network does not need to lay a large number of network cables and install a wireless network transmitting device. At the same time, it also creates a very convenient condition for the later network maintenance, which greatly reduces the cost of network installation and later maintenance.

2.3. Key Technologies of Wireless Network. Wireless network can effectively sense the changes of the external environment and then carry out deeper understanding and learning and effectively adjust and configure the relevant resources within the communication network, so as to meet the changes of the external environment. By fully learning from wireless cognitive network technology, it can not only solve the conflict between the growing demand of spectrum and limited spectrum resources but also effectively solve the problem of spectrum resource shortage and promote the reasonable improvement of spectrum application efficiency [16].

(1) Spectrum sharing: spectrum sharing can help users maximize the application probability of spectrum by managing interference items. Spectrum can be classified from different levels, according to different network architectures divided into component distribution and centralized. Centralization refers to the centralized processing of users' information by the central server and the distributed computing of cognitive terminals to determine the idle spectrum. Through different ways of spectrum allocation, it can be divided into cooperative and noncooperative. In the process of spectrum sharing, the filling sharing method is adopted, which can reduce the interference of primary users to the maximum while spectrum is idle.

(2) Spectrum sensing: in wireless network technology, spectrum sensing is one of the core technologies. This technology can provide valuable spectrum for the majority of users through spectrum hole, time domain, and frequency domain discovery. In essence, there are three kinds of signal detection methods that can detect the primary user autonomously, namely, cyclostationary feature detection, matched filter detection, and energy detection [17]. Among them, the detection of energy has good performance and easy operation, but it is easily restricted by objective factors, which makes the main signal difficult to identify. Detection-matched filter can effectively and quickly detect user information on the basis of clear user information, but in this process, many conditions need to be ensured, such as special receiver, frequency, and synchronous timing. The detection of cyclostationary feature can identify the noise energy and detect the main signal, but the calculation process is complex [18].

(3) Dynamic access: in wireless network technology, dynamic spectrum access technology can be divided into open sharing mode, multilayer access mode, and dynamic special application mode. Among them, in the dedicated dynamic mode, the primary user can completely control the spectrum and at the same time can choose the technology and service mode at will. Open sharing mode can share a variety of systems, and there is no interference between them. Compared with the above two modes, it is found that the multilayer access mode can completely get rid of the impact of this user's transmit power, which can not only effectively expand the scope of application but also further improve the information capacity and throughput [19].

2.4. Wireless Multiterminal Audio System. Wireless multiterminal audio system mainly includes three modules: control point (CP), digital media render (DMR), and digital media server (DMS). Among them, the control points are generally

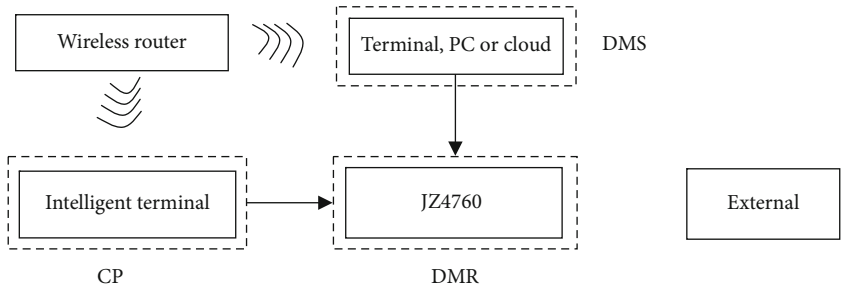


Figure 1: Connection diagram of each module of wireless multiterminal audio system.

mobile phones, tablet computers, and other intelligent terminals. DMR refers to the development board with wireless WiFi function. This paper uses Junzheng development board of MIPS processor. DMR can be either a home computer or an intelligent terminal. The connection of each module of wireless multiterminal audio system is shown in Figure 1.

It can be seen from Figure 1 that there is a logical relationship between the modules of the wireless multiterminal audio system and the wireless network plays a role in data output. The control point (mobile terminal) is first added to the same LAN with the wireless audio chip, and then the self-developed software is opened. The mobile terminal will find all the available devices on the LAN. Several devices can be selected to join the same group. After joining the group, the audio resources on the mobile terminal or other servers can be selected to play [20].

3. Music Multiterminal Audio Authentication Algorithm

3.1. Basic Framework of Music Multiterminal Audio Authentication Algorithm. A music multiterminal audio authentication algorithm is mainly divided into two stages: protection stage and authentication stage. The basic framework of the music multiterminal audio authentication algorithm is shown in Figure 2.

In the protection phase, firstly, an effective note segmentation algorithm is used to segment the music signal into a series of unequal length note segments. Then, based on each note segment, half order chroma features containing rich music semantic information are extracted. Finally, these feature vectors are transformed into binary hash authentication codes by a nonuniform quantization method, which are stored in a trusted second-order third-party certification center for future music certification. In the authentication stage, the music to be authenticated first needs to go through the same process of note segmentation, feature extraction and hash value calculation as in the protection stage to get the hash sequence of the music. Then, note alignment is carried out to reduce the different effects of note segmentation caused by various distortions in the transmission process [21]. Then, the h-value between the hash sequence of the current music segment and the authentication code is calculated in terms of notes Hamming distance to get an error sequence. Based on the error sequence, three new metrics reflecting hash difference, statistical characteristics, and time distribution characteristics are calculated. According to the three metrics, membership degrees of signals belonging to two different operations are calculated by using a fuzzy classification method, and the final "verification confidence" is obtained. Finally, for the music signals that have not passed the authentication, the system will also detect the tampered area testing [22].

3.2. Music Multiterminal Audio Authentication Algorithm Protection Stage

3.2.1. Music Segmentation. This paper adopts the method of calculating short-term energy [10] and, at the same time, considers the sudden change of the music signal in the high and low frequency parts for onset detection. The specific steps are as follows:

(1) The mute segments at the beginning and end of the music are removed, and the method of calculating short-term energy is used to detect the mute frame

(2) The music is decomposed in each frequency band as Table 1, and five subband signals are obtained

(3) The signal on each subband is divided into frames with a length of 60 ms, and there is a 50% overlap between two adjacent frames. For the medium and high frequency part of the signal (subband 2~5), use the energy difference to define the onset detection function:

$$E_i(n) = \sum_{m=(n-1)h+1}^{nh} |x_i(m)|^2, \quad i = 2, 3, 4, 5, \tag{1}$$

$$\mathrm{ons}_i(n) = E_i(n) - \sum_{a=1}^{10} \frac{E_i(n-a)}{a}, \quad i = 2, 3, 4, 5, \tag{2}$$

in which $x_i(m)$ represents the amplitude value of the i subband signal at time m, $E_i(n)$ represents the n frame energy of the i subband signal, h represents the frame length, and $\mathrm{ons}_i(n)$ represents the onset detection function of the i subband signal. Formula (2) uses the index in the method of differential weighting; the closer the frame to the current time is, the higher the weight [23].

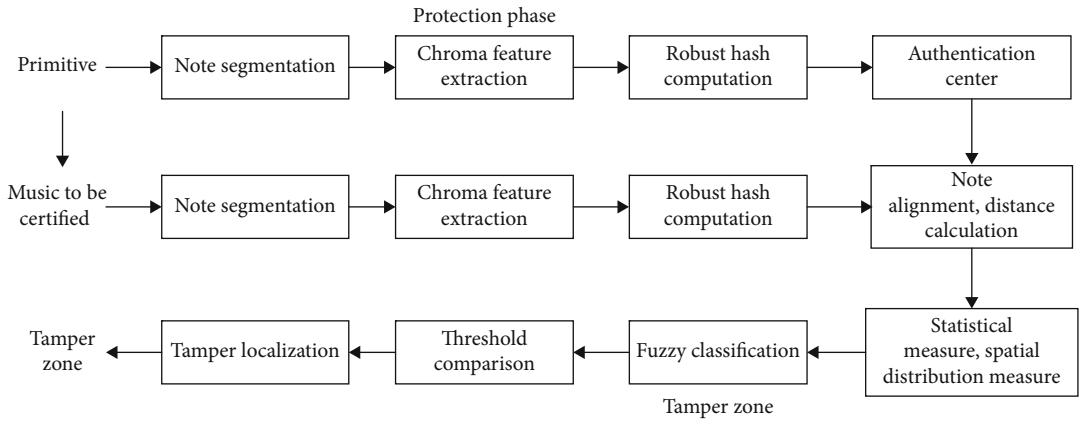


Figure 2: Basic framework of music multiterminal audio authentication algorithm.

Table 1: Frequency range of each subband during subband decomposition.

Subband No.	Freq range (Hz)
1	0~1024
2	1024~2048
3	2048~4096
4	4096~8192
5	8192~22050

For the low frequency part of the signal (subband 1), because the energy change is not obvious, the detection function is defined by using the change of the spectral coefficient [11]:

$$dX_n(k) = X(k, nh) - X(k, (n-1)h),$$
$$\text{ons}_1(n) = \frac{\sum_{k:DX_n(k)>0} dX_n(k)^2}{\sum_{k=1}^{N/2} |X(k, (n-1)h)|^2}, \quad (3)$$

in which $X(k, nh)$ represents the k FFT coefficient of the n frame of the first subband signal and N represents the Fourier transform length

(4) Perform a weighted summation of the detection functions on each subband to obtain the total onset detection function:

$$\text{ons}(n) = \sum_{i=1}^{5} \omega_{0i} \text{ons}_i(n), \quad (4)$$

in which $\omega_0 = [\omega_{01}, \omega_{02}, \omega_{03}, \omega_{04}, \omega_{05}]$ is the weighting coefficient

(5) Find the peak point of the detection function $\text{ons}(n)$ to determine the position of the note onset in the music

According to the above detection results, a music segment roughly corresponding to a note length is formed between two onsets [24].

3.2.2. Chroma Feature Extraction. For a music content authentication system, it is very important to select appropriate features that can fully reflect the semantic information of music. In this paper, we use chroma features, which are widely used in music retrieval, music structure analysis, and other fields [12]. The feature projects the whole spectrum distribution of music signal to 12 half order notes in a complete octave range, which has rich information of scale distribution and melody trend. Therefore, chroma is a 12-dimensional eigenvector, which is calculated frame by frame according to the following formula:

$$\text{Chroma}\left(K', n\right) = \sum_{K:P(K)=K'} X(K, n),$$
$$P(K) = \left[12 \cdot \log 2\left(\frac{K}{\text{NFFT}} \times \frac{f_s}{f_1}\right)\right] \bmod 12, \quad (5)$$

in which $\text{Chroma}(K', n)$ represents the K' dimension data of the chroma feature vector of the signal $x(n)$, $X(K, n)$ represents the K FFT coefficient of the signal $x(n)$, NFFT is the Fourier transform length, f_s represents the sampling frequency, and f_1 is the reference frequency [25].

In the specific implementation, first of all, each segment based on note segmentation is divided into fixed length frames; the frame length is 256. Then, chroma feature is extracted from each frame, and the mean value of chroma feature of each frame in each note segment is taken as the feature of the whole note. So far, each note segment of music signal is represented by a 12-dimensional feature vector.

3.2.3. Robust Hash Computation. Finally, the extracted Chroma features need to be quantized to generate a 36-bit authentication code for each note segment. The specific steps are as follows:

(1) The chroma vector was normalized:

$$p_o(i, j) = \frac{p(i, j)}{\sqrt{\sum_{k=1}^{12} p^2(i, k)}}, \quad (6)$$

in which $p(i,j)$ represents the j component of the feature corresponding to the i note and $p_o(i,j)$ represents the j component of the corresponding normalized feature vector. After the normalization operation, the value of each dimension component is between 0 and 1

(2) The normalized chroma features were quantized nonuniformly:

$$\widehat{p}(i,j) = \begin{cases} \text{floor}(p_o(i,j) \times 10, 0 \le p_o(i,j) < 0.7), \\ 7, 0.7 \le p_o(i,j) \le 1, \quad j = 1, 2, \cdots, 12. \end{cases} \tag{7}$$

In formula (7), $\widehat{p}(i,j)$ represents the j component of the quantized vector, and $\text{floor}(x)$ returns the largest integer not greater than x

(3) Express each $\widehat{p}(i,j)$ with the corresponding 3-bit binary, namely, $\widehat{p}(i,j) = [b_2 b_1 b_0]_2$, and connect all the binary bits to form a 36-bitHash code $h(i)$

Quantization of music features can not only reduce the storage space occupied by authentication information but also improve the robustness of various signal processing. The original music hash sequence is stored in a trusted third-party certification center for future music certification.

3.3. Music Multiterminal Audio Authentication Algorithm Authentication Stage

3.3.1. Note Alignment. In this paper, a dynamic time warping algorithm is used to obtain the most reasonable correspondence between two note sequences. The distance between every two notes is defined as the Hamming distance of their corresponding Hash value:

$$D(i,j) = \|h_o(i) - h_1(j)\| = \frac{1}{36}\sum_{k=1}^{36}[h_o(i,k) \oplus h_1(j,k)], \tag{8}$$

in which $D(i,j)$ represents the Hamming distance between the i note of the original music and the j note of the music to be authenticated, $h_o(i)$ is the Hash code of the i note of the original music, and $h_1(j)$ is the j note of the music to be authenticated Hash code. On the basis of note alignment, the Hamming distance of each two corresponding note Hash values is calculated, and the content integrity determination is further performed according to the sequence of these distance values.

3.3.2. Measurement Index. On the basis of note alignment, the Hamming distance between the hash values of every two corresponding notes is calculated, and all the distance values form a sequence diff. Define the possible tampering points in diff as those points whose distance value exceeds the set threshold T, denoted as PMP, whose index value is stored in the vector POS, as shown in

$$\text{PMP} = \left\{ \text{diff}(i) \middle| \text{diff}(i) \ge T, 1 \le i \le N' \right\},$$

$$\text{POS} = \{\text{pos}(j) | \text{diff}(\text{pos}(j)) = \text{PMP}(j), \quad j = 1, 2, \cdots, \|\text{PMP}\|\},$$

$$T = \begin{cases} \max(PMP) \times 0.5, \max(PMP) \ge T_0, \\ \text{median}(PMP), \max(PMP) \ge T_0, \end{cases} \tag{9}$$

in which $\text{diff}(i)$ represents the distance value corresponding to the i note, T is an adaptive threshold, and its setting takes into account the value of diff under acceptable operations and malicious tampering. The threshold T_0 is used to roughly judge whether the signal has serious distortion. Based on the above concepts, three measurement indicators reflecting the characteristics of diff sequence statistics and time distribution are defined. They have strong distinguishing ability for maintaining content operations and malicious tampering.

(1) Average distortion (AD): the average distortion AD of a signal to be authenticated is defined as the average bit error rate of all PMP points:

$$\text{AD} = \frac{1}{L_{\text{pmp}}}\sum_{j=1}^{L_{\text{pmp}}} \text{PMP}(j), \tag{10}$$

in which $L_{\text{pmp}} = \|\text{PMP}\|$. The average amount of distortion reverses the degree of change in the music content. Obviously, malicious tampering usually has a larger AD value compared to allowable operations.

(2) Uniform measure (UD): the uniformity metric is aimed at reflecting the uniformity of the error distribution. Let $\text{DIS} = \{\text{dis}(j) | \text{dis}(j) = \text{pos}(j+1) - \text{pos}(j), j = 1, 2, \cdots, \|\text{PMP}\| - 1\}$ denote the number of notes between adjacent PMP points, and define the uniformity metric as the standard deviation of DIS:

$$\text{UD} = \left[\frac{1}{N_{\text{dis}}}\sum_{j=1}^{N_{\text{dis}}} \left(\text{dis}(j) - \frac{1}{N_{\text{dis}}}\sum_{j=1}^{N_{\text{dis}}} \text{dis}(j) \right)^2 \right]^{1/2}, \tag{11}$$

in which $N_{\text{dis}} = \|\text{DIS}\|$. Obviously, the distribution of PMP points is more uneven by malicious tampering, the corresponding UD value is also larger, and the UD value corresponding to the allowable operation is smaller.

(3) Maximum connected area size (MC): a group of continuous dense points in the diff sequence form a connected area, and the size of the connected area is defined as the number of all points in the area. MC refers to the number of points contained in the largest connected area. Generally speaking, the MC value of malicious tampering is much larger than the allowable operation. This is because under malicious

tampering, the error tends to be very tightly concentrated in some local areas, while for the allowable operation, the error is often scattered on the entire timeline.

3.3.3. Music Content Certification. In the fuzzy classification method, firstly, it is necessary to define the membership functions for these three indexes, respectively, and describe the degree of their values to the two fuzzy sets of "large" and "small." Define the membership function of the average distortion AD as follows:

$$X_{As}(\mathrm{AD}) = \begin{cases} 1, & 0 \le \mathrm{AD} \le th_1, \\ \dfrac{1}{th_1 - th_2}(\mathrm{AD} - th_2), & th_1 < \mathrm{AD} \le th_2, \\ 0, & \mathrm{AD} > th_2, \end{cases}$$
$$X_{Al}(\mathrm{AD}) = \begin{cases} 0, & 0 \le \mathrm{AD} \le th_1, \\ \dfrac{1}{th_2 - th_1}(\mathrm{AD} - th_1), & th_1 < \mathrm{AD} \le th_2, \\ 1, & \mathrm{AD} > th_2, \end{cases} \tag{12}$$

in which $X_{As}(\mathrm{AD})$ and $X_{Al}(\mathrm{AD})$, respectively, represent the membership degrees of the smaller AD value and th_1 and th_2 are the two thresholds discussed above. Define the membership function of the uniform metric UD as

$$X_{Us}(\mathrm{UD}) = 1 - \frac{1}{1 + e - \alpha(\mathrm{UD} - \beta)},$$
$$X_{Ul}(\mathrm{UD}) = \frac{1}{1 + e - \alpha(\mathrm{UD} - \beta)}, \tag{13}$$

in which $X_{Us}(\mathrm{UD})$ and $X_{Ul}(\mathrm{UD})$, respectively, represent the degree of membership for which the value of UD is small and large. The β value of the parameter is the mean value of the music signal UD calculated under a series of content preservation and malicious tampering, and the parameter α control the entire S curve, especially the change speed at the sudden change point $\mathrm{UD} = \beta$. In this article, α and β are set to 25 and 0.08, respectively. The membership function that defines the maximum connected region size MC is as follows:

$$X_{Ms}(\mathrm{MC}) = \begin{cases} 1, & \mathrm{MC} \le \mu_1, \\ e^{-\left((\mathrm{MC}-\mu_1)^2\right)/2\sigma^2} & \mathrm{MC} > \mu_1, \end{cases}$$
$$X_{Ml}(\mathrm{MC}) = \begin{cases} 1, & \mathrm{MC} \ge \mu_2, \\ e^{-\left((\mathrm{MC}-\mu_2)^2\right)/2\sigma^2} & \mathrm{MC} < \mu_2, \end{cases} \tag{14}$$

in which $X_{Ms}(\mathrm{MC})$ and $X_{Ml}(\mathrm{MC})$, respectively, represent the membership degrees of the smaller MC value and the parameters μ_1 and μ_2 are set to $\mu_1 = 5$, $\mu_2 = 105$, respectively. Given a set of metric values $m = [\mathrm{AD}, \mathrm{UD}, \mathrm{MC}]$, its membership degree of fuzzy class C_i is expressed as follows:

$$X_{C_i}(m) = \sum_{j=1}^{3} w_j X_{C_{ij}}(m_j), \quad i = 1, 2, \cdots, 8, \tag{15}$$

in which $X_{C_i}(m)$ represents the degree of membership of m belonging to class iC_i, $X_{C_{ij}}(m_j)$ represents the degree of membership of m_j belonging to class C_{ij}, and the weight vector w_j reflects the importance of each index for content integrity authentication, which is determined by experiments:

$$D_y = \sum_{i=1}^{8} w_{yi} X_{C_i}(m),$$
$$D_n = \sum_{i=1}^{8} w_{ni} X_{C_i}(m), \tag{16}$$

in which D_y and D_n represent the final authenticity measure and nonauthenticity measure of the music signal, respectively, and w_{yi} and w_{ni} represent the weight contributions of various types of authentication passing and failing, respectively. Finally, the authentication credibility measure $authRatio = D_y/D_n$ is defined. If $authRatio > 1$, the music content is more likely to be authentic, and its authenticity increases with the increase of $authRatio$ value. If $authRatio < 1$, the music content is more likely to have been maliciously tampered with, and the possibility of tampering increases as the value of $authRatio$ decreases. If $authRatio = 1$, the system cannot make a decision.

3.3.4. Tamper Detection. For nonauthentic signals that fail to pass the music content authentication, all connected areas whose amplitude values are not less than a given threshold T are marked in the diff sequence. These areas correspond to the tampered part of the test signal. In this paper, the connected area is defined by the concept of 8 neighborhoods; therefore, the error of tamper detection is about 4 notes in the worst case (generally 1~2 s). Through the above steps, the music multiterminal audio authentication is completed.

4. Experimental Analysis

4.1. Experimental Environment and Data. In order to verify the effectiveness of the multiterminal audio authentication algorithm for music based on wireless networks, this article uses the voice set of the TIMIT voice library, 1280 segments of 4 s speech, of which there are 600 segments of English and 680 segments of Chinese, and the used speech parameters are as follows: sampling rate of 16000 Hz, bit rate of 256 kbps, single channel number audio channel, sampling precision of 16 bits, WAV format, frame length of 20 ms, and frame shift of 10 ms. The experimental hardware platform is Inter Core i3 processor, 2450 M, 8.00G memory, and 800G hard disk, and the experimental environment is Matlab R2012b under Windows7 operating system.

4.2. Tamper Detection and Analysis. When the music is judged as authentication failure, it is often necessary to detect the tampered area. Taking a classical music "the Blue

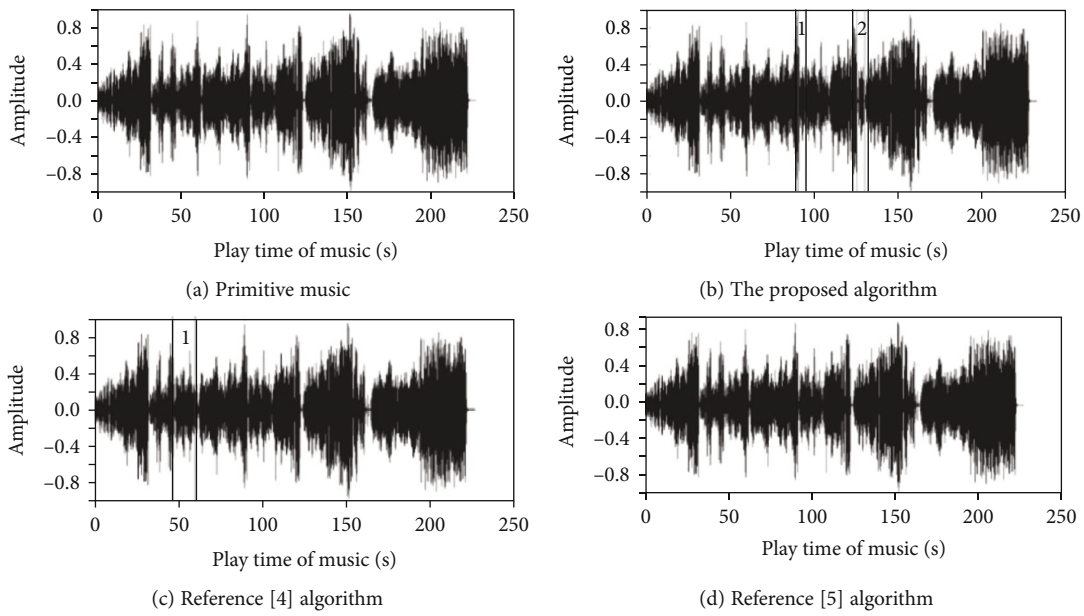


Figure 3: Comparison of detection results of different algorithms.

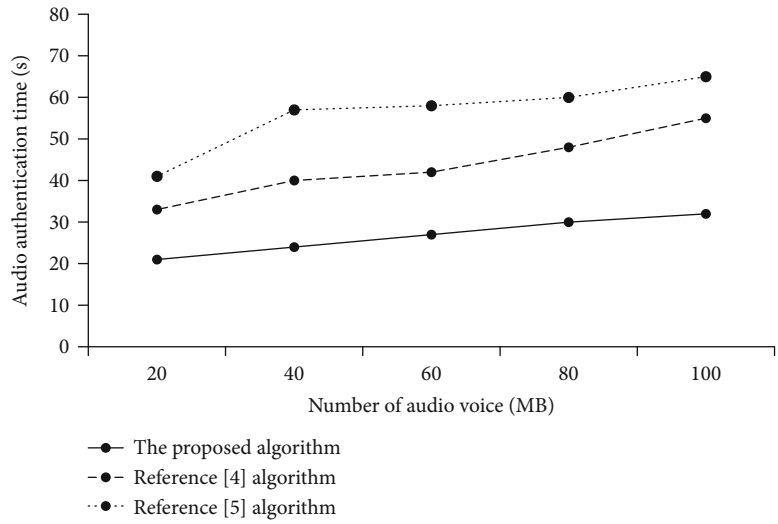


Figure 4: Comparison results of audio authentication time of different algorithms.

Danube" as an example, reference [4]'s algorithm, reference [5]'s algorithm, and the proposed algorithm are used to detect the tampered area, respectively, and the tampered area detection results of different algorithms are shown in Figure 3.

According to Figure 3, the tamper detection ability of reference [5]'s algorithm is poor, and the tampered area is not detected, followed by the tamper detection ability of reference [4]'s algorithm, which can locate one tampered area. The proposed algorithm has good tamper detection ability, which can effectively distinguish the admissible operation and malicious tampering. At the same time, most of the tampered areas can be accurately located by taking notes as the smallest unit, which has high accuracy in tamper detection.

4.3. Audio Authentication Time Analysis. Randomly extract 100 segments of speech from the speech database for operation, respectively, and use reference [4]'s algorithm, reference [5]'s algorithm, and the proposed algorithm for audio authentication and count the audio authentication time of different algorithms as shown in Figure 4.

According to the simulation results in Figure 4, as the number of audio voices increases, the audio authentication time of different algorithms will increase. The algorithm proposed in this paper is more sensitive to audio changes, and compared with the algorithm proposed in Reference [4] and Reference [5], it is more relaxed and the authentication time is shorter. When the audio voice volume reaches 100 MB, the audio authentication time of the algorithm in reference [4] is 67 s, the audio authentication time of the

algorithm in reference [5] is 59 s, and the audio authentication time of the proposed algorithm is only 32 s. Therefore, the audio authentication time of this algorithm is short.

5. Conclusion

This paper presents the research of multiterminal audio authentication for music based on a wireless network. According to the characteristics and key technologies of wireless networks, a wireless multiterminal audio system was established. The short-term energy calculation method is used to consider sudden changes in music signals. Segment the music signal, extract chroma features, and calculate a robust hash value. The dynamic time warping algorithm is used to align the notes, count the measured index values, use fuzzy classification to calculate the membership of the signal, obtain the authentication confidence, and detect the tampered area to realize the music multiterminal audio authentication. It can be seen from the experimental simulation that the algorithm proposed in the paper can effectively improve the tamper detection ability and shorten the audio authentication time. However, the research content of the paper still has some limitations. For example, the algorithm proposed in the paper does not consider different types of audio signals, and whether the algorithm has different performance results under different signals, these are the directions of future research. Therefore, in future research, the proposed algorithm will be extended to other types of audio signal content authentication, and it is hoped that an algorithm suitable for multiple signals can be obtained.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] H. Fang, X. Wang, and S. Tomasin, "Machine learning for intelligent authentication in 5G and beyond wireless networks," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 55–61, 2019.

[2] D. Chen, N. Zhang, R. Lu, N. Cheng, K. Zhang, and Z. Qin, "Channel precoding based message authentication in wireless networks: challenges and solutions," *IEEE Network*, vol. 33, no. 1, pp. 99–105, 2019.

[3] W. I. Khedr, K. M. Hosny, M. M. Khashaba, and F. A. Amer, "Prediction-based secured handover authentication for mobile cloud computing," *Wireless Networks*, vol. 26, no. 6, pp. 4657–4675, 2020.

[4] A. Ferdowsi and W. Saad, "Deep learning for signal authentication and security in massive internet-of-things systems," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1371–1387, 2019.

[5] A. Abdelaziz, R. Burton, F. Barickman, J. Martin, J. Weston, and C. E. Koksal, "Enhanced authentication based on angle of signal arrivals," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4602–4614, 2019.

[6] R. Diamant, P. Casari, and S. Tomasin, "Cooperative authentication in underwater acoustic sensor networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 954–968, 2019.

[7] T. Wang, K. Hu, X. Yang, G. Zhang, and Y. Wang, "A trust enhancement scheme for cluster-based wireless sensor networks," *Journal of Supercomputing*, vol. 75, no. 5, pp. 2761–2788, 2019.

[8] B. Groza, L. Popa, and P. S. Murvay, "Highly efficient authentication for CAN by identifier reallocation with ordered CMACs," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6129–6140, 2020.

[9] S. Li, M. Cheng, Y. Chen et al., "Enhancing the physical layer security of OFDM-PONs with hardware fingerprint authentication: a machine learning approach," *Journal of Lightwave Technology*, vol. 38, no. 12, pp. 3238–3245, 2020.

[10] K. Machat, "Gadgets/gizmos, inventions, and technology in Toronto," *Journal of the Audio Engineering Society*, vol. 67, no. 4, pp. 236–236, 2019.

[11] J. Zhang, H. Zhong, J. Cui, M. Tian, Y. Xu, and L. Liu, "Edge computing-based privacy-preserving authentication framework and protocol for 5G-enabled vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7940–7954, 2020.

[12] H. Liu, P. Zhang, G. Pu, T. Yang, S. Maharjan, and Y. Zhang, "Blockchain empowered cooperative authentication with data traceability in vehicular edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4221–4232, 2020.

[13] H. Vasudev, V. Deshpande, D. Das, and S. K. Das, "A lightweight mutual authentication protocol for V2V communication in internet of vehicles," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6709–6717, 2020.

[14] X. Sun, "Simulation of internet of things terminal dynamic authentication based on cluster architecture," *Computer Simulation*, vol. 37, no. 4, pp. 312–316, 2020.

[15] G. Bansal, N. Naren, V. Chamola, B. Sikdar, N. Kumar, and M. Guizani, "Lightweight mutual authentication protocol for V2G using physical unclonable function," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7234–7246, 2020.

[16] Z. Ye, C. Hu, L. He, G. Ouyang, and F. Wen, "The dynamic time-frequency relationship between international oil prices and investor sentiment in China: a wavelet coherence analysis," *Energy Journal*, vol. 41, no. 1, 2020.

[17] Z. He, F. Zhou, X. Xia, F. Wen, and Y. Huang, "Interaction between oil price and investor sentiment: nonlinear causality, time- varying influence, and asymmetric effect," *Emerging Markets Finance and Trade*, vol. 55, no. 12, pp. 2756–2773, 2019.

[18] C.-H. Chen, "An arrival time prediction method for bus system," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4231-4232, 2018.

[19] H. M. Liu, X. H. Bi, Z. F. Ye, and W. L. Wang, "Arc promoting inpainting using exemplar searching and priority filling," *Journal of Image and Grapgics*, vol. 21, no. 8, pp. 993–1003, 2016.

[20] J. Y. Lin, D. X. Deng, J. Yan, and X. Lin, "Self-adaptive group based sparse representation for image inpainting," *Journal of Computer Applications*, vol. 37, no. 4, pp. 1169–1173, 2017.

[21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural

similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[22] T. Y. Fu, L. Q. Jin, Z. Lei, and Z. Q. Li, "Face super-resolution method based on key points layer by layer," *Journal of Signal Processing*, vol. 32, no. 7, pp. 834–841, 2016.

[23] L. I. Gang, L. I. Haifang, S. H. A. N. G. Fangxin, and G. U. O. Hao, "Noise image segmentation model with local intensity difference," *Journal of Computer Applications*, vol. 38, no. 3, pp. 842–847, 2018.

[24] L. WANG and C. PAN, "Robust level set image segmentation via a local correntropy-based K-means clustering," *Pattern Recognition*, vol. 47, no. 5, pp. 1917–1925, 2014.

[25] Z. Chen, H. Cai, Y. Zhang et al., "A novel sparse representation model for pedestrian abnormal trajectory understanding," *Expert Systems with Applications*, vol. 138, p. 112753, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5532754, 12 pages
https://doi.org/10.1155/2021/5532754

Research Article

TSN: Performance Creative Choreography Based on Twin Sensor Network

Yufeng Wu,[1] **Longfei Zhang**,[1] **Gangyi Ding**,[1] **Dapeng Yan**,[1] **and Fuquan Zhang**[2]

[1]*Key Laboratory of Digital Performance and Simulation Technology, Beijing Institute of Technology, Beijing 100081, China*
[2]*Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, China*

Correspondence should be addressed to Longfei Zhang; longfeizhang@bit.edu.cn

Received 29 January 2021; Revised 23 February 2021; Accepted 10 March 2021; Published 26 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Yufeng Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The purpose of this paper is to improve the efficiency of performance creative choreography (PCC). Our research work shows that we can realize the model integration and data optimization for PCC in complex environments based on the combined architecture of sensor network (SN) and machine-learning algorithm (MLA). In order to explain the process and content of this research better, this paper designs a specific problem description framework for PCC, which mainly includes the following content: (1) a twin sensor network (TSN) architecture based on digital twin information interaction is proposed, which defines and describes the acquisition method, classification (creative data, rehearsal data, and live data), and temporal and spatial features of performance data. (2) Proposed a mobile computing method based on director semantic annotation (DSA) as the core computing module of TSN. (3) A spatial dynamic line (SDL) model and a creative activation mechanism (CAM) based on DSA are proposed to realize fast and efficient PCC of dance with the TSN architecture. Experimental results show that the TSN architecture proposed in this article is reasonable and effective. The SDL model achieved significantly better performance with little time increase and improved the computability and aesthetics of PCC. New research ideas are proposed to solve the computational problem of PCC in complex environments.

1. Introduction

In recent years, the cultural industry has become increasingly prosperous, and people's aesthetic of performing arts has continued to increase. The ability to present a wonderful performance creative has become a key factor in gaining audience recognition according to Lee [1]. Due to the limitation of time and space, the traditional creative process of performance has systemic problems such as low creative efficiency and unsatisfactory creative effect. The specific ones can be summarized as the following three points: First of all, in the traditional performance process, the performance data generated at each stage from creativity to implementation cannot be effectively implemented and fed back. It is mainly reflected in the inability of effective training and implementation of the creative data of the creators in the early stage and the sensory data of actors and audiences without effective feedback channels, and it is even difficult to attract the attention and reference of the creators. Secondly, due to the limitation of time and space, it is difficult to achieve a complete and effective implementation of the multielement performance choreography, such as stage design, actor training, and stage scheduling. It is mainly reflected in the difficulty of performing stage training for actors at the performance site. The implementation of stage design and performance creativity on site is difficult to achieve complete reproduction in the alternate training venue, which limits the effective implementation of performance creativity in the space-time dimension. Finally, there is no quick and effective method to evaluate the effect of performance creativity, resulting in a decrease in the director's ability to recreative performance, mainly reflected in the lack of objective and high-quality evaluation data to evaluate and describe performance creativity, fuzzy performance evaluation standards, diversified evaluation data, and low evaluation efficiency. Directors lack accurate feedback mechanisms and evaluation features, and it is difficult to organize effective recreation. From this point of view, the key reasons for restricting the

performance creative choreography (PCC) are the lack of effective decision-making closed loops and the lack of data feedback capabilities at different stages of the performance.

In order to solve the above problems, this paper starts from the basic principles of the world that constitute the physical domain, energy domain, and information domain, using computer technology to analysis and processing of the information domain to feed back and evaluate the energy domain and physical domain. Based on this concept, this paper treats PCC as a complex system and transforms the director experience of creative performance into computable complex system problems. In order to solve complex system problems, Darema [2] proposed a dynamic data-driven application system (DDDAS), which uses dynamic feedback and control between modelling simulation and experiment, making the simulation system execute in the process, and the data generated by the actual system can be dynamically obtained and responded. Grieves [3–5] proposed the concept of digital twin in 2003 to solve the interaction problem between real and virtual systems around the product lifecycle management (PLM), use data to connect the physical world and the virtual world, and realize multidisciplinary, multiscale, and multiprobability complex system simulation through data transfer between models. In 2011, the US Air Force Research Laboratory (AFRL) and the National Aeronautics and Space Administration (NASA) proposed to build a digital twin of the future aircraft [6]. Since then, the digital twin model system has been continuously expanded and extended. In 2014, Cerrone et al. [7] proposed the finite element model of the digital twin. Tavares et al. proposed an MVV model based on digital twins [8]. Tao et al. [9] extended the three-dimensional model of the digital twin and proposed a five-dimensional model of the digital twin, that is, physical entities, virtual entities, services, twin data, and the connection between various components, and proposed a digital twin system standard. It can be seen that the digital twin technology has a great advantage in solving the information domain problem of the physical world and the virtual world.

The physical and energy domain are also the key factors in solving complex system problems. In recent years, the Internet of Things (IoT) technology has been applied to many areas of our lives [10, 11], radio frequency identification (RFID) and sensor networks have emerged as the times require, and the new social system of the Internet of Everything is changing the human attitude and traditional cognition of society. With the support of 5G communication technology, the emergence of new computing methods such as edge computing [12] and fog computing [13] has greatly improved the interaction between IoT and smart objects, and a large number of objects have emerged, networking technical specifications and new semantic mobile computing modules [14, 15]. The complex IoT network that integrates large-scale smart sensors and the ubiquitous computing platform has become a powerful guarantee for semantic mobile computing [16, 17]. The rapid development of these emerging technologies provides new forms and possibilities for the data description and definition of the physical and energy domain of complex systems. In the traditional performance field, the data description and feedback mechanism between performance elements have become a core factor restricting the PCC. How to define and describe the complex system of PCC is the key to solving the problem.

In summary, this paper regards sensor networks, semantic mobile computing, and digital twins as key technologies and research objects in the physical domain, energy domain, and information domain in the complex system of PCC. The performance data is used as the information description of creative choreography, the sensor network is introduced into the information interaction framework of the digital twin, and the "twin sensor network" (TSN) is established as the performance data acquisition and calculation node.

The main contributions of this study are summarized as follows:

(i) A twin sensor network (TSN) architecture based on digital twin information interaction is proposed, which defines the acquisition method, classification (creative data, rehearsal data, and live data), and spatiotemporal features of performance data and describes the feedback mechanism and symbiosis between performance data in the complex environment

(ii) A mobile computing method based on director semantic annotation (DSA) is proposed, which establishes the mapping relationship between director experience and performance data for the first time and reduces the difference of performance data feature distribution under the conditions of experience constraints

(iii) A spatial dynamic line (SDL) model based on DSA is proposed for dance creative choreography, and a creative activation mechanism (CAM) is established to achieve fast and efficient performance with the TSN architecture

This paper is organized as follows: Section 2 reviews the related works. Section 3 details the methods proposed in our research. Section 4 presents the experiment results. Section 5 summarizes and looks forward to the future work.

2. Related Works

In this chapter, we briefly summarize the recent research work on sensor networks and digital twin technology and propose the research motivation and ideas for constructing TSN.

2.1. Affective Computing Based on Sensor Network. With the rapid development of smart mobile devices, the use of sensors has become widespread. At present, the application of sensors is not only sensing, but accurate and fast data calculation and research, making full use of the advantages of lightweight and mobile computing of sensor nodes to provide users with accurate and effective intervention measures. For example, mobile devices such as smart phones, in addition to the conventional physical positioning based on GPS, also integrate a large number of built-in multimodal sensors. These sensors can quickly perceive changes in the human

body state and play a positive role in human health and emotion prediction [18–20]. Sleep quality detection based on time characteristics has been used in the medical field to assist in the treatment of diseases [21, 22]. Obtaining the physiological characteristics of the human epidermis to determine vital signs and abnormal state monitoring has gradually become a hot spot for physiology and psychology experts [20, 23]. The noninvasive EEG signal analysis technology played a decision-making role in grasping emotional changes and emotional cognition and provided a variety of new possibilities in judging the human body's stress response after stimulation and psychological adjustment strategies [24, 25]. Wang et al. [24] used GSR signals to simultaneously monitor the emotional state of 15 subjects in live performances. The research conducted a cluster analysis of all the audiences through the obtained GSR values and found that the data of 10 subjects were closely related. Martella et al. [25] analysed the audience's feedback in live performances, used a three-axis accelerometer to calculate the acceleration of the subject's body movements, and integrated complex emotional experiences such as "Enjoyment" or "Immersion" for quantification. By calculating the dynamic changes of acceleration, the accuracy of the study on whether the audience is in an "enjoying state" has reached nearly 90%.

Loprinzi et al. [26] proposed a cognitive emotional model to assess physical activity based on experience and to evaluate the exercise habit and intensity of adults. The emotional changes can be seen after the acute exercise, and it is related on the exercise intensity and exercise cycle of the volunteers. At the same time, personal health will also have an impact on emotional changes, which shows that this measurement method has inevitable flaws in its universality. A method of using emotional calculation to evaluate football players was proposed by Liu et al. [27], which combines the text information of the postmatch report and emotional calculation to measure the performance quality of the players. The authors established a player performance evaluation model based on LSTM; however, this method still has some problems to be solved. For example, it is difficult to achieve accurate statistics and quantification of specific behaviours in the game, which has a greater impact on the collection and evaluation of key information. Pinto et al. [28] processed electrocardiogram, electromyography, and dermal electrical activity to find a physiological model of emotion. Using samples of 55 healthy subjects, Pinto et al. used single-peak and multipeak methods to analyse which signals or combinations of signals can better describe emotional responses. Zhang [29] proposed a multimodal emotion recognition method using deep autoencoders for facial expressions and EEG interactions. The decision tree is used as the target feature selection method. Then, based on the facial expression features recognized by the sparse representation, the solution vector coefficients are analysed to determine the facial expression category of the test sample. After that, the bimodal depth autoencoder was used to fuse EEG signals and facial expression signals.

Affective computing based on sensor networks has made great progress in audience emotion detection, athletes' physiological state feedback, and auxiliary medical care. It can be seen from this that with the support of the IoT technology, the application of sensor networks has opened up a new approach for the traditional information industry and established an analysis method and quantitative model based on emotion and cognition in physical information monitoring and feature acquisition. In the traditional performance process, creative information has not yet received effective recognition and feedback.

2.2. Digital Twin Modelling. Digital twin technology is more and more widely used in system engineering. Since 2002, it has been transformed from a concept to a system engineering model, and the status and operating data of the system are continuously updated and maintained throughout the life cycle [30]. With the continuous development of 5G technology, it has promoted the further innovation of smart cities including the digital twin technology of IoT, and digital infrastructure has become the basis for the transformation of traditional industries [31]. People began to create a mirrored world, which can copy and manipulate physical objects (material domains) and control the real world through the resulting data changes [32]. Create digital twins with existing software and operating systems and establish new ideas for simulation and solving complex system problems [33]. Ding et al. [34] proposed a digital twin-based cyber-physical production system (DT-CPPS) and discussed the operating mechanism of supporting mechanisms and real-time data-driven control methods, including status data perception, intelligent interaction, and autonomous decision-making. Wang et al. [35] further analysed the monitoring, management, and durability of the human-machine collaboration system and proposed a human-computer collaboration modelling method based on digital twins to provide a comprehensive perception of products and technologies. This method has greatly inspired the traditional performance field, and it has proposed new forms and focus on performance data acquisition and feature fusion.

Autiosalo et al. [36] introduced a common method to analyse and construct digital twins in various applications. We have identified the common characteristics of digital twins from earlier documents and proposed an analysis method based on these characteristics to compare digital twin examples. This method is used to verify the presence of features and can be further enhanced. Jones et al. [37] provided the characteristics of digital twin, recognition gap knowledge, and necessary areas for future research. When characterizing digital twins, concepts, key terms, and the state of related processes are identified, discussed, and merged to produce 13 features (physical entity/twin, virtual entity/twin, physical environment, virtual environment, state, realization, metrology, twin, twin rate, physical to virtual connection/twin, virtual to physical connection/twin, physical process, and virtual process) and a complete framework of digital twin and its operating process. A quantitative model of sustainable design was established by He et al. [38]. Used the entropy method to determine the weight of the subindices, established the decision matrix of the indicators to be evaluated, reflected the obtained quantitative indicators on the digital twin data flow system, and improved the comprehensive performance

of the product in sustainable design through data fusion, and proved the design and the feasibility of the use process.

A large amount of performance data is generated at different stages of the performance, and these data can only be used as accessories for different stages at present. The PCC cannot be realized without the formation of effective data features and data-driven effective feedback. The results of on-site training cannot be clearly fed back, and there is an information barrier between the director and the actors. The emergence of digital twin modelling technology provides technical support for the loop and feedback of performance data. Performance data acquisition, feature extraction, feature fusion, and data-driven modelling methods will form new and reliable feedback values between the physical world and the virtual world.

2.3. Space Trajectory in Performance Choreography. In recent years, more and more drone performances have become the focus of people's attention. Path planning is the primary issue of UAV collaborative control. Due to the complexity and diversity of UAV cluster path planning, a lot of research has focused on UAV cluster path planning models. In addition to the commonly used planning space representation, there are also methods such as target planning and conditional constraints. In addition to successfully avoiding collisions, the formation of drones must also ensure the smoothness and smoothness of the flight path according to Xu et al. [39], Dong et al. [40], and Ren et al. [41]. Xu et al. [42] and Gu and Deng [43] tried to transform the formation of the crowd into key frame matching and path planning between people. However, in stage performances, this key frame matching and movement method still lacks a certain degree of integrity and aesthetics for expressing artistic effects.

In the process of creative choreography of dance, the most important thing is to enhance the effect of choreography through its own spatial movement and the beauty of the movement trajectory to realize the choreography creativity. The dance choreography mentioned in this article mainly refers to the stage space scheduling choreography of single dance, double dance, and group dance. The planning of stage plays and musicals needs to consider the development of the story and the overall planning of the language interaction between the characters, which is quite different from the choreography mode of dance, so it is not the research scope of this article. In this process, you will definitely encounter various space constraints and behaviour constraints in the performance scene and draw a motion curve with artistic beauty as much as possible without trajectory conflict. Xu et al. [39] proposed a scheme that can automatically generate a visually beautiful crowd key frame motion trajectory. Although it can control a large-scale crowd from the perspective of a cluster, it has certain defects in detail optimization and aesthetic enhancement. Ren et al. [41] and Xu et al. [44] proposed many different methods to simulate the conversion of crowd formations, and Takahashi et al. [45] proposed a spectrum-based crowd movement route control method, but these methods have computational memory limitations on the identification and designation of the key frame.

Wang et al. [46] proposed a real-time obstacle avoidance path tracking algorithm, which enables the UAV to avoid large unknown static and moving obstacles in time and fly to the target after avoiding obstacles during path tracking. Wang et al. [47] proposed a path following method based on virtual force, which compensates the influence of the reference curvature for the circle and the curve of the curvature over time. The design idea of this proportional-differential controller is on the stability and convergence of the trajectory movement. It has certain practical significance and lacks subjective cognition and aesthetic laws in the planning of crowd movement on stage. In the research of crowd simulation, in addition to the study of the shape template and the shape constraint of the controller by Chang and Li [48] and Gu and Deng [49], more attention is paid to the efficiency of crowd reaching the destination according to Golas et al. [50] and Zhao et al. [51].

In the past few years, the technology of space-time mapping of people in a virtual environment has become a mainstream research method, especially in the fields of movies, games, and simulation applications. Ramirez et al. [52] proposed an optimization-based method to generate a smooth collision-free trajectory for the crowd, and the cubic spline curve method played a good role in smoothing the trajectory. Jordao et al. [53] fill people with visual characteristics in a large virtual environment, provide a solution for the artist's spatial composition design, and introduce a specific stage after each iteration to avoid using the global pathfinding process to create local cycle. Discretize the environment into regular patches and meet the overall density and traffic restrictions by updating the direction of the agent in the patch. This optimization method has a certain driving effect on artistic creation. Barnett et al. [54] use Reeb graph to represent the topology of the environment and calculate the maximum capacity of each path in the graph. Lightweight algorithms are used to perform crowd path planning, and advanced parameters such as collaboration and congestion levels can be used to control crowds, saving a lot of computing time in a comprehensive system for rapid crowd interaction control. Toll [55] uses a multilayer navigation network to construct a path planning algorithm in a complex virtual environment, which provides a new idea for large-scale dynamic crowd simulation and a crowd relationship diagram describing multilayer complex environments.

This paper believes that the core problem of PCC for dance is to calculate the trajectory of the actors in the stage space and optimize the trajectory to meet the standards of dance aesthetics. This poses a challenge to the ways and methods of PCC for dance: first, how to draw the trajectory of the actors, and secondly, how to evaluate the drawn trajectory quantitatively. The method of PCC determines whether the original creative can be realized quickly and effectively, especially the PCC for dance. It is necessary to gather high-quality experience and digital choreography to meet the needs of performance creative. Among them, the stage space scheduling of dance performances is the most important part of the choreography, but the traditional space scheduling method relies on the choreographer and actors performing repeated manual rehearsals. This process caused a lot of waste of resources and time consumption, and the final

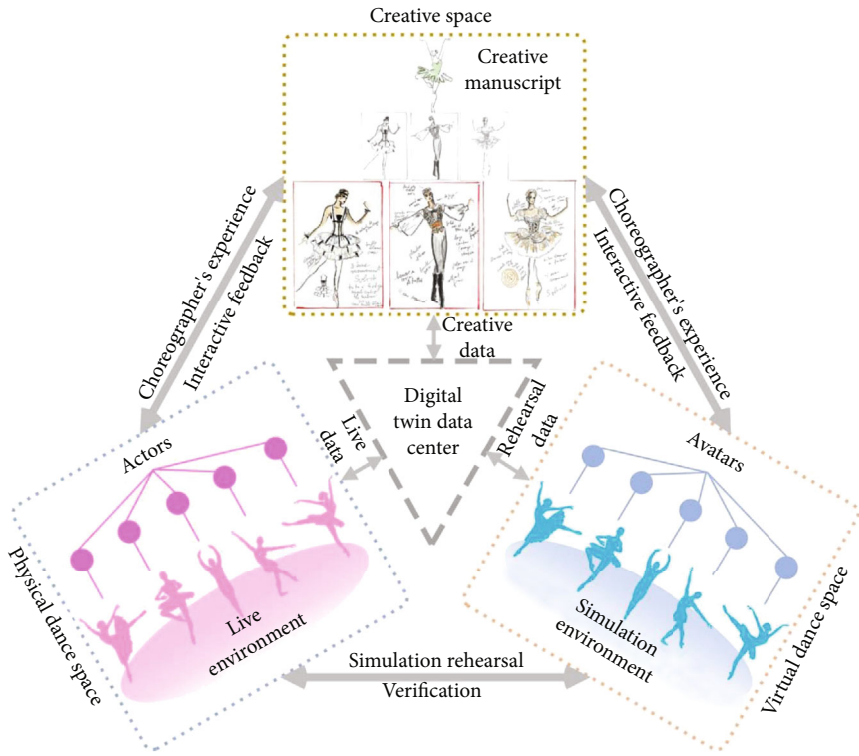


Figure 1: Dance digital twin space (DDTS).

orchestration effect cannot be determined as the optimal solution. So, the PCC for dance begins to show its research value and practical significance.

3. Materials and Methods

3.1. Dance Digital Twin Space (DDTS). This paper conducts a detailed analysis and research on the concept and development of digital twins, especially the system framework of the digital twin workshop [56]. Combining the traditional choreography process and choreographer's feedback, a digital twin space for creative choreography is proposed (see Figure 1).

As shown in Figure 1, DDTS is composed of a physical dance choreography space (PDS), a virtual choreography workshop (VDS), a choreographer creative space (CS), and a digital twin data center (DTDC).

The specific operating mechanism of the digital twin framework: Firstly, in the creative space, the choreographer can draw the manuscript based on experience and creative needs to form visual original creative data. Then, use the experience and creative data of the choreographer as the input of the virtual choreography space and perform virtual choreography of the choreographer's manuscript in the virtual environment. In this space, you can provide choreography feedback to the creative space at any time and assist the choreographer to redesign the creative plan. After that, the virtual choreography space uses the simulation rehearsal data as the input of the physical choreography space to guide the training and choreography of real dancers, and the display data of the physical choreography space verifies the feasibility of the simulation rehearsal data. Finally, the live data of the physical choreography space is fed back to the creative space and provides feedback and redesigns reference for live performance effects for choreographers.

In this closed-loop twin computing framework, three different digital spaces are gathering the generated data in the digital twin data center. In this data center, the process data and redesign data of the entire dance choreography are accumulated and integrated to choreograph. Provide strong data support and computability verification for each link. The behavioural interaction and data interaction of the entire dance digital twin space will continue to occur until the choreographer believes that the design plan has the performance conditions in terms of feasibility and aesthetic imitation, and the closed-loop iterative operation of the framework can be terminated.

3.2. Twin Sensor Network (TSN). In this paper, the sensor network is regarded as a key link in solving the complex system problem of PCC, and performance data is used as the information description of creative choreography. The sensor network is introduced into the information interaction framework of the digital twin. For the DDTS proposed above, the twin sensor network (TSN) is established as a performance data acquisition and calculation node. The design architecture of TSN is shown in Figure 2.

The core of digital twin technology is to connect data and information between the physical world and the virtual world. The TSN architecture mainly includes four modules:

(1) The sensor node, for obtaining performance data, collects the raw data of the performance, including

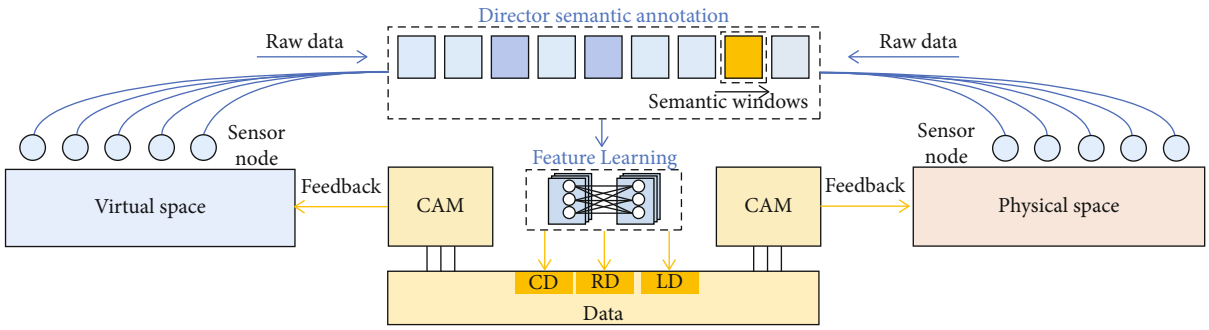


Figure 2: Twin sensor network (TSN).

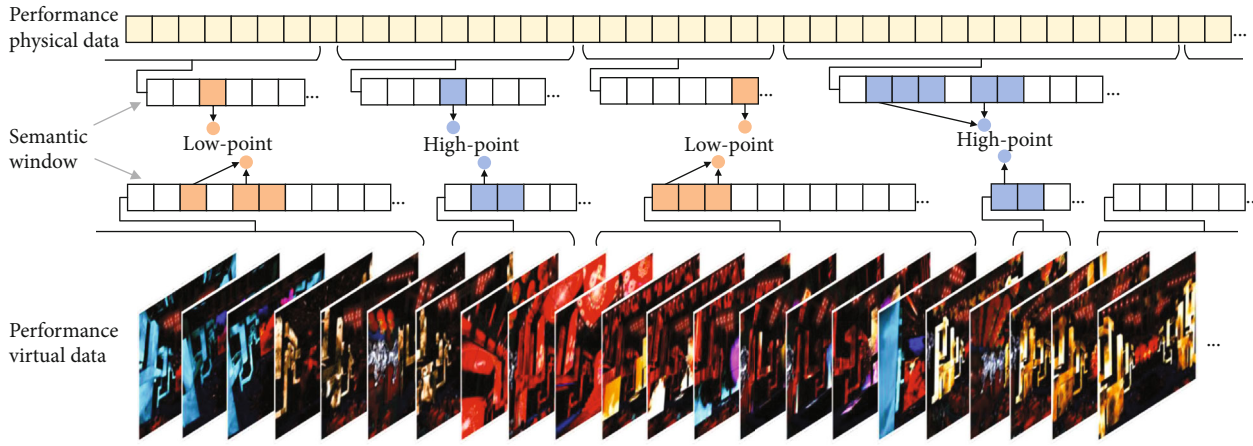


Figure 3: Director semantic annotation (DSA).

multimodal data types such as performance environment, director creativity, actor training, and audience feedback. These data describe the performance from different subjective and objective perspectives, which include the following: EEG signal acquisition uses the whole brain induction head-mounted device, which consists of 14 channel sensors and 2 bipolar reference electrodes, samples at a rate of 128 Hz. The data collection of the audience's facial expressions uses a conventional high-definition camera with an image resolution of 1920*1080, a smart bracelet to collect the audience's heart rate. The device uses a PPG heart rate sensor, a three-axis acceleration sensor, and a three-axis gyroscope

(2) Director semantic annotation (DSA) is one of the innovations proposed in this paper. Aiming at the professionalism and particularity of performance data, DSA uses a supervised model training method to label the training set and test set to achieve the most effective training results. According to the creative intention and aesthetic experience of the performance work, the director defined and marked the high point, normal point, and low point of the performance creative. Among them, high point corresponds to positive and wonderful ideas, normal point corresponds to general reasonable ideas, and low point corresponds to negative failed ideas (see Figure 3)

(3) After the DSA preprocesses the raw sensor data, the feature learning module extracts features and assigns different time tags to the corresponding data type window [16, 57] (see Figure 4).

(4) The creative activation mechanism (CAM) is explained in Section 3.3.

The digital twin technology integrates the data, models, calculation logic, simulation, and sensor feedback of the physical world and the virtual world to draw a complete twin computing system and twin life mechanism [58, 59]. We map the physical space and virtual space of creative choreography to three choreography stages: creative stage, rehearsal stage, and live stage, and propose a computable framework for PCC based on digital twins (see Figure 5).

Among them, the virtual space in the creative stage reflects the creative motivation and experience of the choreographer, including creative motivation, aesthetic experience, theme selection, and creation rules. The creative data feedback from the virtual space includes manuscript data, emotion data, and copyright data. These data are mainly the concrete definition of the choreograph through the creative manuscripts and literary descriptions of the

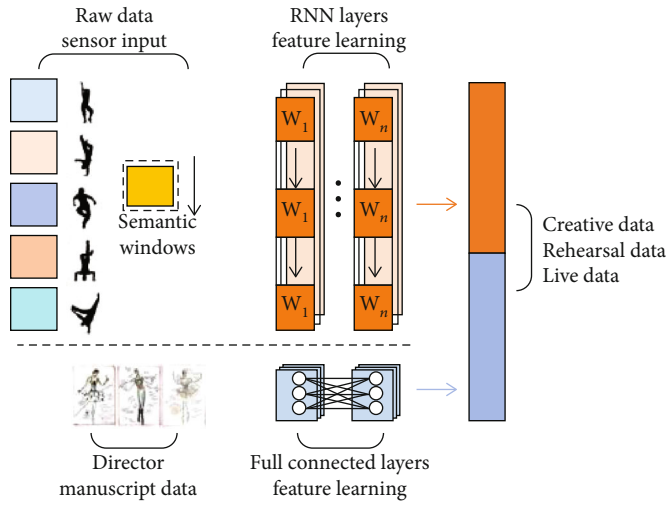


Figure 4: Feature learning.

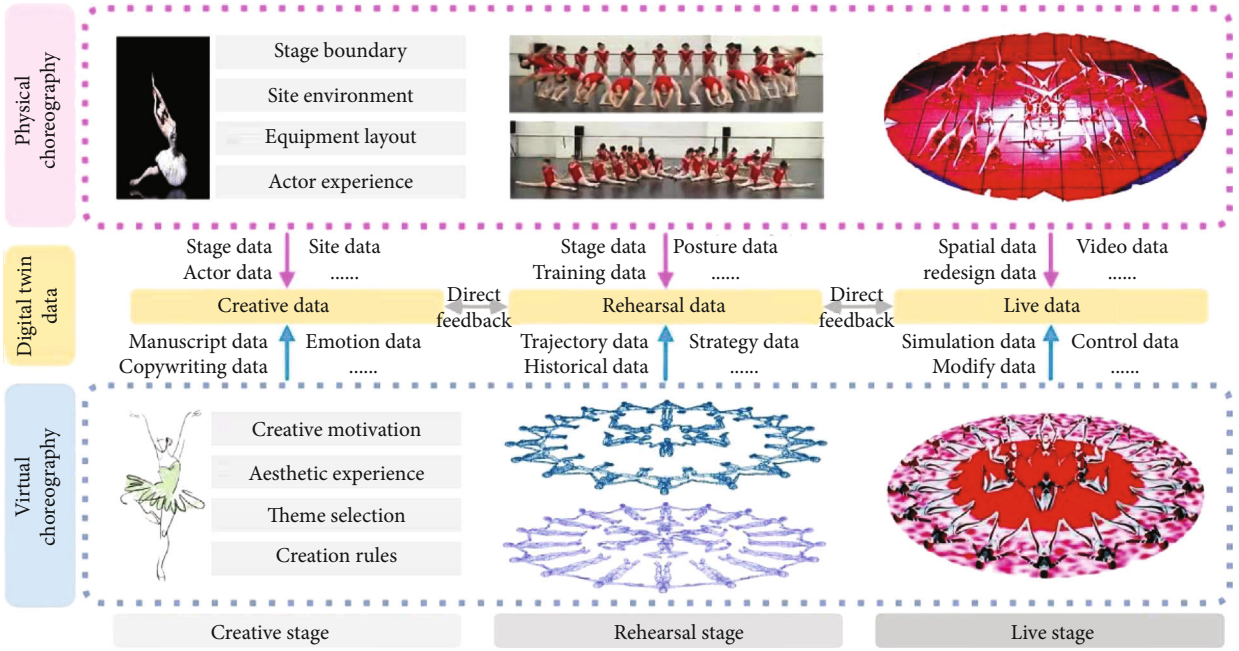


Figure 5: Computable framework for PCC based on digital twins.

choreographer. The relevant design elements of the physical space in the creative stage include stage boundary, site environment, equipment layout, and actor experience. The creative data feedback from the physical space includes stage data, site data, and actor data. These data are the description of the objective characteristics of the physical space. It reflects the objective attributes of the physical choreography space. These characteristics and attributes provide a perceptual experience for the creative ideas of the choreographer. But at the same time, there are corresponding boundary restrictions and constraints on the creative choreography process.

In the rehearsal stage, the virtual space normalizes and standardizes the creative manuscript data of the editor and choreographer to form rehearsal data suitable for training, including trajectory data, strategy data, and historical data. The physical space mainly provides a suitable environment for the training of actors. During the training process of the physical space, the corresponding rehearsal data is fed back, including stage data, posture data, and training data. These data have extremely high feedback value and redesign basis for the rehearsal process.

In the live stage, the virtual space performs full-element simulation of the complete dance performance, including stage sound, light, electricity, and other dance environment data, actor's body movement data, and space scheduling data. Provide simulation data, control data, and modify data for live data. The physical space provides a real-stage environment for dance performances and real spatial data, video data, and redesign data for live data. These data provide a good redesign standard and basis for the rehearsal process

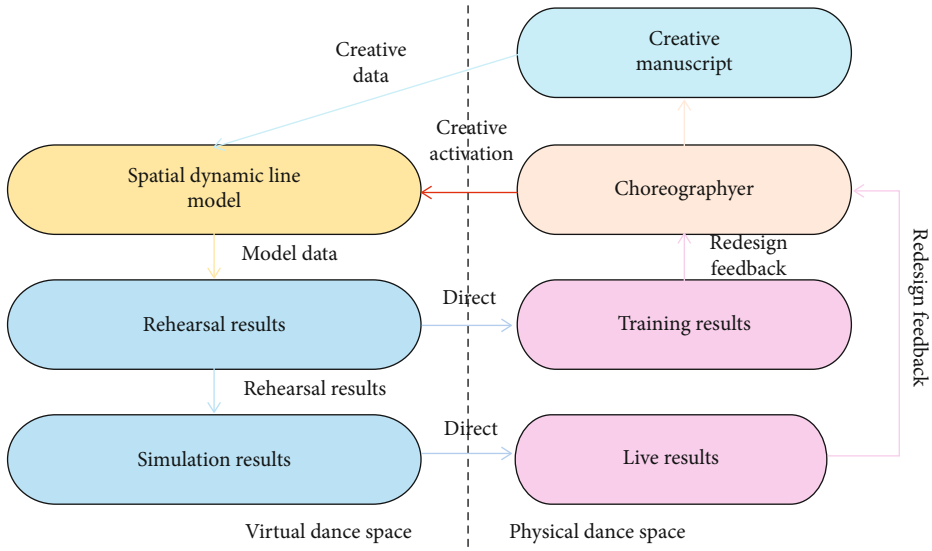


Figure 6: Spatial dynamic line (SDL) model.

and provide real environmental constraints and performance status feedback to the creative realization of the choreographer. In digital twin data, creative data, rehearsal data, and live data are standardized corresponding to three different stages. Creative data provides creative concepts for rehearsal data, and rehearsal data provides training standards for live data. On the contrary, live data puts forward the basis of redesign for rehearsal data, and rehearsal data puts forward the constraint conditions for optimizing creativity for creative data. The computable framework for PCC proposed in this paper double-maps the physical and virtual spaces of the different stages of dance choreography and connects and integrates the performance characteristics of each stage through the iterative symbiosis relationship of twin data. This is the first time in the field of artistic creation that creative choreography has been transformed into a computable digital twin framework.

3.3. Spatial Dynamic Line (SDL) Model. In this work, we dynamically incorporate feedback based on mutual information into SFM. We briefly described the concept of mutual information and how it can be applied to measure the trajectory of actors [39]. Mutual information (MI) is a well-known concept in the field of information theory. It is designed to quantify the interdependence between two random variables as the Equation (1).

$$I(G,H) = \sum_{i,j} p(g_i, h_j) \log_n \frac{p(g_i, h_j)}{p(g_i)p(h_j)}. \quad (1)$$

Among them, $p(g_i)$, $p(h_j)$, and $p(g_i h_j)$ are the individual probability distribution and joint probability distribution of g and h, respectively. It is important to insert more than two key frame structures. In the current framework, users are allowed to insert new intermediate key frame structures. We use the method of [45] to allow the director to generate new key frame data to activate and interfere with the existing motion trajectory as the Equation (2).

$$p(t) = c_{-1}p_{n-1} + c_0 p_n + c_1 p_{n+1} + c_2 p_{n+2}. \quad (2)$$

The spatial dynamic line (SDL) model proposed in this article is shown in Figure 6.

In our proposed model structure, the choreographer can obtain feedback on the choreography effect through training results and live results. The core idea of the SDL model is to accept the creative activation of the choreographer at any time, fine-tune the spatial dynamic line calculated by the model, and redesign the average curvature, composition, and performance in the spatial.

This article proposes that the advantage of CAM compared to [53–55] is that the performance data is formed into a partial loop during the rehearsal stage, and the director's intention is transformed into recreation in the multiple rounds of rehearsal data loops. To meet the needs of the overall creative effect, the optimization and realization of the overall plan can be achieved by changing the limitations and computing power of the local specific stages.

4. Results and Discussion

4.1. Experimental Design. This article selects the traditional folk dance work "Cutting Flowers" to verify the spatial movement model and the digital twin dance creative choreography simulation framework. The core idea of the experiment is to use the actors' initial and target positions drawn by the choreographer, calculate the spatial trajectory from the initial point to the target point by the SDL, and generate the rehearsal effect through the digital twin choreography framework until the space the trajectory of the spatial dynamic line reaches the artistic effect expected by the choreographer. In this iterative creative process, when the choreographer thinks it is necessary to redesign the rehearsal effect, the creative

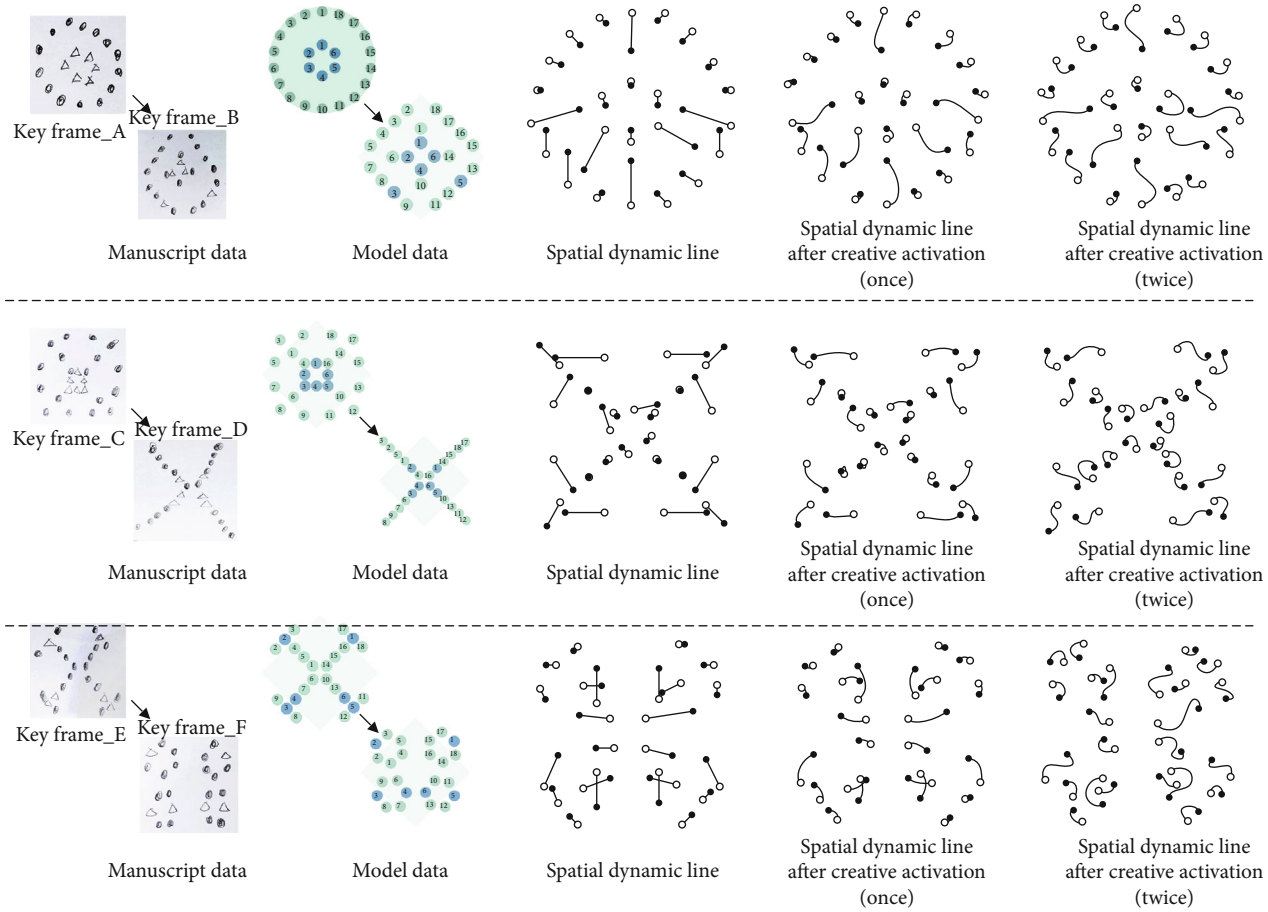


Figure 7: Spatial dynamic line effect by SDL in the dance work "Cutting Flowers."

Table 1: The performance of the SDL with no creative activation, once, and twice in the trajectory experiment from key frame A to key frame B.

Creative activation	Time (s)	Average curvature	Composition	Dance performance
None	1.37	0	1	2
Once	1.51	15.4	2	3
Twice	1.83	37.9	4	5

Table 2: The performance of the SDL with no creative activation, once, and twice in the trajectory experiment from key frame C to key frame D.

Creative activation	Time (s)	Average curvature	Composition	Dance performance
None	1.19	0	1	2
Once	1.35	10.9	3	3
Twice	1.58	27.8	4	5

Table 3: The performance of the SDL with no creative activation, once, and twice in the trajectory experiment from key frame E to key frame F.

Creative activation	Time (s)	Average curvature	Composition	Dance performance
None	1.21	0	0	2
Once	1.47	9.4	2	3
Twice	1.69	34.1	4	5

manuscript can be fine-tuned and used as the creative activation of the SDL to generate a spatial dynamic line that is more with creative and aesthetic needs.

The manuscript data used in the experiment was hand-drawn by the choreographer who decomposed the key frame and used two different graphic shapes to represent the group of actors in the group dance. We use time and average curvature between two points as objective evaluation indicator, composition, and dance performance (range from zero to five) as subjective evaluation indicator which scored by choreographer.

4.2. Results Analysis. This paper conducts experimental verification on dance "Cutting Flowers." Firstly, three sets of key frame transformations are decomposed from the dance structure. Starting from the manuscript of the choreographer, through the model's autonomous calculation, the choreographer's first activation and the choreographer's second activation generated a spatial dynamic line effect

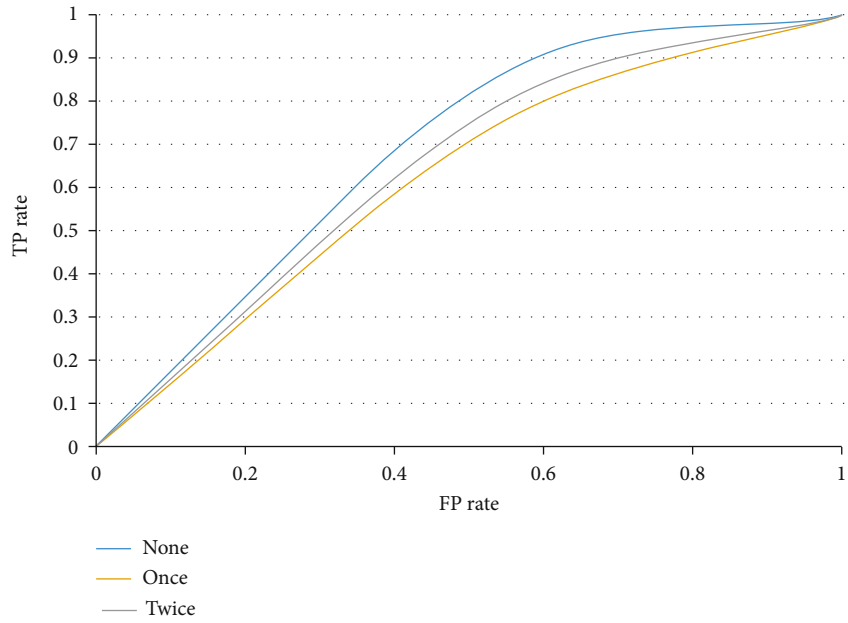


Figure 8: ROC curve of model.

(see Figure 7). Each line represents the performance of spatial dynamic line with no creative activation, one creative activation, and two creative activations. From Tables 1–3, we can see that our model obtains a significantly better performance with very few time increase than the method without creative activation.

The ROC curve of model is shown in Figure 8 during different circumstances (no creative activation, one creative activation, and two creative activation).

From the experimental results, the TSN, SDL, and CAM proposed in this paper are effective and have reached the research expectations of this paper. From Tables 1–3, we can further conclude that with the increase in the number of creative activations, the calculation time has increased slightly, but the average curvature of the spatial trajectory of dance scheduling has been greatly improved, and the dance structure has been powerful which performed. Secondly, from the performance of the SDL model, when the key frame change area is not obvious (key frames E to F), the average curvature increase rate is 2.034% higher than when the key frame change area is more obvious (key frames A and B). This shows that the SDL and CAM proposed in this article have great assistance and driving force for the creative promotion of dance choreographers and the efficiency of on-site scheduling.

At present, the experimental environment and manuscript data of this article still have limitations to a certain extent. For example, the manuscript data used in this article does not have a certain stage boundary in the screen, and it is a single-tone hand-drawn sample. These factors will be improved in the next experiment, adding clear stage boundaries and conditional constraints boundaries and discussing the influence of the multicolor hand-painted background on the SDL model and the improvement of the director's creative emotions.

5. Conclusions

The purpose of this article is to use the digital twin framework to perform creative choreography. To achieve this goal, we first build a dance digital twin space (DDTS) for creative choreography of dance performances. Mapped the choreography in physical and virtual spaces; after that, we proposed the twin sensor network (TSN) as a performance data acquisition and calculation node. Through the symbiosis relationship of performance data (creative data, rehearsal data, and live data), the whole process of PCC is described, which solves the waste of resources caused by traditional nondigital choreography. Finally, we proposed a spatial dynamic line (SDL) model and creative activation mechanism (CAM), allowing the choreographer to achieve creative activation and redesign of dance choreography through a TSN architecture.

Experimental results show that the TSN architecture proposed in this article is reasonable and effective. The SDL model achieved significantly better performance with little time increase and improved the computability and aesthetics of PCC. New research ideas are proposed to solve the computational problem of PCC in complex environments.

Data Availability

The data used to support the findings of this study have not been made available because commercial reasons.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work has been supported by the National Key R&D Program of China under Grant No. 2018YFB1403905.

References

[1] I.-S. Lee, "Performing arts in the age of transmedia," *Journal of Acting Studies*, vol. 17, no. 2, pp. 17–32, 2020.

[2] F. Darema, "Dynamic data driven applications systems: a new paradigm for application simulations and measurements," in *International Conference on Computational Science*, pp. 662–669, Berlin, Heidelberg, 2004.

[3] M. W. Grieves, "Product lifecycle management: the new paradigm for enterprises," *International Journal of Product Development*, vol. 2, no. 1/2, pp. 71–84, 2005.

[4] M. W. Grieves and M. Tanniru, "PLM, process, practice and provenance: knowledge provenance in support of business practices in product lifecycle management," *International Journal of Product Lifecycle Management*, vol. 3, no. 1, pp. 37–53, 2008.

[5] M. Grieves, "Digital twin: manufacturing excellence through virtual factory replication," *White paper*, vol. 1, pp. 1–7, 2014.

[6] E. Glaessgen and D. Stargel, "The digital twin paradigm for future NASA and U.S. Air Force vehicles," in *53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA*, Honolulu, Hawaii, 2012.

[7] A. Cerrone, J. Hochhalter, G. Heber, and A. Ingraffea, "On the effects of modeling as-manufactured geometry: toward digital twin," *International Journal of Aerospace Engineering*, vol. 2014, 10 pages, 2014.

[8] P. Tavares, J. A. Silva, P. Costa, G. Veiga, and A. P. Moreira, "Flexible work cell simulator using digital twin methodology for highly complex systems in industry 4.0," in *Iberian Robotics conference*, pp. 541–552, Cham, 2017.

[9] F. Tao, X. Ma, T. Hu, Z. Huang, J. Cheng, and Q. Qi, "Research on digital twin standard system," *Computer Integrated Manufacturing Systems*, vol. 25, no. 10, pp. 2405–2418, 2019.

[10] T. Su, S. Shao, S. Guo, and M. Lei, "Blockchain-based internet of vehicles privacy protection system," *Wireless Communications and Mobile Computing*, vol. 2020, no. 5, pp. 1–10, 2020.

[11] M. S. Mahmoud and A. A. Mohamad, "A study of efficient power consumption wireless communication techniques/modules for internet of things (IoT) applications," *Advances in Internet of Things*, vol. 6, no. 2, pp. 19–29, 2016.

[12] Z. Li, R. Xie, and L. Sun, "A survey of mobile edge computing," *Telecommunications Science*, vol. 34, no. 1, pp. 87–101, 2018.

[13] A. Arif and A. Ejaz, "A survey on mobile edge computing," in *10th IEEE international conference on intelligent systems and Control*, Coimbatore, India, 2016.

[14] F. Li, D. Wang, Y. Wang et al., "Wireless communications and mobile computing blockchain-based trust management in distributed internet of things," *Wireless Communications and Mobile Computing*, vol. 2020, no. 5, pp. 1–12, 2020.

[15] M. A. Imran, A. Zoha, L. Zhang, and Q. H. Abbasi, "Grand challenges in IoT and sensor networks," *Frontiers in Communications and Networks*, vol. 1, 2020.

[16] I. Perez-Pozuelo, D. Spathis, E. A. Clifton, and C. Mascolo, "Wearables, Smartphones, and Artificial Intelligence for Digital Phenotyping and Health," *Digital Health*, pp. 33–54, 2021.

[17] Z. Tian, D. Liu, W. Chen, W. Wang, S. Pan, and N. Guo, "Analysis and optimization of asymmetric wireless power transfer in concrete," *Wireless Communications and Mobile Computing*, vol. 2020, 12 pages, 2020.

[18] R. Wang, M. S. Aung, S. Abdullah et al., "Crosscheck: toward passive sensing and detection of mental health changes in people with schizophrenia," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 886–897, Heidelberg Germany, 2016.

[19] X. Zhang, W. Li, X. Chen, and S. Lu, "Moodexplorer: towards compound emotion detection via smartphone sensing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, p. 176, 2018.

[20] N. Jaques, S. Taylor, A. Sano, and R. Picard, "Multimodal autoencoder: a deep learning approach to filling in missing sensor data and enabling better mood prediction," in *International Conference on Affective Computing and Intelligent Interaction (ACII)*, San Antonio, Texas, 2017.

[21] S. Servia-Rodríguez, K. K. Rachuri, C. Mascolo, P. J. Rentfrow, N. Lathia, and G. M. Sandstrom, "Mobile sensing at the service of mental well-being: a large-scale longitudinal study," in *Proceedings of the 26th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee*, pp. 103–112, Perth Australia, 2017.

[22] Y. Suhara, Y. Xu, and A. Pentland, "Deepmood: forecasting depressed moodbased on self-reported histories via recurrent neural networks," in *Proceedings of the 26th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee*, pp. 715–724, Perth Australia, 2017.

[23] N. Jaques, O. O. Rudovic, S. Taylor, A. Sano, and R. Picard, "Predicting tomorrow's mood, health, and stress level using personalized multitask learning and domain adaptation," in *Proceedings of IJCAI 2017 Work-shop on Artificial Intelligence in Affective Computing, volume 66 of Proceedings of Machine Learning Research*, pp. 17–33, Melbourne, Australia, 2017.

[24] C. Wang, E. N. Geelhoed, P. P. Stenton, and P. Cesar, "Sensing a live audience," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1909–1912, Toronto, Canada, 2014.

[25] C. Martella, E. Gedik, L. Cabrera-Quiros, G. Englebienne, and H. Hung, "How was it? Exploiting smartphone sensing to measure implicit audience responses to live performances," in *Proceedings of the 23rd ACM International Conference on Multimedia*, pp. 201–210, Brisbane Australia, 2015.

[26] P. D. Loprinzi, S. Pazirei, G. Robinson, B. Dickerson, M. Edwards, and R. E. Rhodes, "Evaluation of a cognitive affective model of physical activity behavior," *Health Promotion Perspectives*, vol. 10, no. 1, pp. 88–93, 2020.

[27] W. Liu, X. Xie, S. Ma, and Y. Wang, "An improved evaluation method for soccer player performance using affective computing," in *2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, Chengdu, China, 2020.

[28] G. Pinto, J. M. Carvalho, F. Barros, S. C. Soares, A. J. Pinho, and S. Brás, "Multimodal emotion evaluation: a physiological model for cost-effective emotion classification," *Sensors*, vol. 20, no. 12, p. 3510, 2020.

[29] H. Zhang, "Expression-EEG based collaborative multimodal emotion recognition using deep autoencoder," *IEEE Access*, vol. 8, pp. 164130–164143, 2020.

[30] A. Madni, C. Madni, and S. Lucero, "Leveraging digital twin technology in model-based systems engineering," *Systems*, vol. 7, no. 1, 2019.

[31] Z. Allam and D. S. Jones, "Future (post-COVID) digital, smart and sustainable cities in the wake of 6G: Digital twins, immersive realities and new urban economies - ScienceDirect," *Land Use Policy*, vol. 101, 2021.

[32] S. Yrjölä, *Connecting You with Your Digital Twin: 6G will Bridge the Physical with the Virtual*, 2020.

[33] M. H. Alsharif, A. H. Kelechi, M. A. Albreem, S. A. Chaudhry, M. S. Zia, and S. Kim, "Sixth generation (6G) wireless networks: vision, research activities, challenges and potential solutions," *Symmetry*, vol. 12, no. 4, p. 676, 2020.

[34] K. Ding, F. T. Chan, X. Zhang, G. Zhou, and F. Zhang, "Defining a digital twin-based cyber-physical production system for autonomous manufacturing in smart shop floors," *International Journal of Production Research*, vol. 57, pp. 6315–6334, 2019.

[35] T. Wang, J. Li, Z. Kong, X. Liu, H. Snoussi, and H. Lv, "Digital twin improved via visual question answering for vision-language interactive mode in human–machine collaboration," *Journal of Manufacturing Systems*, vol. 58, no. 4, p. 24, 2020.

[36] J. Autiosalo, J. Vepsalainen, R. Viitala, and K. Tammi, "A feature-based framework for structuring industrial digital twins," *IEEE Access*, vol. 8, pp. 1193–1208, 2020.

[37] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the digital twin: a systematic literature review," *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, 2020.

[38] B. He, X. Cao, and Y. Hua, "Data fusion-based sustainable digital twin system of intelligent detection robotics," *Journal of Cleaner Production*, vol. 280, p. 124181, 2021.

[39] M. Xu, Y. Wu, Y. Ye, I. Farkas, H. Jiang, and Z. Deng, "Collective crowd formation transform with mutual information-based runtime feedback," *Computer Graphics Forum: Journal of the European Association for Computer Graphics*, vol. 34, no. 1, pp. 60–73, 2015.

[40] X. Dong, B. Yu, Z. Shi, and Y. Zhong, "Time-varying formation control for unmanned aerial vehicles: theories and applications," *IEEE Transactionson Control Systems Technology*, vol. 23, no. 1, pp. 340–348, 2015.

[41] J. Ren, X. Jin, and Z. Deng, "Crowd Formation Generation and Control," in *Handbook of Human Motion*, Springer, 2017.

[42] M. Xu, Y. Wu, and Y. Ye, "A smooth and efficient crowd transformation," in *Proceedings of the 20th ACM international conference on Multimedia*, pp. 1189–1192, Nara Japan, 2012.

[43] Q. Gu and Z. Deng, "Formation sketching: an approach to stylize groups in crowd simulation," in *Proceedings of the Graphics Interface 2011Conference*, pp. 1–8, Canada, 2011.

[44] M. Xu, H. Jiang, X. Jin, and Z. Deng, "Crowd simulation and its applications: recent advances," *Journal of Computer Science and Technology*, vol. 29, no. 5, pp. 799–811, 2014.

[45] S. Takahashi, K. Yoshida, T. S. Kwon, K. H. Lee, J. H. Lee, and S. Y. Shin, "Spectral-based group formation control," *Computer Graphics Forum*, vol. 28, no. 2, pp. 639–648, 2010.

[46] Y. Wang, X. P. Zhu, Z. H. Zhou, and H. Zhang, "UAV path following in 3-D dynamic environment," *Robot*, vol. 36, no. 1, pp. 83–91, 2014.

[47] X. Wang, D. Zhang, and L. Shen, "A virtual force based path following approach for unmanned aerial vehicles," *Robot*, vol. 38, no. 3, pp. 329–336, 2016.

[48] J. Chang and T. Li, "Simulating virtual crowd with fuzzy logics and motion planning for shape template," in *IEEE Conference on Cybernetics & IntelligentSystems*, Chengdu, China, 2008.

[49] Q. Gu and Z. Deng, "Generating freestyle group formations in agent-based crowd simulations," *IEEE Computer Society Press*, vol. 32, no. 1, pp. 20–31, 2013.

[50] A. Golas, R. Narain, S. Curtis, and M. Lin, "Hybrid long-range collision avoidance for crowd simulation," *IEEE Transactions on Visualization & Computer Graphics*, vol. 20, no. 7, pp. 1022–1034, 2014.

[51] M. Zhao, J. Zhong, and W. Cai, "A role-dependent data-driven approach for high density crowd behavior modeling," in *The 2016 annual ACM Conference*, pp. 1–25, ACM Transactions on Modeling and Computer Simulation (TOMACS), 2016.

[52] J. G. Ramirez, D. Lange, P. Charalambous, C. Esteves, and J. Pettré, "Optimization-based computation of locomotion trajectories for crowd patches," in *Proceedings of the Seventh International Conference on Motion in Games*, pp. 7–16, Playa Vista California, 2014.

[53] K. Jordao, P. Charalambous, M. Christie, J. Pettré, and M. P. Cani, "Crowd art: density and flow based crowd motion design," in *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*, pp. 167–176, Paris France, 2015.

[54] A. Barnett, H. P. Shum, and T. Komura, "Coordinated crowd simulation with topological scene analysis," *Computer Graphics Forum*, vol. 35, no. 6, pp. 120–132, 2016.

[55] W. G. Tollvan, *Navigation for Characters and Crowds in Complex Virtual Environments*, 2017.

[56] I. Errandonea, S. Beltrán, and S. Arrizabalaga, "Digital twin for maintenance: a literature review," *Computers in Industry*, vol. 123, p. 103316, 2020.

[57] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys*, vol. 46, no. 3, pp. 1–33, 2014.

[58] B. Moya, A. Badías, I. Alfaro, F. Chinesta, and E. Cueto, "Digital twins that learn and correct themselves," *International Journal for Numerical Methods in Engineering*, 2020.

[59] E. Yildiz, C. Møller, and A. Bilberg, "Virtual factory: digital twin based integrated factory simulations," *Procedia CIRP*, vol. 93, pp. 216–221, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5591387, 9 pages
https://doi.org/10.1155/2021/5591387

Research Article

Automatic Classification Algorithm for Multisearch Data Association Rules in Wireless Networks

Cailing Li and Wenjun Li

Linfen Vocational and Technology College, Linfen Shanxi 041000, China

Correspondence should be addressed to Cailing Li; licailing001@mail.sdufe.edu.cn

Received 29 January 2021; Revised 24 February 2021; Accepted 6 March 2021; Published 18 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Cailing Li and Wenjun Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to realize efficient data processing in wireless network, this paper designs an automatic classification algorithm of multisearch data association rules in a wireless network. According to the algorithm, starting from the mining of multisearch data association rules, from the discretization of continuous attributes of multisearch data, generation of fuzzy classification rules, and the design of association rule classifier and other aspects, automatic classification is completed by using the mining results. Experimental results show that this algorithm has the advantages of small classification error, good real-time performance, high coverage rate, and high feasibility.

1. Introduction

Along with the computer science and technology, especially the rapid development of database technology, as well as the expansion of the scope of human activities and life rhythm speeding up, people can more quickly and more easily, in a more cheap way, obtain and store data, making people increase the ability of generating and collecting data; this makes the data and information indices increase [1]. However, in the face of these extremely inflated numbers of data, it is difficult to see the real valuable information contained in the search, so that people constantly feel the huge pressure from "information explosion" and "data surplus." The proliferation of data hides a lot of information that plays a great role in the progress of human society. Wang proposes a secure data aggregation algorithm for a multidata query, which solves the problem of data missing by tracking the query data and safely tracking multiple data [2]. If these massive data are not used effectively, they will only become "data garbage." Therefore, how to find really useful information from a large number of wireless network multisearch data has become the focus of attention. Yogarajan and Revathi proposed an optimal mobile data collection method for wireless sensor networks. This method studies the classification and collection of mobile data, solves the problem of data classification, and simultaneously searches and classifies data globally, and provides technical support for data confidentiality [3].

In the face of "data explosion and lack of knowledge," data mining technology arises at the historic moment and develops vigorously, showing its strong vitality day by day. Data mining is the process of extracting or "mining" knowledge from a large amount of data, that is, extracting hidden information and knowledge that people do not know beforehand but are potentially useful from a large number of incomplete, noisy, fuzzy, and random data. Data mining bridges the gap between data and knowledge. It can not only describe the development process of historical data but also infer the current data and further predict the development trend of the future. Compared with traditional decision tree algorithms, association rules have higher classification and prediction accuracy. Therefore, in the field of data mining, association classification has been widely concerned. At present, China, Canada, the United States, and other countries have established the National Natural Science Foundation to study the classification of association rules. There are more scholars engaged in this field. With the deepening of research, association rules have become one of the cornerstones of data

mining and have been widely used in decision analysis, business management, and other fields. For example, today's database vendors, Oracle, IBM, and Microsoft, have also inherited the correlation analysis functionality in their products.

At present, the relevant research in China started relatively late, beginning in the early 1980s. At present, the research on text classification in China mainly has three stages,: the first is the feasibility study, the second is the auxiliary classification, and the last is the automatic classification. At the beginning, there were few researches on Chinese text classification, which mainly applied English classification technology to the Chinese research field. At the end of the 1990s, the taxonomy of textual texts began to be studied gradually. However, in the specific research, the characteristics of the Chinese text should be considered comprehensively in order to form a Chinese classification system. There are still many institutions in our country, such as Tsinghua University, Shanghai Jiao Tong University, Harbin Institute of Technology, Northeastern University, Institute of Computing Technology, Chinese Academy of Sciences, and Fudan University, which have done a lot of work in this area. Since 1951, many automatic classification systems have been developed in China. For example, Wu Jun et al. at Tsinghua University used weighted words as features of textual data and then used classification algorithms to construct classifiers. Using the SVM method, Zou Tao et al. from Nanjing University designed an automatic Chinese text classification system CTDCS. Liu Zhengying of Shanxi University and others developed an automatic financial classification system. Based on the neural network optimization algorithm, Wang Yongcheng et al. from Shanghai Jiao Tong University studied the Chinese text classification system. The important role of data mining technology in various fields has been increasingly prominent. Classification is an important analysis method in the field of data mining, while association rule mining and classification are an important research direction in the field of data mining. As two highly active research fields in data mining, they share certain similarities. Since then, the combination of these two important technologies, that is, the mining of association rules for the task of classification, will open a new journey of data classification, namely, association classification. Literature [4] puts forward the concept of "mutual confidence," highlighting the research results of association rules, making the research results of one-dimensional association rules more meaningful and attracting the attention of data mining workers. On the basis of studying the theory of negative association rules, this paper uses reverse thinking to improve the data mining algorithm based on negative association rules and gives examples, ideas, and experimental verification. The proposed algorithm is a useful supplement to the study of negative association rules. In literature [5], a new learning method called RMTFL is proposed. Firstly, a matrix is used to explain the correlation between multitasks and different features, and the feature space of related tasks is calculated by Group Lasso. According to the results, isolated tasks can be obtained. Aiming at the cross-domain text classification, Jin Xiaoming et al. from Tsinghua University proposed the cross-domain active learning method. Zheng Mingling et al. from Southeast University described the correlation between class criteria by referring to the Bayesian network, so that the learning problem of multiple class criteria was changed into a series of classification problems of single-class criteria, which made the mining performance of this algorithm on multiple data sets exceed the performance of existing methods. By Tsinghua University's Wang Jianyong discriminant model by studying the uncertainty data mining problems, in literature [6], the HARMONY algorithm is proposed; the algorithm has time-consuming features that do not need to choose, and from the database, directly find the discriminant model, which makes HARMONY a classic uncertainty than the SVM classification algorithm; the usage effect of the algorithm has larger ascension.

The innovation of this paper lies in the fact that association classification is based on association rules, which not only reflects the application characteristics of knowledge but also reflects the intrinsic correlation characteristics of knowledge. This process is mainly embodied in two aspects: the mining method of multisearch data association rules in a wireless network and how to analyze and classify the mining rules. Based on the above background, this paper designs an automatic classification algorithm for multisearch data association rules in wireless networks.

The research contributions of the thesis include the following:

(1) This paper designs an automatic classification algorithm for multiple search data association rules in wireless networks

(2) According to the automatic classification algorithm, the discretization of the continuous attributes of the multisearch data starts from mining the association rules of the multisearch data

(3) Fuzzy classification of fuzzy data and related design of association rule classifier

The organization structure of this article is as follows.

The first section discusses the introduction part of the paper; the second section discusses the classification algorithm; the third section conducts an experimental analysis; and the fourth section summarizes the paper.

2. Algorithm Design

2.1. Multisearch Data Association Rule Mining. Association rule mining technology is an important research content in the field of data mining. Its purpose is to find the correlation between various data items in the multisearch data set of a given wireless network, that is, to find out the frequently occurring items or attributes in the data set [7, 8]. Generally speaking, the mining process of association rules mainly includes two steps: first, generate all the frequent item sets in the wireless network multisearch data set, and second, generate association rules with the generated frequent item sets. Among them, generating frequent item sets is a key step that affects the quality of association rule mining results.

Figure 1 shows a basic association rule mining process.

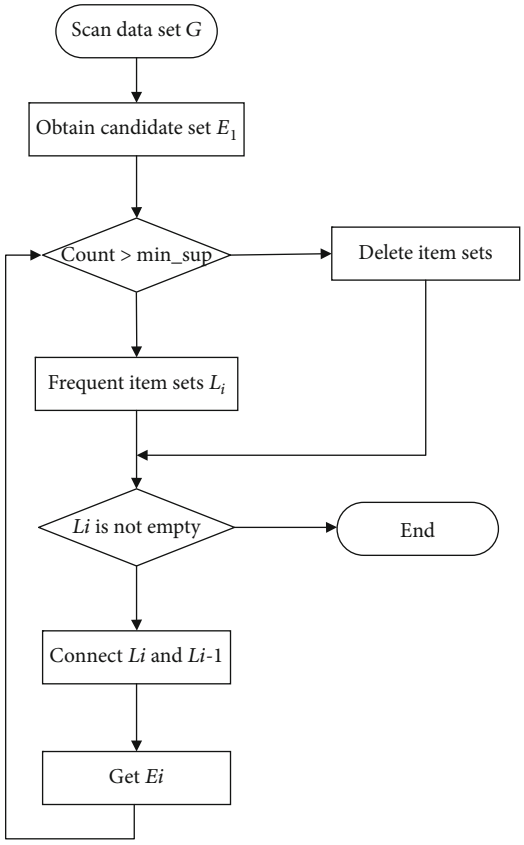


Figure 1: Basic association rule mining process.

Excavation was carried out in two steps:

(1) Firstly, the wireless network multisearch database G is scanned to store the identity TID of each item set, and a candidate set E_1 is generated. Then, the item set that does not meet the minimum support threshold is deleted from E_1 to generate the frequent item set L_1

(2) The loop executes until L_{i-1} is empty. Then, L_i and L_{i-1} are connected first to generate E_i, and a new identifier list can be obtained through the intersection of the identifier list. The count of the item set can be obtained from E_i, and then, the size between the count of E_i and the minimum support threshold min_sup is compared. The item sets greater than or equal to min_sup are kept, the rest are deleted, and the final output is frequent item set L

In the process of generating frequent item sets of the Apriori algorithm, there are two factors that affect the performance of the algorithm [9, 10]. First of all, in the process of iteration, the frequent i item set generated each time needs to scan the original wireless network multisearch database, so the database is scanned too many times, resulting in the algorithm performance decline. Secondly, in the process of tree pruning, it is necessary to scan the frequent $i-1$ set for the occurrence of candidate item sets, so the frequent item sets need to be scanned for many times, resulting in the decrease of algorithm efficiency.

```
Input: database (G), minimum support (min_ sup);
Output: frequent item set (L).
E1 = find_candidate_1-itemsets (G);
Int count = the number of TID in G;
For each item set s of E1 {.
   s.item-set = s;
   s.count = count of s in E1;
   s.Tid-list = the set of all TID includes s;
   If s.count<min_sup* count.
        delete s in E1;
}
L1 = E1;
for (i = 2; Li-1 ≠ ∅; i++){.
   for each item set 11 in Li-1.
      for each item set 12 in Li-1{.
         c.Tid-list =11,Tid-list;
         c.count = count TID in c.Tid-list;
}
      if c.count> = min_sup* count.
   add c to Ei;
   Li = Ei.
}
```

Algorithm 1: I-Apriori algorithm.

In view of the above problems, this study improves the process of frequent item set generation and proposes an improved Apriori algorithm, which is called as follows.

In the I-Apriori algorithm, in the process of generating candidate set E_i every time, in addition to storing item-set and support count, it is more important to add identifier list Tid-list. After the connection operation between item sets is completed, the list of identifiers and the count of item sets can be obtained directly through the Tid-list, and there is no need to rescan the wireless network multisearch database, thus effectively improving the performance of the algorithm.

2.2. Automatic Classification of Multisearch Data Association Rules. As the name suggests, association rule classification refers to the association rules used to distinguish or predict the tags of an instance class [11, 12]. It reflects not only the application characteristics of knowledge, namely, classification or prediction. It also reflects the intrinsic relevance of knowledge. General association rules can only be used to describe concepts, while the role of classification association rules has two concepts description and classification. Different from most other classification methods, association classification has better classification ability and better description ability.

Associative classification techniques generally involve two steps. The first step is to find all the classification association rules whose right part is the class label. The second step is to select the higher priority rule from the found classification association rules to classify the test set. The priority of rules is usually evaluated according to the confidence, support, rule length, or general quality criteria of classification association rules [13–15].

In this study, we propose an improved fuzzy associative classification algorithm, the fuzzy C-means clustering algorithm for continuous attribute fuzzy interval, obtaining high-quality fuzzy classification rules, and on this basis, to join a new pruning strategy to avoid generating useless rules, rules at the same time, using a new importance measure to fusion of multiple fuzzy classification rules, in order to improve classification accuracy.

2.2.1. Discretization of Continuous Attributes of Multisearch Data. The association classification algorithm can only deal with the discrete data directly, and the continuous attributes of the multisearch data need to be preprocessed and fuzzy. The fuzzy C-means clustering algorithm can be used to divide the fuzzy interval according to the distribution characteristics of the data. It is a clustering algorithm based on partition. Its idea is to make the similarity between objects divided into the same cluster maximum, while the similarity between different clusters minimum [16]. The effect of text classification is largely related to the characteristics of the data set itself. In practical research and application, it is generally believed that there is no certain method that can be suitable for data with various characteristics. So in practical application, to apply association classification to the field of text classification, we still have to face the following two problems. One is to use the characteristics of the document itself to improve the speed and performance of classification. This includes the use of document content and organization rules to select appropriate data sources and abstract description of problems. The second is to consider the form and structure of the document itself. This includes analyzing the content of the document. Generally, the document will contain information such as title, author, and text, while the scientific and technological document will contain some key words, and the news manuscript will have a specific time, and WEB documents will contain hyperlinks, tags, and other information. The complaint information will contain the identification of the handling unit of the complaint information. So in practical application, we should decide how to preprocess according to the form and structure of the document itself and specifically study how to improve the algorithm to improve the speed and accuracy of text classification. In this paper, we will rationalize the abstract description of the research problem and then optimize the algorithm according to the actual situation, so as to better realize the text classification problem based on association rules [17–20].

When fuzzy C-means clustering algorithm is used for fuzzy classification of attributes, it is necessary to set the number of fuzzy intervals, which is often given manually and has a great impact on classification accuracy. In order to achieve better fuzzy partition effect, the validity index of the PBM index is adopted to select the optimal interval number:

$$\mathrm{PBM}(K) = \left(\frac{1}{K} \times \frac{E_1}{E_i} \times G\right)^2, \quad (1)$$

wherein K represents the number of fuzzy intervals. On this basis, the PBM index is calculated iteratively, and the maximum PBM index is selected as the final result of the fuzzy interval discretization.

2.2.2. Generate Fuzzy Classification Rules. After fuzzy C-means clustering algorithm is used to fuzzy the continuous attributes of wireless network multisearch data, each continuous attribute is related to a group of fuzzy sets and has a corresponding fuzzy degree. From the definition of fuzzy classification rules, it can be seen that the rules used for fuzzy association classification are only a small subset of the whole set of fuzzy association rules, and the following items only have a single language value as a category label. In order to improve the efficiency of the algorithm, the following pruning strategies were added in this study:

(1) When the candidate frequent 2-item set is connected by the Apriori algorithm, the candidate set that does not contain classification label is deleted. At the same time, the classification label of each frequent 2-item set is moved to the head of each frequent 2-item set when the frequent 2-item set is connected by the Apriori algorithm

First, this policy ensures that each classification rule contains a unique classification label. When generating a candidate frequent 2-item set, this strategy first checks whether each generated candidate frequent 2-item set contains a classification tag, and if it does not, it removes it, thus ensuring that useless rules that do not contain a classification tag are removed. Then, for the generated frequent 2-item set, you reorder it, moving each category label to the head of the rule. According to the attributes of the Apriori algorithm, the frequent item sets generated after that will have at least one classification label, avoiding the generation of useless rules and thus effectively reducing the number of candidate item sets.

(2) When generating the candidate frequent i-item set, the item set that has more than one language values for an attribute in the candidate frequent i-item set is deleted. This policy ensures that each attribute in the candidate frequent i-item set corresponds to only one language value. If a certain attribute in a classification rule corresponds to multiple language values, it is easy to cause uncertainty in classification

The goal of the above two strategies is to reduce the generation of useless candidate sets as early as possible, thereby directly generating useful classification rules. They can greatly reduce the number of candidate sets, thus reducing the database scan time and ensuring the generation of effective fuzzy classification rules.

The process of mining fuzzy association classification rules is as follows:

(Step 1) get frequent 1-item set L_1

(Step 2) the candidate frequent 2-item set is generated and pruned

(Step 3) scanning the wireless network multisearch database, computing the support of candidate

frequent i-item sets, deleting the items that do not conform to the minimum support, and obtaining frequent i-item sets

(Step 4) generate candidate frequent item set E_i from frequent item set L_{i-1}

(Step 5) pruning and reordering candidate frequent item sets

(Step 6) return to Step 3 until no new candidate i-item sets are generated

(Step 7) delete frequent item sets whose trust level does not meet the minimum trust level set by the user.

There are many algorithms for association rules, but there is a classic algorithm, the Apriori algorithm, and most other algorithms are refinements or deductions of this algorithm. In essence, the Apriori algorithm is analyzed and considered to be a wide priority algorithm. The essence of the whole algorithm is that it needs to scan our transaction database, and then, in the scanning process, frequent transaction items are found, that is, frequent transaction item sets. It is worth noting that the algorithm is to find all the frequent item sets. Therefore, there will be multiple scans, and in each scan, only a set of items of the same length will be considered. That is, the number of items contained in the item set. This is what is called the k-item set. To give a concrete example, for example, during the first scan, the algorithm first finds all the item sets in the database that meet the minimum support of 1 item set and therefore has length 1 frequent item sets. Everybody knows it is L_1. After the first step, we need to further mine out the frequent 2-item set based on L_1, i.e., L_2. Continue the loop with this method to find the frequent 3 item sets until the frequent N item sets. The algorithm works until it can no longer find frequent item sets. One problem with this approach is that each time a frequent item set is generated, the entire transaction database needs to be scanned.

2.2.3. Wireless Network Multisearch Data Association Rule Classifier. After generating fuzzy classification rules, we need to use these rules to build a classification system. The key problem of classifier construction is how to determine the importance of different rules and finally set the label of the data to be classified.

The classification rule with the highest membership degree and the highest confidence degree can be simply regarded as the most important rule, so as to make use of this rule for classification, because confidence is a probability estimate of the importance of classification rules, and membership can be seen as the matching degree of the data to be classified to the rules. However, such algorithms adopt a single criterion or use a single rule to classify data, which may lead to paranoia about some categories, resulting in a decline in classification accuracy.

Taking the rule in Table 1 as an example, for the multisearch data I of a wireless network to be classified, rule C and rule D have the highest membership degree, and the confidence degree of rule D is higher than that of rule C.

Table 1: Examples of fuzzy association classification rules.

Rules	Degree of confidence	Support	Membership	Category
A	70	70	0.8	y_1
B	80	70	0.7	y_2
C	30	80	0.9	y_3
D	40	60	0.9	y_4

Table 2: Experimental environment hardware and software statistics table.

	Name	Parameter
Hardware	The processor	Dual core 3.4 GHz
	Memory	8 GB
	Hard disk space	10 TB
	Switches	100 M
Software	The operating system	Ubuntu 13.4
	Hadoop	1.1.8
	Java	jdk 7u21 linux i586
	IDE platform	Eclipse 4.1

Table 3: Comparison results of classification errors of the three algorithms (%).

Number of experiments (times)	Algorithm of reference [4]	Algorithm of reference [5]	Algorithm of this paper
1	3.9	4.2	1.0
2	3.2	4.1	1.5
3	2.8	3.8	1.8
4	2.7	3.5	1.7
5	2.5	4.5	1.2

Therefore, you can use rule E to divide data I into class y. However, although the membership degree of rules A and B is low, their confidence and support degree are higher than that of rule D. Therefore, it seems more reasonable to use rule A or B to classify wireless network multisearch data I into y_1 or y_2. In order to make more reasonable use of multirules for classification judgment, a new fuzzy classification rule judgment metric is introduced, as shown in the following formula:

$$\omega_{(I,R)} = F_{(I,R)} \times S_R \times (C_R)^2, \tag{2}$$

where $F_{(I,R)}$ represents the membership degree, which represents the matching degree of rule R for the pending wireless network multisearch data association rules. S_R is support, and C_R is trust.

In this study, the three parameters were used to measure the importance of fuzzy classification rules. On the one hand, the new measurement integrated membership, support, and

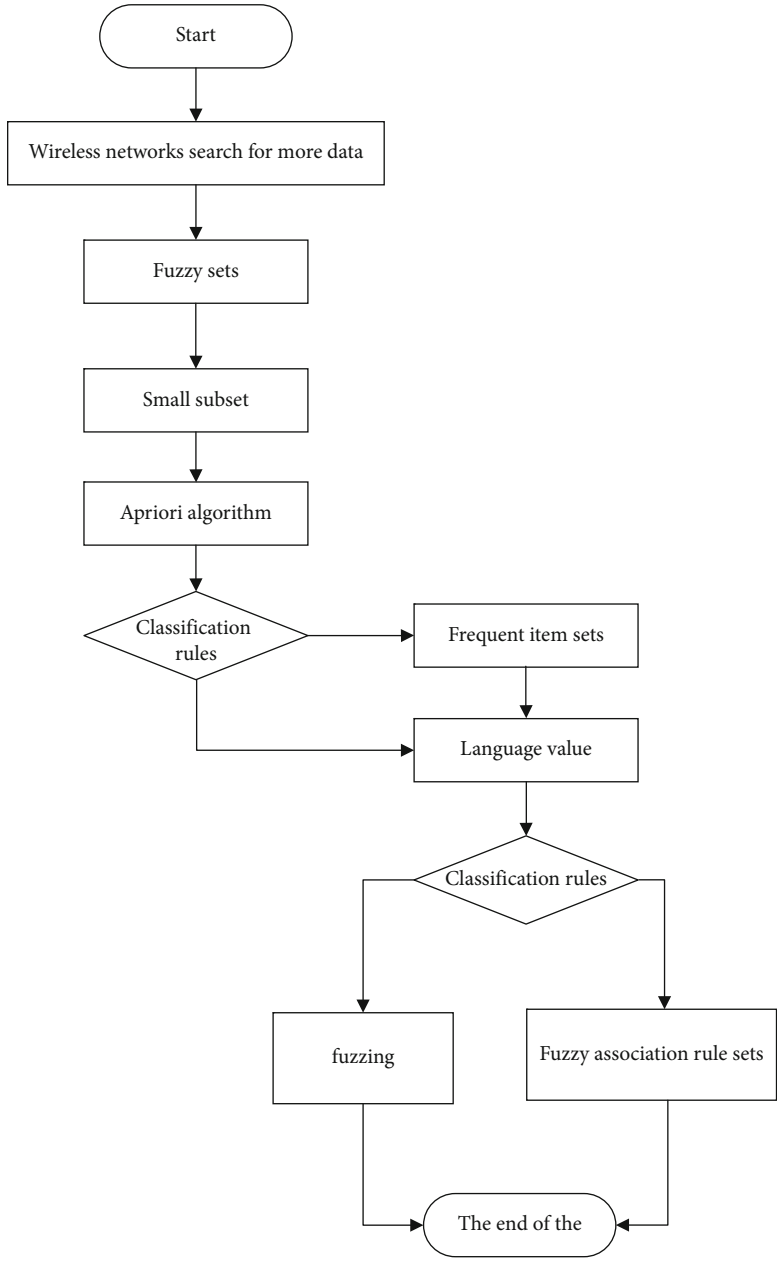


Figure 2: Mining of fuzzy association rule set.

trust to measure the importance of fuzzy classification rules, rather than simply relying on a certain parameter, which can better avoid the paranoid about some categories. On the other hand, on the basis of this determination measure, the classification data can be classified by using the whole fuzzy association classification set, which can effectively improve the classification accuracy of the classifier and avoid the occurrence of overfitting problem.

According to the new determination measure obtained above, rule B has the highest importance $\omega_{\langle I,R\rangle} = 0.3136$, so the data to be classified is classified as y_2.

After using Equation (2) to calculate the importance of each rule, the following formula is used to calculate the average weight of each category. Finally, the category with the largest average weight is selected as the classification category to complete the classification process.

$$\delta_{yj} = \frac{\sum_r \omega_{\langle I,R\rangle}}{N}, \text{where r} \in \text{R and cons}(r) = y_j. \quad (3)$$

In conclusion, the automatic classification algorithm for multisearch data association rules in wireless network is designed.

3. Experiment and Analysis

In order to verify the feasibility of the automatic classification algorithm of multisearch data association rules in wireless network, a simulation experiment is designed to prove it.

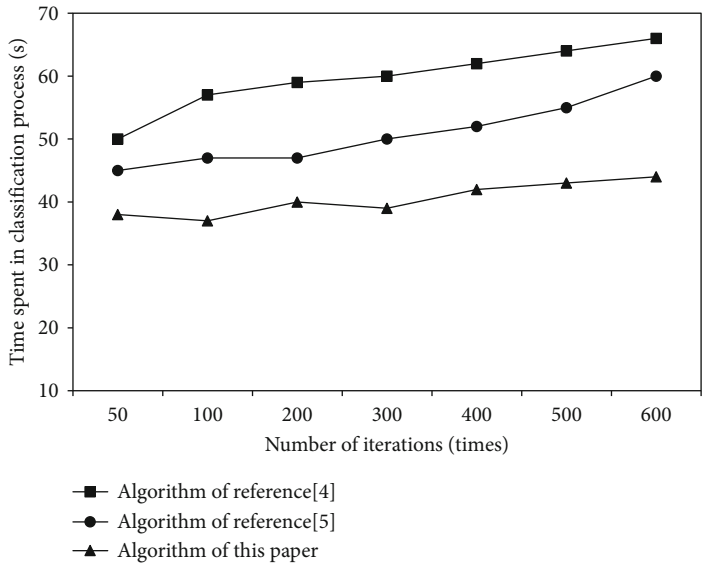


Figure 3: Comparison of time spent in the classification process of different algorithms.

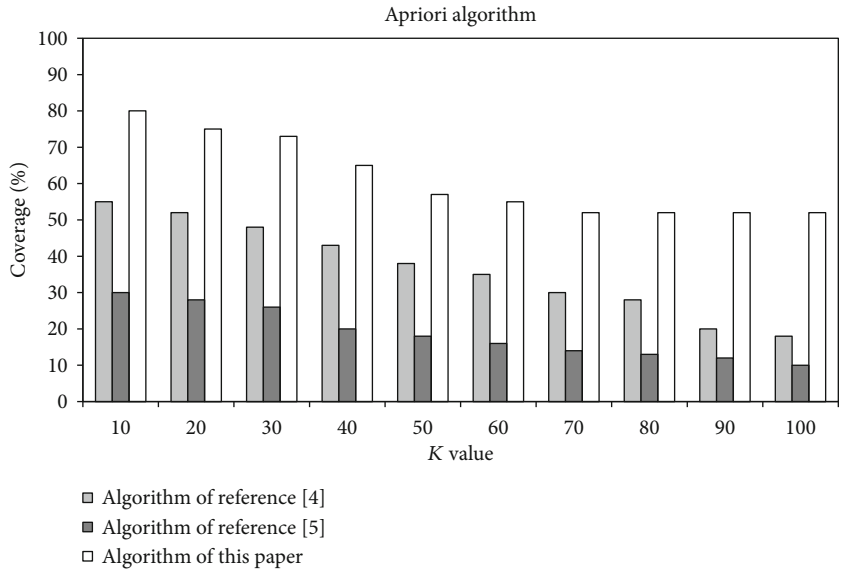


Figure 4: Comparison of coverage of the three algorithms.

3.1. Experimental Environment Design. The software and hardware information of the simulated experimental environment is shown in Table 2.

In order to avoid the uniformity of experimental results, the association rule mining classification algorithm based on FP-growth algorithm in reference [4] was compared with the bit table-based association rule mining and association classification algorithm in reference [5], and the performance verification was completed together with the algorithm in this paper.

3.2. Comparative Analysis of Classification Errors. The classification errors of the three algorithms are compared and analyzed, and the results are shown in Table 3.

As can be seen from Table 2, the classification error of the reference algorithm [4] is lower than that of the reference algorithm [5], but higher than that of the proposed algorithm. The maximum classification error of the algorithm in reference [5] is 4.5%, while the maximum classification error of the proposed algorithm is only 1.8%, which can prove that the proposed algorithm is more effective in classification.

3.3. Comparison of Time Spent in the Classification Process. The classification process time of the three algorithms was compared and analyzed, and the results were shown in Figure 2.

As can be seen from Figure 3, the classification process time of the three algorithms increases with the increase of

the number of experimental iterations. Among them, through the application of the algorithm in this paper, the upward trend of classification process time is not obvious, and it always keeps a relatively low state. The other two methods are obviously less stable than the method in this paper. This indicates that the proposed algorithm has high timeliness of classification and lays a good foundation for the subsequent multisearch data processing in wireless networks.

3.4. Coverage Ratio. Coverage can reflect whether the classification results are widely distributed in the multisearch data set of wireless network. The coverage rates of the three algorithms are compared and analyzed, and the results are shown in Figure 3.

According to Figure 4, with the increase of K value, the coverage rates of the three algorithms all show a downward trend. When the value of K is greater than 60, the data curve of the algorithm in this paper is gradually stable, while the other two comparison algorithms do not show a trend of data stability, and the curve is not stable. The average coverage rate of the proposed algorithm is higher than that of the two methods.

4. Conclusion

This paper designs an automatic classification algorithm for multisearch data association rules in wireless networks. The algorithm starts from the mining of multisearch data association rules and uses parameters to measure the importance of fuzzy classification rules. According to the new measure, the importance of fuzzy classification rules is measured by integrating membership, support, and trust. Automatic classification is accomplished by discretizing continuous attributes of multisearch data with mining results, generating fuzzy classification rules, and designing association rule classifier. Through experiments, this study also proves that the algorithm has low classification error, high classification timeliness, and high coverage rate, indicating that the algorithm can effectively realize automatic classification of data association rules.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Azmi, G. C. Runger, and A. Berrado, "Interpretable regularized class association rules algorithm for classification in a categorical data space," *Information Sciences*, vol. 52, no. 7, pp. 183–196, 2019.

[2] Y. Wang, "Secure data aggregation algorithm for multi-data query in wireless sensor networks," *International Electronic Elements*, vol. 27, no. 7, pp. 138–143, 2019.

[3] G. Yogarajan and T. Revathi, "Nature inspired discrete firefly algorithm for optimal mobile data gathering in wireless sensor networks," *Wireless Networks*, vol. 24, no. 8, pp. 2993–3007, 2018.

[4] Y. Chen, "Study on association rule mining algorithm based on FP-GROWTH algorithm," *Wireless Internet Technology*, vol. 19, pp. 118–121+124, 2017.

[5] J. Dong, *Research on Association Rule Mining and Association Classification Based on Bit Table*, vol. 42, no. 8, 2019Dalian University of Technology, 2019.

[6] Z. Hong-Sen, L. I. Liang, A. O. Chen, P. Zhang, N. Li, and Z. Wang, "Task classification-aware data aggregation scheduling algorithm in wireless sensor networks," *Science Technology and Engineering*, vol. 27, no. 3, pp. 156–170, 2019.

[7] U. Majeed, "Blockchain-assisted ensemble federated learning for automatic modulation classification in wireless networks," in *Proceeding of the KIISE Korea Computer Congress*, USA, Germany, 2020.

[8] X. Wang, Q. Liu, Z. Pan, and G. Pang, "APT attack detection algorithm based on spatio-temporal association analysis in industrial network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 45, no. 12, pp. 688–697, 2020.

[9] S. Mabu, T. Higuchi, and T. Kuremoto, "Semi supervised learning for class association rule mining using genetic network programming," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 15, no. 3, pp. 397–404, 2020.

[10] B. Subbulakshmi and C. Deisy, "An improved incremental algorithm for mining weighted class-association rules," *International Journal of Business Intelligence and Data Mining*, vol. 13, no. 1/2/3, pp. 291–308, 2018.

[11] H. Hongben, "Research on feature classification method of network text data based on association rules," *International Journal of Computers & Applications*, vol. 17, no. 4, pp. 2235–2243, 2018.

[12] L. Jing, "Analysis and improvement of classification based on multiple association rules," in *Proceedings of the 2018 2nd International Conference on Artificial Intelligence: Technologies and Applications (ICAITA 2018)*, USA, Germany, March 2018.

[13] C. An, Y. K. An, S. M. Yoo, and B. E. Wells, "Efficient data association to targets for tracking in passive wireless sensor networks," *Ad Hoc Networks*, vol. 75-76, no. 1, pp. 19–32, 2018.

[14] Z. Shao, Y. Li, X. Wang, X. Zhao, and Y. Guo, "Research on a new automatic generation algorithm of concept map based on text analysis and association rules mining," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 2, pp. 539–551, 2020.

[15] R. Agarwal and M. Mittal, "Inventory classification using multi-level association rule mining," *International Journal of Decision Support System Technology*, vol. 11, no. 2, pp. 1–12, 2019.

[16] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Transactions on Information and Systems*, vol. E102.D, no. 7, pp. 1374–1383, 2019.

[17] F. Wen, F. Min, Y. J. Zhang, and C. Yang, "Crude oil price shocks, monetary policy, and China's economy," *International Journal of Finance & Economics*, vol. 24, no. 2, pp. 812–827, 2019.

[18] Z. Tingting, "Application of data mining technology based on association rules in music classification," *Modern electronic technology*, vol. 43, no. 1, pp. 107–109, 114, 2020.

[19] A. Poordavoodi, M. Reza, H. Haj, A. M. Rahmani, and M. Izadikhah, "Toward a More Accurate Web Service Selection Using Modified Interval DEA Models with Undesirable Outputs," *Computer Modeling in Engineering & Sciences*, vol. 123, no. 2, pp. 525–570, 2020.

[20] L. Mingjiang, "Automatic classification technology of optical network fault data based on support vector machine," *Laser Magazine*, vol. 40, no. 9, pp. 111–115, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5594553, 9 pages
https://doi.org/10.1155/2021/5594553

Research Article

Multiobjective Sensor Ontology Matching Technique with User Preference Metrics

Hai Zhu[1], **Xingsi Xue**[2], **Chengcai Jiang**[3] **and He Ren**[3]

[1]*School of Network Engineering, Zhoukou Normal University, Zhoukou, Henan 466001, China*
[2]*Intelligent Information Processing Research Center, Fujian University of Technology, Fuzhou, Fujian 350118, China*
[3]*College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan 030024, China*

Correspondence should be addressed to Xingsi Xue; jack8375@gmail.com

Received 28 January 2021; Revised 19 February 2021; Accepted 7 March 2021; Published 16 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Hai 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.

Due to the problem of data heterogeneity in the semantic sensor networks, the communications among different sensor network applications are seriously hampered. Although sensor ontology is regarded as the state-of-the-art knowledge model for exchanging sensor information, there also exists the heterogeneity problem between different sensor ontologies. Ontology matching is an effective method to deal with the sensor ontology heterogeneity problem, whose kernel technique is the similarity measure. How to integrate different similarity measures to determine the alignment of high quality for the users with different preferences is a challenging problem. To face this challenge, in our work, a Multiobjective Evolutionary Algorithm (MOEA) is used in determining different nondominated solutions. In particular, the evaluating metric on sensor ontology alignment's quality is proposed, which takes into consideration user's preferences and do not need to use the Reference Alignment (RA) beforehand; an optimization model is constructed to define the sensor ontology matching problem formally, and a selection operator is presented, which can make MOEA uniformly improve the solution's objectives. In the experiment, the benchmark from the Ontology Alignment Evaluation Initiative (OAEI) and the real ontologies of the sensor domain is used to test the performance of our approach, and the experimental results show the validity of our approach.

1. Introduction

Nowadays, sensors are playing a more and more vital role in the distributed systems, and the sensor network has been broadly used in many areas of society, such as military, industrial, smart home, and many other fields [1, 2]. However, the lack of annotations on the sensor information with the semantic meaning is one of the main obstacles to implement the sensor infrastructure [3]. Due to the data heterogeneity issue in the sensor networks, the communications among different sensor network applications are seriously hampered. To integrate the information in different sensor networks, the Semantic Sensor Network (SSN) [4] is proposed, which provides the unified and standard sensor information styles to enhance the semantic connectivity, share sensor data, and improve interoperability. The Semantic Web (SW) provides a knowledge representation of a conceptual reference model, known as an ontology, to restrict domain terms [5, 6], and SSN [7] combines the traditional sensor network with SW's knowledge representation, reasoning, and organizational capabilities. As the key component of SSN, sensor ontology is regarded as the advanced knowledge model for exchanging sensor information.

For the last few years, many sensor ontologies are being progressed, among which the most famous one is SSN ontology [8]. SSN ontology describes the function and properties of sensors, measurement process, and sensor deployment, as well as the observation value of sensors and the method to get the observation value. There have been many kinds of research on sensor data query based on SSN ontology, such as environmental detection systems based on the Internet of Things (IoT) and agricultural ontology systems [9, 10]. SOSA (Sensor, Observation, Sample, Actuator) [11] provides SSN with a lightweight core that aims at expanding the target audience and application areas. In addition, SOSA defines

common classes and attributes as the minimum level of interoperability fallback, that is, data for these common classes and attributes can be securely exchanged across all SSN, SSN modules, and SOSA usage. These sensor ontologies are able to describe the sensor's capabilities, performance, and use conditions, which allow different data to be found in different contexts and used for different purposes. However, since different sensor ontologies are built by domain experts with different backgrounds, the same concept might be represented by different names in different domains, or the same name may be depicted by different concepts in different domains, which generates the sensor ontology heterogeneity problem. The sensor ontology heterogeneity problem is the biggest barrier that hampers the interaction and collaboration between smart applications based on sensor ontologies. Sensor ontology matching is an effective way in solving the sensor ontology heterogeneity problem, which aims at confirming the corresponding relationship between entities of different ontologies. When there are thousands of entities in the ontologies, matching these ontologies manually is a long-time and error-prone task. Therefore, many automatic and semiautomatic matching systems have been developed, which utilize the similarity measures to determine the ontology alignment.

The similarity measure is critical to an ontology matching technique, which can be divided into three types of categories, i.e., string distance-based similarity measure, linguistic-based similarity measure, and structure-based similarity measure [12]. A single similarity measure is not able to effectively measure the similarity in various scenarios with different heterogeneity, so aggregating multiple similarity measures is a must to improve the result's confidence [13, 14]. A popular way of aggregating similarity measures is the weighted sum strategy, where different similarity values are integrated to form a comprehensive similarity. The process of determining the suitable aggregating weights is referred to as ontology metamatching.

The ontology metamatching problem as a nonlinear problem often has many locally optimal solutions, and from this point of view, Evolutionary Algorithms (EA) can be used as a good way to address it [15, 16]. GOAL (Genetics for Ontology Alignment) [17] is the first one that uses Genetic Algorithms (GA) to deal with similarity integration problems. Given a Reference Alignment (RA), GOAL determines the optimal parameter set through GA to integrate different similarity measures to obtain the optimal ontology matching results. Because GOAL needs to obtain the reference matching result in advance, so it is difficult to apply it in practical applications. To overcome this drawback, this work proposes user preference-based evaluation metrics on sensor ontology alignment's quality, which is able to work without the RA.

Besides, many EA-based ontology matching techniques focus on a signal optimization objective, which ignores different users' preferences on the solutions. To meet different users' requirements, this paper proposes a multiobjective sensor ontology matching technique, which uses the Multiobjective Evolutionary Algorithm (MOEA) to simultaneously optimize the inflection point solutions [18], i.e., three solutions with the best recall, precision, and f-measure [19]. In particular, the main contributions of this paper are summarized as follows:

(i) The evaluating metrics on sensor ontology alignment's quality is proposed, which takes into consideration user's preferences and do not need to use RA

(ii) A sensor ontology optimization model is constructed to formally define the sensor ontology matching problem

(iii) An MOEA-based multiobjective sensor ontology matching technique is proposed, which uses a selection operator to improve the solution's objectives

The rest of the paper is organized as follows: Section 2 devoted to introducing the basic concepts on the sensor ontology and the evaluation metrics on ontology alignment; Section 3 gives an optimization model and describes how to use MOEA/D to select the user's three preferred solutions; Section 4 presents the experimental results and analysis; finally, Section 5 draws the conclusion and presents the future work.

2. Sensor Ontology Matching Problem and Similarity Measure

2.1. Sensor Ontology Matching Problem. The sensor ontology is defined as a triple $O = (C, P, I)$, where C is the set of classes, P is the set of properties, and I is the set of instances. Typically, class, property, and instance are collectively known as entities. Figure 1 shows the sensor ontology matching process, where O_1 and O_2 are two sensor ontologies to be matched, respectively; A is the standard results of ontology matching, namely, ontology alignment, and it is optional; A' is the ontology aligned by the ontology matching system, which is a set of mapping elements; r is an external resource; and p is a set of parameters, including the weights of the various similarity measures and the threshold used to filter out matching pairs with low confidence. Each element of A' is a 4-tuple (e, e', n, r), where e and e' are entities in different ontologies, respectively, n is the confidence of similarity between two entities, and r is the relationship between two entities, typically the equivalence.

2.2. Similarity Measure

2.2.1. Lexical-Based Similarity Measure. The lexical-based similarity measure, or the string-based similarity measure, determines the semantic similarity between entities by calculating the morphological similarity of the entities. In particular, for sensor ontology modeled by OWL language, the terms that are used to calculate the similarity are as follows: (1) the ID of the entity in the sensor ontology, also known as the local name; (2) the tag of the ontology entity with OWL syntax "rdfs: label"; and (3) the comment of the ontology entity with OWL syntax "rdfs: comment." Typically, the ID, label, and comment of an entity within the sensor ontology are called the textual contents. There are many lexical-based similarity measures, such as N-gram [20], Similarity Measure for Ontology Alignment (SMOA) [21], and Levenshtein [22]. According to [23, 26], N-gram and SMOA have better performance when solving ontology matching

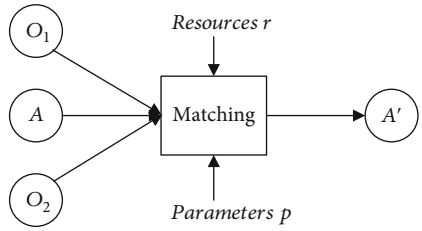


Figure 1: Sensor ontology matching process.

problems; therefore, they were used in this paper, which is, respectively, defined as follows:

$$\text{N-gram}(s_1, s_2) = \frac{2 \times \text{comm}(s_1, s_2)}{N_{s_1} + N_{s_2}}, \tag{1}$$

where s_1 and s_2 are two strings and split into substrings with three characters, $\text{comm}(s_1, s_2)$ is the number of the same substrings between s_1 and s_2, and N_{s_1} and N_{s_2}, respectively, represent the number of the substring in the two strings.

Unlike N-gram, SMOA takes into account the difference and same between the two strings, which is defined as follows:

$$\text{SMOA}(s_1, s_2) = \text{comm}(s_1, s_2) - \text{diff}(s_1, s_2) + \text{winklerImpr}(s_1, s_2), \tag{2}$$

where s_1 and s_2 are the two strings, $\text{comm}(s_1, s_2)$ is a measure of the similarity between s_1 and s_2, $\text{diff}(s_1, s_2)$ is a measure of their difference, and $\text{winklerImpr}(s_1, s_2)$ is an improved approach proposed in [24].

2.2.2. Linguistic-Based Similarity Measure. The linguistic similarity measure is used for determining the semantic distance between two words by considering the synonym or the relation between the hypernym and hyponym. It requires the utilization of external resources, such as WordNet [25]. Unlike other electronic dictionaries, WordNet is an English electronic dictionary based on cognitive linguistics, which not only ranks words but also forms a "web of words" according to their meanings. In WordNet, all words are constructed in hierarchical relationships, with the same depth representing a closer relationship. Conversely, the greater the depth difference, the more alienated the relationship between the two words.

This work uses the Wup similarity measure, which is provided by Wu and Palmer [26] and works based on this principle that the path length is from the closest common ancestor node to the root node. Given two strings s_1 and s_2, their similarity computing method is as follows:

$$\text{Wup}(s_1, s_2) = \frac{2 \times \text{depth}(\text{LCA}(s_1, s_2))}{\text{depth}(s_1) + \text{depth}(s_2)}, \tag{3}$$

where $\text{LCA}(s_1, s_2)$ stands for the closest common parent concept between s_1 and s_2 in WordNet, and $\text{depth}(s)$ is the hierarchy depth of s in the WordNet.

2.2.3. Structure-Based Similarity Measure. Structure-based similarity measure makes use of the context information to calculate two classes' similarity. This work uses the SimRank model [27] to calculate the structure-based similarity value, which is a graph-based topological information model for measuring similarity between two classes. The motivation behind SimRank is that if two classes are referenced by similar classes (that is, they have similar upper and lower nodes), they are also similar, i.e., the similarity is contagious in SimRank. This work uses the matrix form of SimRank [28] which is defined as follows:

$$\mathbf{S} = C \cdot \left(\mathbf{W}^T \cdot \mathbf{S} \cdot \mathbf{W}\right) + (1 - C) \cdot \mathbf{I}. \tag{4}$$

$\mathbf{S}$ is the SimRank similarity matrix whose elements and two variables $\mathbf{S}$ are, respectively, an updated one (on the right-hand side) and an old one (on the right-hand side); $[\mathbf{S}]_{a,b}$ is the distance $s(a, b)$ between note a and note b from two ontologies, respectively; $\mathbf{W}$ is a column-normalized adjacency matrix whose element is $[\mathbf{W}]_{a,b} = 1/|\mathfrak{T}(b)|$ where the mark $\mathfrak{T}(b)$ denotes the set of nodes that point to a node, if there is a directed edge (a, b), otherwise 0; $\mathbf{I}$ is an identity matrix, while C is an attenuation factor, which usually takes 0.8. The similarity value is initialized as follows:

$$s_0(a, b) = \begin{cases} 1, a = b, \\ 0, a \neq b. \end{cases} \tag{5}$$

Equation (4) is used for updating the value of the similarity value in the matrix, and the number of cycles $k = 5$.

A single similarity measure is not able to ensure its effectiveness in all matching tasks, which is due to the complex intrinsic of entity heterogeneous features. To enhance the confidence of the similarity value, it is necessary to aggregate multiple similarity measures. Next, we formally define the sensor ontology metamatching problem, which studies how to determine the optimal aggregating weights for various similarity measures.

3. Sensor Ontology Metamatching Problem

In general, the optimization problem can be divided into the unconstrained optimization problem and constrained optimization problem according to whether there are constrained conditions. The unconstrained optimization problem is referred for solving the optimal objective under the condition of infinite resources, while the constrained optimization problem is to solve the optimal objective under the condition of limited resources. The unconstrained optimization problem as a special circumstance of the constrained optimization problem is short of restricted condition. In this work, the sensor ontology matching problem is regarded as a constrained optimization problem, and the following three points shall be considered before determining the optimization model:

(1) Decision variables, which refer to the undetermined constants and variables related to the constraints

and objective functions involved in the sensor ontology metamatching optimization problems

(2) The objective function, which refers to the function related to a variable and for which the extreme value (maximum or minimum) is to be found

(3) Constraint conditions, which refer to the conditions that variables must meet when finding the extreme value of an objective function

3.1. Decision Variable. The process of ontology metamatching is displayed in Figure 2, where $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$, and $\mathbf{M}_4$ are the similarity matrices corresponding to four different similarity measures, respectively, and w_1, w_2, w_3, and w_4 are, respectively, their aggregating weights. The final similarity matrix $\mathbf{M}$ can be obtained by integrating different similarity matrices and filtered by threshold v. In this work, we need to optimize four aggregating weights and one threshold, which consists of the decision variables.

3.2. Objective Function. The traditional evaluating metrics on alignment's quality are recall, precision, and f-measure, respectively, [17], which are defined as follows:

$$\text{recall}(A, R) = \frac{|R \cap A|}{|R|}, \tag{6}$$

$$\text{precision}(A, R) = \frac{|R \cap A|}{|A|}, \tag{7}$$

$$f\text{-measure}(A, R) = \frac{\text{recall}(A, R) \times \text{precision}(A, R)}{\alpha \times \text{recall}(A, R) + (1 - \alpha) \times \text{precision}(A, R)}, \tag{8}$$

where A is the RA which contains a set of standard entity pairs and R is the matching results we found which contains a set of founded matching pairs. The recall is used to test the completeness of alignment, and precision of the correct rate of found pairs. To get a more practical solution, recall and precision are integrated to form the f-measure metric. The α in Formula (8) is a preference coefficient, and when α is close to 1.0, the f-measure is closer to recall, and vice versa to precision. However, it is unrealistic to obtain the standard matching results in advance, especially when the ontology boasts a huge amount of data.

In this work, a group of evaluating metrics is proposed on the alignment's quality without using RA, which is taken as the optimization objective. There are three user preferences on the alignment, i.e., completeness, accuracy, and unbiasedness [29], which are defined, respectively, as follows:

$$\text{completeness}(\mathbf{M}) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} \Gamma([\mathbf{M}_{ij}])}{\min(m, n)}, \tag{9}$$

$$\Gamma([\mathbf{M}_{ij}]) = \begin{cases} 1, & \text{when element} - \boldsymbol{M}_{ij} \text{ are largest in } i\text{th line and } j\text{th column in } \mathbf{M} \text{ matrix,} \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

where $\mathbf{M}$ is the similarity matrix after aggregation and $[\mathbf{M}_{ij}]$ is the similarity of the ith row and the jth column from the $\mathbf{M}$.

$$\text{accuracy}(\mathbf{M}) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} \varphi([\mathbf{M}_{ij}])}{\min(m, n)}, \tag{11}$$

$$\varphi([\mathbf{M}_{ij}]) = \begin{cases} [\mathbf{M}_{ij}], & \text{when element} - \mathbf{M}_{ij} \text{ are largest in } i\text{th line and } j\text{th column in } \mathbf{M} \text{ matrix,} \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

After the synthesis matrix $\mathbf{M}$ is obtained by integrating various similarity measurement methods, if the accuracy value of the matrix $\mathbf{M}$ is higher, it can indirectly reflect the higher matching accuracy.

$$\text{unbiasedness}(\mathbf{M}) = 2 \times \frac{\text{completeness} \times \text{accuracy}}{\text{completeness} + \text{accuracy}}. \tag{13}$$

Finally, the objective to be optimized are defined as follows:

$$\max\ F(\mathbf{M}) = (\text{completeness}(\mathbf{M}), \text{accuracy}(\mathbf{M}), \text{unbiasedness}(\mathbf{M}))^T. \tag{14}$$

3.3. Constraint Condition. In this work, there are two constraints, which are on the aggregating weights and the threshold, respectively. According to the weighted sum strategy, the aggregating weight w_i corresponding to the ith similarity measure should comply with the constraint $w_i \in [0, 1]$, $\sum_i w_i = 1$. When the threshold's upper limit is higher than 0.8 or the lower limit is smaller than 0.1, it is not easy to ensure the alignment's quality. Thus, threshold v's range should be $[0.1, 0.8]$.

4. Interactive Multiobjective Sensor Ontology Matching Technique

4.1. Multiple-Objective Evolutionary Algorithm with Decomposition. Multiple-Objective Evolutionary Algorithm with Decomposition (MOEA/D) [30] as a popular MOEA algorithm decomposes the multiobjective optimization problem into a variety of single-objective optimization subproblems and optimizes them at the same time. Because of the decomposition operation, MOEA/D has a great advantage in keeping the distribution of the solution. However, when moving the solutions to the Pareto Front (PF), it might sacrifice some objectives. To uniformly improve the solution's objectives, this work proposes a selection operator, which prevents excessive deterioration of the final solution. The flowchart of MOEA/D is shown in Figure 3.

4.2. Objective Decomposition. In this work, there are two objectives that need to optimize, i.e., maximize the completeness and accuracy of the alignment. To this end, we define the number of the subproblems as N, and their distributed weight vectors are $(\alpha_i, \beta_j)^T = ((i-1)/(N-1), (N-i)/(N-1))^T$, $i = 2, 3, \cdots, N-1$. If the two weight vectors $\lambda_i = (\alpha_i, \beta_i)^T$ and $\lambda_j = (\alpha_j, \beta_j)^T$ are close to each other which depends on

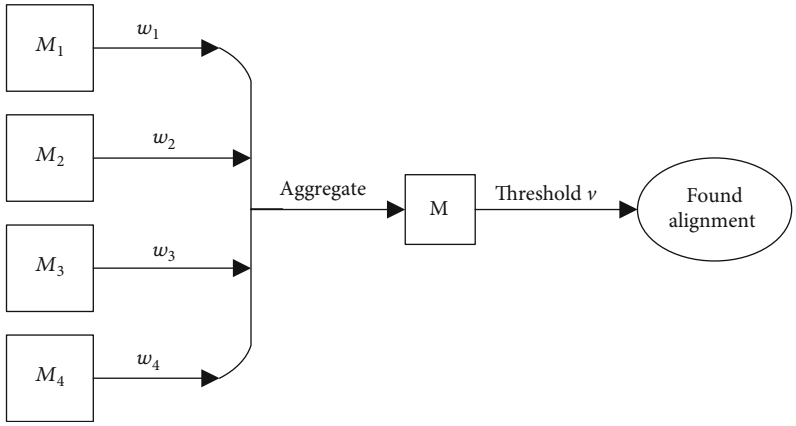


Figure 2: Similarity matrix aggregation.

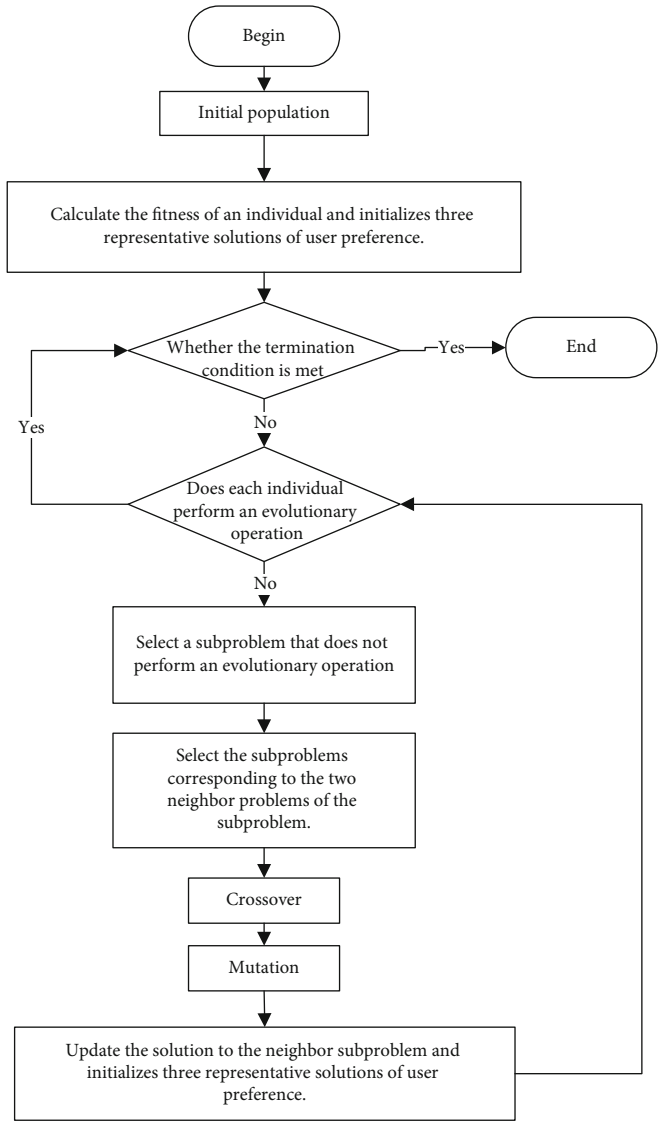


Figure 3: The flowchart of MOEA/D.

Euclidian distance, ith and jth subproblems are called neighbor problems, which can help each other to achieve optimization during the evolving process. The weighted integration strategy adopted in this paper is the Tchebycheff approach [28], which is able to convert the approximation problem of Pareto Front into several scalar optimization problems:

$$\begin{cases} \min\ g^{te}(x \mid \lambda, \varkappa^*) = \max\{\lambda_i | f_i(x) - \varkappa_i^* |\} \\ \text{s.t.} x \in \Omega, \end{cases} \tag{15}$$

where $\varkappa^* = (\varkappa_1^*, \varkappa_2^*)^T$ is the reference point, and we set $\varkappa^* = (1.0,1.0)^T$; $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)^T$ is a weight vector, and $\sum_{i=1}^{2} \lambda_i = 1$; Ω is the feasible solution space. Equation (15) is the criterion for updating the neighbor solution, i.e., supposing the newly generated solution is y and the original solution is x, if $g^{te}(y \mid \lambda, \varkappa^*) \le g^{te}(x \mid \lambda, \varkappa^*)$, x should be replaced by y.

4.3. Encoding Mechanism. This paper adopts the binary coding mechanism, in which a chromosome is divided into two parts. The first one is the combined weights of various similarity measures, respectively, and the second part is the threshold. The decoding process is shown in Figure 4. Let p be the number of similarity measures, the group of segmentation points can be expressed as $c' = \{c_1', c_2', \cdots, c_{p-1}'\}$. When decoding, firstly, the elements in c' are arranged in ascending order in this form $c' = \{c_1, c_2, \cdots c_{p-1}\}$, and then, the weight is calculated according to Equation (16). In particular, the second part does not need to execute the special decoding process.

$$w_k \begin{cases} c_1, & k = 1, \\ c_k - c_{k-1}, & 1 < k < p, \\ 1 - c_{p-1}, & \text{other}. \end{cases} \tag{16}$$

4.4. Genetic Operator. EA's selection operator aims at selecting individuals with high fitness and removing individuals with low fitness. According to MOEA/D, the selection operator works basing on the neighborhood of an individual in the population. First, the completeness and accuracy values of the current individual's neighbor individuals are obtained; then, the distance (d_i) of the solution of the neighbor individual from the origin point (see also Figure 5) is calculated; and finally, two neighbor individuals with a relatively long distance are selected for cross operation. In this work, the selection operator dedicates to improve the solution's quality evenly in terms of two objectives. The new selection strategy will be described in more detail below.

Because the user preference information is considered in this paper, it is necessary to select a better trade-off solution as an alternative and then perform crossover and mutation operations on such an alternative. In terms of the trade-off between solutions, uniform weight and solution are combined into an objective function. Firstly, the corresponding solutions and their neighbor solutions are selected, and then, the scores of the three solutions under the uniform weights are calculated. The calculation of the score is as follows: If there is a uniform weight vector $[h, 1-h]^T$, the score corresponding to the three solutions is score = $[h, 1-h]^T \times$ [completeness, accuracy]. Through the above method, the scores of each solution under a uniform weight can be calculated, and the two solutions with high scores are selected as the candidate solutions for the crossing and mutation operator.

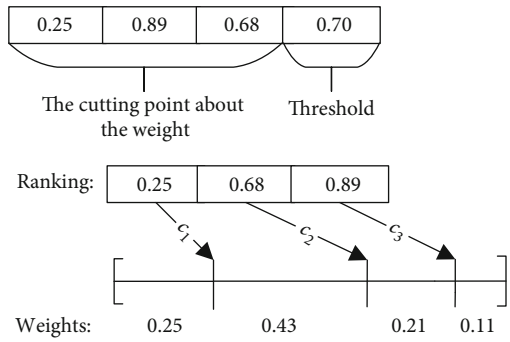


Figure 4: Decoding operation.

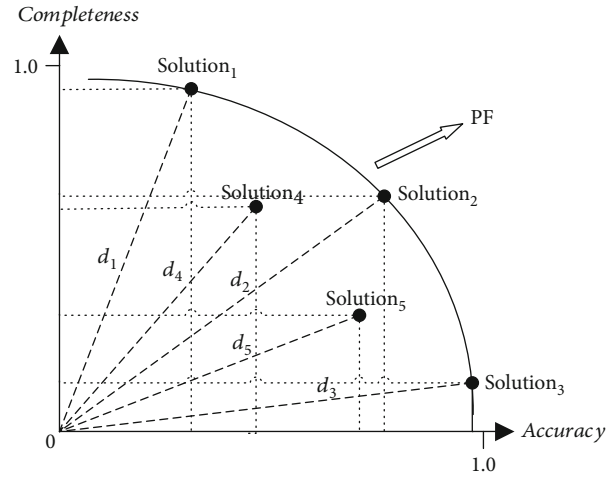


Figure 5: Process of selecting solutions.

The single point crossover and locus mutation were used as crossover and mutation strategy, respectively.

5. Experiment

5.1. Experiment Configuration. In this work, the testing cases we used consist of Ontology Alignment Evaluation Initiative (OAEI)'s (http://oaei.ontologymatching.org) benchmark and three real sensor ontologies, to test the effectiveness of our approach, whose brief descriptions are shown in Table 1.

The comparison systems in the experiment are from OAEI's participants and can be found on OAEI's official website, while the results of our approach are the mean value of 30 independent runs. The configuration of our approach is as follows:

(i) Similarity measures: N-gram, SMOA, WuAndPamer, and SimRank

Table 1: The brief description of testing cases.

ID	Brief description
101-104	The target ontology is identical to the source or different with some OWL constraints.
201-210	The linguistic features are different between the target ontology and the source ontology, but their structures are the same.
221-247	The conceptual structure is different between the target ontology and the source ontology, but their linguistic features are the same.
248-266	The target ontology and source ontology have different lexical, linguistic, and structural features.
301-304	Both the target ontology and the source ontology are derived from the ontology used in reality.
Old SSN	The descriptions of sensors, data, and concepts.
New SSN	There is some improvement on the old SSN ontology, making it more applicable.
SOSA	The lightweight core of SSN.

Table 2: Comparison in terms of precision.

	edna	AML	LogMap	LogMapLt	XMap	LogMapBio	IOMap	CroLOM	Our approach
101	0.64	1	0.94	0.56	0.94	0.5	1	1	1
201-202	0.259	1	0.962	0.307	0.959	0.343	1	1	0.994
221-247	0.754	1	0.889	0.6	0.97	0.518	1	1	1
248-266	0.285	1	0.959	0.348	0.975	0.256	1	1	0.898
Average	0.336	1	0.947	0.389	0.972	0.312	1	1	0.927

Table 3: Comparison in terms of f-measure.

	edna	AML	LogMap	LogMapLt	XMap	LogMapBio	IOMap	CroLOM	Our approach
101	0.78	0	0.95	0.71	0.97	0.62	0	0	1
201-202	0.315	0.231	0.474	0.345	0.391	0.267	0	0	0.745
221-247	0.831	0.754	0.814	0.726	0.961	0.598	0	0	1
248-266	0.318	0.36	0.436	0.338	0.413	0.202	0	0	0.556
Average	0.41	0.41	0.51	0.409	0.51	0.28	0	0	0.656

(ii) Number of subproblems: 3

(iii) Population scale: 40

(iv) Crossover probability: 0.5

(v) Mutation probability: 0.01

We use the above configuration empirically, which ensures our approach achieves the highest average results in all testing cases.

5.2. Results and Analysis

5.2.1. Testing on OAEI's Benchmark. The result we compared is from OAEI's participants in Tables 2–4, where the testing cases are divided into four parts, i.e., 101, 201-202, 221-247, and 248-266, according to the heterogeneity features of the ontologies. In Table 5, it is obvious that both edna and LogMapLt perform well in the testing cases 101 and 221 to 247, and our method's results are good as well. In the testing cases 201 to 202, none of the matching systems outperform others while our method's results are close to 0.7, which shows that our method is effective than other competitors in terms of recall. The testing cases in 201-210 are heterogeneous in terms of both lexical and linguistic. All the testing cases own different heterogeneous characteristics, which require cooperation among different similarity measures. As to why the IOMap and CroLOM systems have achieved such poor results, our analysis is as follows. The IOMap is a system to solve the cross-lingual problem. Similarly, the full name of the CroLOM system is the Cross-Lingual Ontology Matching System and it uses Yandex translators, NLP techniques, and similarity methods based on word and synonym categories. Both systems have achieved excellent results in dealing with cross-language ontology, but they cannot achieve satisfactory

Table 4: Real sensor ontology alignment.

	Recall	Precision	f-measure
SSN (old)-SSN (new)	0.97	1.0	0.98
SSN (old)-SOSA	0.97	1.0	0.98
SSN (new)-SOSA	1.0	1.0	1.0

Table 5: Comparison in terms of recall.

	edna	AML	LogMap	LogMapLt	XMap	LogMapBio	IOMap	CroLOM	Our approach
101	1	0	0.96	0.99	1	0.56	0	0	1
201-202	0.404	0.139	0.372	0.396	0.291	0.228	0	0	0.683
221-247	1	0.669	0.858	0.995	0.959	0.71	0	0	1
248-266	0.395	0.264	0.323	0.387	0.314	0.181	0	0	0.461
Average	0.505	0.317	0.426	0.498	0.429	0.28	0	0	0.582

results in the benchmark test set because the major of IOMap and CroLOM is cross-lingual. From the experimental results, we can see that our approach is able to better aggregate different similarity measures to determine more correct correspondences. This argument is further supported by the average recall.

In Table 2, we compare our approach in terms of precision. It can be seen that AML, IOMap, and CroLOM are able to determine the alignments with precision value 1.0. However, they achieve this goal by sacrificing recall. There is a contradiction between recall and precision, and when a matching technique tries to optimize one metric, it might deteriorate the other one. Since our approach is able to trade-off recall and precision, and although our approach's precision values are not the best, they are in general high. In particular, our approach's mean precision is only a bit lower than other competitors.

Table 3 presents all competitor's f-measure values. As can be seen from the table, IOMap and CroLOM, which have high precision and low recall, own low f-measure, and so do AML. Our approach is far ahead of other matching systems, which has the best results on the testing cases 101, 201 to 202, 221 to 247, and 248 to 266. In particular, our approach's results in the testing cases 101 and 221 to 247 are almost perfect, and even in the highly heterogeneous groups (201 to 202 and 248-266), our approach's f-measure values are still high. In contrast, none of the other matching systems' f-measure values are higher than 0.5. It can be seen from the average value that the alignments obtained by our method are much better than other matching systems. The reason behind this is that the suitable selection operator makes the nondeteriorating solution filtered out in the evolution process, which improves the result's quality. Furthermore, the solution selection strategy of inflection point based on user preference makes our approach be able to better trade-off the solution's recall and precision.

5.2.2. Testing on Real Sensor Ontologies. Three real sensor ontologies used in this work are, respectively, SSN (both new and old versions) and SOSA. SSN ontology is used to describe the processes and characteristics of the sensors and their observations, which follows a horizontal and vertical modular architecture. SSN and SOSA vary in scope and degree of the axiom and support a variety of applications, such as satellite imagery, large-scale scientific monitoring, industrial, smart-home infrastructure, social perception, citizen science, observation-driven ontology engineering, and the Internet of Things (IoT). The new SSN is different from the original SSN; the new SSN simplifies the relationship between device, platform, and system classes on the old SSN. New SSN supports humans and other animals as agents better permitting all major classes to be virtual.

Three pairs of real sensor ontology alignment are shown in Table 4. It can be seen that our approach can determine the perfect alignment when matching the new SSN (new) with SOSA. Since our method can only find the one-to-one matching relationship between the two ontologies, its recall and precision are affected on the other matching tasks. In general, our approach can determine high-quality sensor ontology alignments.

6. Conclusion and Future Work

Sensor ontology matching as a challenging problem needs to be solved, a new optimization model is constructed, and a multiobjective optimization algorithm based on user preference is used to find the appropriate solution. Our method is aimed at simultaneously determining different alignments for the users with different preferences. In order to test the performance of our approach, the benchmark provided by OAEI and three pairs of sensor ontologies are used. The experimental results show that our method is effective.

In tomorrow's work, we are interested in further lifting the performance of the algorithm by getting the expert involved. Besides, our method is not scalable, whose performance significantly drops with the growing ontology scale. A feasible solution would be the introduction of some divide-and-conquer strategy to improve the scalability of our approach.

Data Availability

The data used to support this study can be found in http://oaei.ontologymatching.org.

Conflicts of Interest

The authors declare that they have no conflicts of interest in the work.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Nos. 61801527 and 61103143) and the Natural Science Foundation of Fujian Province (No. 2020J01875).

References

[1] A. Bröring, J. Echterhoff, S. Jirka et al., "New generation sensor web enablement," *Sensors*, vol. 11, no. 3, pp. 2652–2699, 2011.

[2] D. J. Russomanno, C. R. Kothari, and O. A. Thomas, "Building a sensor ontology: a practical approach leveraging ISO and OGC models," in *Proceedings of the 2005 International Conference on Artificial Intelligence*, pp. 637–643, Las Vegas, Nevada, USA, 2005.

[3] A. Bröring, K. Janowicz, C. Stasch, and W. Kuhn, "Semantic challenges for sensor plug and play," in *Web and Wireless Geographical Information Systems. W2GIS 2009*, pp. 72–86, Springer, 2009.

[4] A. Sheth, C. Henson, and S. S. Sahoo, "Semantic sensor web," *IEEE Internet Computing*, vol. 12, no. 4, pp. 78–83, 2008.

[5] C. W. B. Smith, "Formal ontology and information systems," *Telemedicine & Virtual Reality*, vol. 3, no. 5, pp. 309–321, 1993.

[6] W. Kuhn, "Semantic engineering," in *Research Trends in Geographic Information Science*, Springer, Berlin Heidelberg, 2009.

[7] X. Xue, "A compact firefly algorithm for matching biomedical ontologies," *Knowledge and Information Systems*, vol. 62, no. 7, pp. 2855–2871, 2020.

[8] M. Compton, "What now and where next for the W3C semantic sensor networks incubator group sensor ontology," in *International Conference on Semantic Sensor Networks*, pp. 1–8, Bonn, Germany, 2011.

[9] F. Lécué, S. Tallevi-Diotallevi, J. Hayes et al., "Smart traffic analytics in the semantic web with STAR-CITY: scenarios, system and lessons learned in Dublin City," *Journal of Web Semantics*, vol. 27-28, pp. 26–33, 2014.

[10] A. Khamparia and B. Pandey, "Performance analysis on agriculture ontology using SPARQL query system," in *2014 International Conference on Data Mining and Intelligent Computing (ICDMIC)*, pp. 1–5, Delhi, India, 2014.

[11] K. Janowicz, A. Haller, S. J. D. Cox, D. Le Phuoc, and M. Lefrançois, "SOSA: a lightweight ontology for sensors, observations, samples, and actuators," *Journal of Web Semantics*, vol. 56, pp. 1–10, 2019.

[12] J. David, F. Guillet, and H. Briand, "Matching directories and OWL ontologies with AROMA," in *Proceedings of the 15th ACM international conference on Information and knowledge management - CIKM '06*, pp. 830-831, Arlington Virginia, USA, 2006.

[13] X. Xue and J. Chen, "Optimizing ontology alignment through hybrid population-based incremental learning algorithm," *Memetic Computing*, vol. 11, no. 2, pp. 209–217, 2019.

[14] X. Xue, C. Yang, C. Jiang, P. W. Tsai, G. Mao, and H. Zhu, "Optimizing ontology alignment through linkage learning on entity correspondences," *Complexity*, vol. 2021, 12 pages, 2021.

[15] X. Xue and Y. Wang, "Optimizing ontology alignments through a memetic algorithm using both MatchFmeasure and unanimous improvement ratio," *Artificial Intelligence*, vol. 223, pp. 65–81, 2015.

[16] X. Xue, X. Wu, C. Jiang, G. Mao, and H. Zhu, "Integrating sensor ontologies with global and local alignment extractions," *Wireless Communications and Mobile Computing*, vol. 2021, 10 pages, 2021.

[17] M. G. Jorge, E. Alba, and J. F. Aldana-Montes, "Optimizing ontology alignments by using genetic algorithms," in *Proceedings of the First International Workshop on Nature Inspired Reasoning for the Semantic Web*, pp. 1–15, Karlsruhe, Germany, 2008.

[18] S. Bechikh, L. B. Said, and K. Ghédira, "Searching for knee regions of the Pareto front using mobile reference points," *Soft Computing*, vol. 15, no. 9, pp. 1807–1823, 2011.

[19] V. Rijsberge, *Information Retrieval*, University of Glasgow, 1975.

[20] V. Mascardi, A. Locoro, and P. Rosso, "Automatic ontology matching via upper ontologies: a systematic evaluation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 5, pp. 609–623, 2010.

[21] X. Xue and J. S. Pan, "An overview on evolutionary algorithm based ontology matching," *Journal of Information Hiding and Multimedia Signal Processisng*, vol. 9, no. 1, pp. 75–88, 2018.

[22] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions and reversals," *Soviet Physics Doklady*, vol. 1965, no. 10, pp. 707–710, 1965.

[23] G. Stoilos, G. Stamou, and S. Kollias, "A string metric for ontology alignment," in *Proceedings of 4th Int. Semantic Web Conference*, pp. 623–637, Karlsruhe, Germany, 2005.

[24] W. E. Winkler, "The state of record linkage and current research problems," Tech. Rep. US Census Bureau, Suitland-Silver Hill, MD, USA, 1999.

[25] G. A. Miller, "WordNet," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[26] Y. Wang, "Verb semantics and lexical selection," *Computer Science*, vol. 14, no. 101, pp. 325–327, 1994.

[27] J. Glen and J. Widom, "SimRank: a measure of structuralcontext similarity," in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 538–543, Edmonton, Canada, 2002.

[28] C. Li, J. Han, G. He et al., "Fast computation of simrank for static and dynamic information networks," in *Proceedings of the 13th International Conference on Extending Database Technology - EDBT '10*, pp. 465–476, Lausanne, Switzerland, 2010.

[29] X. Xue, J. Lu, and J. Chen, "Using NSGA-III for optimizing biomedical ontology alignment," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 3, pp. 135–141, 2019.

[30] Q. Zhang and H. Li, "MOEA/D: a multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2008.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5537471, 8 pages
https://doi.org/10.1155/2021/5537471

Research Article

The Design of University Physical Education Management Framework Based on Edge Computing and Data Analysis

Wenze Sun and Yuchi Gao

Department of Physical Education, Teaching Center of Fundamental Courses, Ocean University of China, Qingdao 266000, China

Correspondence should be addressed to Yuchi Gao; sunwenze@ouc.edu.cn

Received 6 January 2021; Revised 9 February 2021; Accepted 23 February 2021; Published 13 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Wenze Sun and Yuchi 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.

Educational informatization has become the only way for mankind to enter the information age, and the application of educational resource library is a key issue in the integration of information technology disciplines. Aiming at the shortcomings of existing college physical education management technology, this article is dedicated to developing a new online management platform for physical education. Firstly, the introduction of edge computing technology in the architecture design of college physical education system can fully improve the efficiency of the management system. Secondly, based on data cleaning, the BP neural network algorithm modified by particle swarm is used to conduct in-depth analysis of the data of the university sports teaching management system and obtain corresponding optimization measures. Finally, it is verified in actual tests that the system can quickly perform data search, entry, modification, deletion, and other operations, which can ensure the realization of the school sports department's networked, scientific, standardized, and digitalized management of physical education resources.

1. Introduction

Information technology is changing the production, life, learning, and education methods of human society [1, 2]. It not only brought a brand-new information environment to the education field, but also caused a comprehensive innovation of educational concepts, and had a profound impact on educational thinking, teaching methods and teaching methods [3]. The physical fitness of current students is declining year by year, which is closely related to the public physical education in colleges and universities [4, 5]. In order to improve the level and effectiveness of university teaching, the research of public sports is very important, and we will carry out related research.

At present, many industries have moved towards office standardization, automation, and networking development, of course, including the physical education management department [6]. While encouraging teachers and students of various colleges and universities to engage in physical fitness activities, the education department also found that colleges and universities have corresponding drawbacks in sports management, such as the shortage of sports venues and unreasonable resource allocation [7]. In recent years, with the rapid development of computers, more and more companies, universities, and related departments use network and information management models [8]. Modern management has long become the trend of efficient operation in various industries. While reducing the pressure of managers, this model also provides convenience for people's lives [9]. While the construction of sports Informa ionization is carried out, information resources are also valued by sports organizations, and the development of an information-based sports teaching management system is imperative [10]. Only in this way, sports resources will be distributed, and the needs of organizations and individuals at all levels for sports information will also be met. To achieve informatization and standardization of physical education management in colleges and universities, it must develop a physical education management system. Through the development of the student physical education management system, the physical education management system of colleges and universities realizes management Informa ionization and standardization [11, 12]. By improving the efficiency of sports information management, this will not only greatly

improve the efficiency of student sports management information in colleges and universities but also further improve the physical fitness of students [13, 14].

Educational informatization has become the only way for mankind to enter the information age, and the application of the educational resource library is a key issue in the integration of information technology disciplines. Aiming at the shortcomings of existing college physical education management technology, this article is committed to developing a new online management platform for physical education. Develop a new university sports teaching management system by combining edge computing and particle swarm-modified BP neural network algorithm technology. Through this sports management system, the sports resource management of the school sports department can be scientific, standardized, and digitized. The system can quickly perform operations such as data search, entry, modification, and deletion.

2. Overview of Related Technologies

2.1. Overview of College Physical Education Management. Physical education classroom management is an important part of physical education activities and an important factor affecting the quality of classroom teaching. Herbart once said that "if you do not grasp the constraints of management firmly and gently, teaching any homework is impossible." From this, we can know the importance of scientific management in teaching [15, 16]. Unlike other subjects, public physical education is taught outdoors to impart sports knowledge and skills so that students can form a lifelong physical education class [17]. In order to better complete the physical education teaching goals and improve the quality of teaching, physical education teachers control the impact of various potential factors in the classroom to complete the teaching goals, which requires teachers to have outstanding management capabilities [18]. Schematic diagram of the concept of college physical education management is shown in Figure 1.

In the process of public sports classroom teaching management, it is often impossible for both teachers and students to consider multiple tasks at the same time [19, 20]. We must realize the psychological compatibility between teachers and students in the teaching process [21, 22]. In the classroom teaching process, how teachers manage their work tasks and their attitudes towards students as well as how students manage their own learning tasks and their attitudes towards teachers' work are very important. In the teaching process, we learn from each other, combine the advantages of all aspects, and slowly optimize classroom teaching management. But while changing the concept of management methods, certain principles should be implemented [23]. In order to achieve informatization and standardization of physical education management in colleges and universities, it must develop a physical education management system. Through the development of the student physical education management system, the physical education management system of colleges and universities realizes management Informa ionization and standardization [24].

2.2. Edge Computing Technology. With the unprecedented popularity of mobile devices such as smart phones, cloud-based mobile sensors, tablets, and wearable devices, the Internet of Things has been extensively developed, thereby promoting the development of smart environments [25]. The Internet of Things enables mobile users to experience smart applications and enjoy various services with high-quality service quality, such as mobile online games and augmented reality. These services usually require many mobile devices with limited size and low power consumption to perform computationally intensive and time-sensitive tasks [26]. However, due to the low computing power and limited battery life of mobile devices, it is challenging to perform these services [27, 28]. Edge computing is an advanced version of cloud computing that reduces latency by bringing the service closer to the end user. Edge computing reduces the load of the cloud by providing resources and services in the edge network. However, edge computing will increase cloud computing by reducing the sensitivity of sensitive applications to end users. Like the cloud, edge service providers can provide end users with applications, data computing, and storage services [29].

Mobile edge computing (MEC) and wireless power transmission (WPT) are two promising technologies to solve the above challenges. In recent years, because MEC can significantly improve the computing power of mobile devices in a cost-saving and energy-saving manner, it has attracted increasing attention from the industry and academia. It allows mobile devices to offload some or all their computationally intensive tasks to MEC servers located at the edge of the wireless network, such as cellular base stations and access points. Unlike traditional cloud computing, MEC servers are deployed near end users [30]. Therefore, MEC has the potential to provide low-latency services, save energy for mobile users, and achieve high security. So far, many leading companies have regarded MEC as a promising technology for future wireless communication networks [31].

3. Analysis of the Needs of College Physical Education Management

We understand the background and significance of the physical education management system and the relevant research status at home and abroad. This chapter will analyze the needs of the physical education management system from the perspective of system function and nonfunctional needs analysis based on the content of the chapters. The overall framework of the system is planned, the system development platform is selected, and the relevant technology required for the development system is selected [32]. The college physical education management system uses information technology to efficiently manage the actual situation of physical education. It mainly includes competition management, physical fitness test management, physical education management, and user management. Through continuous investigation, in line with the principles of practicability, safety, and scientific, the demand analysis diagram of the college physical education management system is obtained, as shown in Figure 2.

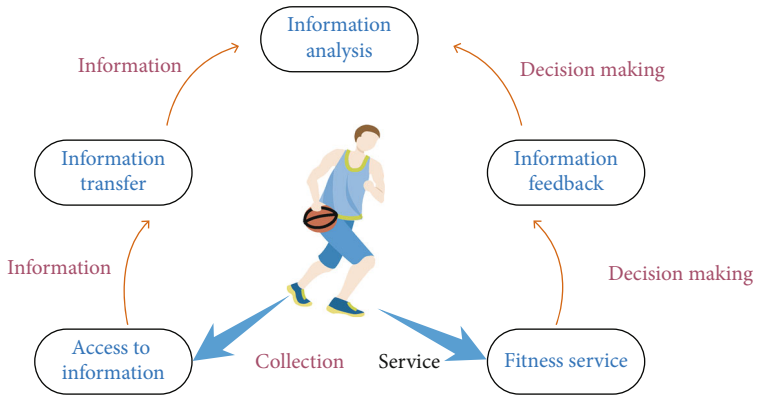


Figure 1: Schematic diagram of the concept of college physical education management.

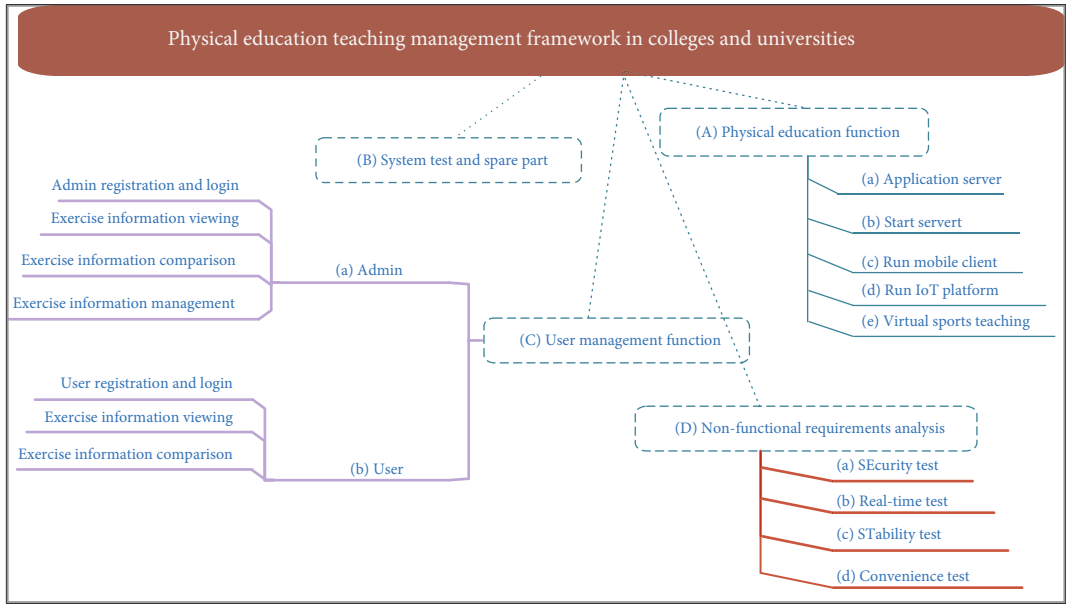


Figure 2: Distribution diagram of demand analysis of physical education teaching management in universities.

By strengthening and standardizing the work of physical education administrators, the fairness, orderliness, and efficiency of physical education management are guaranteed [33]. The demand analysis of the university sports management system in Figure 2 shows practical, safe, and scientific principles. From the current situation of the construction of public sports information and education resources in colleges and universities, the existing sports facilities and related sports electronic media resources have been greatly improved [34]. Some college sports departments or departments have certain experience in case handling, sports multimedia course teaching, and sports network distance teaching. The relevant sports webpages of the colleges and universities are also very refined and perfect. Some schools have also shared their own excellent sports courses with brother colleges and universities on the sports website page. However, due to the different levels of emphasis on different schools, there are still many colleges and universities in a backward situation in the update of sports information resources and video library webpages, and there are very few colleges and universities that realize the joint construction and sharing of public sports education resources between schools. Facing the needs of college physical education in public schools, there is an urgent need to develop a public sports resource sharing application platform system to enable the exchange of resources among colleges and universities, reduce redundant construction, and promote the informatization of college sports. The task system of the complete public sports resource sharing application platform differs from the nature of the unit. The system can be divided into several major functional modules through the summary and subtask system. Each part has specific functions and is divided into 3 major modules. The five relatively small parts of the part are teacher management module, student learning module,

and administrator module. The small modules are sports online learning, student world, physical education teaching, teacher world, and administrator system.

4. Analysis on the Management Data of Physical Education in Colleges and Universities

Facing the needs of college physical education in public schools, there is an urgent need to develop a public sports resource sharing application platform system to enable the exchange of resources among colleges and universities, reduce redundant construction, and promote the informatization of college sports. In order to be able to compare with the original scheme, a simulation model based on simulation is introduced. The task system of the complete public sports resource sharing application platform varies with the nature of the unit. The hidden nodes of BP neural network are usually determined by repeated forward and backward propagation methods. By modifying or constructing training methods to change the number of hidden nodes, the corresponding initial weights and thresholds will also change, thereby affecting the network performance convergence and learning efficiency.

In order to reduce the impact, a neural network model based on particle swarm optimization is used to optimize the adjustment of weights and thresholds, thereby speeding up the convergence of the network and improving the learning efficiency of the network.

$$f1(x) = \frac{1}{1 + e^{-\lambda x}}, \tag{1}$$

$$f1'(x) = \lambda f1(x)[1 - f1(x)]. \tag{2}$$

Suppose the neural network is trained to get a network with different weights and thresholds. After each training, the sample calculates the training error of the test sample and the test error of the test sample according to the formula.

$$E(w) = \frac{1}{2} \sum_{k \in \text{outputs}} (t_k - o_k)^2, \tag{3}$$

$$F(X_{n+1}|J_n) \neq F(X_{n+1}|J_n - Y_n). \tag{4}$$

After each training, BP algorithm and particle swarm optimization algorithm will jointly affect the adjustment of weights. Therefore, the basic BP network weight adjustment formula should also add a weight adjustment formula.

$$E(X_{n+1}|J_n) \neq E(X_{n+1}|J_n - Y_n), \tag{5}$$

$$X_t = \sum_{i=1}^{k} \alpha_i X_{t-i} + \sum_{i=1}^{k} \beta_i Y_{t-i} + u_{1t}. \tag{6}$$

By substituting the two coefficients into the equation of the second problem, we can get

$$x^{(j+1)} = x^{(j)} + s_j d^{(j)}, \tag{7}$$

$$\Delta f_r = \left[\frac{\partial f_r}{\partial x_1}, \frac{\partial f_r}{\partial x_2}, \cdots, \frac{\partial f_r}{\partial x_n}\right]. \tag{8}$$

If the upper layer contains factor, the layer weight values are $b_1, b_2, b_3 \cdots b_n$, respectively. In addition, we will calculate the composite weight of each layer element to the target layer.

$$\begin{aligned} \left|f^{(j)} - f^{(j-1)}\right| &\leq \tau, \\ \left|f^{(j)} - f^{(b)}\right| &\leq \tau, \end{aligned} \tag{9}$$

$$F = \frac{(SSE_r - SSE_u)/k}{SSE_u/(T - 2k)} \sim F_{(k,T-2k)}, \tag{10}$$

$$F(X_{n+1}|\Omega_n) \neq F(X_{n+1}|\Omega_n - Y_n). \tag{11}$$

Database design refers to the purpose of meeting data storage and processing requirements in a specific application environment and building a certain database model to effectively save data. In order to better complete the physical education teaching goals and improve the quality of teaching, physical education teachers control the impact of various potential factors in the classroom to complete the teaching goals, which requires teachers to have outstanding management capabilities. Database design is the core of the establishment of the system and also the basis for the establishment of the required database and application system. Combined with the actual situation of the physical education management of college, MySQL Server is used as a tool to build a database.

5. College Physical Education Teaching Management Test and Evaluation

Whether the university education and teaching management system can exert system efficiency, reduce production costs, and improve the system of management depend on the feedback after using the system. After design and development, the system enters the testing phase. Software testing is an important activity that is aimed at finding defects and reducing risks to ensure software quality, only through examples to test the system, according to the results of the operation, to evaluate the performance of the system. Types of physical education teaching management tests in colleges and universities-standard 1 are shown in Figure 3.

In the early and midstage of the construction of the college sports teaching management cluster, that is, when the data in the cluster has not reached the scale of fully training the diagnostic model, the fault diagnosis model can be provided by the real-time data management system in the cluster to provide some real-time collected bearing fault data set. Types of physical education teaching management test in colleges and universities-standard 2 are shown in Figure 4.

But it is still sufficient to complete the full training of the model. Currently, it is necessary to use migration learning to make edge computing end servers work together. Transfer learning between cloud and edge computing can greatly improve the accuracy of diagnosis. Through

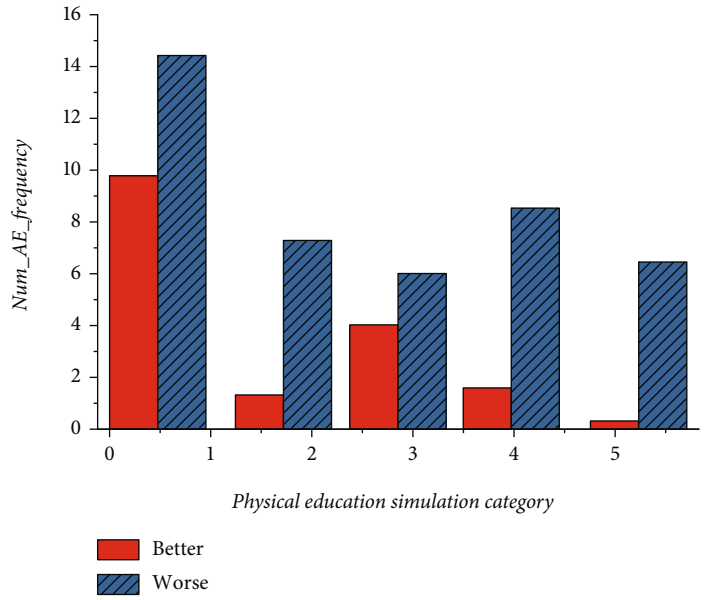


Figure 3: Types of physical education teaching management tests in colleges and universities-standard 1.

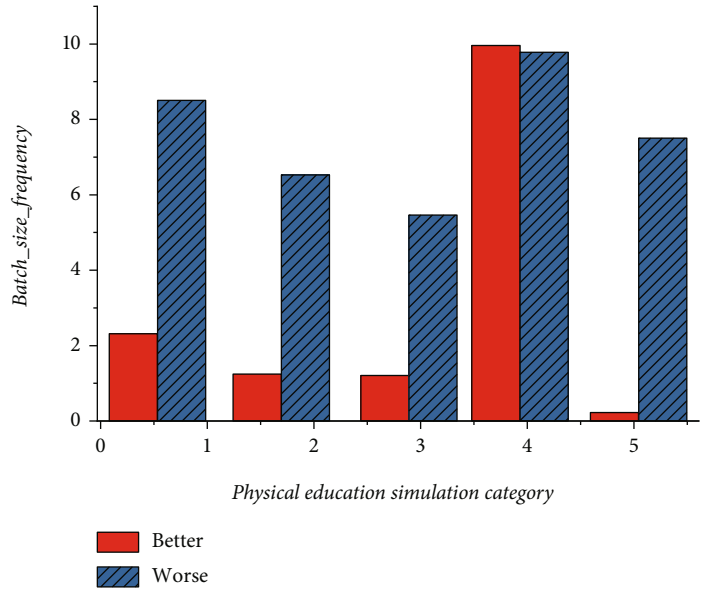


Figure 4: Types of physical education teaching management test in colleges and universities-standard 2.

exploration and experimentation, we can draw conclusions and use a small amount of training data in the initial stage of cluster construction. Changes with the number of MEC frequencies-energy consumption are shown in Figure 5.

At the edge end, only the local model needs to be updated with the model parameters trained on the fog end to accurately complete the real-time diagnosis of the bearing status. Changes with the number of MEC frequencies-number of people served are shown in Figure 6.

Carry out necessary tests for the design and implementation of the public sports education resource sharing platform system, and test the corresponding components. The build test is divided into two aspects: software and hardware to test. Build an independent test environment. Changes with the number of MEC frequencies-the number of packets transmitted are shown in Figure 7.

The running environment of the test needs to meet the public conditions. Next, conduct corresponding preinstallation tests on the public sports education resource sharing platform system. Perform individual tests on individual main test items. Changes with the number of MEC frequencies-excellent evaluation are shown in Figure 8.

Through the above tests, it is found that most of the functions of the system can meet the system requirements. That

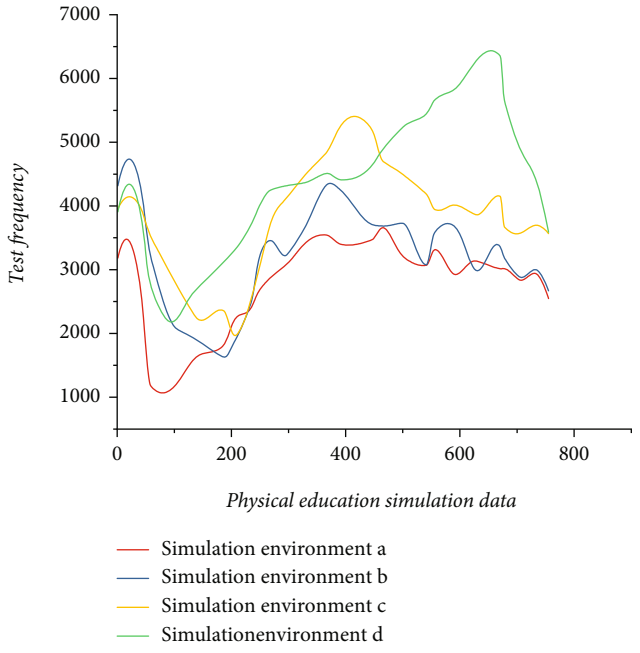


Figure 5: Changes with the number of MEC frequencies-energy consumption.

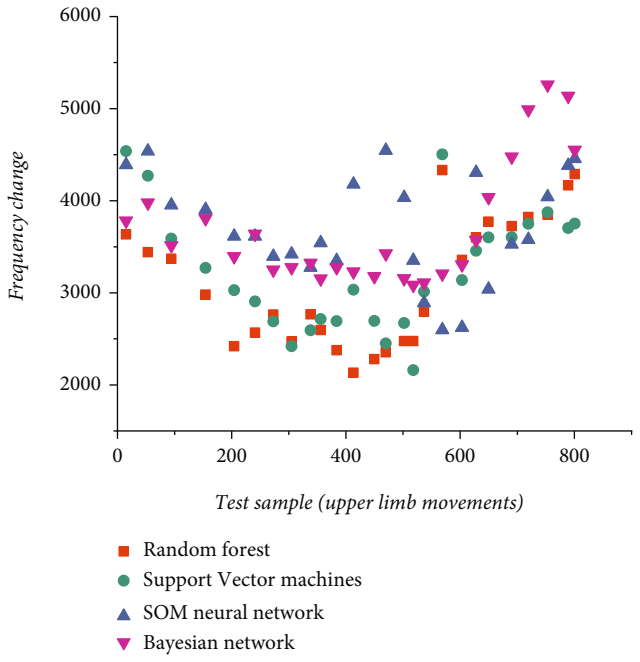


Figure 6: Changes with the number of MEC frequencies-number of people served.

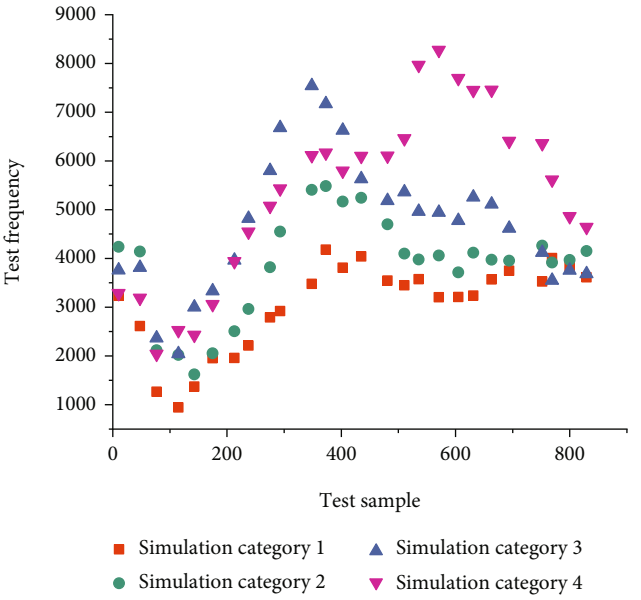


Figure 7: Changes with the number of MEC frequencies-the number of packets transmitted.

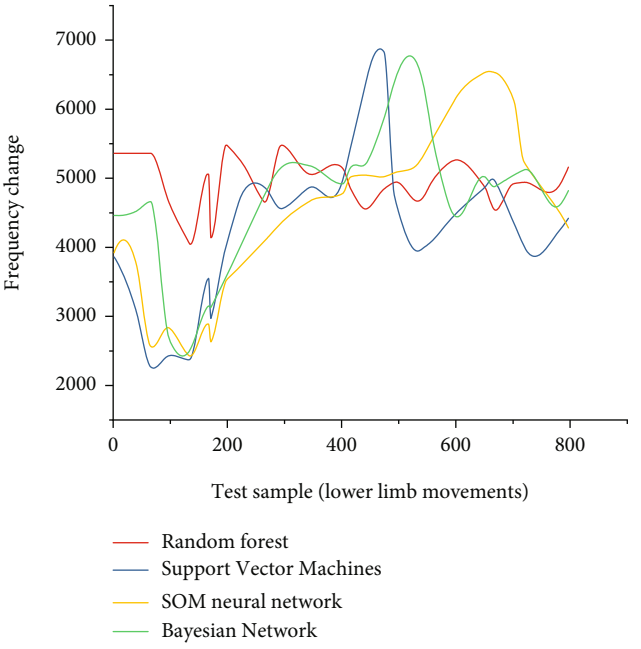


Figure 8: Changes with the number of MEC frequencies-excellent evaluation.

is, the basic function of the system can satisfy the management of physical education. In terms of sports event registration, score entry, and related sports event publicity, the value of the sports education management system is gradually reflected, because many sports event publicity can be directly published to the system with the help of the system. Promoting the informatization of university sports plays an important role. Facing the demand for physical education in public schools and universities, there is an urgent need to develop a public sports resource sharing application platform system to realize the exchange of resources between universities and reduce redundant construction. For example, the feedback information could not be fed back to the administrator in time, and there was a phenomenon that the registration could not be successful once during the registration

process. But for the current test and use, no other errors have been found. We will also further analyze the related design process and debugging progress.

6. Conclusion

The college sports management system has realized management Informa ionization and standardization. This has not only greatly improved the efficiency of colleges and universities in student sports management information. This will also greatly improve the efficiency of sports information managers. Educational Informa ionization has become the only way for mankind to enter the information age, and the application of educational resource libraries is a key issue in the integration of information technology disciplines. Aiming at the deficiencies of existing university sports management techniques, this article is dedicated to developing a new sports online management platform. Combine edge computing and particle swarm-improved BP neural network algorithm technology to develop a new type of university sports teaching management system. The system can quickly perform operations such as data search, entry, modification, and deletion. The current sports management system still has certain shortcomings. For example, the feedback information cannot be fed back to the administrator in time, and there is a phenomenon that the registration is unsuccessful once during the registration process. But for the current test and use, no other errors were found. We will further analyze the related design process and debugging progress. In the future, we will further devote ourselves to sports management research, so as to ensure the realization of the network of these school sports departments, the scientific, standardized, and digitalized sports resource management.

Data Availability

We already included the analysis data in our manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] X. Zhang, J. Zhang, Z. Liu, Q. Cui, X. Tao, and S. Wang, "MDP-based task offloading for vehicular edge computing under certain and uncertain transition probabilities," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3296–3309, 2020.

[2] W. Zhang, Z. Zhang, S. Zeadally, H. C. Chao, and V. C. M. Leung, "Energy-efficient workload allocation and computation resource configuration in distributed cloud/edge computing systems with stochastic workloads," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 6, pp. 1118–1132, 2020.

[3] K. Zhang, Y. Hu, F. Tian, and C. Li, "A coalition-structure's generation method for solving cooperative computing problems in edge computing environments," *Information Sciences*, vol. 22, no. 2, pp. 536–543, 2020.

[4] J. Zhang, L. Zhou, F. Zhou et al., "Computation-efficient offloading and trajectory scheduling for multi-UAV assisted mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2114–2125, 2020.

[5] X. Xu, X. Liu, X. Yin, S. Wang, and L. Qi, "Privacy-aware offloading for training tasks of generative adversarial network in edge computing," *Information Sciences*, vol. 6, no. 11, pp. 345–356, 2020.

[6] X. Xiong, K. Zheng, L. Lei, and L. Hou, "Resource allocation based on deep reinforcement learning in IoT edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 6, pp. 1133–1146, 2020.

[7] X. Wang, Z. Ning, S. Guo, and L. Wang, "Imitation learning enabled task scheduling for online vehicular edge computing," *IEEE Transactions on Mobile Computing*, vol. 12, pp. 11–24, 2020.

[8] T. Wang, Y. Liang, Y. Yang et al., "An intelligent edge-computing-based method to counter coupling problems in cyber-physical systems," *IEEE Network*, vol. 34, no. 3, pp. 16–22, 2020.

[9] M. C. I. Van Schalkwyk, M. Mckee, J. V. Been, C. Millett, and F. T. Filippidis, "Analysis of tobacco industry pricing strategies in 23 European Union countries using commercial pricing data," *Tobacco Control*, vol. 12, no. 4, pp. 126–132, 2019.

[10] Z. Sun and M. R. Nakhai, "An online learning algorithm for distributed task offloading in multi-access edge computing," *IEEE Transactions on Signal Processing*, vol. 99, p. 1, 2020.

[11] G. Sole, "University of Otago: towards specialisation with a Masters in Physiotherapy, endorsed in Sports Physiotherapy (continuing professional development series)," *British Journal of Sports Medicine*, vol. 51, no. 5, pp. 481-482, 2017.

[12] P. V. Rosen, I. M. Dohrn, and M. Hagströmer, "Association between physical activity and all-cause mortality: a 15-year follow-up using a compositional data analysis," *Scandinavian Journal of Medicine & Science in Sports*, vol. 30, 2019.

[13] S. Rahimirad, M. Navaderi, S. Alaei, and M. H. Sanati, "Identification of hsa-miR-106a-5p as an impact agent on promotion of multiple sclerosis using multi-step data analysis," *Neurological Sciences*, vol. 3, no. 3, pp. 16–22, 2021.

[14] A. Mann, A. Mann, C. H. Tator, and J. D. Carson, "Investigation of knowledge and attitude about concussion diagnosis and management among family medicine residents," *British Journal of Sports Medicine*, vol. 51, no. 11, pp. A53.2–A5A53, 2017.

[15] J. Liu, H. Guo, J. Xiong, N. Kato, J. Zhang, and Y. Zhang, "Smart and resilient EV charging in SDN-enhanced vehicular edge computing networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 217–228, 2020.

[16] Z. Li, L. He, X. Tang, W. Han, and J. Wang, "Design and implementation of real assets analysis and evaluation system of power grid based on big data," *Electrical Engineering*, vol. 33, no. 4, pp. 134–147, 2019.

[17] S. Li, S. Lin, L. Cai, W. Li, and G. Zhu, "Joint resource allocation and computation offloading with time-varying fading channel in vehicular edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3384–3398, 2020.

[18] F. W. Leung, M. Koo, H. Jia et al., "Water exchange (WE) and quality improvement—enhanced advanced adenoma detection: a pooled data analysis of 6 randomized controlled trials," *Journal of Clinical Gastroenterology*, vol. 23, no. 3, pp. 34–47, 2020.

[19] L. Yang, C. Zhong, Q. Yang, W. Zou, and A. Fathalla, "Task offloading for directed acyclic graph applications based on edge computing in Industrial Internet," *Information Sciences*, vol. 540, pp. 51–68, 2020.

[20] D. Kim, H. Lee, H. Song, N. Choi, and Y. Yi, "Economics of fog computing: interplay among infrastructure and service providers, users, and edge resource owners," *IEEE Transactions on Mobile Computing*, vol. 99, p. 1, 2019.

[21] M. Ke, Z. Gao, Y. Wu, X. Gao, and K. K. Wong, "Massive access in cell-free massive MIMO-based Internet of things: cloud computing and edge computing paradigms," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 3, pp. 321–330, 2021.

[22] D. L. Johnson, "Management of the anterior cruciate ligament deficient knee," *Clinics in Sports Medicine*, vol. 36, no. 1, pp. 654–662, 2017.

[23] J. Alent, D. M. Narducci, B. Moran, and E. Coris, "Sternal injuries in sport: a review of the literature," *Sports Medicine*, vol. 33, no. 4, pp. 134–147, 2019.

[24] J. M. Hung, X. Li, J. Wu, and M. F. Chang, "Challenges and trends in developing nonvolatile memory-enabled computing chips for intelligent edge devices," *IEEE Transactions on Electron Devices*, vol. 99, pp. 1–10, 2020.

[25] J. H. Huh, "Reefer container monitoring system using PLC-based communication technology for maritime edge computing," *Journal of Supercomputing*, vol. 33, no. 4, pp. 234–247, 2018.

[26] M. P. Hosseini, T. X. Tran, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Multimodal data analysis of epileptic EEG and rs-fMRI via deep learning and edge computing," *Artificial Intelligence in Medicine*, vol. 104, article 101813, 2020.

[27] M. A. Guillén, A. Llanes, B. Imbernón, R. Martínez-Espaa, and J. M. Cecilia, "Performance evaluation of edge-computing platforms for the prediction of low temperatures in agriculture using deep learning," *Journal of Supercomputing*, vol. 77, no. 9, pp. 145–153, 2021.

[28] M. Gairing and R. Savani, "Computing stable outcomes in symmetric additively separable hedonic games," *Mathematics of Operations Research*, vol. 12, no. 3, pp. 321–330, 2019.

[29] J. J. Fyfe and J. P. Loenneke, "Interpreting adaptation to concurrent compared with single-mode exercise training: some methodological considerations," *Sports Medicine*, vol. 48, no. 2, pp. 289–297, 2018.

[30] R. Du, Y. Liu, L. Liu, and W. Du, "A lightweight heterogeneous network clustering algorithm based on edge computing for 5G," *Wireless Networks*, vol. 26, no. 3, pp. 145–153, 2020.

[31] P. Derbeko, S. Dolev, and E. Gudes, "Wavelet-based dynamic and privacy-preserving similitude data models for edge computing," *Wireless Networks*, vol. 27, no. 1, pp. 351–366, 2021.

[32] C.-H. Huang, L.-I. Hsu, T.-K. Chang et al., "Stress distribution of the patellofemoral joint in the anatomic V-shape and curved dome-shape femoral component: a comparison of resurfaced and unresurfaced patellae," *Knee Surgery Sports Traumatology Arthroscopy*, vol. 25, no. 1, pp. 263–271, 2017.

[33] P. E. Boyle, M. D. Richardson, M. C. Savin, D. E. Karcher, and D. A. Potter, "Ecology and management of earthworm casting on sports turf," *Pest Management Science*, vol. 75, no. 8, pp. 145–153, 2019.

[34] H. Guo, X. Yang, and Y. Zhu, "Bloch theory-based gradient recovery method for computing topological edge modes in photonic graphene," *Journal of Computational Physics*, vol. 379, pp. 403–420, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5560652, 7 pages
https://doi.org/10.1155/2021/5560652

Research Article

Research on Data Optimization of CT System for Patients with Severe Pneumonia Based on Filtered Back Projection and Internet of Things

Shaojie Qin and Wentao Wang

Intensive Care Unit of The Second Affiliated Hospital, Zhengzhou University, Zhengzhou 450003, China

Correspondence should be addressed to Wentao Wang; zzdxdff@zzu.edu.cn

Received 17 January 2021; Revised 9 February 2021; Accepted 24 February 2021; Published 13 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Shaojie Qin and Wentao 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.

CT technology can clearly and intuitively display the internal structure, composition, material, and defect status of the detected object under the condition of no damage to the detected object. This paper is based on the filtering back projection and the Internet of Things technology to carry out the data optimization of the CT system of patients with severe pneumonia. First of all, the Internet of Things technology is used in the structural design and operation process of the CT system for patients with severe pneumonia, which improves the advancement and convenience of the original system. Secondly, the system data is optimized through the filtering back projection algorithm, which improves the data accuracy and transmission efficiency. Finally, simulation tests have verified the effectiveness of the established CT system data optimization method for patients with severe pneumonia, which is helpful for medical comparison, analysis, and diagnosis.

1. Introduction

Computed tomography (CT) is a medical diagnosis method that combines electronic computer technology and X-ray imaging technology. Due to the rapid development of electronic computer technology in recent years, this method has become one of the most common medical clinical diagnostic methods [1, 2]. Severe pneumonia is a common respiratory disease. The various organs of the elderly are gradually failing, the immune system is weakened, and they are more likely to suffer from severe pneumonia [3]. The high-resolution CT image can accurately detect the subtle differences in the radiation attenuation characteristics between the tissues on the cross-section of the human body to be measured and display them in the form of digital images and can finely distinguish the different densities [4]. Therefore, CT technology is widely used in cross-sectional tissues for medical purposes [5]. The disease has the characteristics of insidious onset, critical illness, and rapid progress. The initial symptoms are local inflammation gradually worsening to systemic inflammation, causing severe infectious complications such as septic shock, sepsis, and multiple organ dysfunction, which poses a great threat to the life safety of patients and requires effective treatment intervention as soon as possible [6, 7].

Although the above-mentioned medical imaging modes have a wide range of important applications, due to factors such as the performance of their own equipment and external conditions, medical images still have many unavoidable defects [8, 9]. CT technology can clearly and intuitively display the internal structure, composition, material, and defect status of the inspected object without damaging the inspected object [10, 11], for example, the image is blurred or the image boundary is not obvious. In addition, in the process of forming and transmitting medical images, its image quality will also be affected to a certain extent [12]. All these have caused certain difficulties and obstacles to doctors' image interpretation and accurate diagnosis. In order to enhance the quality of medical images, enable doctors to observe and analyze the lesions and anatomical structures more clearly, and improve the accuracy of diagnosis, postprocessing of medical images becomes necessary

[13, 14]. This paper is based on the filtering back projection and the Internet of Things technology to carry out the data optimization of the CT system of patients with severe pneumonia. The established CT system data optimization method for patients with severe pneumonia is helpful for medical comparison, analysis, and diagnosis.

2. CT Tomography Technology

CT is computed tomography. CT imaging uses the attenuation characteristics of X-rays as a diagnostic basis to scan through a certain thickness of a certain part of the human body [14, 15]. The high-sensitivity receiver receives the X-rays that pass through this layer and converts them into visible light, which is converted into electrical signals by photoelectric conversion, and then converted into digital signals by an analog-digital converter. Each number in the digital matrix is converted into small squares from black to white (i.e., pixels) by a digital-to-analog converter and arranged into a matrix to obtain CT images [16, 17]. After entering the computer, the selected layer is divided into multiple voxels, and the scanned information is calculated to obtain the X-ray attenuation coefficient or absorption coefficient of each voxel and arranged in a digital matrix.

X-rays have varying degrees of penetrating ability to various substances. When a beam of X-rays passes through the human body, it will be attenuated due to absorption and scattering [18, 19]. However, due to the different density and atomic composition of various tissues of the human body, the degree of absorption and scattering is also different, that is, different tissues of the human body have different attenuation coefficients for X-rays [20, 21]. In general, when X-rays with the same energy are used for incidence, the higher the density of human tissues, the greater the attenuation of X-rays, and the lower the density of human tissues, the lower the attenuation of X-rays. CT images express the absorption degree of X-rays by organs and tissues in different gray scales, which is consistent with X-ray imaging. The black areas indicate low absorption areas, i.e., low-density areas, such as lungs; white areas indicate high absorption areas, i.e., high-density areas [22, 23].

However, compared with X-ray images, CT is completely dependent on the projection data of the attenuation value of the object in the inspected layer. It has nothing to do with the structure of the inspected layer. It is a complete tomographic image with more accurate layer thickness and clearer image. The density resolution is higher. In addition, the tomographic images obtained by CT scanning can also be processed by computer software to obtain multidirectional cross-sectional images, such as coronal images and sagittal images, for easy observation [24]. The above characteristics provide an important guarantee for the research to obtain clear, high-resolution, and economical geometric information of cerebral aneurysms. The reconstructed image can be displayed on the monitor in different Di levels or stored in the computer hard disk in digital form.

3. Overview of IoT Technology

The Internet of Things has entered all aspects of people's lives, and as a technical support for economic development, it has shown great potential as a promoter of world economic growth in the 21st century [25, 26]. The Internet of Things is a systematic fusion of multiple technologies. To understand the Internet of Things, we should start from its system architecture and fully understand it to grasp the overall situation from the top level and lay a solid foundation for specific application design [27, 28]. A sensor conduction network is composed of probes, remote sensing components, and various terminal devices [29, 30]. The sensing layer is composed of different forms of sensing components, QR codes, RFID tags, and barcode readers distributed in various building groups. A schematic diagram of IoT technology architecture is shown in Figure 1.

The access network is a fusion of the communication network and the Internet composed of various wired and wireless nodes and fixed and mobile gateways and can be accessed through the Internet, the broadcasting network, and the communication network. Different from classic information systems, IoT requires large-scale data access and transmission capabilities [31–33]. The role of the middleware layer is to shield heterogeneity, achieve two-way interaction, and data preprocessing. At the same time, various system platforms and technical resources with different advantages are integrated and shared to form a multifunctional integrated system with diverse resource allocation and complete functions. The middleware layer can provide a variety of services to the upper layer, integrate multiple applications on the basic system, and provide differentiated services for different user needs.

4. Optimization of CT System Data for Patients with Severe Pneumonia

4.1. Filtering Back Projection Algorithm. Interpolation here refers to a model-based method used to estimate continuous data from discrete data within a known coordinate range. Classical linear interpolation techniques include nearest neighbor interpolation, bilinear interpolation, and bicubic interpolation [34, 35]. In the back-projection process of the FDK algorithm, due to the discreteness of the data, the projection address of the pixel will appear "misaligned," which usually requires interpolation, especially for the edges of the image. The nearest neighbor interpolation will produce blackness and jaggedness, while bilinear interpolation and bicubic interpolation will blur the edges of the image. Therefore, researchers have introduced many nonlinear interpolation methods, such as local adaptive interpolation methods.

In an ideal situation, the projection is the integral of the specific parameters of the object along a straight line passing through the object, and the measured object is expressed as a two-dimensional function, that is, the angle between the ray normal and the light.

$$\rho = x \cos\theta + y \sin\theta. \tag{1}$$

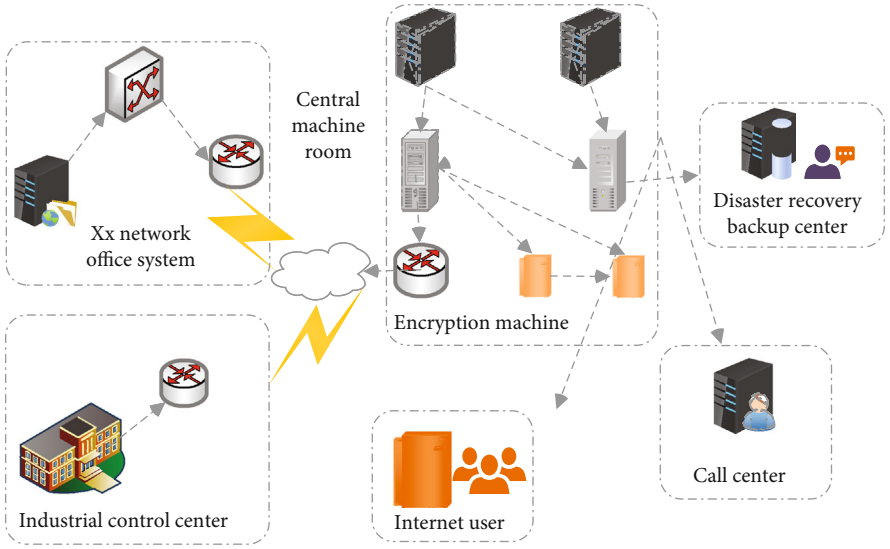


Figure 1: Schematic diagram of IoT technology architecture.

According to the conversion relationship between polar coordinates and coordinates, $\beta = \arctan(y/x)$ and $r = x^2 + y^2$, the formula can be rewritten as

$$\rho = r\cos(\theta - \beta). \tag{2}$$

A schematic diagram of the filtered projection geometric ray is shown in Figure 2. According to the geometric properties of the parallel projection, after the ray scan, the complete image of the original image should be obtained, and a series of projection data are measured [36]. The projections of all rays passing through a certain pixel are accumulated in the projection domain. This pixel can be reconstructed:

$$f(x, y) = \int_0^{\pi} G(x\cos\theta + y\sin\theta, \theta)d\theta. \tag{3}$$

The geometric relationship between these pixels is shown in the figure. According to the geometric relationship, the angle between the ray and the pixel is

$$f(r, \alpha) = \int_0^{\pi} G(r\cos\varphi, \theta)d\theta. \tag{4}$$

A schematic diagram of the filtered projection polar coordinate geometry is shown in Figure 3. It can be seen from the above formula that the projection length is no longer dependent on this angle, so the area to be solved for the image is transformed from the plane with XY as the horizontal and vertical coordinates to the polar coordinate plane with the length as the rotation angle [37]. In the process of CT image reconstruction, it is first converted from the XY coordinate domain to the polar coordinate domain and then reconstructed [38].

4.2. Optimization of CT System Data for Patients with Severe Pneumonia Based on Filtered Back Projection. What is obtained after reconstructing the image in the polar coordinate domain is not the final result. It is necessary to use the coordinate conversion relationship $x = r\cos\alpha$ and $y = r\sin\alpha$ and map it to the XY coordinate system to truly reflect the original image. MATLAB software is used to run the program to get the image shown in Figure 4.

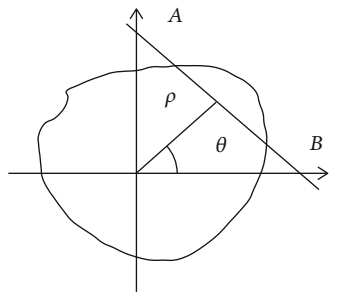


Figure 2: Schematic diagram of filtered projection geometric ray.

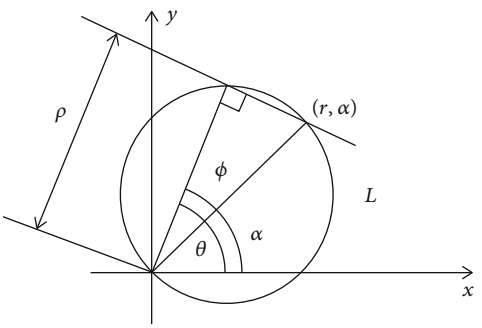


Figure 3: Schematic diagram of filtered projection polar coordinate geometry.

In this paper, after processing the absorption matrix with MATLAB, the absorption rate image of the object can be obtained, as shown in Figure 1. The figure shows that the object is uniform. When X-rays penetrate an object, the rays

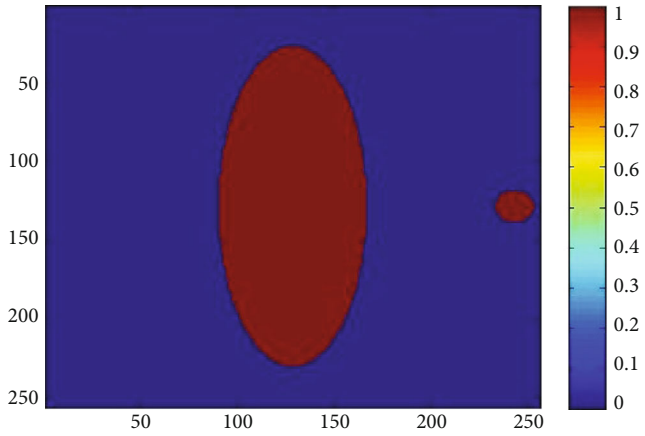


Figure 4: The geometry of the filtered projection medium.

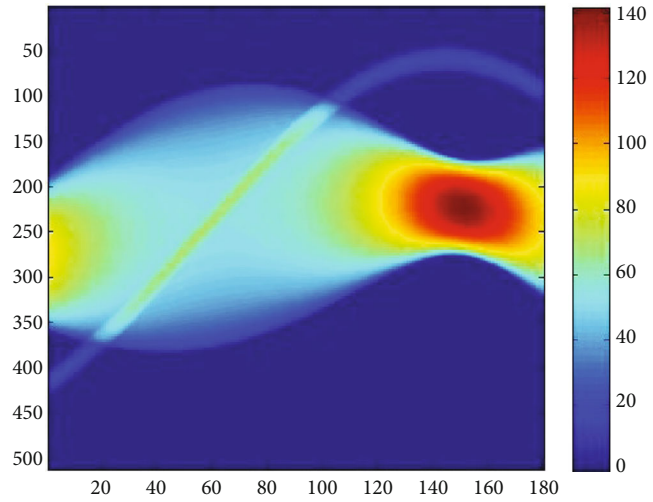


Figure 5: Fitting diagram of filtered projection received information data.

will be absorbed and scattered by the material, which will weaken the intensity of the X-rays. It can be known from Beer's theorem that when a single light is emitted, the attenuation of the X-ray of the object in a single uniform material will change, and the attenuation of its intensity satisfies formula (1).

$$I = I_0 e^{-\mu \Delta x}. \tag{5}$$

Assuming that the test object is segmented and uniform, the linear attenuation coefficients of each segment are as shown in formula (2):

$$\mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3 + \cdots = \ln\left(\frac{I_0}{I}\right). \tag{6}$$

Take an ideal ray source, which emits extremely thin pen beam rays, and place a detector on the opposite side. Measure the ray receiving signal g and the intensity of the rays reaching the detector after being attenuated by the object, thus obtaining a set of data. Then, rotate at an angle to obtain another set of data under the new angle. Repeat this until 180 sets of data are obtained after 180 rotations. The fitting diagram of the filtered projection received information data is shown in Figure 5.

According to Radon transformation, suppose the absorption function of a beam of parallel light is shown in formula (3):

$$f(x,y)\begin{cases} \neq 0, & (x,y) \in \Omega, \\ = 0, & (x,y) \notin \Omega. \end{cases} \tag{7}$$

Among them, Ω is the medium, and the attenuation rate of the intensity is $dI/I = -f(x,y)du$. Formula (4) is solved:

$$I = I_0 \exp\left[-\int_L f(x,y)du\right]. \tag{8}$$

The relationship between the coordinate system X-Y established with the pallet as the reference system and the coordinate system s-u established with the detector as the reference system is shown in

$$\begin{cases} s = X\cos\theta + Y\sin\theta, \\ u = -X\sin\theta + Y\cos\theta. \end{cases} \tag{9}$$

From the data fitting diagram of the received information of the template, it can be seen that the curve in the figure is approximately fitted to a sine function. According to the symmetry of the sine function, the abscissa of the origin can be calculated as 62, and the abscissa of the lowest point of the sine function is 144. Corresponding to the distance between the abscissas of the two points, the radian corresponding to the entire curve can be calculated, that is, the entire CT system rotates counterclockwise around the center of rotation, and a total of 180 sets of received information are obtained. It can be calculated that the CT system receives information every time it passes.

$$g(s,\theta) = Rf = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - s) \cdot dxdy, \quad -\infty < s < +\infty, 0 \le \theta \le \pi. \tag{10}$$

Calculate the expression of the sine function, and then obtain $\phi = 3\pi/4$, that is, the initial angle is $3\pi/4$, so the 180 directions of the X-ray used by the CT system are$135^\circ + 1.0980N^\circ$, where N represents the number of rotations.

5. Optimization of CT System Data for Patients with Severe Pneumonia

5.1. Data Source and Test Environment. In order to verify the effectiveness of this method on CT images, a simulated severe pneumonia CT image was used as a test sample, the similarity

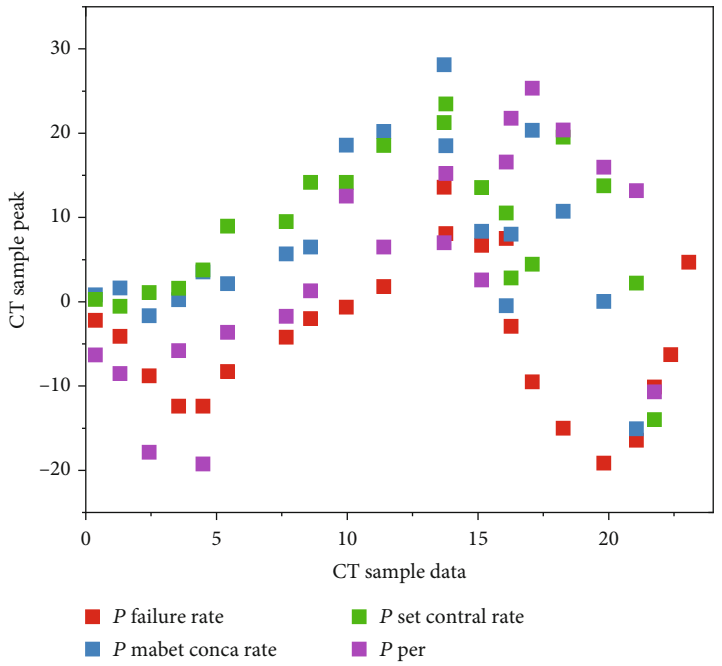


Figure 6: Peak signal-to-noise distribution of CT samples under different conditions.

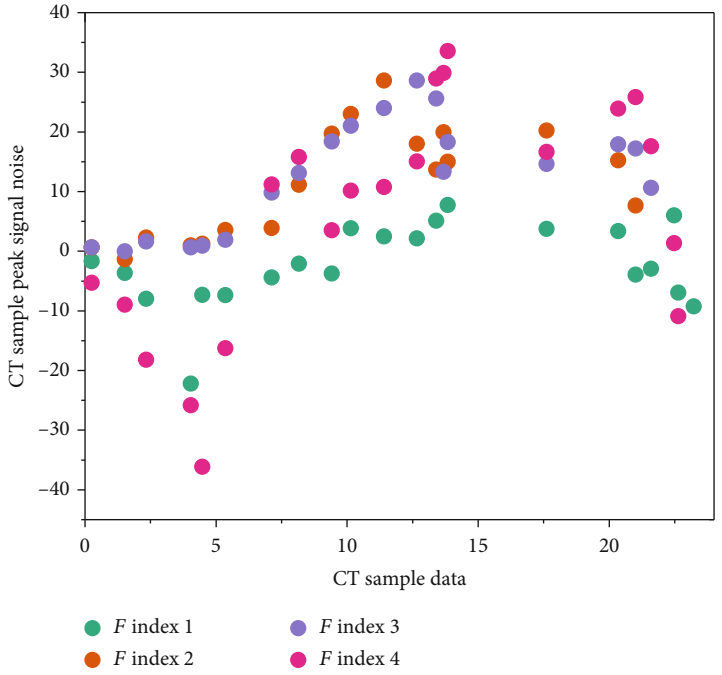


Figure 7: Peak distribution of CT samples under different search window radii.

window radius was selected as 3, the search window radius was selected as 7, and the smoothing coefficient was selected as 0.2, and the algorithm comparison experiment was performed.

On the basis of the description of the filtering back projection algorithm, the Radon inverse transformation is performed on the projection image generated by the data again, and the filtering back projection algorithm is established. The sample image is added with Gaussian noise interference with a mean value of 0 and a variance of 0.01 before proceeding. In terms of noise processing, we use the filtered back projection algorithm, the DCT self-similar algorithm and the DCT threshold de-coefficient self-similar algorithm. The result of comparing the denoising effect is shown in Figure 6, which compares the peak signal-to-noise ratio and the signal-to-noise ratio. In addition, the experimental results of the three result images are shown in Figure 7. After acquiring a digital image, the first thing to do is to preprocess the

image to improve the image quality or reduce the amount of calculation. The specific content of image preprocessing depends on the specific system requirements. Edge extraction first detects the discontinuities of the local characteristics of the image and then connects these discontinuous edge pixels into a complete boundary. In the image, the areas where the gray level changes drastically, that is, the intensity discontinuity, correspond to the edges. Edge extraction is to detect the intensity discontinuities and to determine their precise position in the image. It can be seen that the peak signal-to-noise ratio and signal-to-noise ratio of the simulated CT image obtained by the DCT self-similar algorithm and the DCT threshold decoefficient self-similar algorithm are larger than those of the original algorithm. Therefore, through experimental simulation comparison, we believe that the improved filtered back projection algorithm has better results in CT image processing.

6. Conclusion

CT technology can clearly and intuitively display the internal structure, composition, material, and defect status of the inspected object without damaging the inspected object. In order to improve the quality of medical images, enable doctors to observe and analyze lesions and anatomical structures more clearly, and improve the accuracy of diagnosis, post-processing of medical images becomes necessary. Based on the filtering back projection and the Internet of Things technology, this paper optimizes the data of the CT system of patients with severe pneumonia. The established data optimization method of the CT system for severe pneumonia is helpful for medical comparison, analysis, and diagnosis. Although the above-mentioned medical imaging modes have extensive and important applications, medical images still have many unavoidable defects due to factors such as the performance of their own equipment and external conditions. For example, the image is blurred or the image boundary is not obvious. In addition, in the process of forming and transmitting medical images, the image quality will also be affected to a certain extent. We will further carry out research in related fields to improve the data optimization level of the CT system.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All authors agree to submit this version of the manuscript.

References

[1] Q. Zhu, B. Zhang, Q. Chen et al., "Model reductions for multi-scale stochastic optimization of cooling water system equipped with closed wet cooling towers," *Chemical Engineering Science*, vol. 224, 2020.

[2] T. Zhang, Y. Xing, L. Zhang, X. Jin, H. Gao, and Z. Chen, "Stationary computed tomography with source and detector in linear symmetric geometry: direct filtered backprojection reconstruction," *Medical Physics*, vol. 47, no. 5, pp. 2222–2236, 2020.

[3] M. Tozakidou, I. Apine, K. U. Petersen, J. M. Weinrich, and J. Herrmann, "Comparison of different iterative CT reconstruction techniques and filtered back projection for assessment of the medial clavicular epiphysis in forensic age estimation," *International Journal of Legal Medicine*, vol. 134, no. 1, pp. 431–443, 2019.

[4] O. Hermine, X. Mariette, P. L. Tharaux et al., "Effect of tocilizumab vs usual care in adults hospitalized with COVID-19 and moderate or severe pneumonia: a randomized clinical trial," *JAMA Internal Medicine*, vol. 181, no. 1, pp. 32–40, 2021.

[5] K. Smart, A. M. Bröske, D. Rüttinger et al., "PK/PD mediated dose optimization of emactuzumab, a CSF1R inhibitor, in patients with advanced solid tumors and diffuse-type tenosynovial giant cell tumor," *Clinical Pharmacology & Therapeutics*, vol. 108, no. 3, pp. 616–624, 2020.

[6] J. van Sluis, R. Boellaard, R. A. Dierckx, G. Stormezand, A. W. J. M. Glaudemans, and W. Noordzij, "Image quality and activity optimization in oncologic 18F-FDG PET using the digital biograph vision PET/CT system," *Journal of Nuclear Medicine*, vol. 61, no. 5, pp. 764–771, 2020.

[7] L. Shahrin, M. J. Chisti, B. Brintz et al., "Clinical and laboratory predictors of 30-day mortality in severe acute malnourished children with severe pneumonia," *Tropical Medicine & International Health*, vol. 25, no. 11, pp. 1422–1430, 2020.

[8] M. Schloz, T. C. Pekin, Z. Chen, W. van den Broek, D. A. Muller, and C. T. Koch, "Overcoming information reduced data and experimentally uncertain parameters in ptychography with regularized optimization," *Optics Express*, vol. 28, no. 19, pp. 28306–28323, 2020.

[9] Y. Qichun, Z. Xuesong, E. James, M. Almendinger, and X. Huang, "Microlocation for smart buildings in the era of the Internet of Things–a survey of technologies, techniques, and approaches," *Environmental Pollution*, vol. 16, 2019.

[10] C. Michie, I. Andonovic, C. Davison et al., "The Internet of Things enhancing animal welfare and farm operational efficiency," *Journal of Dairy Research*, vol. 87, no. S1, Supplement 1, pp. 20–27, 2020.

[11] S. N. Matheu, J. L. Hernández-Ramos, A. F. Skarmeta, and G. Baldini, "A survey of cybersecurity certification for the Internet of Things," *ACM Computing Surveys*, vol. 53, no. 6, pp. 1–36, 2021.

[12] M. Tozakidou, I. Apine, K. U. Petersen et al., "Comparison of different iterative CT reconstruction techniques and filtered back projection for assessment of the medial clavicular epiphysis in forensic age estimation," *International Journal of Legal Medicine*, vol. 134, no. 1, pp. 355–361, 2020.

[13] S. Luo, L. Zheng, S. Luo, N. Gu, and X. Tang, "Data sustained misalignment correction in microscopic cone beam CT via optimization under the Grangeat epipolar consistency condition," *Medical Physics*, vol. 47, no. 2, pp. 498–508, 2020.

[14] J. Liu, H. R. Zhao, H. L. Wei et al., "Efficacy of bronchoalveolar lavage as adjunct therapy in the treatment of neonatal severe pneumonia: a prospective case–control study," *Journal of Tropical Pediatrics*, vol. 66, no. 5, pp. 528–533, 2020.

[15] F. Liu, L. Lei, L. Haitao, S. Liping, and P. Longkai, "The level of plasm donor-derived cell-free DNA in kidney transplant patients with severe pneumonia," *Transplantation*, vol. 104, no. S3, p. S317, 2020.

[16] F. Kassem and G. Kang, "Security and privacy in the Internet of Things," *Computer*, vol. 43, no. 13, pp. 765–778, 2019.

[17] H. Jiang, Z. Xiao, Z. Li, J. Xu, F. Zeng, and D. Wang, "An energy-efficient framework for Internet of Things underlaying heterogeneous small cell networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 2, pp. 232–246, 2020.

[18] J. Huang, J. Guo, H. Li, W. Huang, and T. Zhang, "Efficacy and safety of adjunctive corticosteroids therapy for patients with severe community-acquired pneumonia: a systematic review and meta-analysis," *Medicine*, vol. 98, no. 13, 2019.

[19] P. Guilpain, C. L. Bihan, V. Foulongne, P. Taourel, and V. L. Moing, "Rituximab for granulomatosis with polyangiitis in the pandemic of covid-19: lessons from a case with severe pneumonia," *Annals of the Rheumatic Diseases, annrheumdis*, vol. 80, no. 1, 2020.

[20] G. Zaharchuk and G. Davidzon, "Artificial intelligence for optimization and interpretation of PET/CT and PET/MR images," *Seminars in Nuclear Medicine*, vol. 51, no. 2, pp. 134–142, 2021.

[21] L. Fichera, G. Li-Destri, and N. Tuccitto, "Fluorescent nanoparticle-based internet of things," *Nanoscale*, vol. 12, no. 17, pp. 9817–9823, 2020.

[22] E. Hittinger and P. Jaramillo, "Internet of Things: energy boon or bane?," *Science*, vol. 364, no. 6438, pp. 326–328, 2019.

[23] A. M. Dimmitt, J. A. Pelz, M. E. Albertson et al., "Evaluation of adaptive statistical iterative reconstruction-V reconstruction algorithm vs filtered back projection in the detection of hypodense liver lesions: reader performance and preferences," *Journal of Computer Assisted Tomography*, vol. 43, no. 2, pp. 200–205, 2019.

[24] Z. Deng, Y. Hu, P. Yang et al., "Diagnosis and treatment of an acute severe pneumonia patient with COVID-19: case report," *Journal of Medical Virology*, vol. 92, no. 10, pp. 1728–1730, 2020.

[25] K. Cho, J. Chae, R. Y. Kwon, S. C. Bong, and K. S. Cho, "The application of the filtered backprojection algorithm to solar rotational tomography," *Astrophysical Journal*, vol. 895, no. 1, 2020.

[26] W. C. Chen, J. S. Niu, I. P. Liu et al., "Study of a palladium (Pd)/aluminum-doped zinc oxide (AZO) hydrogen sensor and the Kalman algorithm for Internet-of-Things (IoT) application," *IEEE Transactions on Electron Devices*, vol. 67, no. 10, pp. 4405–4412, 2020.

[27] J. Chauhan and P. Goswami, "An integrated metaheuristic technique-based energy aware clustering protocol for Internet of Things based smart classroom," *Modern Physics Letters B*, vol. 34, no. 22, 2020.

[28] V. Chang, V. M. Muñoz, and M. Ramachandran, "Emerging applications of internet of things, big data, security, and complexity: special issue on collaboration opportunity for IoTBDS and COMPLEXIS," *Computing*, vol. 102, no. 6, pp. 1301–1304, 2020.

[29] R. Ceipek, J. Hautz, A. de Massis, K. Matzler, and L. Ardito, "Digital transformation through exploratory and exploitative internet of things innovations: the impact of family management and technological diversification∗," *Journal of Product Innovation Management*, vol. 38, no. 1, pp. 142–165, 2021.

[30] D. T. Braddock, P. R. Stabach, K. Zimmerman, D. Kavanagh, C. T. Sauei, and K. J. Yarema, "Protein engineering and glycan optimization improves pharmicokinetics of an enzyme biologic 10-fold," *FASEB Journal*, vol. 30, Supplement 1, 2019.

[31] R. Boudewijns, H. J. Thibaut, S. J. F. Kaptein et al., "STAT2 signaling restricts viral dissemination but drives severe pneumonia in SARS-CoV-2 infected hamsters," *Nature Communications*, vol. 11, no. 1, 2020.

[32] D. Blez, A. Soulier, F. Bonnet, E. Gayat, and M. Garnier, "Monitoring of high-flow nasal cannula for SARS-CoV-2 severe pneumonia: less is more, better look at respiratory rate," *Intensive Care Medicine*, vol. 46, no. 11, pp. 2094-2095, 2020.

[33] M. Beckmann, P. Maass, and J. Nickel, "Error analysis for filtered back projection reconstructions in Besov spaces," *Inverse Problems*, vol. 37, no. 1, 2020.

[34] O. Barker, "Realizing the promise of the Internet of Things in smart buildings," *Computer*, vol. 53, no. 2, pp. 76–79, 2020.

[35] A. Villa-Henriksen, G. T. C. Edwards, L. A. Pesonen, O. Green, and C. A. G. Sørensen, "Internet of Things in arable farming: implementation, applications, challenges and potential," *Biosystems Engineering*, vol. 191, pp. 60–84, 2020.

[36] T. Atefeh, P. Reza, S. K. Mojtaba, M. Asghar, R. Mehdi, and R. Reza, "Pediatric regional Drl assessment in common Ct examinations for medical exposure optimization in Tehran Iran," *Radiation Protection Dosimetry*, vol. 21, no. 3, pp. 563–574, 2020.

[37] M. A. Guillén, A. Llanes, B. Imbernón et al., "Performance evaluation of edge-computing platforms for the prediction of low temperatures in agriculture using deep learning," *Journal of Supercomputing*, vol. 77, no. 1, pp. 818–840, 2021.

[38] R. Du, Y. Liu, L. Liu, and W. Du, "A lightweight heterogeneous network clustering algorithm based on edge computing for 5G," *Wireless Networks*, vol. 26, no. 3, pp. 1631–1641, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9849843, 1 page
https://doi.org/10.1155/2023/9849843

Retraction

Retracted: Application Analysis of New Internet Multimedia Technology in Optimizing the Ideological and Political Education System of College Students

Wireless Communications and Mobile Computing

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

Copyright © 2023 Wireless Communications and Mobile Computing. This is an open access article 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, "Application Analysis of New Internet Multimedia Technology in Optimizing the Ideological and Political Education System of College Students," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5557343, 12 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5557343, 12 pages
https://doi.org/10.1155/2021/5557343

WILEY | Hindawi

Research Article

Application Analysis of New Internet Multimedia Technology in Optimizing the Ideological and Political Education System of College Students

Haibin Yu

Harbin University of Science and Technology, Harbin 150080, China

Correspondence should be addressed to Haibin Yu; quentina_yu@hrbust.edu.cn

Received 26 January 2021; Revised 5 February 2021; Accepted 23 February 2021; Published 10 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Haibin Yu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ideological and political education is the most important part of the daily education management of college students. The methods and methods of ideological and political education are very old, and students' learning efficiency is very low. How to use ideological and political education technology combined with modern technology for teaching has become a current research hotspot. Based on this background, this article proposes to use a new type of Internet multimedia technology interactive teaching. In the ideological and political network multimedia teaching system, use occasions suitable for IP multicast applications, such as broadcast teaching, group discussions, and on-demand courseware. Among these functions, the IP multicast mechanism is appropriately used. The adoption of the extended conversation node scales each conversation group to a multicast group, and the members of the multicast group represent the participants, which brings convenience and ease of management. Through case study analysis, it can be seen that this method can reduce the burden on the system and improve efficiency, and the number of multicast members is unlimited, which has a very good auxiliary effect on course learning. Through the Internet multimedia technology, the innovation of ideological and political education has been realized, the learning environment of students has been improved, the ways of ideological and political education have been broadened, and the education system has been better optimized.

RETRACTED

1. Introduction

In recent years, with the rapid development of information technology, digital technology, and mobile communication technology, multimedia has increasingly penetrated into many aspects, such as politics, economy, culture, and social life [1]. It has become an important field closely related to the future of the country in the wave of information technology [2]. In particular, the typical representative of the new media—the Internet has entered every classroom, every dormitory, into the daily study of college students, and in their daily lives, college students have obviously become the most active group using the new media [3]. College students acquire a great deal of knowledge and information through new media, which also affect their way of life, study, and thinking. The ideological and political education of students brings both new opportunities and great challenges [4].

In the 3G era, the CPC Central Committee and the State Council have attached great importance to making good use of the mobile Internet to carry out ideological and political education and have repeatedly stressed the need to actively explore new ways and methods of ideological and political education for college students under the new situation [5]. Take the initiative to occupy the new position of network ideological and political education and firmly grasp the initiative of network ideological and political education [6]. As an important place for ideological and cultural exchanges and collisions, colleges and universities should make good use of new media technology, which is a favorable medium to carry out ideological and political education for college

students, and how to deal with the adverse effects of new media on college students [7]. To improve the effectiveness of ideological and political education for college students has become a major issue we are facing [8].

In the mature network communication environment of mobile Internet, we should correctly understand the main characteristics of mobile Internet and its position and role in ideological and political education of college students [9]. Actively explore the psychological and behavioral characteristics of college students' use of mobile Internet, solve the problems in ideological and political education of college students under mobile Internet environment, and innovate the ways and methods of ideological education work [10]. It is the mission of ideological education in colleges and universities in the new situation to seize the new network platform, give play to its greatest advantage, and reform its unfavorable factors [11–13].

Based on the existing research results in related fields, this paper attempts to explore this topic from three aspects [14]:

First, based on the existing research results of mobile Internet and network multimedia teaching, this paper defines the concepts of mobile Internet and ideological and political education and generalizes the application and construction of mobile Internet teaching model.

Second, this paper investigates the use of mobile Internet in college students and opportunities and challenges faced by ideological and political education of college students under mobile Internet environment. Mobile Internet as a double-edged sword, on the one hand, is conducive to expand the time and space of ideological and political education of college students and to enrich the content of ideological and political education of college students [15]. It provides a new communication platform for the ideological and political education of college students and helps to improve the effectiveness of the ideological and political education of college students. On the other hand, the mobile Internet brings a series of negative influences to the values and behavior of college students; the authority of ideological and political educators is challenged; the traditional ways and methods of ideological and political education are challenged [16].

Third are the thoughts on optimizing the construction of ideological and political education for college students under mobile Internet environment. This part is the focus and difficulty of the full text, but also the foothold of the full text. This paper gives some suggestions from five aspects: the construction of wireless network platform, students' self-education, educators, colleges and universities, and the state [17].

The research contributions of this article include the following:

(1) This article proposes a method of interactive teaching using new Internet multimedia technology

(2) In the ideological and political network multimedia teaching system, use occasions suitable for IP multicast applications, such as broadcast teaching, group discussions, and on-demand courseware

(3) The use of extended conversation nodes extends each conversation group to a multicast group, and the members of the multicast group represent participants, which brings convenience and ease of management

The rest of this paper is organized as follows. Section 2 discusses related work, followed by opportunities, and challenges faced by college students' ideological and political education under mobile Internet environment are discussed in Section 3. Countermeasures of ideological and political education for college students under mobile Internet environment are discussed in Section 4. Section 5 concludes the paper with summary and future research directions.

2. Related Work

What is the new media? Many scholars believe that the "new media" we are talking about is born in the 1960s, when P. Goldmark (P Goldmark) proposed a commercial development plan in 1967. He is here to refer to electronic video as the new media (new media). However, the popular view is not accurate, and the concept of "new media" dates back at least to 1959. In March of that year, Marshall McLuhan gave a speech at a national conference on higher education called "the Electronic Revolution: new revolutionary influence of the media". So, even since Marshall McLuhan's speech, the word "new body" has been in the public eye for at least half a century. But the fact is that since the 1960s, as the public has seen, the term "new media" has become popular in the United States, and the word "new media" has since been accepted worldwide [18].

After more than ten years of scientific research exploration and research accumulation, many achievements have been made in the future network research [19]. There are several important directions in network architecture: information-centered network, supporting mobile network, the fusion network of integration of computing perception and storage, and the network of separating control plane and data plane, etc. Information-centric network (ICN) represents projects such as Netlnf [20], NDN, and DONA. The purpose of this project is to represent the current "location-centric" network structure [21]. Become the network structure centered on "information" changes the current network "end-to-end" transmission mode, and the content becomes the most important entity of ICN network, and ICN makes the network focus on the information itself rather than the location of the data [22]. It can solve the current network content and repeat transmission and other issues. The network supporting mobility is mainly represented by Mobility First [23]. At present, the network is based on fixed end-to-end links and cannot meet more and more requirements of mobile devices and mobile services. Mobility First is a mobile scene oriented network architecture [24]. It is able to provide services such as car networking strong, credible mobility support. The integration of computing, perception, and storage can widely introduce perception, computation, and storage into the network, so that computing leads transmission, storage helps transmission, and breaks through the single simple data transmission function of current network. Enhance the integrated service

RETRACTED

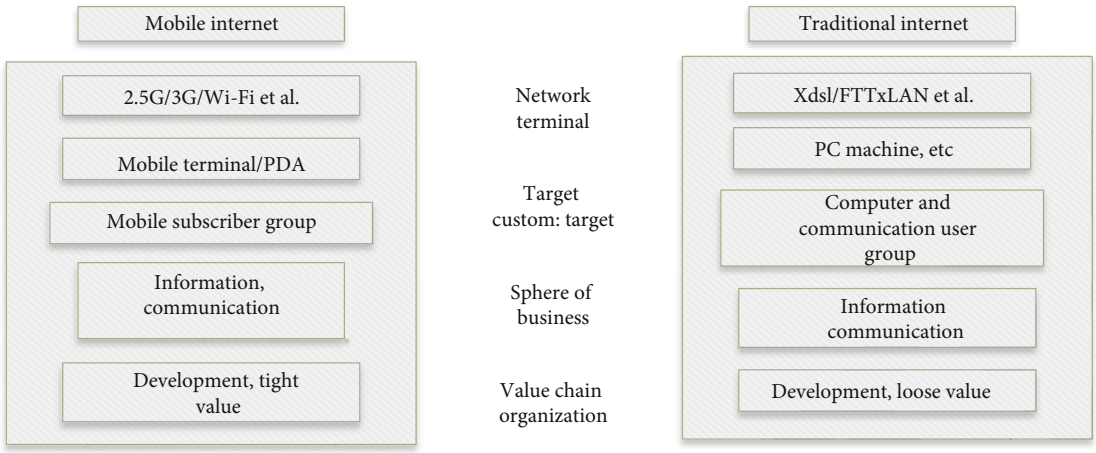


Figure 1: Comparison between mobile Internet and traditional Internet.

capability of the converged network. Figure 1 shows a comparison between the current mobile Internet and the traditional Internet [25].

3. Design of Network Multimedia Teaching System Model

In order to effectively control in multicast mode and give full play to the function of the teacher machine, we use the conference control mode of H.243: the chairman control mode [26]. This model is mainly used in group discussion, where the teacher machine is equivalent to the chairman node, controls the teaching process, coordinates the students' learning process, and avoids the possible confusion in the teaching process (screen switching, group discussion) [27]. In this mode, the usual processing process is as follows:

(1) The master initializes and listens for the message in the group

(2) A student who wishes to speak makes a request to the teacher's machine for a right to speak [28]

(3) The teacher machine records the mark (address, port) of the student machine in the list of group member attributes (Table 1) and continues to listen

(4) The teacher machine arbitrates and multicasts the result to the student machine in the group

(5) The students who have the right to speak multicast information to other members

(6) The students multicast the information to the group members

(7) The teacher machine removes the student machine from the group member attribute list and continues to select the student machine that wishes to speak. List of member attributes is shown in Table 1

Its status diagram is shown in Figure 2.

Table 1: List of member attributes.

Number	Priority	Local ID (local IP)	Group ID (multicast group IP)	Data port	Control port
1	0	61.14	224.0.0.8	5050	6666
2	3	61.14	224.0.0.8	5050	8888

It can be seen from Figure 2 that the state of ideological and political learning has a good relationship with tool technology, and each link can affect the effect and efficiency of learning.

3.1. Application of IP Multicast in Network Multimedia Teaching System. IP multicast is the latest technology in the field of IP. The main way to implement it is to broadcast multicast with class D address on the net. IP multicast adopts the concept of group address, and a class D IP address represents a group of hosts. Put the users who need the data into the user group and pass the data to the user who really needs it. The network for establishing a multicast environment is shown in Figure 3.

In the network multimedia teaching system, there are not few occasions suitable for the application of IP multicast, such as broadcast teaching, group discussion, and courseware on demand. In these functions, the IP multicast mechanism is used appropriately [29]. Undoubtedly, it can make the system make full and reasonable use of the limited bandwidth. It expands the scale of the session node to map each session group with a multicast group, whose members represent the participants, which can bring convenience in implementation and easy to manage. It can reduce the system's burden and improve efficiency, and the number of multicast members is unlimited. IP's A characteristic determines the possibility that the node size of the group has almost unlimited growth.

3.2. Multimedia Features of the Internet

3.2.1. Immediacy, Openness, and Convenience Are the Characteristics of New Media. Generally speaking, the

RETRACTED

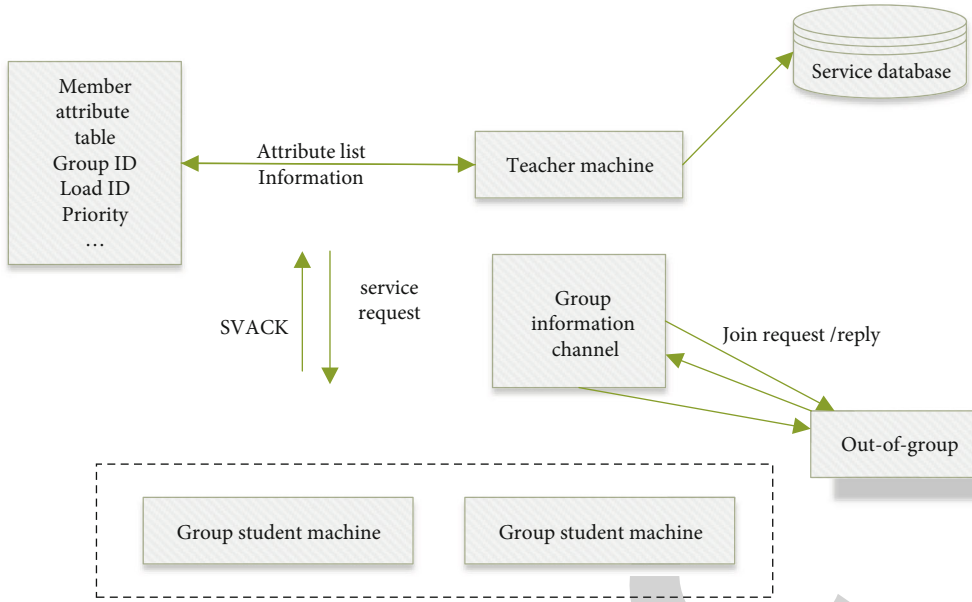


Figure 2: State diagram.

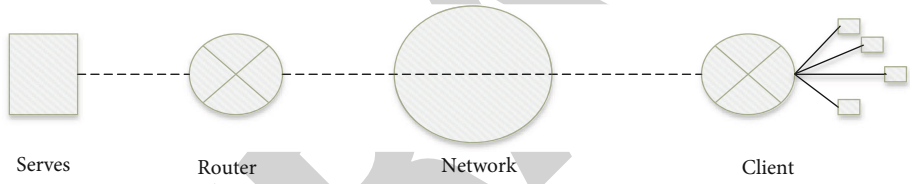


Figure 3: Multicast communication model.

publishing, broadcasting, and feedback of traditional media have a certain period, and the layout and duration are strictly limited, so their dissemination will be limited by time and space. New media based on digital technology can send and receive information at anytime and anywhere, breaking the boundaries of time and space completely. The traditional media information dissemination needs "gatekeeper" examination, and the audience must rely on "information collection and compilation center". The dissemination of traditional media is limited to countries and regions, while the new media use the Internet and communications satellites, completely breaking the time and the space. With the restrictions on speed and time, people can even use "net names" and other means to conceal personal real information, thus using the corresponding devices as "virtual" through Weibo, WeChat, and other platforms without any obstacles to their own articles and viewpoints. Video and other information sent and received the real realization of space and information openness. And the new media has the powerful information reserve function, and the use of the retrieval function can quickly and accurately find the information, which is rich in means of communication and content to meet the needs of different audiences. People can also have their preference-personalized choice, as far as possible to show themselves, which greatly enriched the work of people, study, and life. The conference video codec and transmission flow is shown in Figure 4.

3.2.2. Rich Content, Various Forms, and Multiple Values Are the Important Features of New Media. Under the new media environment, the information source is more diverse, and the new media is not limited by the layout and broadcast time, and its information capacity is infinite from the whole mankind. His message covers almost all aspects of human history and can span several continents and reach users all over the world at the same time. People can use new media to send text, pictures, video, audio, and so on and express their views to forward other people's information and so on. But at the same time, the same information has different values for different people because of the different subject. Sometimes it is profit-making information to the unscrupulous plotters, but harmful to the victims. Information or junk information, which make the value of information, vary from one person to another. In short, it can be seen from the analysis of the characteristics of new media that multimedia technology has been applied in all aspects of life. New media has not only changed people's life world. Multimedia technology has also changed the way people accept new things.

3.3. Measurement of Available Network Bandwidth. In networks, bandwidth refers to the rate of data transmission over

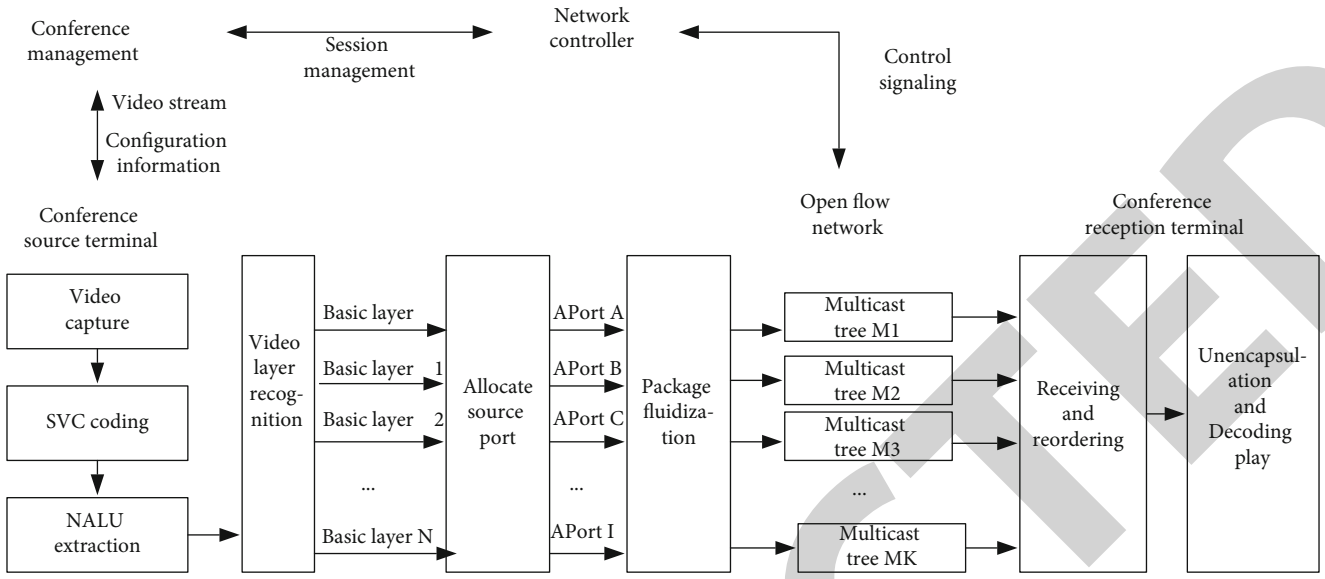


Figure 4: Flow chart of conference video codec and transmission.

network links or end-to-end paths, which can be measured by bit rate (bit per second, bps). The bandwidth resources in the network are usually limited. The current multimedia applications such as videoconferencing and IPTV will consume a lot of network bandwidth resources, and the available bandwidth of the application program will directly affect the performance of the application program. It is very possible for various applications to share network bandwidth resources to influence and compete with each other. The continuous increase of data traffic may eventually result in the total data traffic exceeding the maximum load of the network bandwidth, and then, the network congestion will occur. For bandwidth tube, the basic purpose of the rational and traffic control strategy is to ensure the efficient utilization of network resources while maintaining an acceptable level of service for specific services. The measurement of the bandwidth usage of network links is the basis of bandwidth management and traffic control. Link capacity refers to the maximum capacity that the network can provide for data transmission, and link capacity is usually determined. In the data link layer of the network, the data is normally sent and received at a constant rate, which depends on the characteristics of the transmission medium and the transceiver hardware and calculates the link capacity at the network layer. Need to consider the impact of additional overhead such as packet encapsulation in layer 2, specifically. Assuming the nominal transmission rate of a link at the data link layer, the transmission n time for a network layer packet of length Li3 is

$$\Delta L3 = \frac{LL3 + HL2}{CL2}. \tag{1}$$

Where the additional protocol overhead of encapsulating data packets into layer 2 data frames, the link capacity at the network layer (^ 3) can be obtained as follows:

$$CL3 = \frac{LL3}{\Delta L3} = CL2 \cdot \frac{1}{1 + HL2/LL3}. \tag{2}$$

As you can see, the ratio of link capacity is related, so we select a lower Li3 value to increase the < ^.

3.3.1. The Average Available Bandwidth at Intervals Can Be Expressed.

$$\bar{A}(t, t+\tau) = \frac{1}{\tau} \cdot \int_{t}^{t+\tau} A(\chi) d\chi. \tag{3}$$

The OpenFlow protocol provides the statistical function of sending and receiving port data, which can be used to measure the use of network bandwidth. First, the controller sends the statistics request message to the OpenFlow switch with the time period r, and then, the OpenFlow switch sends the counter value of the report port of the statistics reply message to the controller; by this method, the controller can get the average load of the link z over the time period (f, t+r) from the following formula:

$$\bar{\mu} i(t, t+\tau) = \frac{Ni(t+\tau) - Ni(t)}{\tau}. \tag{4}$$

In OpenFlow networks, link capacity is usually known and specific values can be obtained by querying device information. After obtaining the average load within + fZ (I+r), you can calculate the average available bandwidth of link I over the software defined network (Z, I+T):

$$\bar{A} i(t, t+\tau) = CL3 - \overline{\mu 2}(t, t+\tau). \tag{5}$$

The available bandwidth of the network link can be

RETRACTED

obtained by 33.5-inch mode, thus providing the reference for the management of QoS [3] in the aspect of network bandwidth resources.

3.4. Current Situation of Multimedia Development in China. With the continuous evolution of multimedia and Internet technology, the number of multimedia users, the number of multimedia applications, and the network traffic occupied by multimedia services are increasing with each passing day. At home, the 39th China Internet Development Statistics report by China Internet Network Information Center (CNNIC) and the China Internet Information Center show that by the end of 2016, the number of Internet video users in China had increased by 40.64 million. It has reached 545 million (as shown in chart 1. 1), an increase of 8.1% over the same period last year. The usage rate of network video users has risen to 74.5%, the number of mobile phone network video users has reached nearly 500 million, an increase of 23.4% over the same period last year, and the utilization rate of mobile phone network video has increased to 71.9%. In addition, the number of live webcast users has increased to 47.1% of the total number of Internet users, up 19.32 million, 1% from mid-2016. Globally, according to a white paper issued by Cisco, the global IP traffic will reach 2.3ZB/year by 2020, and IP video traffic will account for 82% of the total Internet user traffic. There will be millions of minutes of video content per second over the Internet, video surveillance, virtual reality, Internet television, and VOD, and other multimedia applications will grow dramatically. It can be seen that multimedia service has become one of the most important services in the network. Size and usage of Internet Video/Mobile Network Video users from December 2014 to December 2016 are shown in Figure 5.

3.5. The General Situation of the Use of the Mobile Internet among University Students. With the development of digital technology, new media has become an indispensable part of college students' life. According to the survey of Chinese teenagers, Figure 5 shows that "College students in most of the use of Internet applications are high, among which the use of search engines and network news reached 92.3% and 85.4% respectively." The utilization rate of network music and video was 88.8% and 78.9% respectively, and that of instant messaging and email was 95.1% and 82.7%, respectively. The usage rate of online shopping, online payment, and online banking was 58.6%, 58.6%, and 4%, respectively. It can be seen that new media has become an important channel for college students to obtain information, exchange ideas, and express their personal will. At present, the proportion of college students using mobile phone Internet access reached 97.6%. Figure 5 shows that 68.8% of the students said they used their mobile phones to surf the Internet every day, and 27.6% of the students used their mobile phones more than four times a day or more. Most of the time college students use their cell phones to surf the Internet every day is within 15 minutes. Figure 5 is the main application direction of student's mobile phone.

4. Research on Ideological and Political Education of College Students in the Environment of Mobile Internet

In recent years, with the development of wireless network technology, mobile Internet has become an important part of college students' study and life. With its unique characteristics of portability, immediacy, privacy, interactivity, compulsion, and fragmentation of time utilization, the mobile Internet has brought unprecedented new opportunities for the development of ideological and political education in colleges and universities. But at the same time, it also brings great challenges to the ideological and political education in colleges and universities. Frequency map of using mobile phone for college students is shown in Figure 6.

4.1. Mobile Internet Brings New Opportunities to Ideological and Political Education of College Students. Most students can quickly obtain a large amount of information through mobile phone, network, wireless terminal, and so on and organize and filter the information to their advantage. This plays an important role in college students' self-management, overall development, and liberation of personality freedom. Through such communication platforms as QQ, E-Mail, Faction, Weibo, and WeChat, college students are more relaxed and free in interpersonal communication, which provides a channel for students who are unwilling and not good at interpersonal communication in peacetime to express themselves and communicate with others and makes easy for them to get a sense of accomplishment and satisfaction. The way of thinking of the college students under the new media environment great changes has also taken place in learning methods. Students can search all kinds of learning resources, watch courses online, and turn passive learning into active learning according to their own interests. At the same time, a large number of college students choose to purchase and pay online, becoming the main customer group of online business applications. New media has deeply penetrated into all aspects of college life and has had a great impact on their way of life, way of thinking, mode of behavior, and value orientation. Figure 6 is the main application direction of student's mobile phone.

The traditional educational carrier in classroom includes TV, fixed computer, and print media. Because of its own physical restrictions, even in the classroom, it is very difficult to be fully shared by the students, and not to mention, it is impossible to be enjoyed by students anytime and anywhere after class. Through the mobile Internet network connection, each mobile phone becomes a small personal computer. Students can not only access the Internet through a mobile phone terminal, but also obtain massive amounts of information immediately. Moreover, it can download and store the most valuable information through the mobile memory card, and the mobile Internet can fully realize the sharing of resources between students and teachers. Figures 7 and 8 show, currently, the locations where college students use mobile phones to surf the Internet are mainly in the study rooms (32%), classrooms (31%), dormitories (29%), libraries (11%), and other public places (5%), which cannot be

RETRACTED

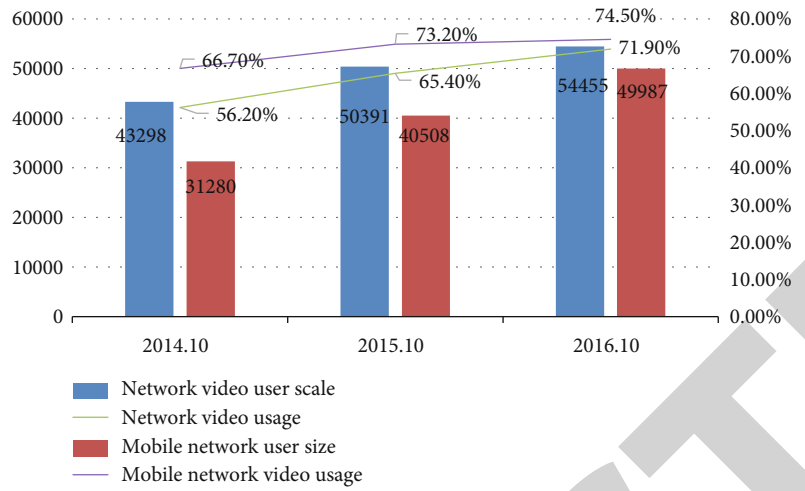


Figure 5: Size and usage of Internet Video/Mobile Network Video users from December 2014 to December 2016.

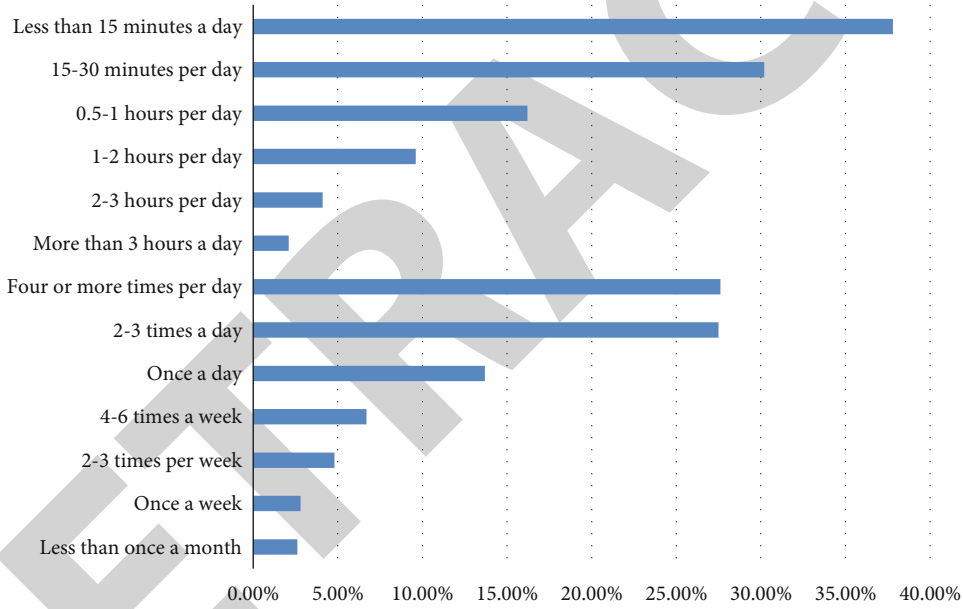


Figure 6: Frequency map of using mobile phone for college students.

covered by traditional networks. Especially in traffic (78.5%), waiting (68.9%), bedtime or early morning (42.3%), these are previously difficult to use the fragmentation of time. As a result of the integration of mobile communication and traditional Internet, mobile Internet not only has the flexibility and mobility of mobile communication, but also makes full use of the openness and interactivity of traditional Internet, so that the network and users can start. Finally, it can expand the time and space of ideological and political education in colleges and universities without limitation of time and space. The main application directions of students using mobile phone to surf the internet are shown in Figure 7. Spatial distribution of students using mobile Internet is shown in Figure 8. Time distribution of students' use of mobile Internet is shown in Figure 9.

4.2. Challenges Faced by Ideological and Political Education of College Students in Mobile Internet Environment. At the same time, new media is also a double-edged sword for university students and their huge social benefits. Internet technology has brought new opportunities for ideological and political education. At the same time, new Internet technology also has many limitations. For example, in the process of Internet technology teaching, filling-in teaching is difficult for learners to accept and teaching tools are limited.

4.2.1. College Students' Values and Behavior Are Affected. The important stage of the formation of human ideology is the adolescence, during which the thought is highly plastic. The new media provides a virtual world for college students, where they can freely express their personal will, show their

RETRACTED

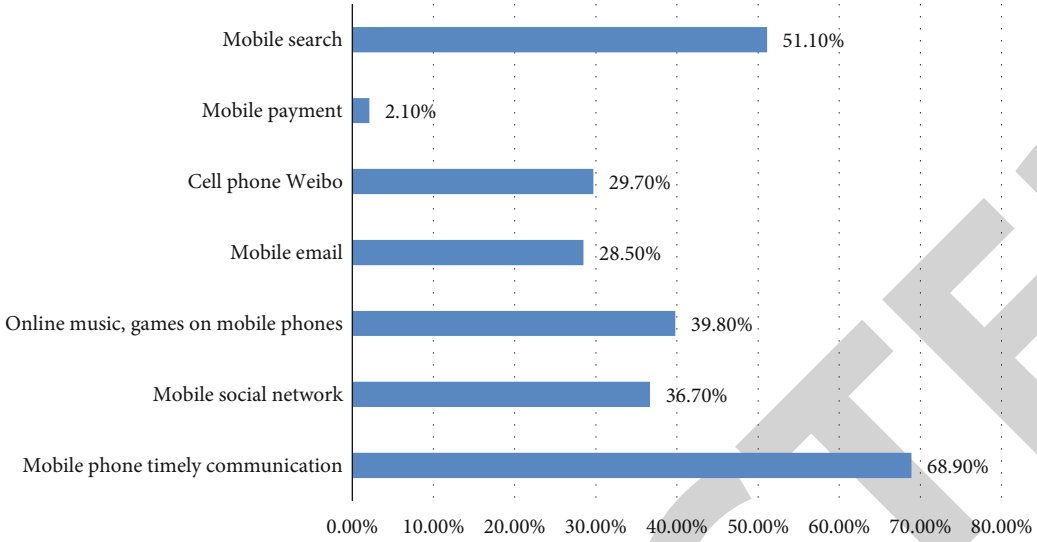


Figure 7: Main application directions of students using mobile phone to surf the Internet.

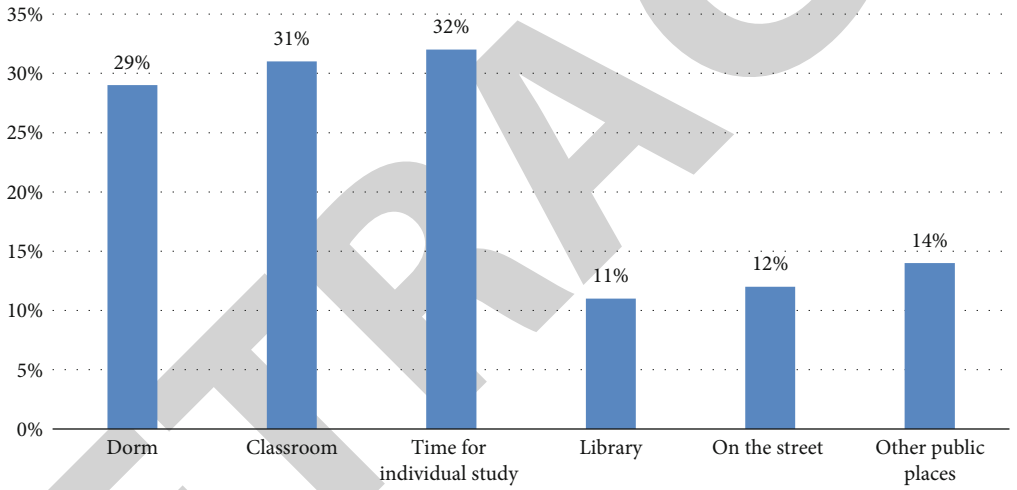


Figure 8: Spatial distribution of students using mobile Internet.

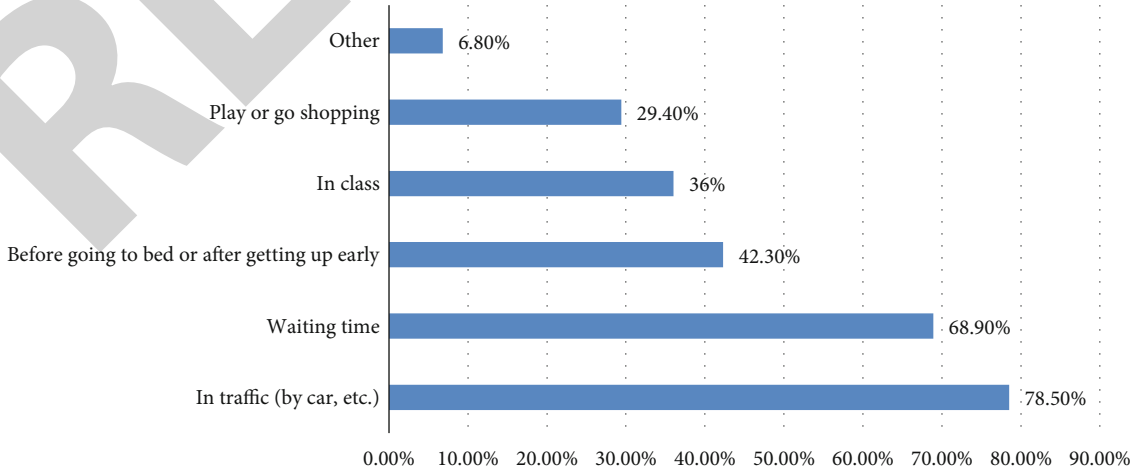


Figure 9: Time distribution of students' use of mobile Internet.

RETRACTED

talents, and better convey their feelings. This coincides with college students' psychology of advocating individuality, equality, and seeking independence. The "accessibility" of the new media has fully aroused the enthusiasm and initiative of college students, who yearn for freedom, independence, and the realization of their own value and are willing to express their views and opinions. The complexity of network information is greatly reduced. It reduces the students' ability to discern information. Figure 9 shows the messages published on the current mobile Internet in which only 7.7% of the students said they could fully identify the truth and falsehood, while 47.3% said they are rarely able to identify the authenticity of information. Students' ability to identify the authenticity of mobile Internet information is shown in Figure 10.

In recent years, with the continuous development of wireless network technology, the mobile Internet has more and more illegal elements that spread colorLove, violence, and other harmful information as an important channel. The Ministry of Public Security, the Ministry of Culture, and other relevant departments aim at mobile phone websites. Machine short message dissemination of spam pornographic information is to carry out a number of rectification activities. Figure 10 shows that at present, only 19.10% college students say they never often receive violence, crime, pornography, etc., when they use their cell phones to surf the Internet. For negative content of the information, most college students have more or less received related bad information. How often students receive negative messages using the mobile Internet is shown in Figure 11.

At present, the popularity of smart phones for criminals to spread a variety of yellow, violent crime, and other bad information of the dynamic picture is more convenient. As soon as some mobile phone websites are opened, they will pop out all kinds of provocative subtitles and dynamic pictures, such as sexy eyebrow and bloody fights, which have a great negative impact on the growth of college students' physical and mental health. Some students cannot resist the temptation or even change from the receiver of bad information to a disseminator.

4.2.2. Traditional Ways of Ideological and Political Education Are Challenged. In the 3G era, the rapid development of mobile Internet technology has profoundly changed people's way of life, mode of production, mode of thinking, and mode of information dissemination. By mobile Internet, this information revolution's influence is more comprehensive and profound. At present, the use of mobile phones to surf the Internet has gradually become an important means of communication and communication between ideological and political educators and college students in colleges and universities and has not only brought about a comprehensive impact on the lives of educators and students. It also puts forward new challenges to the traditional ways and methods of ideological and political education.

The survey shows that 27.10% of the college students are not very interested in the ideological and political education curriculum, and only 11.20% of the students think that the current ideological and political education methods are very

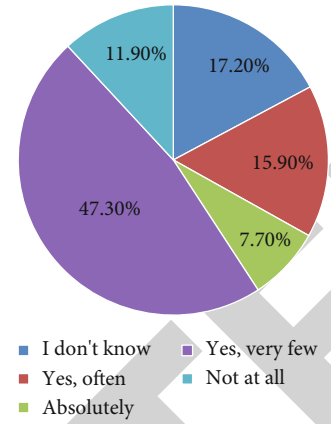


Figure 10: Students' ability to identify the authenticity of mobile Internet information.

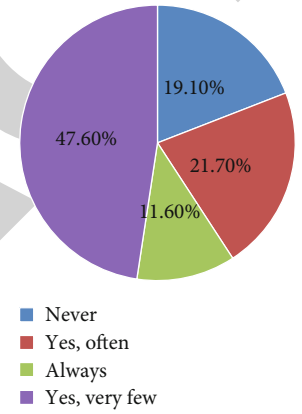


Figure 11: How often students receive negative messages using the mobile Internet.

rich. In the traditional ideological and political education classroom, educators enjoy absolute information advantages and high prestige, and condescending indoctrination education is the main way of education for educators. Moreover, the teaching method is single, the interaction of ideological and political education is not strong, the classroom atmosphere is dull, and the initiative of students to participate in the discussion is not high, with the change of information technology, the interactivity, immediacy, and openness of mobile Internet. The traditional ways and methods of ideological and political education are faced with severe challenges. Degree of interest of students in current ideological and political education is shown in Figure 12.

4.3. Analysis of the Reasons for the Challenges Faced by Ideological and Political Education in the Mobile Internet Environment

(1) Colleges and universities pay less attention to the construction of mobile Internet platform: at present, most colleges and universities in the use of mobile Internet to carry out ideological education work

RETRACTED

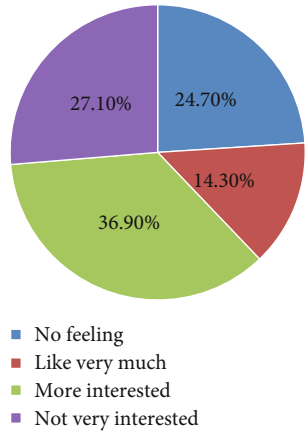


Figure 12: Degree of interest of students in current ideological and political education.

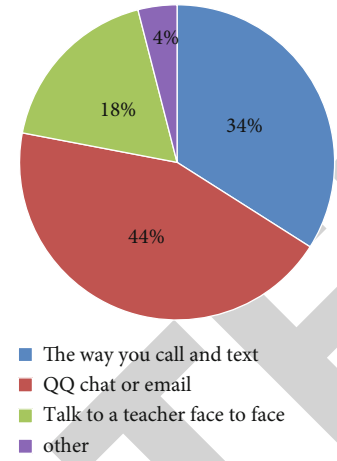


Figure 13: The most important communication between students and educators.

often do not accurately grasp the mobile Internet information dissemination mode and technical characteristics. The lack of comprehensive, scientific, and systematic planning for the use of the mobile Internet for ideological and political education and the failure to fully integrate the various forces have led to insufficient efforts and unclear planning in the use of the mobile Internet to carry out educational work at present. The target is unknown. The use of mobile Internet to carry out ideological and political education work cannot do without certain funds and technical investment, only sufficient capital investment and technical design. Only when the support is provided, can it be operated normally and effectively. At present, an important reason for the low effect of ideological education in colleges and universities using mobile Internet is the lack of funds and technical investment

(2) The ability of educational subjects to use mobile Internet is generally on the low side: mobile Internet brings a new space and environment to ideological and political education in colleges and universities. At the same time, it puts forward a higher demand for the cultivation of talents in the subject of ideological and political education in colleges and universities. If the main body of education wants to use mobile internet to carry out ideological and political education, he must constantly improve his own network knowledge, with a large number of firm political stand, solid professional foundation, deep theoretical foundation, and strong business ability. The main body of ideological and political education which can master and apply the knowledge and skills of mobile interconnection is an important basic condition for the effective development of ideological and political education in the mobile internet. However, at present, the majority of ideological and political education personnel are building, which lags behind the development of mobile Internet. As shown in Figure 12, more and more college students use new communication methods to communicate with educational work. The most important communication between students and educators is shown in Figure 13

(3) The rules and regulations to ensure the effective operation of the mobile Internet are not perfect: the use of the mobile Internet to carry out ideological and political education must be based on strict and perfect rules and regulations. Mobile Internet is a new type of internet in our country. Therefore, with the increasing popularity and penetration of mobile Internet in colleges and universities, it is necessary to make relevant network rules and regulations so as to reduce the negative impact of mobile Internet on college students' education. Actively promote the effective operation of ideological and political education in colleges and universities under the new situation

(4) The theoretical research of mobile Internet lags behind the development of mobile Internet: mobile Internet is a new foreign product of western culture in China, until 2009; when the national 3G license was issued, the mobile Internet began to develop rapidly in China. Realize the true meaning of China Mobile Network Age. As a double-edged sword, on the one hand, the mobile Internet enriches the content of college students' ideological and political education and expands the space and time of college students' ideological and political education by virtue of its advantages of openness, fictitious, instantaneity, and interactivity. On the other hand, the opening up of the mobile Internet is also the western trend of mastering the right to speak on the Internet. It opens the door to carry out ideological infiltration into our country and brings new challenges to our ideological and political education

(5) The construction of "mobile phone platform" in ideological and political education has lagged behind: at present, most colleges and universities have

RETRACTED

established their own mobile phone short message, mobile phone newspaper, mobile phone Fetion, mobile phone web page, and so on to actively carry out ideological and political education work. However, the current ideological and political education mobile platform construction is still in the traditional mode of work, and mobile Internet features are not obvious

4.4. Countermeasures of Ideological and Political Education for College Students under Mobile Internet Environment. The development of mobile Internet brings unprecedented challenges and opportunities to ideological and political education of college students. On the one hand, the traditional ways and methods of ideological education, the ideological and moral quality of educators, the quality of science and culture, the ability to deal with modern information technology, and the values and behavior of college students are all facing severe challenges. On the other hand, as an advanced, revolutionary scientific, and technological force, the mobile Internet will continue to promote the globalization of knowledge economy and accelerate the development and transformation of human society. It opens up new space and creates new channels for the ideological and political education of college students. Facing the new situation, we must change the traditional concept of ideological education work in time, build a new platform for ideological and political education by using the new technology of wireless network, and strengthen the cultivation of college students' ideological education professional team. In innovating the ways and methods of ideological and political education, we should strengthen the construction of mobile internet, perfect the management system of mobile internet, strengthen the censorship of mobile internet information, and purify the network environment.

5. Conclusion

Multimedia technology has been deeply rooted in many aspects of social life and has become an important field closely related to the future of the country in the wave of information technology. The interaction of new media provides a broad communication platform for college students and broadens the communication channels between teachers and students, and the openness of multimedia greatly enriches the content and methods of ideological and political education. The anonymity and convenience of multimedia can also provide a unique perspective to understand the effects of ideological and political education in a timely and comprehensive manner and greatly improve the effectiveness of work. Give full play to the positive role of new media in the multimedia environment, strengthen the construction of campus networks, and take the initiative to occupy a new position in ideological and political education in new media. In the process of ideological and political education, we must make full use of new media to broaden our horizons, build a new platform for ideological and political education in universities, strengthen the construction of campus culture, and create a good learning environment.

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 interest regarding the publication of this paper.

References

[1] D. Liang, "Innovative strategies of ideological and political education for college students based on mobile Internet in 3G era," *Ideological and Theoretical Education*, vol. 1, pp. 82–86, 2010.

[2] W. Ji and Z. Yong, "An analysis of the concept of new media," *New Media Research*, vol. 1, no. 12, 2015.

[3] C. Shi, Y. Li, J. Zhang, Y. Sun, and P. S. Yu, "A survey of heterogeneous information network analysis," *IEEE Transactions on Knowledge & Data Engineering*, vol. 29, no. 1, pp. 17–37, 2017.

[4] A. S. Shafigh, B. Lorenzo, S. Glisic et al., "A framework for dynamic network architecture and topology optimization," *IEEE/ACM Transactions on Networking*, vol. 24, no. 2, pp. 717–730, 2016.

[5] X. D. Liang, "The design and implementation of mutual-help teaching system based on multimedia network," *Procedia Environmental Sciences*, vol. 12, pp. 1231–1236, 2012.

[6] Y. F. Li, "Design and realization of network multimedia teaching system based on B/S model," *Applied Mechanics & Materials*, vol. 687-691, pp. 2447–2450, 2014.

[7] J. Cheng and D. Li, "On ways to infiltrate ecological civilization education into ideological and political education of higher vocational colleges," *Science Education Article Collects*, vol. 56, no. 4, pp. 337–341, 2016.

[8] Q.-Y. Zuo, M. Chen, G.-S. Zhao, C.-Y. Xing, G.-M. Zhang, and P.-C. Jiang, "Research on SDN technology based on OpenFlow," *Journal of Software*, vol. 24, no. 5, pp. 1078–1097, 2013.

[9] M. Yang and W. X. Lim, "Recent development of internet finance in China," *East Asian Policy*, vol. 7, no. 3, pp. 46–60, 2015.

[10] E. Logota, F. B. Saghezchi, H. Marques, and J. Rodriguez, "Cooperative strategies for end-to-end energy saving and QoS control," in *Novel 3D media technologies*, pp. 135–161, Springer, 2015.

[11] A. Wongsriwor, V. Imtawil, and P. Suttisopapan, "Design of rate-compatible LDPC codes based on uniform shortening distribution," *Engineering and Applied Science Research*, vol. 45, no. 2, pp. 140–146, 2018.

[12] M. Baldi, N. Maturo, E. Paolini, and F. Chiaraluce, "On the use of ordered statistics decoders for low-density parity-check codes in space telecommand links," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, pp. 1–15, 2016.

[13] S. van Ginkel, J. Gulikers, H. Biemans, and M. Mulder, "Towards a set of design principles for developing oral presentation competence: a synthesis of research in higher education," *Educational Research Review*, vol. 14, pp. 62–80, 2015.

RETRACTED

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6628068, 13 pages
https://doi.org/10.1155/2021/6628068

Research Article

Analysis of e-Commerce Characteristics Based on Edge Algorithm and COX Model

Jinzhong Zuo

Chaohu University, Chaohu, Anhui 238024, China

Correspondence should be addressed to Jinzhong Zuo; 050003@chu.edu.cn

Received 27 December 2020; Revised 14 January 2021; Accepted 22 February 2021; Published 9 March 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Jinzhong Zuo. This is an open access article 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 C2C (consumer to consumer) e-commerce transactions in China, the problem of information asymmetry in C2C e-commerce transactions is becoming more and more obvious. This paper analyzes the specific performance and influence of information asymmetry in C2C e-commerce transactions and then proposes a strategy to eliminate information asymmetry in C2C e-commerce transactions. Based on the perspective of information asymmetry, C2C e-commerce is taken as the research object, and the research methods such as literature analysis method, qualitative analysis method, and case analysis method are comprehensively used to explore the C2C e-commerce supervision problem in China under the information asymmetry environment. The reason and countermeasures. Secondly, it analyzes the e-commerce government supervision theory in detail and studies the definition, content, and principle of e-commerce government supervision, the necessity of e-commerce government supervision, and the main forms of e-commerce government supervision in China. The status of China's e-commerce supervision is elaborated, including the question of "who is going to supervise," "who is supervising," and "how to supervise." Aiming at the problems of imperfect legal system in China's e-commerce supervision, lack of cross-departmental coordination mechanism, and difficulty in investigation and evidence collection of e-commerce illegal behavior, we put forward targeted and operational policy recommendations.

1. Introduction

With the development and progress of society, human beings have entered the era of e-commerce, which has changed the traditional business model and improved the level of information management, playing an irreplaceable positive role in China's economic development. Based on this, it is very necessary to promote the progress of the e-commerce era. In-depth understanding of the actual situation of e-commerce operation found that information management and information systems have a greater impact on e-commerce. Therefore, in order to adapt to the development needs of the e-commerce era, it is very meaningful to strengthen the information management and information system construction.

e-Commerce refers to various business activities carried out by means of servers and browsers in the context of the current network environment, so that merchants and consumers can realize online transactions and directly realize shopping and transactions through electronic payment and the like. Compared to traditional business, e-commerce is a new service that combines financial, transaction, and business activities.

According to the transaction object of e-commerce, e-commerce can be divided into the following categories: The first category is enterprise-to-business e-commerce (business to business, referred to as "B2B"). It refers to the way in which enterprises conduct transactions through e-commerce, and the sales method relative to B2C is business to customer. For example, a factory that sells disposable tableware, if its main shipment target is a restaurant, this is the B2B model, where the restaurant provides disposable tableware to the user, and the restaurant-to-user model is focused on B2C/B2B [1]: the establishment of an interenterprise network and the stability of the supply chain system. B2C needs to rely on economies of scale to attract purchases and lower prices to increase profits. B2B does not rely on scale, but relies on the establishment of an interenterprise network to

stabilize its sales. The second category is business to consumer e-commerce (business to customer, referred to as "B2C"). This form of e-commerce is generally based on the online retail industry, mainly through the network to carry out online sales activities, and is an electronic retail model. The B2C model is the earliest e-commerce model in China. The third category is consumer-to-consumer e-commerce (customer to customer, referred to as "C2C"). It means e-commerce behavior among consumers. For example, a consumer has a computer, trades through the network, and sells it to another consumer. This type of transaction is called C2C e-commerce. The fourth category is business to government (B2G). It is an e-commerce activity between enterprises and enterprises and government agencies. For example, some local governments in China are currently experimenting with online procurement. The government publishes the purchase list through the Internet and conducts bidding through online bidding. Enterprises and companies bid through electronic means and finally complete the government procurement task electronically [2].

The rest of this paper is organized as follows. "Characteristics of Information Asymmetry in C2C e-Commerce" introduces the information asymmetry characteristics of C2C e-commerce, followed by the game analysis of e-commerce regulation designed in "Public Policy of e-Commerce Supervision in China," the information asymmetry strategy analysis is discussed in "Eliminate the Strategy of Information Asymmetry in C2C e-Commerce Transactions," and "Conclusion" concludes the paper with summary and future research directions.

2. Characteristics of Information Asymmetry in C2C e-Commerce

2.1. Information Asymmetry Performance. Although the state has relevant regulations, the opening of an electronic store must go to the industrial and commercial department for registration, but in practice, this provision is ineffective. No matter who you are, if you want to do online retail, you do not have to go to the industrial and commercial department to register. You can log in to any C2C e-commerce website and provide relevant information and documents for online application. You can open the store free of charge on the C2C e-commerce website. Operating, arrange the storefront as long as you choose the template provided by the website. The displayed products are all displayed by pictures. The sales of the products are managed automatically by the ready-made software. This means that you can open a C2C e-commerce website with a little time, an all-encompassing super department store. Since the cost of input is almost zero, the opening of the C2C electronic store is like a child's play. When do you want to open it, and when do you want to turn it off? When the switch is completely closed, the switch does not need to be considered. Because of this, the credibility and business status of a C2C electronic store is only clear to the merchants themselves, and it is difficult for consumers to obtain relevant information. Consumers are at a serious information disadvantage compared to merchants [3]. e-Commerce construction has been transformed from an industry to include financial cooperation between governments, enterprises, and other industries and institutions, as shown in Figure 1.

In C2C e-commerce transactions, consumers and merchants are engaged in real transactions in a virtual environment without encountering them. Consumers cannot see real merchandise, and the understanding of merchandise depends entirely on the images and text on the web. The pictures and texts of these related products are provided by the merchants themselves, the authenticity of the pictures and texts, and the closeness of the products [4].

Of course, only the business itself knows the degree well, but consumers do not know. Obviously, in the C2C e-commerce transaction, for the quality of goods, the merchant has all the information of the goods and is in an absolute information advantage; the consumer has no knowledge of the situation of the goods and is in a complete information disadvantage. The information on the quality of goods between consumers and businesses is seriously asymmetrical. The characteristics of e-commerce market are summarized in Table 1.

2.2. Impact of Information Asymmetry in C2C e-Commerce Transactions

2.2.1. Damage to the Interests of Consumers. Since C2C e-commerce transactions are conducted under a virtual network environment, consumers and businesses do not meet each other. Consumers' understanding of products depends entirely on the pictures and captions displayed by the merchants on the web page. Because the merchants open a C2C e-commerce store, there is almost no need to invest any cost. Therefore, in the case of severe information asymmetry, the merchant may publish false information on the web page for the purpose of maximizing profits, exaggerating the performance of the commodity, and inducing consumption. In the decision to make a purchase, although some C2C e-commerce websites have provisions for returning goods, because these businesses are the parents of C2C e-commerce websites, C2C e-commerce websites are mostly biased towards "parents" in order to not offend "parents" [5]. In one party, after the consumer finds that he has been defrauded, if he wants to return the goods, he must first negotiate with the merchant to see if the merchant is willing to return, how much money to return, and who will bear the logistics cost of returning the goods. Obviously, in the process of negotiating with the merchant, the consumer is in negotiation. In inferior position, if the negotiation fails, the consumer must complain to the C2C e-commerce website and must provide the identification certificate of the relevant institution, and the appraisal organization must pay the time and money for the appraisal of the goods. If consumers want to take the road of legal rights protection, due to the lack of supervision of online business and the gaps in e-commerce laws, and the litigation costs often far exceed the value of goods purchased by consumers, these issues make legal rights. It has become a kind of "expectation" [6].

2.2.2. Increase Transaction Costs. Because the merchant will conceal the defect of the product and exaggerate the

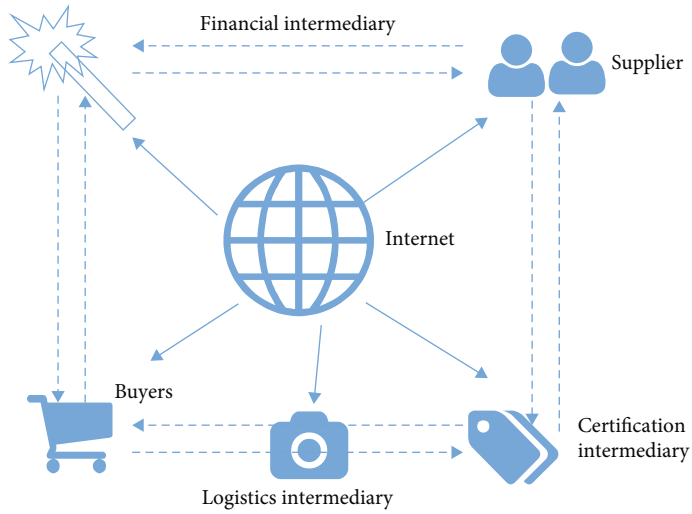


Figure 1: e-Commerce system structure.

Table 1: Market characteristics of e-commerce.

	Complete competition	Monopolistic competition	Oligopoly	Complete monopoly	e-Commerce market
Vendor number	Mass	Mass	Few	Only	A lot of
Completeness of information	Complete information	Incomplete information	Incomplete information	Incomplete information	Price information is variable, quality is relatively poor
Enter condition	Free entry	Free entry	Entry barriers are high and barriers to entry are extremely high	Entry barrier	Low barrier
Product homogeneity	Homogeneous	Differentiation	Homogeneous or differentiated	No similar alternatives	Differentiation

performance of the product, the consumer does not dare to easily believe the description of the product on the web page. In order to protect their legitimate rights and interests and avoid being deceived, consumers must learn more about the goods, which requires a lot of time, energy, and financial resources to collect information. In this way, regardless of whether the final transaction is successful or not, the consumer has already invested in the cost of the time, effort, and financial resources used to collect the product information before the transaction. The cost of unwarranted consumers due to information asymmetry seriously damages the interests of consumers and reduces the efficiency of transactions in the C2C e-commerce market [7].

2.2.3. Initiating "Reverse Selection." Because, in the C2C e-commerce transaction process, the information between the consumer and the merchant is seriously asymmetrical, the merchant knows the true quality of the product, and the consumer does not know the true quality of the product. Therefore, in the case where the quality of the product cannot be confirmed, consumers are only willing to buy goods at low prices for the purpose of protecting their own interests from loss, while high-quality goods are unattended because of higher prices than inferior goods. In order to obtain the corresponding profits, the merchants who manage high-quality goods will certainly cater to the low-cost choices of consumers and operate inferior commodities. In the end, the inferior goods will flood the C2C e-commerce market and form a situation of "inferior goods expelling good goods," resulting in inefficient market. The interests of consumers and businesses are undermined [8].

3. Public Policy of e-Commerce Supervision in China

According to the "China e-Commerce Report" published by the Ministry of Commerce, the amount of e-commerce transactions in the mainland in 2013 exceeded 10 trillion yuan (RMB, the same below), reaching 10.28 trillion yuan, a year-on-year increase of 26.8%. The transaction volume of e-commerce enterprises is shown in Figures 2 and 3.

3.1. Difficulties in China's C2C e-Commerce Tax Collection and Management

3.1.1. The Scale of Collection and Management Is Huge and Lacks Legal Basis. C2C is actually a professional term for e-commerce, which refers to the e-commerce between individuals, where C refers to consumer, which is abbreviated to C because the English word for consumer is customer (consumer), and C to C is abbreviated to C2C because the pronunciation of 2 in English is the same as to. C2C is

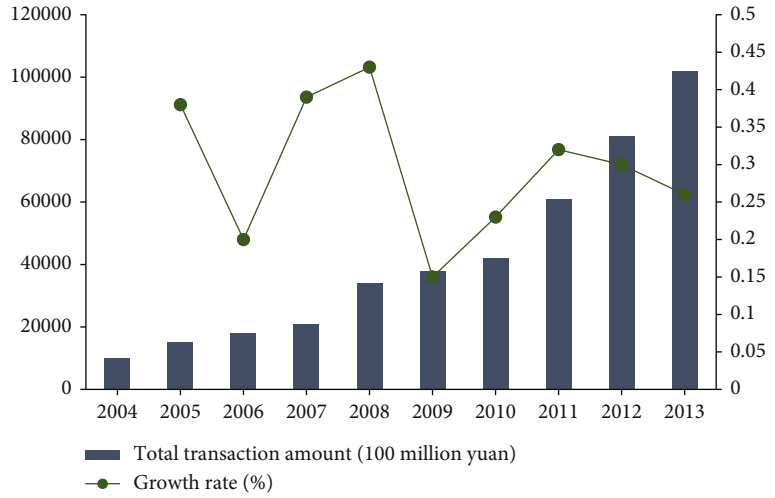


Figure 2: Changes in the total amount and trend of e-commerce transactions in China from 2004 to 2013.

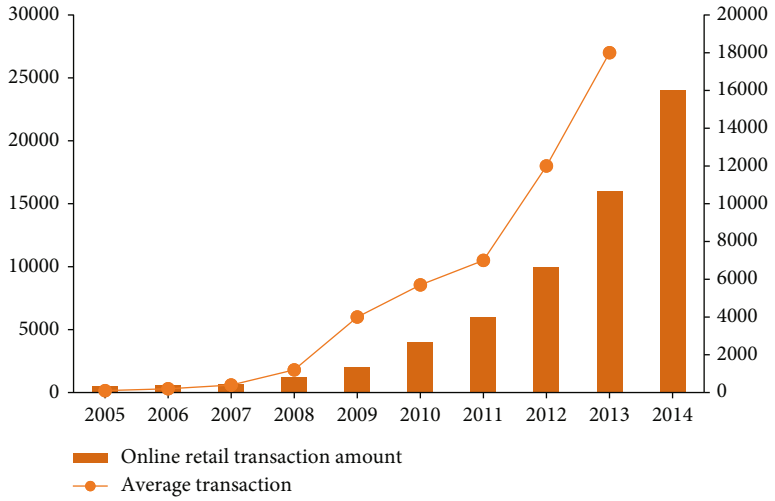


Figure 3: 2005-2014 C2C network retail transaction amount map.

customer to customer. Since 2008, the individual merchants in China's C2C e-commerce have shown a spurt of development, mainly in the scale and quantity of transactions, both of which have significantly improved. As of December 2015, according to relevant data, there are 418 million online payment users, and this growth momentum will not continue to grow. Although we have entered the era of big data and the efficiency of processing tax-related information has been significantly improved, the data is too large to avoid omissions, and the situation is relatively complicated in the actual tax collection and management, relying on the information in China's current tax system. Management ability and resource status are difficult to monitor taxation one by one [9].

3.1.2. The Elements of the Collection and Management Are Difficult to Determine. It is difficult to determine the object of collection and the taxpayer. C2C e-commerce transactions include all kinds of physical goods and labor services, which can be described as all-encompassing. The reason why C2C e-commerce levy objects are difficult to be divided into different categories is that both parties to the transaction are individuals, regardless of the transaction method, payment method, or transaction procedure, and its flexibility is far more than traditional business or B2B, B2C e-commerce, considering that the scope and scale of business between individual merchants are also very different, and the business content of C2C e-commerce is also different, so it is necessary to manage these C2C e-commerce businesses with different sizes and simple business scope. The main body sets the standard and then divides the object of collection into a huge project. The biggest characteristic of this business model is that the transaction information is transmitted through the network. Some individual merchants adopt electronic encryption technology in order to avoid the taxation department's collection and management. The flexibility of the merchants in C2C e-commerce makes the taxation department difficult to determine the object of collection [10].

3.1.3. Tax Location Is Difficult to Determine. Taxation locations are directly related to taxation issues. There are two

main theories about the taxation supervision of C2C e-commerce. One is the theory of regional jurisdiction. This theory is mainly based on the registration place of the online store in real life, the place where the business activity takes place, the place where the buyer consumes the action, or the location of the server. In attributable tax department, however, this theoretical division is not rigorous enough. The scope of the online store's business activities is very limited, and the place where the activity takes place can be changed frequently. To a large extent, there is a complete overlap with the tax jurisdiction of the industrial and commercial registration place. The other is the theory of unified jurisdiction. The central government uniformly uses electronic taxation to supervise all online store taxation in China's internet trading platform and has set up a new electronic taxation department to monitor and manage the online stores of e-commerce platforms. Although the concept of this theory is relatively good, it can greatly reduce the probability of tax evasion and tax evasion. However, in China's overlarge market, the establishment of a unified model consumes relatively large resources, and local taxes are incorporated into the central taxation. In line with China's national conditions, it must be modified on the basis of this theory before it can be applied to practical operations. C2C e-commerce belongs to the new field. It is still necessary to think deeply about applying the current tax system to e-commerce and taxing it [11].

3.2. Performance of Information Asymmetry in C2C e-Commerce Tax Collection and Management. There is also information asymmetry in the traditional tax collection and management, but due to the characteristics of CZC e-commerce and the limitation of technological development level, it is more obvious in the C2C e-commerce tax collection and management. The performance of information asymmetry in C2C e-commerce can be refined to all aspects, between the government and the tax authorities, between the tax authorities, between tax collectors and taxpayers, etc., but the most important and most difficult to solve is the acquisition. The information is asymmetrical between the two parties.

3.2.1. Asymmetry of Taxpayer Tax Information. In the substantive sense, tax collection and management are the management process for the collection, processing, processing, and application of tax information. The broad tax information refers to all kinds of information, data, news, and intelligence about tax economic behavior, status, and effects. In this paper, tax information refers to the narrowly defined tax information, that is, the information required by the tax authorities to implement tax collection and management, including taxpayer basic information and taxpayer business information. The basic information of the taxpayer refers to the taxpayer's identity (type, identification number, and name), taxation location, bank account number, and contact telephone number, and the taxpayer's business information refers to sales, fund collection and payment, procurement, inventory, expenses, and other information [12].

The asymmetry of the taxpayer's tax information means that the taxpayer does not disclose the private information he owns, which makes the tax authorities and taxpayers have a gap in understanding the taxation-related situation and form an information asymmetry. In fact, many tax information is generated in the specific transaction process of e-commerce, so it is possible to study the tax information in e-commerce tax collection and management by analyzing the e-commerce transaction process. The flow of e-commerce transactions and tax information transfer are shown in Figure 4.

Figure 4: e-Commerce transaction process and tax information transfer diagram. As can be seen from Figure 4, the main body of e-commerce transactions includes buyers, sellers, banks, and network intermediaries. The taxes involved in e-commerce are mainly turnover tax and income tax, and taxpayers are usually sellers. The taxpayer has all the tax information; the buyer in the online retail generally realizes the transaction through the network intermediary, so the buyer knows the sales information but does not necessarily know the basic information of the taxpayer and the network intermediary (the third party trading platform, the third party payment platform, and the credit intermediary). As an indispensable chain in online retail e-commerce, it has basic information and business information of taxpayers, such as business flow, information flow, capital flow, and logistics, but the network platform is not an authority, and it is impossible to force users to pay taxes [13]. The tax information owned by the taxpayer is composed of information reported by the taxpayer and audit information.

3.2.2. Information Asymmetry on Tax Regulations. On the one hand, some taxpayers lack the necessary understanding of relevant tax laws and regulations and procedures for fulfilling their tax obligations, and the tax authorities are in an information superior position. In recent years, the scale of online retail sales has expanded dramatically. The number of individual businesses in C2C e-commerce has greatly increased, and there are many taxpayers. Individuals often lack the habit of actively understanding tax laws. The taxpayer's unclear knowledge and inadequate taxation by the tax authorities are likely to cause the taxpayer's ignorant tax noncompliance, which leads to the state of taxation information asymmetry. On the other hand, some taxpayers who have studied the tax law in depth have used the tax law to blindly evade taxes, leaving the tax authorities at an information disadvantage [14]. e-Commerce is a new type of business operation. At present, the tax law legislature has not issued relevant tax policies, and the business level of tax law enforcement personnel is uneven. As a result, taxpayers who have thoroughly studied tax regulations can make full use of legal blind spots and hide true private information. Tax authorities are in a disadvantaged position when making tax returns or false declarations. The consequences of this situation are extremely serious. The taxpayer's sense of honesty and taxation declines, and the tax noncompliance is also raised. The tax collection and management of this part of the taxpayer cannot be effectively implemented, and the seriousness of the tax law cannot be reflected [15].

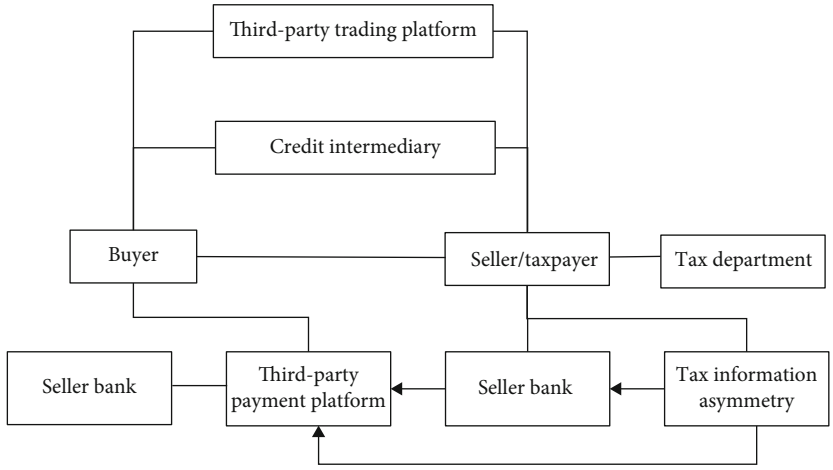


Figure 4: e-Commerce transaction process and tax information transfer diagram.

3.3. Game Analysis. From the above, we can see that C2C e-commerce is not tax-free in China, but it has caused difficulties in the collection and management of tax authorities for various reasons. Through the use of information asymmetry and game theory, the C2C e-commerce tax collection and management model is established. C2C e-commerce tax collection and management are a category of incomplete information dynamic games. Tax authorities do not know taxpayers, and taxpayers know themselves. Through natural selection, the taxpayer acts first, and the tax authorities act accordingly. The tax authorities can observe the taxpayer's actions and cannot observe the type, but the tax authorities can speculate on the types of taxpayers by observing their actions and then choose their own best actions. In addition, there are two inferences: there is information asymmetry between taxpayers and tax authorities, and tax authorities are not perfect for tax audits [16].

3.3.1. Model Premise Hypothesis

(1) Taxpayers and tax authorities involved in the game are rational people, that is, they are seeking to maximize their own interests

(2) The information asymmetry between the two parties is collected, and the effectiveness of the tax authorities' inspection is determined by the degree of information symmetry. Taxpayers have all their tax information, and the smaller the amount of information obtained by tax authorities, the higher the degree of information asymmetry, and the lower the probability of finding results. Record the degree of information asymmetry as J, J (0, 1), and check the validity of the record e

(3) The taxable credit income is K, which is the additional income that the taxpayer can bring in the long-term legal tax payment, such as the cost brought by the tax authorities and the easier access to the bank from the bank to reduce the financing cost

(4) The taxation authority selects the audit record as α and does not check the record $(1 - \alpha)$, the taxpayer evades the tax record as β, and pays tax according to law as $1 - \beta$; the taxpayer expects the payment level to be recorded as $\gamma 1$, the tax authority expects the payment level is $\gamma 2$, the taxpayer's actual taxable income is recorded as E, the declared income is recorded as $F(E \geq F)$, the tax rate is t, the penalty rate is set to m, and the taxpayer's taxation cost increased by the tax authorities is H_1, when the tax authorities inspect the time. The cost of collection and management is H_2 [17].

3.3.2. Game Process between Taxpayers and Tax Authorities. According to the characteristics of incomplete information dynamic game, we can summarize the game behavior of tax authorities and taxpayers as follows [18]:

First, the taxpayer knows all the information about his/her income, and the taxpayer declares taxable income. The tax authorities are unaware of the taxpayer's real income. When the tax authorities speculate that the taxpayer's declared income is equal to the actual income, they will not choose to audit, because it can not only increase the income but also make the behavior pay more costs. When the tax authorities speculate that the income reported by the taxpayer may be lower than the actual income, it will be checked at a certain rate. In the end, the tax authorities can choose to reject or accept the amount declared by the taxpayer based on the results of the inspection [19].

In this game process, the tax authorities have two decisions: the taxpayer's report information that the tax authorities can observe and the tax authorities' acceptance of the taxpayer's declaration. In fact, the tax authorities are unable to monitor the actual taxable income of taxpayers, that is, each type of information is from different taxpayers. Under the condition of information asymmetry, the tax authorities are at an obvious information disadvantage [20].

3.3.3. Model Solution. Taxpayer expected payment level:

$$\gamma 1 = (1-\beta)(Et-K) + \beta[\mathrm{Ft} + \alpha\theta t(1+m)(E-F) + \alpha H1]. \tag{1}$$

The first item on the right side of the middle type indicates the tax payable by the taxpayer in the legal tax selection, the second item.

The third item indicates that the taxpayer is required to pay the tax if the taxpayer evades the tax.

The tax and fines, the fourth item, indicate the taxation costs incurred by the taxpayer during the inspection process by the tax authorities.

Tax authorities expect the level of payment [21–23]:

$$\gamma 2 = \alpha H2 - \alpha[\beta\theta t(1+m)(E-F)] - \mathrm{Ft}. \tag{2}$$

The first item on the right side of the formula (2) indicates the cost incurred by the tax authorities for tax audits. The second item is the tax and fines that the taxpayer chooses to evade and is found to be paid. The third item is the tax paid by the taxpayer.

By deriving partial biases for (1) and (2), respectively, the minimum level of expected tax payment for taxpayers and tax authorities can be obtained [24]:

$$\begin{aligned} \frac{\partial \gamma 1}{\partial \beta} &= \mathrm{K}\text{-}Et + \mathrm{Ft} + \alpha\theta t(1+m)(E-F) + \alpha H1, \\ \frac{\partial \gamma 2}{\partial \alpha} &= H2 - \beta\theta t(1+m)(E-F), \end{aligned} \tag{3}$$

which make $\partial\gamma 1/\partial\beta = 0$, $\partial\gamma 2/\partial\alpha = 0$.

Solutions are as follows:

$$\begin{aligned} \alpha* &= \frac{Et - \mathrm{Ft} - \mathrm{K}}{\theta t(1+m)(E-F) + H1}, \\ \beta* &= \frac{H2}{\theta t(1+m)(E-F)}. \end{aligned} \tag{4}$$

Through the above analysis, three kinds of equilibrium can be obtained.

Equalization One. In order to ensure that the conditions $\alpha * \in (0,1)$, $\beta * \in (0,1)$ need to be met

$$\beta\theta t(1+m)(E-F) + H2 \geq Et - \mathrm{Ft} - \mathrm{K}\theta t(1+m)(E-F) \geq H2. \tag{5}$$

Equilibrium Two. If it is available and if it can be satisfied at this time

$$\theta t(1+m)(E-F) + H2 \leq Et - \mathrm{Ft} - \mathrm{K}, \alpha* = 1; \theta t(1+m)(E-F) \geq H2, \beta* = 1. \tag{6}$$

Balance Three. If $\theta t(1+m)(E-F) \geq H2$, $\theta t(1+m)(E-F) + H2 \leq Et - \mathrm{Ft} - \mathrm{K}$, and if you can meet it at this time, you can get $\alpha * = 0$, $\beta * = 1$.

3.3.4. Equilibrium Analysis of the Model. In the equilibrium one, the taxpayer evades taxes with a certain probability, and the tax authorities conduct audits from a certain probability. When the tax authorities' inspection efficiency, fine rate, taxpayer and tax auditing costs, and taxable integrity gains are relatively high, and the tax authorities' inspection costs are low, a balance can be achieved. At this time, the degree of information asymmetry is not high, the inspection efficiency of the tax authorities is relatively high at this time, the audit cost is relatively low, and the taxpayer's cost of noncompliance and tax credit integrity are relatively high. This shows that the tax laws and regulations are relatively perfect in this case, the information of the two parties is relatively complete, the means of collection and management are more adequate, and the phenomenon of noncompliance with taxation is less. The traditional business model and the B2B collection and management situation are more consistent with this equilibrium. In these models, the degree of information asymmetry is relatively small. The perfect tax legal system and supervision system can effectively help the tax authorities to fully obtain taxpayer tax information and improve the tax authorities. Tax inspection and collection efficiency reduce taxpayers' tax evasion and tax evasion. The simulation results of various models are shown in Table 2.

In the second phase of equilibrium, taxpayers must choose to evade taxes, and the tax authorities must choose to conduct audits. In this equilibrium, the taxation department's fine rate, audit efficiency, taxpayer's audited cost, and honest tax return are relatively low, and the tax auditing cost is also low enough. This shows that the information asymmetry is very high, the taxpayer's noncompliance and tax credits are very low, the tax audit efficiency and cost are relatively low, taxpayers will see the risk of tax evasion is small, and the benefits are high. Taxpayers' tax rates and tax authorities' inspection rates are high. The situation of equilibrium two generally does not exist in real life, because the auditing costs of the tax authorities are not so low that they can benefit from auditing all taxpayers.

In the third phase of equilibrium, taxpayers choose to evade taxes, and tax authorities choose not to conduct audits. In this equilibrium, the degree of information asymmetry is extremely high, and the cost of tax auditing is high. At the same time, the taxpayer's fine rate, the cost of inspection by the tax authorities, and the integrity of tax payment are very low. At this point, the tax authorities are unable to obtain any taxpayer's tax information, the inspection cost is very high, and the taxpayer's tax evasion success rate is extremely high. This situation is in line with the C2C e-commerce model and is also the case of our research. Because the C2C e-commerce tax collection is seriously asymmetrical between the two parties, the tax authorities have almost no knowledge of the tax information of C2C individual businesses, and the auditing costs are relatively high. Therefore, the tax authorities often do not conduct tax collection, and the taxpayers know the tax. The authorities are unable to grasp the tax information, and the cost of tax evasion is very low. Therefore, taxpayers often choose to use online business to evade taxes, resulting in large tax noncompliance [25].

Table 2: Analysis of model results.

	Taxpayer	Taxation authority	Inspection efficiency	Fine rate	Taxpayer and tax inspection cost	Tax integrity income tax	Authority audit cost	Result
Equilibrium one	Tax evasion with a certain probability	Auditing from a certain probability	High	High	High	High	Low cost	Low degree of information asymmetry
Balance 2	Must choose tax evasion	Must choose audit	Low	Low	Low	Low	Low enough	Information asymmetry is very high
Equilibrium three	Select a certain tax evasion	Choose not to audit	Low	Low	Low	Low	High	Information asymmetry is extremely high

In summary, information asymmetry is the main reason for the difficulty in tax collection and management of e-commerce models. The higher the degree of information asymmetry, the higher the taxpayer's tax noncompliance, the greater the difficulty of tax collection and management, resulting in increased tax losses. Especially in China's current C2C e-commerce, this phenomenon is very prominent. The characteristics of the C2C e-commerce model are doomed to a higher degree of information asymmetry than other models. C2C individual merchants have very low tax compliance, and tax authorities often choose to abandon the collection and management.

3.3.5. Analysis and Application of cox Model. Cox Model Structural Style. This model proposed by England statistician D.R.Cox in 1972 that is mainly used in studying each kind of factor (to be called covariate) regarding the lifetime length relations carries on the multifactor analysis. Instead of directly considering the relationship between the survival function and the covariance, the risk function is used as the dependent variable, and its structure form is

$$h(t) = h_0(t)\ \exp\ (\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_m X_m). \tag{7}$$

in which $h_0(t)$ is the base risk rate (nonparametric part) and $B = -1.329$ is the regression coefficient (estimated according to the observation value).

The regression coefficient $\beta_j \geq 0$, covariable value, the higher the risk function $e^{-1.329} = 0.265$ value indicate the individual risk of death.

When regression coefficient $\beta_j = 0$ is expressed, the covariance has no effect on the risk function $h(t)$.

When regression coefficient $\beta \leq 0$, the covariate value is bigger, the risk function $h(t)$ value is smaller, and the expression individual death risk is smaller.

The Basic Characteristics of Cox Model. In the case of unknown, the β in the covariance can still be estimated, so the parameter estimation method of the Cox regression model is called the partial likelihood value method.

In the Cox model, $h_0(t)$ is not modeled, so the Cox model is also called a semiparametric model.

The Cox model of the basic assumes that regardless of the baseline risk, at any point in time, the presence of an exposed individual is not present in the exposed individual events in the risk which is constant. Also, namely two groups of crowds light in any time have the event risk proportion which is constant (or explanation expose for someone in all time to have event function all are same); therefore, the Cox model is called the proportion risk model, and the corresponding covariate parameter estimation must satisfy the proportion risk supposition. The Cox model does not have the intercept, because the intercept is a part, and is cancelled in the partial likelihood value law.

Cox Model Assumptions. The ratio of the risk function of any two individuals, that is, the risk ratio (HR) to maintain a constant proportion, with independent time t is as follows:

$$HR = \frac{h_i(t)}{h_j(t)} = \exp\left[\beta_1\left(X_{i1} - X_{j1}\right) + \beta_2\left(X_{i2} - X_{j2}\right) + \cdots + \beta_p\left(X_{im} - X_{jm}\right)\right]. \tag{8}$$

The efficiency of the covariates in the model does not change with time and change and inspects some covariates whether they satisfy the PH hypothesis; the simple method is observed according to this variable grouping survival curve overlapping, and prompts do not satisfy the PH hypothesis.

Cox Model Parameter Estimation. The Cox model parameter estimation needs to construct the partial likelihood function, calculated by the maximum likelihood method, and estimates the parameters of b_j, and general steps of maximum likelihood estimation method are as follows:

Write likelihood function L

$$L(\theta_1, \theta_2, \cdots, \theta_n) = \begin{cases} \prod_{i=1}^{n} p(x_i; \theta_1, \theta_2, \cdots, \theta_n) \\ \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \cdots, \theta_n). \end{cases} \tag{9}$$

Take the logarithm and the derivation number, obtain the stationary point, the maximum point, and make the conclusion.

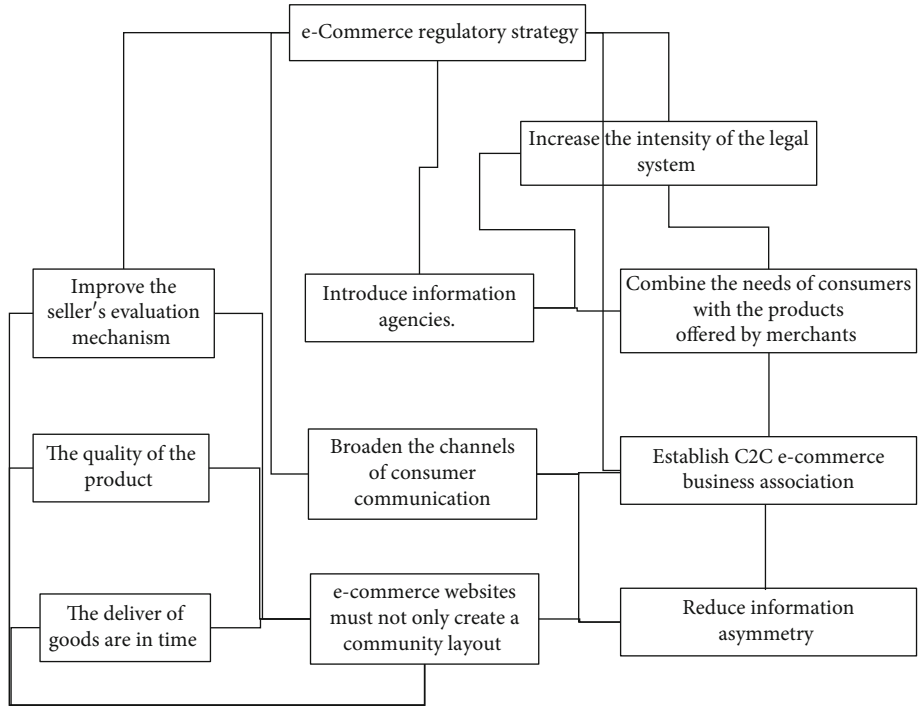


Figure 5: The strategy of improving the e-commerce information system.

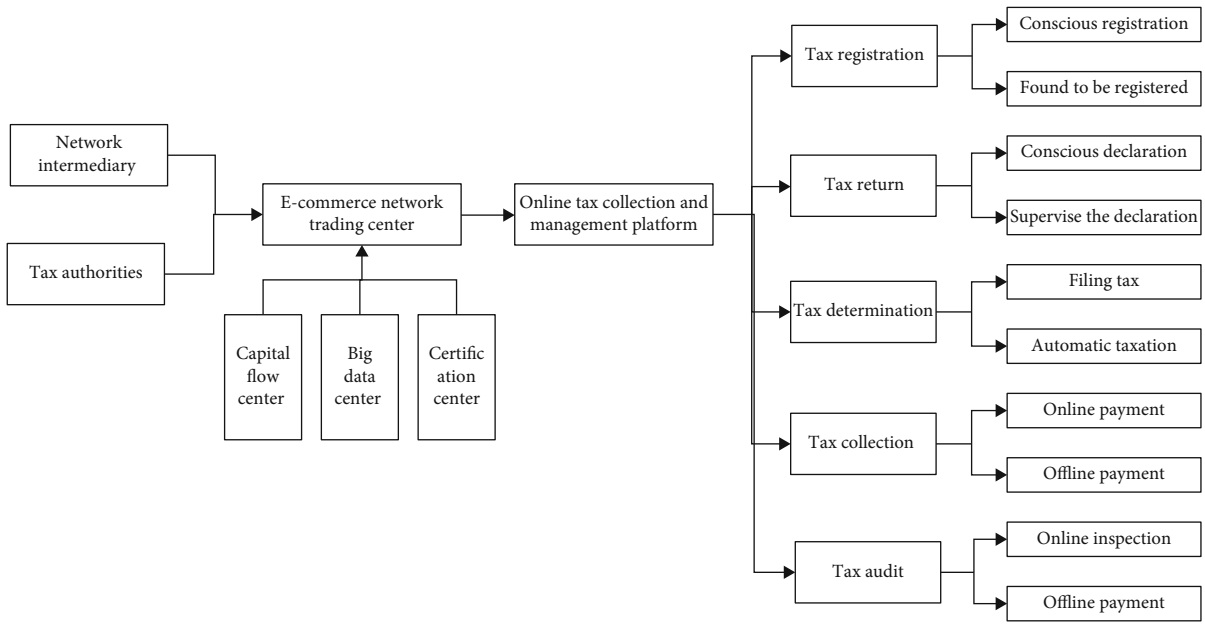


Figure 6: C2C e-commerce tax collection and management system based on network intermediary information.

Table 3: Special data test based on large data statistical model.

Import and export trade	2000	2002	2004	2006	2008	2010	2015
	3.2	+0.6	-0.3	+0.4	+0.7	-0.1	-0.8
Large data statistical model	5.0	6.4	3.5	2.9	4.5	2.5	4.1
Polynomial kernel model	0.5	0.7	0.1	0.4	0.5	0.3	0.4

4. Eliminate the Strategy of Information Asymmetry in C2C e-Commerce Transactions

4.1. Improve the Seller's Credit Evaluation Mechanism. At present, the C2C e-commerce website basically provides a seller credit evaluation mechanism, that is, after the consumer receives the purchased product and uses it, the seller evaluates the merchant according to the actual situation of the product and the service status obtained. Generally, the evaluation content includes the quality of the product, whether the performance of the product is consistent with the description of the merchant on the webpage, whether the service attitude of the merchant is good, and whether the delivery of the commodity is timely. Therefore, the consumer's evaluation forms the credit of the merchant. If the consumer is satisfied with the service and product quality of the merchant, there are many good evaluations, and the credit degree of the merchant is high. On the contrary, if the consumer is not satisfied with the service and product quality of the merchant, the poor evaluation is more, and the credit degree of the merchant is low. In this way, when consumers make purchases, they can understand the past consumers' evaluation of the services and quality of the merchants according to the creditworthiness of the merchants, thereby changing the information disadvantage status of the consumers and minimizing the lack of integrity of the merchants. The framework of e-commerce information system is shown in Figure 5.

4.2. Broaden Consumer Communication Channels. For the quality of goods, consumers have the most say. Therefore, it is necessary to broaden the communication channels of consumers so that consumers can more fully understand the relevant information of products before purchasing goods and maximize the change between consumers and businesses. In information asymmetry, to this end, the C2C e-commerce website not only needs to establish a "community" layout but also provides the necessary space for consumers to discuss the quality of a certain merchant service or a certain product more actively and to effectively classify the content of the community. To facilitate consumer information exchange. At the same time, a search engine should be established, and consumers can quickly find out all the evaluations of a certain product sold by a certain merchant through a search engine.

4.3. Introduce Information Intermediaries. In order to reduce the cost of collecting information by consumers themselves and avoid "reverse selection" resulting in "inferior goods expelling good goods," information brokers can be introduced in C2C e-commerce websites to improve the transaction efficiency of C2C e-commerce market. Intermediaries are required to conduct long-term tracking of the quality of goods operated by merchants on the C2C e-commerce website and conduct fair evaluation and screening. The information intermediary can combine the needs of the consumer with the goods provided by the merchant according to the specific needs of the consumer and facilitate the transaction between the buyer and the seller or provide a good information consultation service for the consumer. The information

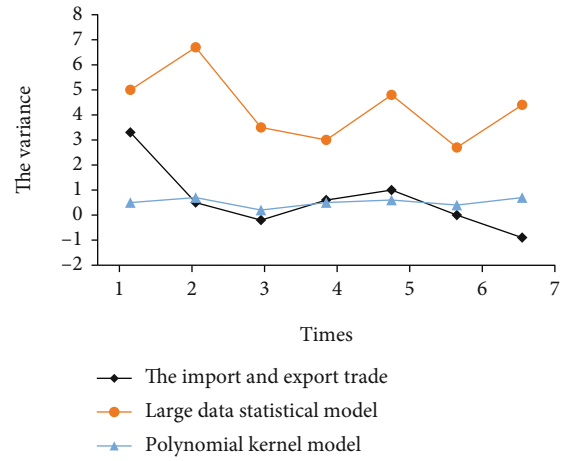


Figure 7: Special data validation of statistical models.

intermediary obtains certain services through the standardized service. In commission, this will not only reduce the blindness of consumers when shopping and reduce the risk of trading but also encourage businesses that manage high-quality goods to continue to provide quality goods, so that the survival opportunities of inferior goods are getting smaller and smaller.

4.4. Strengthen the Construction of the Legal System. The government must not only strictly enforce the legislation but also strictly enforce the law. The complaints against consumers must be dealt with in a timely manner, and the businesses that provide inferior goods should be punished, and the opportunity cost of the reverse selection behavior of the merchants who operate the inferior commodities should be increased to curb the behavior of the species. In occurrence, for those who manage high-quality goods, according to the level of their credibility evaluation, they can subsidize taxation and establish an incentive mechanism for integrity. The C2C e-commerce tax collection and management system based on network intermediary information as shown in Figure 6.

4.5. Establish a C2C e-Commerce Merchant Industry Association. The binding force of the merchants increases the degree of consumer trust in the merchants and the products or services they operate, reduces information asymmetry, and reduces the efficiency losses caused by market failures. It is necessary to give full play to the industry supervision function of C2C e-commerce merchants, strengthen the communication and communication between C2C e-commerce merchants, and use the online instant chat mode and the information publishing function of the network platform to supervise and manage similar businesses. The industry credit evaluation content can be increased in the existing credit evaluation mechanism. The industry credit evaluation is mainly evaluated by the C2C e-commerce merchant industry association based on the credit evaluation of the merchants and the complaints of the consumers over a period of time, so as to convey the correct information of the high-

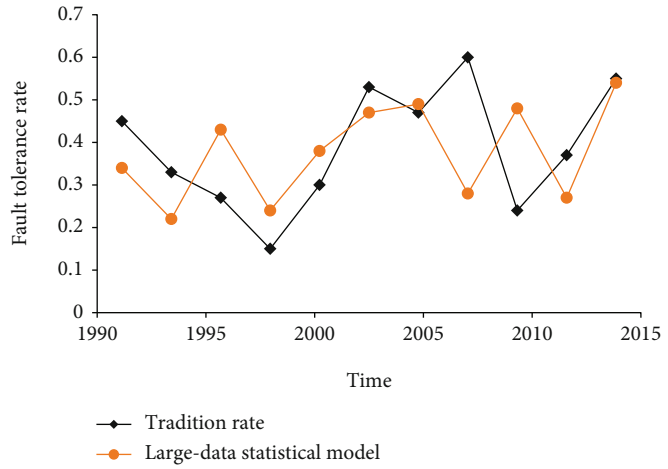


Figure 8: Special data test icons based on large-data statistical models.

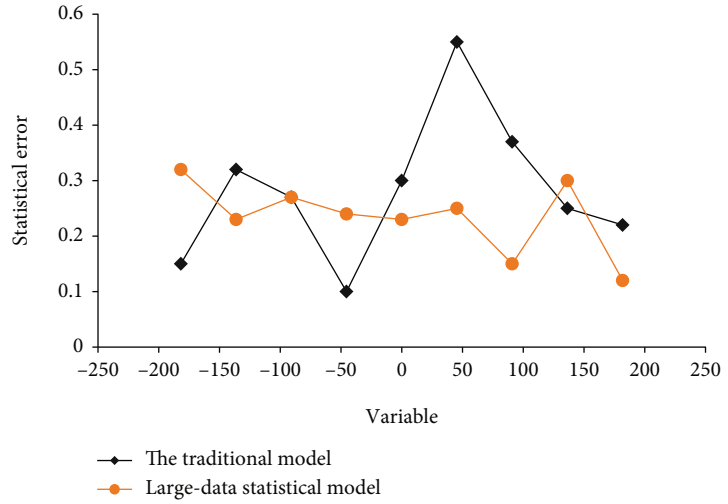


Figure 9: Comparison of special data test and traditional model based on large-data statistical model.

quality goods to the consumers and reduce the price distortion.

4.6. Case Analysis. The test data of total trade amount of e-commerce enterprises in a city are as follows. Three different models analyze and compare the data, the results are shown in Table 3, and the graph is shown in Figure 7.

It is clearly evident from the Figure 7 that our proposed big data statistical model in the application process is of the superiority. In the course of this test, we found that the results of e-commerce business statistics were inaccurate. Data is similar to the traditional model existed in the system's computing process. Therefore, this paper introduces the middleware, first of all to identify the composition of the traditional model code, and then it is classified according to the code, one by one test. Through the test results, it can be found that the existing code features are obvious. Therefore, only the middle plug-in needs to be added to screen the elements of the enterprise e-commerce statistical errors so as to obtain the accurate detection data of the traditional model. The middle plug-in can be used to optimize the system during use and constantly adjust according to the amount of electricity supplier with the exchange. Therefore, during the control process of the program, through constant testing, random e-commerce business statistics are used to carry out practical research, constantly filling it and using the self-improvement and repair functions of big data statistical model to continuously improve accuracy of the data statistics model. The use of this big data statistical model can accurately improve this study and increase the number of tests before it is put into use, so as to increase the guarantee on the original basis. A special data test diagram based on a large data statistics model is shown in Figure 8.

From the HDFS perspective, Hadoop nodes in a big data statistical model are stored by the system and provide a data node serving the positioning block and a distributed file system. The system's name space consists of named nodes. The architecture of the HDFS master-slave architecture greatly simplifies the distributed system architecture. Compared with the traditional model, the design features the big data

statistical model system: centralized management of meta data, decentralized storage of data blocks (64 MB) to ensure data security, data replication (at least 3 backups per data), and high fault tolerance. Traditional database systems should be capable of real-time, but for big data sets, at the terabyte or petabyte level. However, the retrieval speed of data mining decreases sharply. However, the comparison of various models can provide more favorable data analysis results for the C2C e-commerce model. The comparison between special data test and traditional model based on big data statistical model is shown in Figure 9.

Through the above experimental results, it can be observed that the parallel computing of big data sets on a cluster is a typical example of nonrelational data management and analysis techniques. Map reduce abstracts a series of map operations and reduces operations from the data processing tasks to complete data filtering and aggregation, respectively, and manage them through a simple interface. In simple terms, the calculation process and the basic principle are to decompose a big data set into hundreds of small data sets. Each data set is processed by a node in the cluster and generates intermediate results. These intermediate results in turn by a large number of nodes merge to form the final result. On the premise of ensuring the accuracy of the big data statistical model, the computer system can also process the bivariate model according to the analysis. Making corresponding judgments on the statistical errors and making secondary statistics through the automatic rectification of the system ensure that the output of the final output can ensure its objectivity. Thus, the e-commerce statistical theory analysis model can reach the objective statistical results regardless of the single variable or multivariate cases and provide accurate guidance for the development of the enterprise. On the basis of ensuring the national economic stability, the majority of participants seek to maximize the benefits.

5. Conclusion

On the basis of summarizing the hotspots of C2C e-commerce tax research, this paper combines relevant basic concepts and theoretical analysis and inserts case analysis to discuss the current situation of tax collection and management in China. Finally, based on network intermediary information, C2C e-commerce tax collection and management means, related improvement measures, and a brief analysis are established. Although, at present, the tax policy of e-commerce in China is not yet mature, e-commerce tax payment will become an inevitable trend in the future development. In order to ensure the effective implementation of taxation policies, we must not only reduce taxation costs but also vigorously popularize e-commerce taxation policies, cultivate taxpayers' awareness of taxation, improve the professional ethics and quality of taxpayers, implement real-name taxation policies, and raise taxpayers. The professional ethics and quality meet the era requirements of e-commerce taxation under the C2C model. Through the establishment of the game theory model, the analysis of information asymmetry is the main reason for the difficulty in tax collection and management of C2C e-commerce in China. From the information asymmetry in China's C2C e-commerce performance, the reasons for the formation of information asymmetry and its impact and effects are analyzed. In the era of modernization, our government and tax authorities should fulfill the power of taxation according to law to ensure the smooth progress of modernization. In addition, because the level of e-commerce supervision and construction is relatively abstract and complex, involving laws, systems, and technologies, this paper analyzes the issue of e-commerce tax collection and management from the perspective of information asymmetry of tax collection and has certain limitations, in terms of countermeasures, mainly from the improvement of the legal system, tax collection, and management. The improvement of information technology and the updating of the management methods of tax authorities propose the proposal of the construction of e-commerce tax collection and management system. It is worthwhile to further study in order to propose a comprehensive and multiangle construction plan.

Data Availability

The authors approve that data used to support the finding of this study are included in the article.

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 is supported by the Humanities and Social Science Research Project of Anhui University in 2020: "Evaluation and promotion path of rural industry integration development level" research progress and outcomes, Humanities and Social Science Research Project of Anhui University in 2016 (SK2016A0636).

References

[1] M. Chahal, S. Harit, K. K. Mishra, A. Sangaiah, and Z. Zheng, "A survey on software-defined networking in vehicular ad hoc networks: challenges, applications and use cases," *Sustainable Cities and Society*, vol. 35, no. 11, pp. 830–840, 2017.

[2] T. Yang, X. Long, A. K. Sangaiah, Z. Zheng, and C. Tong, "Deep detection network for real-life traffic sign in vehicular networks," *Computer Networks*, vol. 136, no. 8, pp. 95–104, 2018.

[3] Z. Zheng, P. Wang, J. Liu, and S. Sun, "Real-time big data processing framework: challenges and solutions," *Applied Mathematics & Information Sciences*, vol. 9, no. 6, pp. 3169–3190, 2015.

[4] Oxford University Press, "Recommendations for exercise training in chronic heart failure patients," *European Heart Journal*, vol. 22, no. 2, pp. 125–135, 2001.

[5] B. Chen, X. Li, X. Xie, Z. Zhong, and P. Lu, "Fatigue performance assessment of composite arch bridge suspenders based

on actual vehicle loads," *Shock and Vibration*, vol. 2015, pp. 1–3, 2015.

[6] B. D. Anderson, T. Brinsmead, D. Liberzon, and A. Stephen Morse, "Multiple model adaptive control with safe switching," *International Journal of Adaptive Control and Signal Processing*, vol. 15, no. 5, pp. 445–470, 2001.

[7] D. J. Condon, B. Schoene, N. M. McLean, S. A. Bowring, and R. R. Parrish, "Metrology and traceability of U-Pb isotope dilution geochronology (EARTHTIME Tracer Calibration Part I)," *Geochimica et Cosmochimica Acta*, vol. 164, pp. 464–480, 2015.

[8] J. Bo Andersen, B. B. Roldgaard, B. B. Christensen, and T. Licht, "Oxygen restriction increases the infective potential of listeria monocytogenes in vitro in Caco-2 cells and in vivo in Guinea pigs," *BMC Microbiology*, vol. 7, no. 1, p. 55, 2007.

[9] K. Vafaï, "Abuse of authority and collusion in organizations," *European Journal of Political Economy*, vol. 21, no. 2, pp. 385–405, 2005.

[10] L. Lei, X. Shen, M. Dohler, Z. Zhong, and C. Lin, "Queuing models with applications to mode selection in device-to-device communications underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 6697–6715, 2014.

[11] J. Gräsvik, S. Winestrand, M. Normark, L. J. Jönsson, and J. P. Mikkola, "Evaluation of four ionic liquids for pretreatment of lignocellulosic biomass," *BMC Biotechnology*, vol. 14, no. 1, 2014.

[12] N. Dobbs, "Measures with positive Lyapunov exponent and conformal measures in rational dynamics," *Transactions of the American Mathematical Society*, vol. 364, no. 6, pp. 2803–2824, 2012.

[13] T. Sieber, B. E. Weißenberger, T. Oberdörster, and J. Baetge, "Let's talk strategy: the impact of voluntary strategy disclosure on the cost of equity capital," *Business Research*, vol. 7, no. 2, pp. 263–312, 2014.

[14] B. S. Uzochukwu, O. E. Onwujekwe, N. P. Uguru, M. D. Ughasoro, and O. P. Ezeoke, "Willingness to pay for rapid diagnostic tests for the diagnosis and treatment of malaria in Southeast Nigeria: ex post and ex ante," *International Journal for Equity in Health*, vol. 9, no. 1, 2010.

[15] T. L. Saitone and R. J. Sexton, "Product differentiation and quality in food markets: industrial organization implications," *Annual Review of Resource Economics*, vol. 2, no. 1, pp. 341–368, 2010.

[16] L. Donglin and S. O. Economic, "Game analysis of credit problem and countermeasures of c2c e-commerce," *Journal of Zhoukou Normal University*, vol. 36, no. 4, pp. 758–764, 2016.

[17] F. Yang, A. Wang, J. Wu, and L. Tang, "Designing credit supervision mechanism in C2B2C e-commerce based on game theory," *Systems Engineering-Theory & Practice*, vol. 37, no. 8, pp. 2102–2110, 2017.

[18] D. Lague, N. Brodu, and J. Leroux, "Accurate 3D comparison of complex topography with terrestrial laser scanner: application to the Rangitikei canyon (N-Z)," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 82, no. 4, pp. 10–26, 2013.

[19] A. Lohse, "Stability of heteroclinic cycles in transverse bifurcations," *Physica D: Nonlinear Phenomena*, vol. 310, pp. 95–103, 2015.

[20] I. Etxebarria, J. Ajuria, and R. Pacios, "Polymer:fullerene solar cells: materials, processing issues, and cell layouts to reach power conversion efficiency over 10%, a review," *Journal of Photonics for Energy*, vol. 5, no. 1, article 057214, 2015.

[21] J. R. Brown, R. F. Lusch, and C. Y. Nicholson, "Power and relationship commitment: their impact on marketing channel member performance," *Journal of Retailing*, vol. 71, no. 4, pp. 363–392, 1995.

[22] C. Carlsson, J. Carlsson, K. Hyvonen, J. Puhakainen, and P. Walden, "Adoption of mobile devices/services-searching for answers with the UTAUT," in *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, Kauai, HI, USA, 2006.

[23] S. H. Chang, K. Y. Wang, W. H. Chih, and W. H. Tsai, "Building customer commitment in business-to-business markets," *Industrial Marketing Management*, vol. 41, no. 6, pp. 940–950, 2012.

[24] W. Chaouali, I. Ben Yahia, and N. Souiden, "The interplay of counter-conformity motivation, social influence, and trust in customers' intention to adopt internet banking services: the case of an emerging country," *Journal of Retailing and Consumer Services*, vol. 28, no. 2, pp. 209–218, 2016.

[25] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Transactions on Information and Systems*, vol. E102.D, no. 7, pp. 1374–1383, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5582132, 10 pages
https://doi.org/10.1155/2021/5582132

Research Article

A Multilevel Single Stage Network for Face Detection

Kanghua Hui,[1] **Jin Wang**,[2] **Huaiqing He**,[1] **and W. H. Ip**[3]

[1]*Department of Computer Science and Technology, Civil Aviation University of China, 300300, China*
[2]*Department of Air Traffic Management, Civil Aviation University of China, 300300, China*
[3]*Department of Mechanical Engineering, University of Saskatchewan, S7K6P1, Canada*

Correspondence should be addressed to Jin Wang; jinwang_2019@aliyun.com

Received 13 January 2021; Revised 1 February 2021; Accepted 15 February 2021; Published 27 February 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Kanghua Hui 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.

Recently, tremendous strides have been made in generic object detection when used to detect faces, and there are still some remaining challenges. In this paper, a novel method is proposed named multilevel single stage network for face detection (MSNFD). Three breakthroughs are made in this research. Firstly, multilevel network is introduced into face detection to improve the efficiency of anchoring faces. Secondly, enhanced feature module is adopted to allow more feature information to be collected. Finally, two-stage weight loss function is employed to balance network of different levels. Experimental results on the WIDER FACE and FDDB datasets confirm that MSNFD has competitive accuracy to the mainstream methods, while keeping real-time performance.

1. Introduction

Face detection, the basis of face alignment [1, 2], face recognition [3, 4], facial expression analysis [5, 6], and other related facial problems, has always been a hot issue and widely applied in terms of computer vision. As the major breakthrough for face detection, Viola-Jones [7] used the handcrafted features for cascading detection. There have already been lots of researches devoted to exploring the methods for extracting features effectively [8] and designing more complex handcrafted features or higher-efficiency cascade structure based on Viola-Jones detector.

Recently, with the development of deep learning, convolutional neural networks (CNN), great advance has been achieved and practically applied in image classification [9] and semantic segmentation [10], which also inspired the research of face detectors that can be classified into two modes. One integrates CNN into traditional detection structure. For example, Cascade CNN [11] successfully applied CNN to traditional detection methods through a simple to complex cascade classification network. However, most of these methods use the constructing of Gaussian pyramid image and the sliding window to detect thus resulting in unbalanced samples with negative samples increased and causing RAM to be occupied severely and unable to process images with high resolution. The other focuses on the improvement of generic object detection based on the belief that face is merely another specific object to be detected. For example, Face R-CNN [12] exploited several new techniques including new multitask loss function design, online hard example mining, and multiscale training strategy to improve Faster R-CNN [13] in multiple aspects. S3FD [14] proposed a scale-equitable face detection framework to handle different scales of faces well based on SSD [15]. ISRN [16] presented an improved SRN face detector by combining some useful techniques together from generic object detection and so on.

Although all of these methods improved detection accuracy for ordinary faces, they are difficult to handle small faces fast and effectively. There are also some studies on small faces. HR [17] provided an in-depth analysis of image resolution, object scale, and spatial context for the purposes of finding small faces. Bai [18] proposed an algorithm to directly generate a clear high-resolution face from a blurry small one by adopting a generative adversarial network. PyramidBox++ [19] improved each part to further boost the performance, including Balanced-data-anchor-sampling, Dual-PyramidAnchors, and Dense Context Module. These

three methods are recognized as achieving state-of-the-art performance on public dataset but suffer from time-consuming inference. The multilevel single stage network for face detection is inspired by the second mode.

The modern object detectors can be roughly divided into two groups: two-stage face detectors and one-stage detectors. Two-stage detector achieves the better performance but has low time efficiency, for example, SSFD$^+$ [20] focus on achieving comparable performance and simplifying the network architecture for detecting multiscale faces while it spends 582 ms on detecting a picture. Conversely, one-stage detector has faster speed. Face detection has the high demand of speed in real applications, and therefore, one-stage methods are chosen for the backbone. YOLO v3 [21], the version of the YOLO, has fast detection speed, simple structure, low false positive, and strong versatility. The idea of YOLO v3 is to divide an image into an $S \times S$ grid, and each grid cell is parameterized relative to some reference boxes through anchors regressing the bounding box coordinates, an objectness score, and a class probability according to anchors. Although SSD is also an excellent one-stage method and uses multiscale feature map, it is worse for small objects because semantic value for bottom layer is not high at least not higher than YOLO v3. Recently, there are lots of researchers focusing on detecting faces based on YOLO v3. For example, Gurkan [22] achieved 6% improvement on the detection rate of small sized faces, however, with the expense of a longer inference time. Thus, a multilevel single stage network for face detection based on YOLO v3 is designed and gets excellent results especially for small faces on the public datasets.

Although YOLO v3 performs well, it still causes some problems when directly used for face detecting without any adaptation especially for small faces. Lower features are used to detect smaller faces for YOLO v3. However, the lower feature with less semantic information will result in difficult detection for small faces. In addition, many factors, such as sample-unbalanced, severely challenge face detection. The initiative of the research is summarized as follows:

(1) Using multilevel network structure with more anchor scales to detect smaller faces

(2) Adopting enhanced feature module to add contextual and multiscale information to improve the ability of detecting small faces

(3) Balancing the outputs of networks of each level with two-stage weight loss function to optimize the network training

(4) Achieving excellent results on FDDB and WIDER FACE datasets with real-time detection speed

2. Multilevel Single Stage Network for Face Detection

This section will explain MSNFD through following aspects: the basic framework, enhanced feature module, and two-stage weight loss function.

2.1. Basic Framework. YOLO v3 is the preferred choice compared to other network architectures because the backbone of YOLO v3 is simple, and it is the faster one of the methods which have excellent performance for small faces. The basic framework of the method is shown in Figure 1. Some layers which are not critical are not drawn to maintain its visual clarity. First of all, the network is single stage, and the backbone network, Darknet-53, is the same as YOLO v3. Four residual networks used are named ResNet_2, ResNet_8, ResNet_8, and ResNet_4. The residual network is shown at the bottom left of Figure 1. The feature maps from ResNet_2, ResNet_8, ResNet_8, and ResNet_4 of the backbone are separately named Ori_0, Ori_1, Ori_2, and Ori_3 after a series of convolution and concatenating. The module "concat" in Figure 1 indicates multiscale fusion which is the same as YOLO v3. Up-sample small scales to the large scale, and then let them merge into a longer tensor.

Then, the enhanced feature module is added into the basic framework. The obtained feature maps are named Enh_1, Enh_2, and Enh_3, as shown in Figure 1. Ori_1, Ori_2, Ori_3, Enh_1, Enh_2, and Enh_3 constitute the first-level detection network. The first-level network has 6 outputs. The 6 outputs are named 13×13, 26×26, 52×52, 13×13 enhance, 26×26 enhance, and 52×52 enhance, respectively. Each output is encoded as an $S \times S \times (B \times 5 + C)$ tensor. The tensor contains $B = 3$ bounding boxes and $C = 1$ class probability for each one of the $S \times S$ grid. As shown in Figure 1, the 6 outputs have three different grids: 13×13, 26×26, and 52×52, respectively. For each one of the B bounding boxes, it generates coordinates and an objectness score. Each bounding box has one preset anchor. Different grids have bounding boxes with different sizes (width and length) of preset anchors. Each size of anchors is called as anchor scale. Increasing the number of preset anchor scales improves the matching ability between faces and anchors. That is to say, each bounding box is easier to be regressed to generate results.

The second-level network is formed by first-level network merging with low layer feature map, Ori_0. The input image is divided into a smaller and denser 104×104 grid for second-level network. The output of second-level network is named 104×104. The second-level network improves the detection capability of small faces particularly. Firstly, the smaller each grid is, the smaller the anchor scale is. The second-level network is attached with three smaller anchor scales to ensure matching small faces. Furthermore, to address the problem that the lower feature with less semantic information will result in difficult detection for small faces, the first-level network is integrated into the second-level network so that second-level network enriches semantic information greatly. In summary, it is a theoretical analysis that the proposed structure has detection capability particularly for small faces.

2.2. Enhanced Feature Module. The enhanced feature module borrows the idea of dilated convolution [23]. Small faces have little pixels at lower layer of the network. As the convolution goes on, the features left at high layers will be too few to get sufficient semantic information which will make it difficult

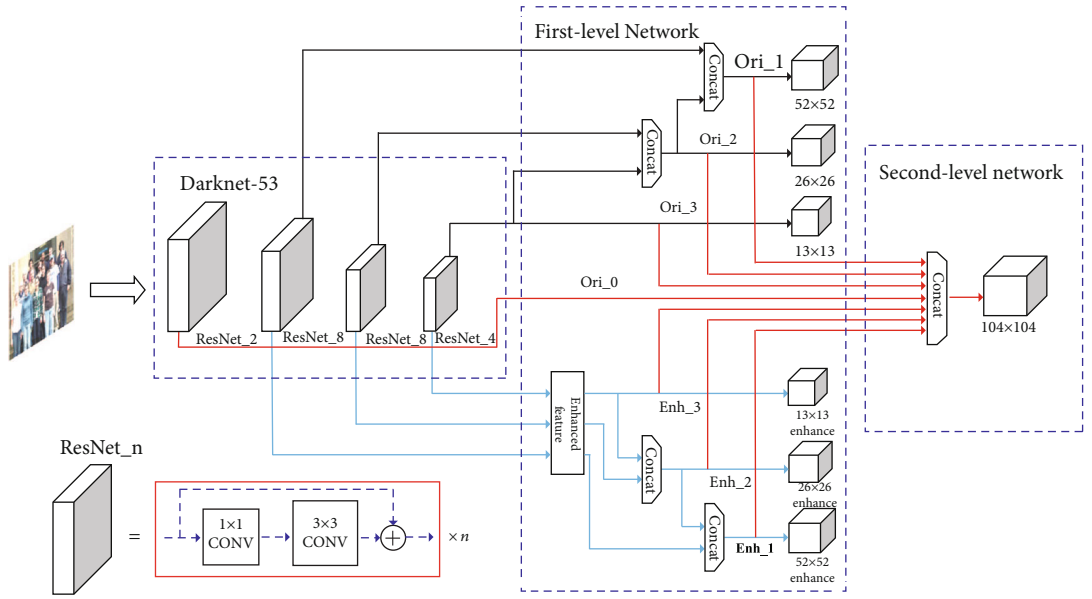


Figure 1: Basic framework of MSNFD. The basic framework consists of backbone network, first-level network, and second-level network. The schematic of ResNet_n is at the lower left corner.

to detect small faces. And adding context always helps for detecting small faces [17], and it is crucial to enlarge the receptive field to add context. As discussed above, enhanced feature module is proposed. The enhanced feature module is composed of three common convolutions (dilated rate = 1) and two dilated convolutions (dilated rate = 2), as shown in Figure 2. This module has three branches. The rectangular blocks are common convolutions, and the dotted blocks are dilated convolutions. Different dilated rates are shown at the bottom left of Figure 2. The first branch of the module is designed to retain information of origin feature map. The common convolution this branch only has is used to reduce the dimension. The second branch combines a dilated convolution with a common convolution. The dilated convolution is shaped by a 3×3 kernel and 2 dilated rate. N dilated rate is considered as inserting n-1 zeros between every two values in the convolution kernel. Downsampling is a common way to enlarge the receptive field, but it results in reducing the feature map and loss of some important information of small faces.

Dilated convolution can also enlarge the receptive field and fuse contextual information effectively [23] especially for small faces with the same number of parameters. It will not change the size of feature map. As shown in Figures 3(a) and 3(b), the context here means information of the face around, such as the neck and clothes, which benefits for detecting small faces. The third branch further enlarges the receptive field and contains more contextual information based on the second branch.

In addition, the enhanced feature module combines information of receptive fields of different scales through by three branches. That is to say, the module combines both texture and contextual information, respectively, from the common and dilated convolution. Information of features on multiple scales will be enriched so that the enhanced feature module can handle multiscale faces to some extent while detecting small faces performs better.

2.3. Two-Stage Weight Loss Function. Second-level network has different anchor scales from first-level network. Second-level network decreases the anchor scale with lots of simple and easily divided negative samples. Joint training for the first-level and second-level networks is adopted. As it is, the sample imbalance causes unbalanced training. If the loss function between first-level and second-level network has not been balanced to an even state, the whole training process will slide overwhelmingly into the second-level network because there are more samples in it than the first-level. In that case, the first-level network gets insufficient training, and the second-level network performs ill because the second-level is generated from the first-level. Thus, two-stage weight loss function is proposed. Considering the match between the anchor and the loss function, the location loss function remains the same as the distance equation used in clustering. The first stage loss function is defined as equations (1)–(3).

$$L_{\text{first_output}}(p, p^*, t, t^*) = \sum_{i=1}^{6}\left[\frac{1}{N_{\text{cls}_i}}\sum L_{\text{cls}_i}(p, p^*) + \frac{1}{N_{\overline{\text{iou}_i}}}\sum p^* L_{\overline{\text{iou}_i}}(t, t^*)\right], \tag{1}$$

$$L_{\text{second_output}}(p, p^*, t, t^*) = \frac{1}{N_{\text{cls}}}\sum L_{\text{cls}}(p, p^*) + \frac{1}{N_{\overline{\text{iou}}}}\sum p^* L_{\overline{\text{iou}}}(t, t^*), \tag{2}$$

$$L_{\text{one_stage}} = \frac{N_{\text{second}}}{N_{\text{first}} + N_{\text{second}}} L_{\text{first_output}} + \frac{N_{\text{first}}}{N_{\text{first}} + N_{\text{second}}} L_{\text{second_output}}. \tag{3}$$

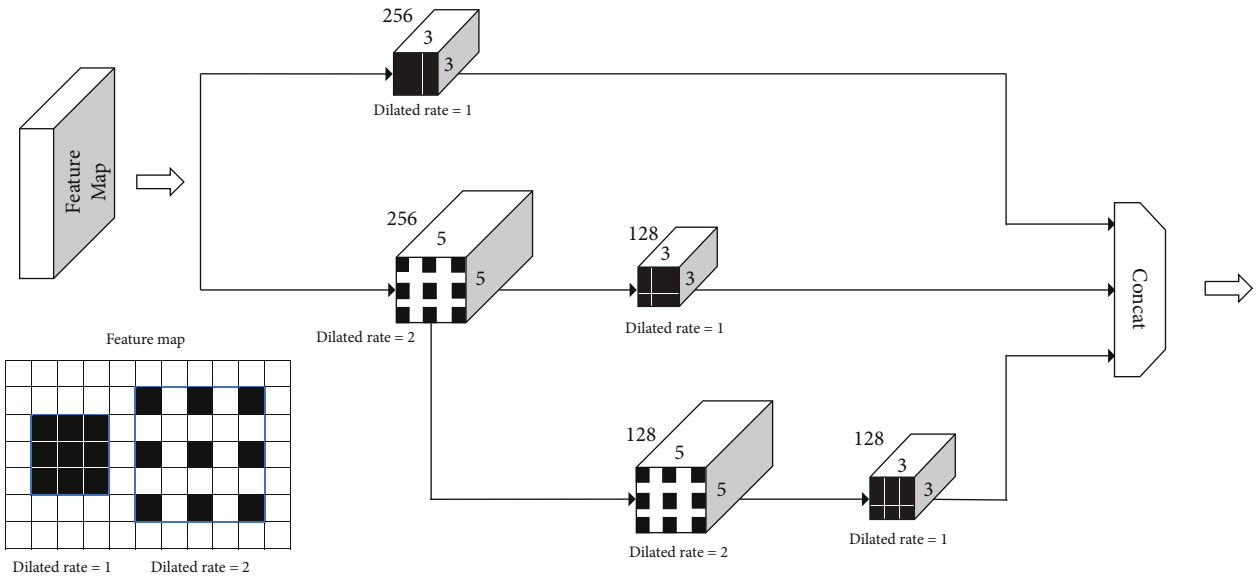


Figure 2: Enhanced feature module. The enhanced feature module is composed of two common convolutions and two dilated convolutions.

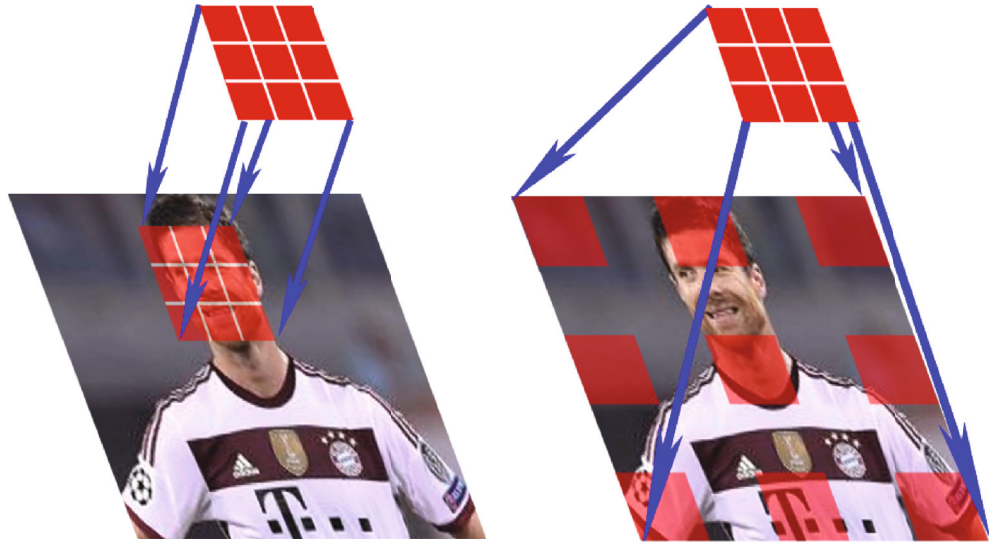


Figure 3: Comparison of different type convolutions: (a) common convolution; (b) dilated convolution.

$L_{\text{first_output}}$ and $L_{\text{second_output}}$ refer to the loss function of first-level and second-level network, p is the predicted probability of face, and $p^* = \{0, 1\}$ is the true probability. When the anchor is positive, the value of p^* is 1, otherwise 0. t means the predicted bounding boxes and t^* the ground truth. $L_{\text{cls_i}}$ is the classification loss over two classes (face vs. background) of the i-th output, and $i = \{1, 2, \cdots, 6\}$ represents six outputs of first-level network. $\text{iou}(t, t^*)$ indicates intersection over union (IOU) between the predicted bounding box (t) and the ground truth (t^*). The larger the IOU is, the more accurately the method locates the face. $L_{\overline{\text{iou_i}}}$ indicates location loss which adopts the equations $1 - \text{iou}(t, t^*)$ used in clustering. Moreover, $L_{\overline{\text{iou_i}}}$ only has effect on positive anchors. L_{cls} and $L_{\overline{\text{iou}}}$ are classification loss and location loss of second-level network, respectively. Equation (3) shows the first-stage loss function, where N_{first} and N_{second} represent the number of anchors in the first- and second-level network, respectively. The second-level network has more anchors, so its proportion in the whole loss function needs to be reduced. Thus, the imbalance of loss function between first- and second-level networks can be eliminated via first-stage training.

After the loss function gets balanced and adjusted through the first-stage training, the first-level network will be fully trained. But it is still difficult to train small faces, so the training of second-level network in the second stage needs to be strengthened by means of increasing the

Table 1: Anchor scale of different methods.

Output	MSNFD	YOLO v3
13×13	(106, 140) (166, 220) (320, 400)	(90, 116) (156, 198) (326, 373)
26×26	(33, 38) (32, 49) (38, 45)	(30, 61) (45, 62) (59, 119)
52×52	(17, 23) (21, 25) (23, 30)	(10, 13) (16, 30) (23, 33)
13×13 enhance	(44, 56) (56, 72) (74, 98)	—
26×26 enhance	(28, 31) (25, 36) (29, 39)	—
52×52 enhance	(13, 13) (13, 17) (15, 18)	—
104×104	(6, 7) (9, 11) (11, 14)	—

proportion of the second-level network. The second-stage loss function is defined as equation (4).

$$L_{\text{two_stage}} = \frac{N_{\text{first}}}{N_{\text{first}} + N_{\text{second}}} L_{\text{first_output}} + \frac{N_{\text{second}}}{N_{\text{first}} + N_{\text{second}}} L_{\text{second_output}}. \tag{4}$$

3. Experiments and Analysis

3.1. Training Details. The training dataset used in the research is the WIDER FACE [24] training set including 12,880 images, 62 daily scenes, and a total of 158,945 faces. These faces are also available on scales of a wide range, from tiny to giant. In order to further enhance the model of training robustness, data enhancement is used during training, such as vertical flipping, random cropping, and random panning. The classification loss adopts focal loss to balance the positive and negative samples, and the location loss uses cross-entropy loss. The optimization uses stochastic gradient descent, and the batch-size is set to 8. In the first stage of training, 3×10^5 steps about 100 epochs are iterated with initial learning rate of 10^{-4}. In the second stage, 9.7×10^4 steps about 30 epochs are iterated with initial learning rate of 10^{-6}, based on the first stage but finely tuned. Pretrained Darknet-53 which obtains 55.3%mAP on COCO [25] is used to initialize the parameters.

3.2. Anchor Analysis. YOLO v3 uses k-means clustering to determine the anchor scale. It is shown in Table 1. K-means is also used to cluster but with more cluster centers to get more anchor scales on WIDER FACE training set in this research. As listed in Table 1, the anchor scales range from 6 pixels to 400 pixels and are divided into 7 types, each for a certain output of the multilevel networks. Within the testing framework based on anchors, the matching degree between anchors and faces determines the quality of the eventual result of the final training. Expanding number of the anchor scale improves the matching degree between anchors and faces. MSNFD does not increment B of the output tensor, $S \times S \times (B \times 5 + C)$ simply but adds the output tensors with the same $B = 3$ as YOLO v3 because the multilevel network has more outputs. In this research, anchor scale ranges from YOLO v3's 9 (3×3 outputs) to 21 (3×7 outputs). The 21 kinds of anchor scales added newly by this research, and the previously existing 9 kinds of anchor-scale of YOLO v3 are, respectively, matched with the face size

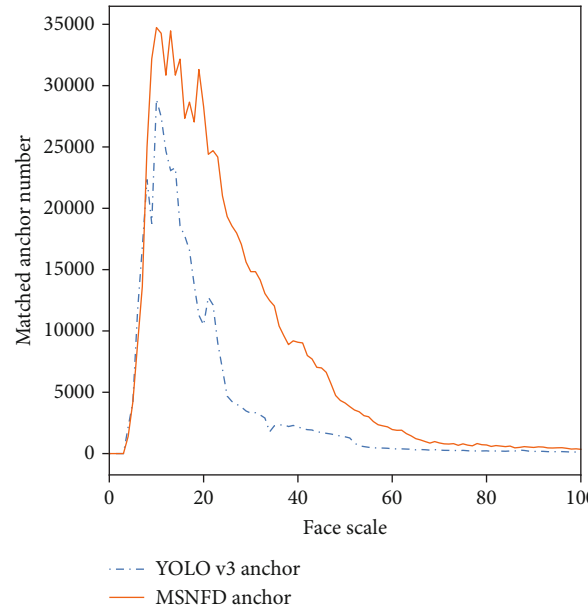


Figure 4: Comparisons on number of anchor matching faces between YOLO v3 and MSNFD.

of the training set. If IOU is greater than 0.4, the matching is successful. The statistics of successful matching on different scales is visualized in Figure 4. As shown in Figure 4, the multilevel network has more anchor scales, and the number of successful matching cases is far more than in YOLO v3, which improves the matching rate between the discrete anchor and the continuous face. In summary, the preset anchor is effective, which benefits the multilevel network being fully learned.

3.3. Model Analysis. In order to prove the effectiveness of the proposed methods, different methods are experimented on WIDER FACE validation dataset and Multiattribute Labelled Faces (MALF) dataset [26]. The WIDE FACE validation set is split into three subsets: easy, medium, and hard. "Hard" subset contains faces with much smaller scales, extreme pose, exaggerated expressions, and large portion of occlusion [24]. "Hard" subset is regarded as a particular set of small faces. MALF chooses the size of 60 and 90 to divide the scale range into small, medium, and large intervals corresponding to small, medium, and large subsets. Both datasets are very suitable for evaluating the method. Different parts are

Table 2: Results on MALF set and WIDER FACE validation set.

Method	WIDER FACE				MALF		
	Easy (%)	Medium (%)	Hard (%)	Time (ms)	Large (%)	Medium (%)	Small (%)
RFAB-f-R-FCN	91.3	89.6	79.4	109	—	—	—
PPN	79.5	77.4	57.7	17	—	—	—
XIE	88.5	85.4	73.8	190	—	—	—
HR	91.9	90.8	82.3	377	—	—	—
VJ	41.2	33.3	13.7	20	45.8	35.7	11.4
YOLO v3	87.5	80.9	47.3	29	75.0	73.1	58.5
MSNFD (enhance feature module)	86.6	79.9	52.3	30	77.5	82.6	65.5
MSNFD (first-level)	91.8	87.8	54.5	31	79.5	83.5	69.3
MSNFD (first-level + second-level)	91.2	87.5	67.1	33	79.7	84.1	77.8
MSNFD (first-level + second-level + two-stage train)	93.5	90.7	71.1	33	82.8	88.8	80.2

gradually added to existing YOLO v3 framework to build up MSNFD. The platform's hardware configuration involves CPU i7-9700 and GPU NVIDIA GTX 1080ti. Calculate average time detecting on WIDER FACE validation dataset and the average precision (AP) on WIDER FACE set and MALF set and the results are shown in Table 2. In this section, RFAB-f-R-FCN [27], PPN [28], XIE [29], HR [17], and VJ [7] are also used for being compared to MSNFD. They are the latest methods published in the past two years or the classical methods.

3.3.1. Enhanced Feature Module. Firstly, add enhanced feature module into YOLO v3 only, the result shows that AP increases from 58.5% to 65.5% on the small set of MALF and from 47.3% to 52.3% on the hard set of wider face, because the enhanced feature module can combine the contextual information of small faces to improve the representation ability of features. However, the detection effect drops on the easy and medium set of WIDER FACE. It is unsatisfactory to add enhanced feature module into YOLO v3 only.

3.3.2. First-Level Network. Secondly, in order to reverse the decline of detection capability of large and medium faces, enhanced feature module and the YOLO v3 together constitute the first-level network which raises AP to 79.5%, 83.5%, and 69.3% on MALF and 91.8%, 87.8%, and 54.5% on WIDER FACE. Attention should be paid that the performance of first-level network on the hard dataset is not outstanding enough compared to only adding the enhanced feature module with AP just raising by 3.8% from 65.5% to 69.3% and 2.2% from 52.3% to 54.5%. This is because that YOLO v3 is not ideal enough for small face detection and performs poorly when combined with the enhanced feature module to form the first-level network.

3.3.3. Second-Level Network. Thirdly, in order to further improve the ability of detecting small faces, the second-level network is introduced. The second-level network not only incorporates the contextual and multiscale information from the first-level network but also contains the texture information from low layers. Moreover, it reduces anchor scale to match smaller faces. The above adjustments help to improve the performance of small faces detection. As listed in Table 2, AP on small and hard set raised to 8.5% and 12.6%, which fully indicates that second-level network benefits small faces detecting. In addition, it can be seen that after adding second-level network, AP drops from 91.8% to 91.2% on easy set. As having analysed earlier, the adding of second-level network will cause imbalance of both samples and the loss functions between networks of different levels. The second-level network has more samples and will occupy the main part of learning process, resulting in the inappropriate deviation of normal distribution and the decline of classification ability of the first-level network.

3.3.4. Two-Stage Weight Loss Function. Finally, to further optimize the network, it is necessary to introduce two-stage weight loss function to balance network of different levels. The first-stage training can balance the loss between networks of different levels and solve the problem of performance declining in detecting large faces after adding the second-level network. The second-stage training intends to further optimize the second-level network, that is, to enhance the detection ability for small faces. As shown in Table 2, after using the two-stage weight loss function, whether the face is large or small and easy or hard, the overall performance is largely improved and achieves AP, 82.8%, 88.8%, and 80.2% and 93.5%, 90.7%, and 71.1%.

3.3.5. Comparison and Runtime Analysis. From Table 2, it is noticed that YOLO v3 does not perform as well as PPN on hard subset of WIDER FACE. However, MSNFD has more excellent results than PPN and YOLO v3. It indicates that the better performance is correlated with the method proposed. The AP of MSNFD is 23.8% on hard subset and 21.7% on small subset higher than YOLO v3, though MSNFD is just 4 ms slower than YOLO v3. The best three methods with higher AP on hard subset of WIDER FACE such as 82.3%, 79.4%, and 71.1% are HR, RFAB-f-R-FCN, and MSNFD. HR is a well-known method designed to detect tiny faces. HR provides much better detection performance about 11.2% higher than MSNFD, but HR runs much slower. HR takes 377 ms to detect an image on wider face while MSNFD only needs 33 ms. In addition, the results of MSNFD

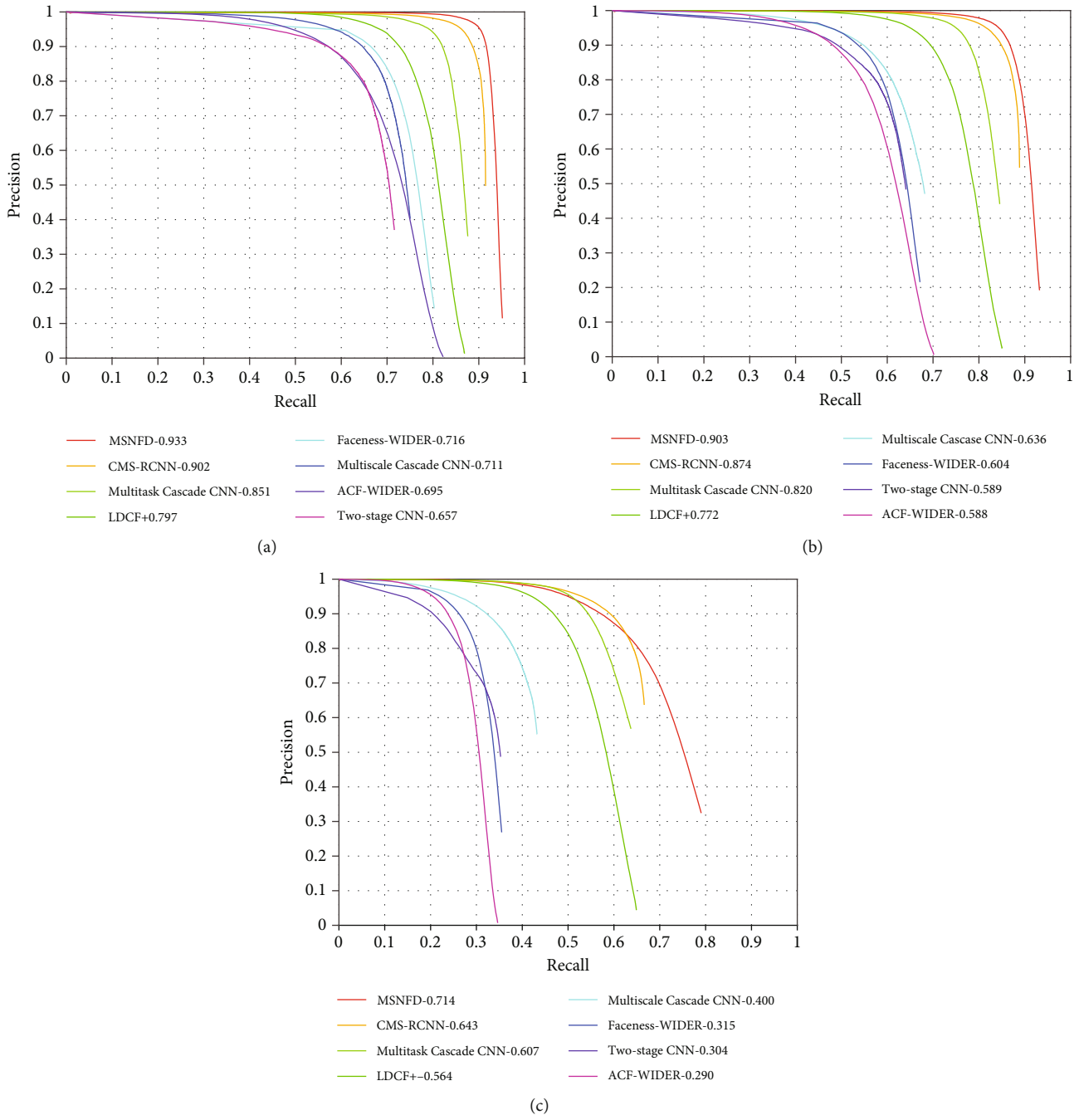


Figure 5: Comparisons on WIDER FACE test set: (a) test: easy; (b) test: medium; (c) test: hard.

are rather similar to HR on easy and medium subsets. MSNFD even surpasses HR on easy subset. RFAB-f-R-FCN is designed for small-scale face published in 2020. The run time of MSNFD is 33 ms which is nearly 3 times smaller than RFAB-f-R-FCN. The two fastest methods are PPN which is published in 2019 and VJ. They only need 17 ms and 20 ms. However, their detection results are extremely disappointing especially for small faces. PPN achieves 57.7%, and VJ achieves 13.7% on hard subsets and 11.4% on small subsets. MSNFD runs slower than PPN and VJ a bit but has much higher AP. Moreover, XIE is a newly proposed method in 2019 which has very closed accuracy to MSNFD, but it runs nearly 6 times slower than MSNFD. In other words, MSNFD can balance accuracy and speed well compared to other methods with similar accuracy. In a word, MSNFD improves accuracy greatly while retaining YOLO V3's speed advantage. Both test code and models are available online at https://github.com/JackJinWang/Face-detection/tree/master.

4. Comparisons with Real-Time Methods

To compare MSNFD with other current popular methods which can detect faces at close to real-time speed, tests were taken on WIDER FACE test set and Face Detection Data

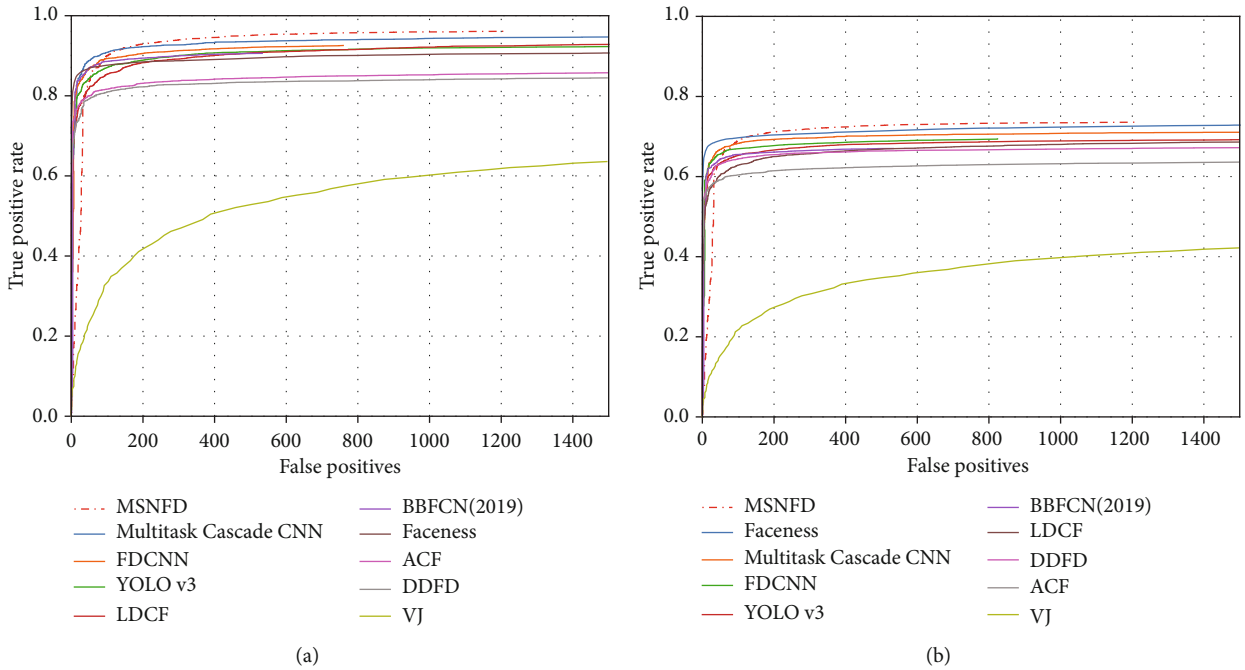


Figure 6: Comparisons on FDDB: (a) discontinuous ROC; (b) continuous ROC.

Figure 7: Detection examples.

Set and Benchmark (FDDB) [30]. Furthermore, the result is further shown via detecting a random selection of images with small faces from the public dataset.

4.1. Evaluate on Benchmark

4.1.1. WIDER FACE Dataset. The WIDER FACE test set has 16,097 images. It is a challenging test set. The results of comparison between MSNFD and real-time methods Multitask Cascade CNN [31], CMS-RCNN [32], LDCF [33], Multiscale Cascade CNN [24], Faceness [34], Two-stage CNN [24], and ACF [35] are shown in Figures 5(a)–5(c), where MSNFD achieves 93.3%, 90.3%, and 71.4%AP on easy, medium, and hard test sets of the WIDER FACE, respectively, much better than other methods.

4.1.2. FDDB Dataset. FDDB is also a pretty challenging face detection dataset with 2,845 images of 5,171 faces in total in a variety of states, such as occlusion, rare poses, low resolution, and out of focus. MSNFD is compared to the popular and real-time methods BBFCN [36], DDFD [37], Multitask Cascade CNN [31], FDCNN [38], LDCF [33], YOLO v3

[21], Faceness [34], VJ [7], and ACF [35] on FDDB. The ROC curves are shown in Figures 6(a) and 6(b), which the higher the ROC is, the better the method performs. The ROC curve of MSNFD is higher than other methods indicating that MSNFD performs much better.

4.2. Detection Examples. This section shows some selected examples of small faces detection from WIDER FACE test set (in Figure 7). Though, there are so small and many faces in these images, MSNFD is able to detect accurately. These successful results definitely prove that MSNFD is able to work perfectly especially for small faces.

5. Conclusion

A multilevel single stage network for face detection with more anchor scales is proposed to raise the matching rate between anchors and faces especially small faces. The enhanced feature module in the first-level network not only integrates the contextual information but also strengthens the ability to process multiscale data, enriching the overall information of small face features. The second-level network has smaller anchor scales. In addition, it is generated from the first-level one and is integrated with texture features of low layers, which enhances the process of semantic facial information, especially for small faces. During the training, the two-stage weight loss function is used to balance the network to optimize the classification effect of different levels. MSNFD improves accuracy greatly while retaining YOLO V3's speed advantage. Tests on the WIDER FACE datasets show that MSNFD achieves 93.3%, 90.3%, and 71.4% AP. In a word, MSNFD has greater reference significance for real-time face detection.

Data Availability

The data used to support the findings of this study are available from the authors 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 Scientific Research Project of Tianjin Education Commission (2020KJ024).

References

[1] Y. Feng, F. Wu, X. Shao, Y. Wang, and X. Zhou, "Joint 3D face reconstruction and dense alignment with position map regression network," in *Computer Vision – ECCV 2018: 15th European Conference*, pp. 557–574, Munich, Germany, 2018.

[2] J. Wan, J. Li, Z. Lai, L. Zhang, and B. du, "Robust face alignment by cascaded regression and de-occlusion," *Neural Networks*, vol. 123, pp. 261–272, 2020.

[3] G. Hermosilla, J. L. Verdugo, G. Farias, E. Vera, F. Pizarro, and M. Machuca, "Face recognition and drunk classification using infrared face images," *Journal of Sensors*, vol. 2018, Article ID 5813514, 8 pages, 2018.

[4] Z. Lei, X. Zhang, S. Yang, Z. Ren, and O. F. Akindipe, "RFR-DLVT: a hybrid method for real-time face recognition using deep learning and visual tracking," *Enterprise Information Systems*, vol. 14, no. 9-10, pp. 1379–1393, 2020.

[5] K. Li, Y. Jin, M. W. Akram, R. Han, and J. Chen, "Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy," *The Visual Computer*, vol. 36, no. 2, pp. 391–404, 2020.

[6] B. Bozorgtabar, D. Mahapatra, and J. P. Thiran, "ExprADA: adversarial domain adaptation for facial expression analysis," *Pattern Recognition*, vol. 100, article 107111, 111 pages, 2020.

[7] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[8] K. Adnan and R. Akbar, "Limitations of information extraction methods and techniques for heterogeneous unstructured big data," *International Journal of Engineering Business Management*, vol. 11, article 1177, 2019.

[9] E. Real, A. Aggarwal, Y. Huang, and Q. V. Le, "Regularized evolution for image classifier architecture search," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 4780–4789, 2019.

[10] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, 2020.

[11] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5325–5334, Boston, MA, USA, 2015.

[12] H. J. Wang, Z. Li, X. Ji, and Y. Wang, "Face r-cnn," 2017, https://arxiv.org/abs/1706.01061.

[13] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[14] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Li, "S^{3fd}: single shot scale-invariant face detector," in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 192–201, Venice, 2017.

[15] W. Liu, D. Anguelov, D. Erhan et al., "Ssd: single shot multibox detector," in *Computer Vision – ECCV 2016: 14th European Conference*, pp. 21–37, Amsterdam, The Netherlands, 2016.

[16] S. Zhang, R. Zhu, X. Wang et al., "Improved selective refinement network for face detection," 2019, https://arxiv.org/abs/1901.06651.

[17] P. Hu and D. Ramanan, "Finding tiny faces," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. , 20171522–1530, 2017.

[18] Y. Bai, Y. Zhang, M. Ding, and B. Ghanem, "Finding tiny faces in the wild with generative adversarial network," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21–30, Salt Lake City, UT, 2018.

[19] Z. Li, X. Tang, J. Han, J. Liu, and R. He, "Pyramidbox++: high performance detector for finding tiny face," 2019, https://arxiv.org/abs/1904.00386.

[20] L. Shi, X. Xu, and I. A. Kakadiaris, "SSFD$^+$: a robust two-stage face detector," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 1, no. 3, pp. 181–191, 2019.

[21] J. Redmon and A. Farhadi, "Yolov3: an incremental improvement," 2018, https://arxiv.org/abs/1804.02767.

[22] F. Gurkan, B. Sagman, and B. Gnsel, "YOLOv3 as a deep face detector," in *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*, pp. 605–609, Bursa, Turkey, 2019.

[23] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, https://arxiv.org/abs/1511.07122.

[24] S. Yang, P. Luo, C.-C. Loy, and X. Tang, "Wider face: a face detection benchmark," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5525–5533, Las Vegas, NV, USA, 2016.

[25] T.-Y. Lin, M. Maire, S. Belongie et al., "Microsoft coco: common objects in context," in *Computer Vision – ECCV 2014: 13th European Conference*, pp. 740–755, Zurich, Switzerland, 2014.

[26] B. Yang, J. Yan, Z. Lei, and S. Z. Li, "Fine-grained evaluation on face detection in the wild," in *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, pp. 1–7, Ljubljana, Slovenia, 2015.

[27] C. Tang, S. Chen, X. Zhou, S. Ruan, and H. Wen, "Small-scale face detection based on improved R-FCN," *Applied Sciences*, vol. 10, no. 12, pp. 4177–4193, 2020.

[28] D. Zeng, H. Liu, F. Zhao, S. Ge, W. Shen, and Z. Zhang, "Proposal pyramid networks for fast face detection," *Information Sciences*, vol. 495, no. 12, pp. 136–149, 2019.

[29] R. Xie, Q. Zhang, E. Yang, and Q. Zhu, "A method of small face detection based on CNN," in *2019 4th International Conference on Computational Intelligence and Applications (ICCIA)*, pp. 78–82, Nanchang, China, 2019.

[30] V. Jain and E. Learned-Miller, *Fddb: a benchmark for face detection in unconstrained settings*, University of Massachusetts, Amherst technical report, 2010.

[31] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[32] C. Zhu, Y. Zheng, K. Luu, and M. Savvides, *CMS-RCNN: contextual multi-scale region-based CNN for unconstrained face detection*, Springer International Publishing, 2017.

[33] E. Ohn-Bar and M. M. Trivedi, "To boost or not to boost? On the limits of boosted trees for object detection," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 3350–3355, Cancun, 2016.

[34] S. Yang, P. Luo, C.-C. Loy, and X. Tang, "From facial parts responses to face detection: a deep learning approach," in *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 3676–3684, Santiago, Chile, 2015.

[35] B. Yang, J. Yan, Z. Lei, and S. Z. Li, "Aggregate channel features for multi-view face detection," in *IEEE International Joint Conference on Biometrics*, pp. 1–8, Clearwater, FL, USA, 2014.

[36] L. Liu, G. Li, Y. Xie, Y. Yu, Q. Wang, and L. Lin, "Facial landmark machines: a backbone-branches architecture with progressive representation learning," *IEEE Transactions on Multimedia*, vol. 21, no. 9, pp. 2248–2262, 2019.

[37] S. S. Farfade, M. J. Saberian, and L.-J. Li, "Multi-view face detection using deep convolutional neural networks," in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, pp. 643–650, Shanghai, China, 2015.

[38] D. Triantafyllidou, P. Nousi, and A. Tefas, "Fast deep convolutional face detection in the wild exploiting hard sample mining," *Big Data Research*, vol. 11, pp. 65–76, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 8842396, 8 pages
https://doi.org/10.1155/2021/8842396

Research Article

Depth and Width Changeable Network-Based Deep Kernel Learning-Based Hyperspectral Sensor Data Analysis

Jing Liu,[1] **Tingting Wang**,[2] **and Yulong Qiao**[1]

[1]*College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China*
[2]*School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China*

Correspondence should be addressed to Yulong Qiao; qiaoyulonghrbeu@126.com

Received 25 September 2020; Revised 28 November 2020; Accepted 20 January 2021; Published 22 February 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Jing Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sensor data analysis is used in many application areas, for example, Artificial Intelligence of Things (AIoT), with the rapid developing of the deep neural network learning that promotes its application area. In this work, we propose the Depth and Width Changeable Deep Kernel Learning-based hyperspectral sensing data analysis algorithm. Compared with the traditional kernel learning-based hyperspectral data classification, the proposed method has its advantages on the hyperspectral data classification. With the deep kernel learning, the feature is mapped through many times mapping and has the more discriminative ability. So, the deep kernel learning has the better performance compared with the multiple kernels learning. And it has the ability to adjust the network architecture for hyperspectral data space, with the optimization equation of the span bound. The experiments are implemented to testified the feasibility and performance of the algorithms on the hyperspectral data analysis, with the classification accuracy of hyperspectral data. The comprehensive analysis of the experiments shows that the proposed algorithm is feasible to hyperspectral sensor data analysis and its promising classification method in many areas data analysis.

1. Introduction

With the developing of the deep learning theory, the researchers have proved the deep architecture is feasible and effective in several applications. Deep kernel learning is widely used in image analysis, and the results show it is the effective framework [1, 2], including image annotation [3], classification [4], segmentation [5], and anomaly object detection [6]. The previous works present the deep multiple kernels learning [7] and the novel objective function [8], and in this deep kernel learning framework, the adaptive back propagation is used to update coefficients and weights of the network to improve the ability of representation [9–12].

In the deep kernel framework, the accuracy and speed of visual recognition are effectively improved by using random gradient descent, which provides a research framework for the follow-up deep kernel learning. The researcher proposed a multilayer kernel machine based on the deep confidence network in 2012, which describes the structure from the support vector machine to deep multilayer more completely [12]. From the perspective of popular learning, Brahma et al. analyzed the principle of improving the performance of kernel learning under the deep neural network from the perspective of popular learning and gave a relatively perfect theoretical framework of deep kernel learning [13]. For the classification problem, Wong et al. proposed a regularized depth Fisher kernel mapping method in 2011. The learning sample space was explicitly mapped to the feature space, and the deep neural network was used to improve the separability of features according to the Fisher criteria [14]. Based on the traditional KPCA, The research results show that FKNN and DLF effectively improve the image classification performance [15]. To solve the clustering problem, the researcher proposed a k-means algorithm based on RBF kernel based on the nonlinear kernel and multilayer architecture [16], the researchers used the kernel k-means method to detect the change of the satellite image [17], and realized the pixel representation of a pseudo training set of the nonlinear clustering partition by deep learning areas. For the recursive problem, the researcher proposed a local deep

kernel learning algorithm to accelerate the prediction performance of nonlinear support vector machines and realize the function of optimizing the spatial tree structure [18]. In view of the problem of the deep model selection, Strobl et al. proposed an idea based on multicore deep learning and effectively improved the performance of deep kernel learning through learning multilayer and multiple kernel functions [8]. For the improvement of the deep kernel learning algorithm with different application backgrounds, Li et al. proposed a data recognition algorithm combining CNN and multicore learning for emotional recognition, which effectively improved the recognition performance [19]. On tumor classification of ultrasound images, a research framework based on multicore deep learning was proposed [20]. On the problem of discourse level analysis, Poria proposed a data recognition algorithm based on the multicore model and deep CNN. For the problem of protein-protein interaction in biological information, Yang proposed a classification algorithm based on deep kernel learning [21]. Deep kernel learning has to be applied to many areas, for example, image recognition [33] and pattern prediction [34]. The structure variable deep kernel learning architecture is very feasible, because the width and depth are adjusted. The features are extracted with the different kernel networks, and the performance are different with the different parameters. These parameters are adjusted through the data training. The number of training data is different in the different applications.

In our recent work [35], the novel multiple kernel learning-based deep neural network algorithm work was proposed for the data classification. The work presented the algorithm framework, and it has the advantage on the nonlinear data classification. For the hyperspectral data classification, the algorithm [35] has its ability for the data analysis, but it must be optimized for the hyperspectral data analysis. The hyperspectral data analysis is different from the general data classification in [35]. We have the analysis on the kernel learning with the advantage on the hyperspectral data analysis in the previous work [36], but the algorithm in [36] is the traditional kernel learning framework, with less performance than the deep learning. So, it is effective to improve the performance of kernel learning on the hyperspectral learning. The similar results are testified on the SAR image classification in the recent work [37], only many kernel functions are combined with some combination weights, and the features are mapped with only one time of projection compared with the traditional kernel, but the structure in [37] is unchangeable, and the performance is limited.

So, motived by the work [35–37], we propose the novel structured changeable deep kernel for hyperspectral sensing data analysis. Different from the work [35], the parameter of structure is optimized with the optimization algorithm, so compared with the algorithm in [35], the algorithm has the ability of adjusting the architecture for the more performance on the hyperspectral data analysis compared with the algorithm in [36, 37]. In the framework, the architecture parameters are solved with the optimization equation using the span bound. The performance on the hyperspectral sensor data is testified and evaluated for hyperspectral applications.

2. Proposed Variable Structure-Based Deep Learning Network

2.1. Framework. As an excellent method of nonlinear feature extraction and classification, the kernel learning method can be extended to the deep learning network architecture and enhance the ability of multilayer feature extraction. Although the deep learning network and kernel learning algorithm have been combined and the algorithm is presented in the previous work [37], it cannot achieve the adaptive adjustment of the network structure. So, the network is not adaptive to heterogeneous data.

For the hyperspectral data analysis, the hyperspectral data $D = \{(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)\}$ where the hyperspectral vector $x_i \in R^n$ with the labels vector $y_i \in \{-1, 1\}$, $i = 1, 2, \cdots, m$. With the kernel function $k(x_i, x_j)$, the function is calculated with $k(x_i, x_j) = \phi(x_i)\phi(x_j)$ with the nonlinear mapping features of the input hyperspectral data. The hyperspectral data matrix, K, is calculated with the following equation. And the classification criterion function is shown as

$$f(x) = \text{sign}\left(\sum_{i=1}^{m} \alpha_i y_i k(x_i, x) + b\right), \tag{1}$$

where α_i is the coefficient vector and b is the bias vector for the classification criterion $f(x)$. For the hyperspectral data classification application, the optimization equation is listed as follows:

$$\begin{aligned} \min_{(\alpha,b,\xi)} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m} \xi_i s.t. \mathrm{y}_i(\alpha_i k(x_i, x) + b) \\ \geq 1 - \xi_i, \xi_i \geq 0, \mathrm{C} > 0, \quad i = 1, 2, \cdots, m, \end{aligned} \tag{2}$$

where $\omega = \sum_{i=1}^{m} \alpha_i y_i x_i$, and the parameters ξ_i and C are chosen with the crossvalidation method; then, the algorithm can be defined as

$$K(x_i, x) = \sum_{i=1}^{M} \theta_i k_i(x_i, x), s.t. \theta_i \geq 0, \sum_{i=1}^{M} \theta_i = 1, \tag{3}$$

where M is the number of the different kernel functions, and θ_i is the parameter of combining the kernel functions. Then, under the multiple kernel functions, the classification decision criterion is transformed as

$$f(x) = \text{sign}\left(\sum_{k=1}^{M} \theta_k \sum_{i=1}^{m} \alpha_i y_i k(x_i, x) + b\right). \tag{4}$$

And with this learning model, the combined kernel of each layer in the DWS-MKL algorithm is a linear combination of multiple kernel functions. The general framework is shown in Figure 1. M is the total number of basis kernel functions, and θ_i is the combination coefficient of the combined kernel function. In the previous work [35], we have present the MKL framework.

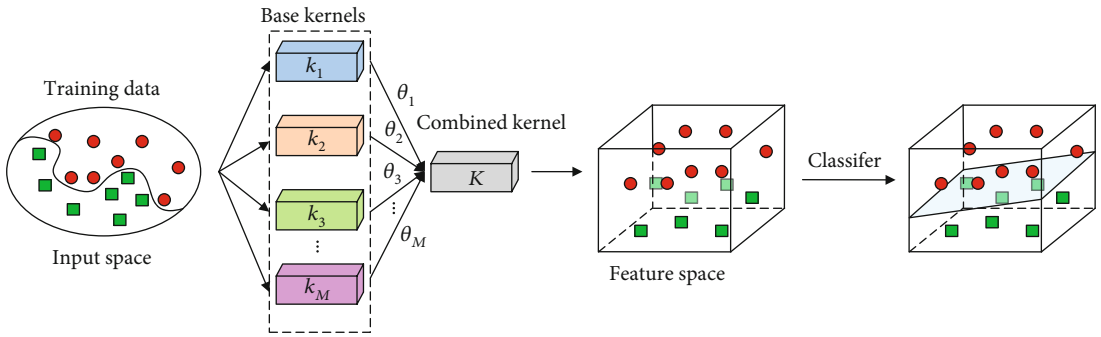


Figure 1: The framework of multiple kernel learning.

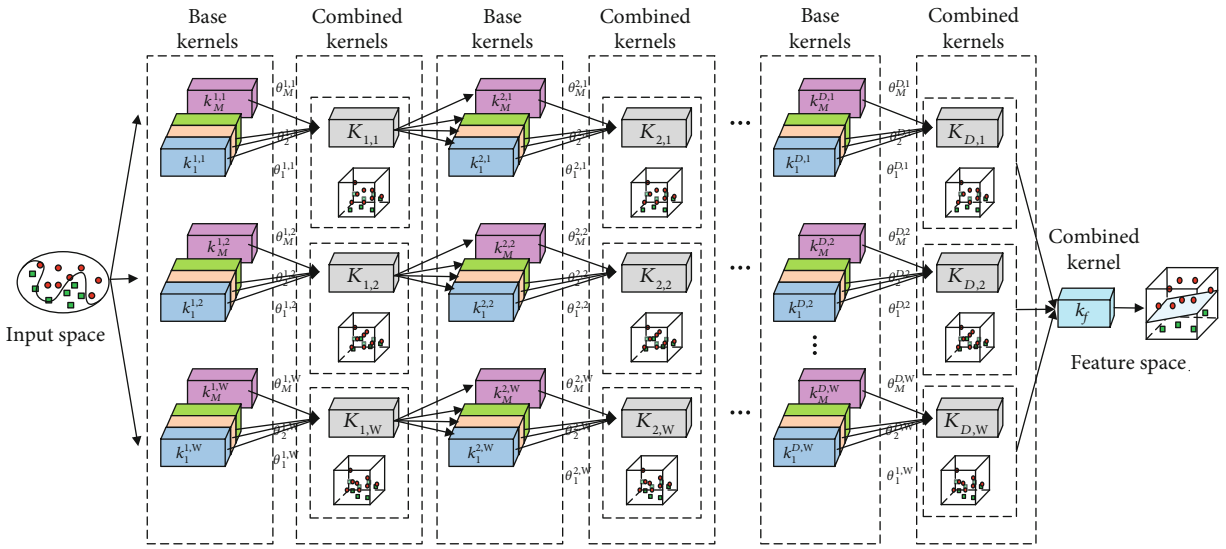


Figure 2: The framework of depth-width-scaling multiple kernel learning.

The overall architecture of the DWS-MKL algorithm's variable depth and width combined structure is shown in Figure 2. Since the layers of the DWS-MKL algorithm are directly cascaded, the output of the combined kernel function of the previous layer is the input of the base kernel function of the next layer, and the channels are independent of each other. Therefore, it is easy to adjust the depth and width combination of the algorithm. For the same data set, the classification accuracy of the DWS-MKL algorithm under different depth and width combined architectures is different. For data sets of different scales, if the depth and width combination structure are fixed, the classification accuracy rate under all data sets cannot be maximized. Therefore, it is necessary and reasonable to determine the complexity of the model according to the complexity of the problem to be solved. Adjust the structure of the algorithm flexibly according to the data set, which is the main feature of the DWS-MKL algorithm. $K_{D,m}$ is the first combined core of the last layer except K_f for the deep and wide combined structure of the DWS-MKL algorithm. $\theta_m^{d,w}$ is the combination coefficient corresponding to the base kernel function.

The DWS-MKL algorithm uses SVM as a classifier to solve basic data classification problems. The input of the SVM classifier is the output feature of the combined kernel K_f. The specific architecture of the classifier is shown in Figure 3. The decision function of the SVM is

$$f(x) = \text{sign}\left(\sum_{i=1}^{n} \alpha_i y_i (x_i \cdot x) + b\right), \tag{5}$$

Where α_i is the dual coefficient and b is the bias of the decision function $f(x)$. The optimization problem of SVM is

$$\begin{aligned} \min_{(\alpha,b,\xi)} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m} \xi_i s.t. \mathrm{y}_i(\alpha_i k(x_i, x) + b) \\ \geq 1 - \xi_i, \xi_i \geq 0, \mathrm{C} > 0, \quad i = 1, 2, \cdots, m, \end{aligned} \tag{6}$$

where $\omega = \sum_{i=1}^{m} \alpha_i y_i x_i$, ξ_i is the slack variable, and C is the regularization coefficient. The basic structure of the DWS-MKL algorithm is a multiple combined kernel, and each combined kernel is a weighted linear combination of multiple basic kernel functions. Therefore, the decision function of the algorithm can be written as

$$f(x) = \text{sign}\left(\sum_{k=1}^{M} \theta_k \sum_{i=1}^{m} \alpha_i y_i k(x_i, x) + b\right), \tag{7}$$

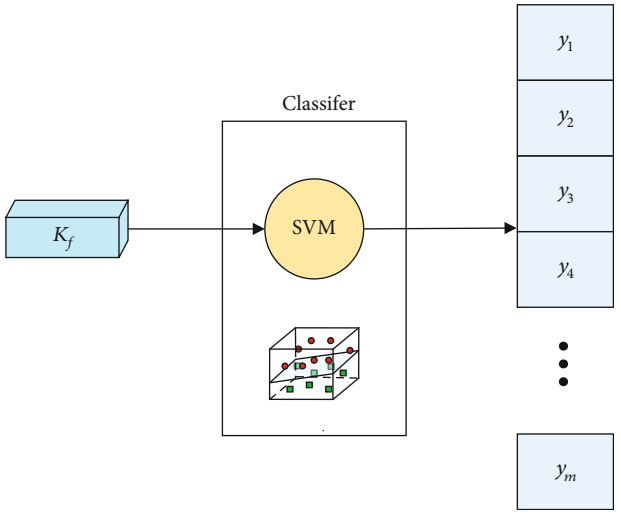


Figure 3: Classification network.

Table 1: Formulation of 3-order kernel functions.

Kernel function	Formulation
Linear kernel	$k^{(3)}(x,y) = k^{(2)}(x,y) = k(x,y) = x \cdot y$
Polynomial kernel	$k^{(2)}(x,y) = \left(\alpha(k(x,y)+\beta)^p\right)$
	$k^{(3)}(x,y) = \left(\alpha\left(k^{(2)}(x,y)+\beta\right)^p\right)$
Gaussian kernel	$k^{(2)}(x,y) = e^{-\left(1/\sigma^2\right)\left(1-k(x,y)\right)}$
	$k^{(3)}(x,y) = e^{-\left(1/\sigma^2\right)\left(1-k^{(2)}(x,y)\right)}$

where θ_k is the combination coefficient of the deep and width combination architecture of the DWS-MKL algorithm. Without loss of generality, the decision function of the DWS-MKL algorithm is uniformly written as

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{m} \alpha_i y_i k(x_i, x; \theta_i) + b\right), \tag{8}$$

where combination parameters θ, dual coefficients α, and offsets b are learned through the DWS-MKL algorithm.

2.2. Algorithm. Extending the combined kernel cascade to L -layers, the combined kernel function of the first layer of a deep multiple kernel learning algorithm can be expressed as

$$\begin{aligned} K^{(L)}(x,y) = {} & \phi^{(L)}\left(\phi^{(L-1)}\left(\cdots\phi^{(1)}(x)\right)\right) \\ & \cdot \phi^{(L)}\left(\phi^{(L-1)}\left(\cdots\phi^{(1)}(y)\right)\right), \end{aligned} \tag{9}$$

where x and y are the input vectors of the algorithm. $\phi^{(L)}$ is the nonlinear mapping kernel function. In addition to cascading multilayer combined kernel, the DWS-MKL algorithm extends multiple deep structures to multiple channels, and the channels are independent of each other. Finally, the feature output of each channel is input to a combined kernel K_f according to the sum and average rule. The combined kernel formula is as follows:

Input: $I = \{(x_i, y_i) \mid i = 1, 2, \cdots .m\}$, $\theta^0 1/W$, D, W, lr, C.
Output: Parameter matrix.
1 Initialize k, compute K;
2 for $i \leftarrow 1$ **to** N **do**
3 compute the optimal vector with K_f;
4 **for** $w \leftarrow 1$ **to** W **do**
5 **for** $d \leftarrow 1$ **to** D **do**
6 compute $\tilde{K}_{sv}$, B, $\partial S_p^2/\partial\theta$;$\partial T_{span}/\partial\theta$
7 update $\theta^{i-1} \leftarrow \theta^i + lr \cdot \partial T_{span}/\partial\theta$;
8 **end**
9 **end**
10 end

Algorithm 1: Model = DWS-MKL (I, θ^0, D, W, lr, C).

$$K_f = \frac{1}{W}\sum_{m=1}^{W} K_{D,m}, \tag{10}$$

where K_f is the combined kernel of the last layer except for the deep and wide combined structure of the DWS-MKL algorithm. The total number of combined kernel in each layer is determined by the number of independent channels of the algorithm. The combined kernel $K_{D,\mathrm{W}}$ of the DWS-MKL algorithm in each layer, and channel is defined as

$$\begin{aligned} K_{d,w} = {} & \sum_{m=1}^{M} \theta_m^{d,w} k_m^{d,w},\ d = 1, 2, \cdots, D,\ w = 1, 2, \cdots, W \\ & s.t. \sum_{m=1}^{M} \theta_m^{d,w} = 1, \theta_m^{d,w} > 0 \quad m = 1, 2, \cdots, M, \end{aligned} \tag{11}$$

where $k_m^{d,w}$ is the base kernel function of the layer channel, and $\theta_m^{d,w}$ is the combination coefficient corresponding to the base kernel function.

There are many choices for the basis kernel function of the combined kernel, such as the linear kernel function, polynomial kernel function, RBF kernel function, Laplace kernel function, and sigmoid kernel function. The basic kernel function of the DWS-MKL algorithm can be combined according to specific practical applications. In the example test of the DWS-MKL algorithm of the present invention, three kernel functions are selected as the basis kernel functions of each combined kernel, namely, linear kernel, polynomial kernel, and RBF kernel. The specific formula is shown in Table 1.

2.3. Procedure. The depth and width of the algorithm are limited to $D \in [1, 2, 3]$, $\mathrm{W} \in [1, 2, 3]$. After that, we input the training data for iterative training and use a crossvalidation algorithm to determine the penalty coefficient C and the optimal depth D and width W. After training, we obtain the SVM classification model with classification capability and the combination coefficient of the variable depth and width

Table 2: Performance evaluation on the deep kernel learning.

Targets	Training/test samples	2-DNN [37]	3-DNN [37]	Proposed
2S1&BRDM_2	1000/1000	89.0%	92.9%	93.8%
BRDM_2&ZSU234	1000/1000	89.7%	90.4%	91.5%
2S1&BRDM_2	400/400	91.7%	92%	93%
D7&T62	400/400	100%	100%	100%
T62&2S1	400/400	85.0%	85.5%	86.5%

Table 3: Detection results on deep kernel learning.

Targets	Training/test samples	2-DNN [37]	3-DNN [37]	Proposed
2S1	1000/1000	80.5%	83.2%	84.5%
BRDM_2	1000/1000	85.6%	85.2%	86.4%
ZSU234	1000/1000	85.0%	89.2%	90.5%
D7	1000/1000	65.3%	67.0%	68.0%
ZIL131	1000/1000	68.5%	72.0%	73.5%

architecture of the algorithm. We represent the training set data as $I = \{(x_i, y_i) \mid i = 1, 2, \cdots .m\}$. lr is the learning rate.

3. Experiments

3.1. Algorithm Feasibility Evaluation. Firstly, we evaluate the feasibility of the deep kernel learning, on the MSTAR data.

In the experiment, the same SAR image target as the deep belief network experiment is used as the verification object, and ten kinds of chariot targets under different angles are selected as the data set. Firstly, the original images in MSTAR data set are preprocessed and cut to the same size. Then, 1000 images are selected as the training set and 1000 images as the test set. In the experiments, in order to test the influence of the number of samples in the data set on the results of feature extraction, a comparative experiment was carried out with 400 training sets and 400 test sets as the data objects under the condition that other conditions remain unchanged.

Different data sets are constructed to verify the effect of feature extraction in the two modes, and the target classification results after feature extraction of depth kernel mapping are compared with other common methods. Firstly, the classification task is used to verify the effect of feature extraction. The ordinary SVM only supports two classification methods. Therefore, this part carries out pairwise combination for different categories of chariot targets, then constructs data sets according to the combination structure, and verifies the effect of deep kernel mapping structure feature extraction in classification under two different data set sizes. The results are shown in Table 2.

Through five groups of two classification tasks, the feature extraction results of depth kernel mapping are verified for different objects and different size data sets. The performance of the same group of network parameters in the classification task of different targets is slightly different, but the overall accuracy rate has been improved with the depth of the structure, and the increase of the amount of data contained in the data set will improve the classification accuracy, but the calculation time will also increase accordingly. Therefore, in the follow-up research, we need to focus on the actual application of the indicators of operation speed and algorithm performance. The network structure is optimized and adjusted. After verifying the classification results, some classification methods are implemented to evaluate the algorithms, and some results are shown in Table 3.

Table 4: Accuracy performance on the different learning methods.

Method	RBF-SVM	DBN	CNN	3-kernel deep [37]	Proposed
Accuracy	83.5%	88.2%	92.3%	95.9%	96.5%

The detection accuracy of five kinds of targets is different, which shows that the differences between each category of targets and other categories are different, and the performance of features extracted by different targets is also different. However, with the increase of structure depth, the detection performance is also improved. It shows that the depth structure can improve the ability of kernel mapping in remote sensing image feature extraction, but for all kinds of specific targets and detection, it is also necessary to optimize the structural parameters to improve its performance. Under the same computing resources and data objects, we compare the classification performance of the common target feature extraction methods in the target classification task. The support vector machine with RBF kernel is selected in the single kernel mapping, the four-layer network structure with good effect is selected for the deep confidence network, and the common AlexNet model is selected for the convolution network. The classification results of each method are shown in Table 4.

By comparing the classification results of feature extraction based on depth kernel mapping and other classification methods, it can be found that the feature extraction effect of deep multicore mapping is the best. Compared with the common single kernel mapping algorithm, the accuracy has been greatly improved, which shows that the depth structure can improve the feature extraction performance of the kernel mapping algorithm. In this paper, we achieve the better performance with the depth structure of the foundation, the existing loss function, and the number and category of the kernel function. So, it is effective for the mapping structure with the parameter optimization algorithm.

Secondly, we have the evaluations on the UCI dataset, to test the performance with different depths and depth widths.7 public data sets from the UCI benchmark are chosen for the binary and multiple classification tasks. For the comparisons, we have implemented other algorithms, including EasyMKL [22], SimpleMKL [23], SM1MKL [24], and L2MKL [25, 26]. These algorithms with the optimal parameter in the references and 5-fold cross validation method are used to choose the parameter for our algorithm. Some results are shown in Table 5.

From a statistical point of view, there are 8 frameworks that are combined by 3-width in our method regardless of the depth. This work proves that a wider width can improve

Table 5: The classification performance by comparing with existing MKL algorithm.

Dataset	EasyMKL[1][22]	SimpleMKL[2][23]	SM1MKL[3][24]	L_2MKL[4][25, 26]	Our method (D,W)
Australian	84.63 (3)	84.30 (4)	85.79 (1)	84.05 (5)	85.50 (3,3) (2)
Breast	69.06 (5)	69.78 (4)	70.50 (1)	70.50 (1)	70.50 (1,1) (1)
Cleve	81.08 (4)	80.40 (5)	81.75 (2)	81.75 (2)	84.45 (3,3) (1)
Cloud	69.78 (4)	75.43 (2)	73.29 (3)	61.20 (5)	99.80 (1,3) (1)
Diabetes	74.21 (5)	74.73 (4)	76.56 (2)	76.56 (2)	77.34 (3,3) (1)
Pima	78.38 (1)	78.12 (3)	77.34 (4)	77.34 (4)	78.38 (3,3) (1)
Sonar	86.66 (1)	76.69 (3)	72.38 (5)	74.28 (4)	84.76 (1,3) (2)
Average rank	3.27	2.14	2.57	3.29	1.28

[1]Avaliable at https://github.com/okbalefthanded/EasyMKL. [2]Avaliable at https://github.com/maxis1718/SimpleMKL. [3]Avaliable at https://sites.google.com/site/xinxingxu666/. [4]The implementation is obtained in footnote 4.

Table 6: Performance evaluation on different kernel learnings.

Datasets	Indian pines dataset		Pavia University dataset	
Methods	OA (%)	KC (%)	OA (%)	KC (%)
RBF-SKL [27]	73.23	63.73	75.43	68.76
POLY-SKL [27]	75.77	66.28	78.08	72.07
Mahal-RBF-SKL [31]	76.92	67.79	76.26	69.87
Mahal-Poly-SKL [31]	77.64	68.62	79.07	73.24
SK-CV-RBF-SKL [30]	67.03	64.23	75.71	69.14
SK-POLY-SKL [30]	69.37	66.96	77.62	71.37
EasyMKL [22]	68.13	65.53	76.82	72.14
SimpleMKL [23]	69.22	66.78	73.56	66.76
SM1MKL [24]	77.34	69.62	79.98	74.34
L2MKL [25, 26]	77.37	74.85	79.67	75.87
NMF-MKL [32]	67.48	64.81	71.57	64.42
KNMF-MKL [32]	68.22	65.63	72.80	65.81
MKL1 [28]	74.23	69.63	78.69	72.11
MKL2 [29]	76.07	72.84	79.24	72.92
QMKL	79.28	75.85	80.96	75.34
Proposed deep	81.53	78.34	83.14	78.76

the performance of the framework when combined with the appropriate depth. In summary, DWS-MKL outperforms the MKL algorithm and improves the classification performance of the algorithm.

3.2. Performance on the Hyperspectral Data Classification. In this paper, we have the experiments on the Indian Pines dataset, Pavia University data, and these datasets are the same to our previous work [36] for the comparisons. The Kappa coefficient (KC) and Overall accuracy (OA) are used to evaluate the algorithms. 10 basis kernels are used to learning, and [0.01, 2] is the scale parameter σ. On the classifier settings, a standard multiclass SVM is used to classification, and the optimal parameters are chosen with the crossvalidation method. The original spectral features are reduced with the PCA algorithm. From the comparisons, the 15 methods also are implemented on two hyperspectral databases, including RBF-SKL [27], POLY-SKL [27], Mahal-RBF-SKL [31], Mahal-Poly-SKL [31], SK-CV -RBF-SKL [30], SK-POLY-SKL [30], EasyMKL [22], SimpleMKL [23], SM1MKL [24], L2MKL [25, 26], NMF-MKL [32], KNMF-MKL [32], Euclidean-MKL1 [28], Euclidean-MKL2 [29], and QMKL [36]. The detail algorithm description can refer to work [36].

As shown in Table 6, compared with other kernel learning methods, the proposed method achieves the highest recognition rate. The results show that the deep structure can effectively improve the discrimination ability of data. In the comparison of the algorithms, including single kernel learning, multicore learning, and variable structure kernel learning, the variable structure deep kernel learning proposed in this paper achieves the best results, which shows that the algorithm is feasible. The database used in the experiment is actually collected database, which has good practicability. It can be shown that the method proposed in this paper can be used in practical applications.

3.3. Discussion. From the above experimental results, we can see that the framework of deep kernel learning is effective and can effectively solve the problem of classification. The results on hyperspectral data sets show that the proposed variable depth kernel learning architecture is also suitable for practical data classification problems and can improve the accuracy of classification. In addition, this method has the characteristics of small amount of calculation and small scale of network parameters, which is especially suitable for application scenarios with limited computing resources. In the paper, three kinds of data have their own characteristics, so the UCI data, radar data, and hyperspectral data have the representative meaning for the practical application. So, in the practical applications, the algorithms should have the corresponding performances to the results. Moreover, the proposed algorithm can be implemented in the limited computing ability of platform, for example, DSP and FPGA. So, it can be used for the many computing practical application areas.

4. Conclusion

In this paper, we propose the novel deep kernel learning for the hyperspectral data classification, and the solve method with the gradient projection is used to optimize the error of the classifier. The Depth and Width Changeable Deep Kernel Learning-based hyperspectral sensing data analysis

algorithm is proposed. Some experiments show that the proposed method has its advantages on the hyperspectral data classification. The experiments are implemented to testified the feasibility and performance of the algorithms on the hyperspectral data analysis. In the future research, the computing efficiency should be improved, because it can be used in the video recognition and image processing with small computing ability.

Data Availability

We have not used specific data from other sources for the simulations of the results. The two popular hyperspectral datasets in this paper, Indian Pines dataset and Pavia University data, are free download with the website: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes. The proposed algorithm is implemented in python.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

We would like to thank Dr. Li Li, Prof. Junbao Li, and Prof. Yanfeng Gu to provide the kernel-based learning programs of their papers for comparison in the experiments. This work is supported by the National Science Foundation of China under Grant No. 61871142.

References

[1] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2017.

[2] G. Wu, M. Kim, Q. Wang, B. C. Munsell, and D. Shen, "Scalable high-performance image registration framework by unsupervised deep feature representations learning," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1505–1516, 2016.

[3] M. Jiu and H. Sahbi, "Nonlinear deep kernel learning for image annotation," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1820–1832, 2017.

[4] C. Wang, J. Shi, Q. Zhang, and S. Ying, "Histopathological image classification with bilinear convolutional neural networks," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4050–4053, South Korea, 2017.

[5] W. Li, X. Qian, and J. Ji, "Noise-tolerant deep learning for histopathological image segmentation," in *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 3075–3079, Beijing, China, 2017.

[6] N. Kumar, A. V. Rajwade, S. Chandran, and S. P. Awate, "Kernel generalized gaussian and robust statistical learning for abnormality detection in medical images," in *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 4157–4161, Beijing, China, 2017.

[7] J. Zhuang, I. W. Tsang, and S. C. H. Hoi, "Two-layer multiple kernel learning," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 909–917, Ft. Lauderdale, FL, USA, April 2011.

[8] E. V. Strobl and S. Visweswaran, "Deep multiple kernel learning," in *2013 12th International Conference on Machine Learning and Applications*, pp. 414–417, Miami, FL, USA, 2013.

[9] I. Rebai, Y. BenAyed, and W. Mahdi, "Deep multilayer multiple kernel learning," *Neural Computing and Applications*, vol. 27, no. 8, pp. 2305–2314, 2016.

[10] I. Rebai, Y. BenAyed, and W. Mahdi, "Deep architecture using multi-kernel learning and multi-classifier methods," in *2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*, pp. 1–6, Marrakech, Morocco, 2015.

[11] I. Rebai, Y. BenAyed, and W. Mahdi, "Deep kernel-SVM network," in *2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 1955–1960, Vancouver, BC, Canada, 2016.

[12] L. Le, J. Hao, Y. Xie, and J. Priestley, "Deep kernel: learning kernel function from data using deep neural network," in *Proceedings of the 3rd IEEE/ACM international conference on big data computing, applications and technologies*, pp. 1–7, New York, NY, USA, 2016.

[13] P. P. Brahma, D. Wu, and Y. She, "Why deep learning works: a manifold disentanglement perspective," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 10, pp. 1997–2008, 2016.

[14] W. K. Wong and M. Sun, "Deep learning regularized fisher mappings," *IEEE Transactions on Neural Networks*, vol. 22, no. 10, pp. 1668–1675, 2011.

[15] S. Wang, Y. Jiang, F. L. Chung, and P. Qian, "Feedforward kernel neural networks, generalized least learning machine, and its deep learning with application to image classification," *Applied Soft Computing*, vol. 37, pp. 125–141, 2015.

[16] K. Ni and R. Prenger, "Learning features in deep architectures with unsupervised kernel k-means," in *2013 IEEE Global Conference on Signal and Information Processing*, pp. 981–984, Austin, TX, USA, 2013.

[17] V. Harikrishnan and A. Paulose, "A deep learning kernel K-means method for change detection in satellite," *International Journal of Science and Research*, vol. 3, no. 12, pp. 1220–1226, 2014.

[18] C. Jose, P. Goyal, P. Aggrwal, and M. Varma, "Local deep kernel learning for efficient non-linear SVM prediction," in *Proceedings of the 30th International Conference on Machine Learning*, vol. 28, pp. 486–494, Atlanta, Georgia, USA, June 2013.

[19] W. Li, F. Abtahi, and Z. Zhu, "A deep feature based multi-kernel learning approach for video emotion recognition," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pp. 483–490, New York, NY, USA, 2015.

[20] X. Liu, J. Shi, and Q. Zhang, "Tumor classification by deep polynomial network and multiple kernel learning on small ultrasound image dataset," in *Proceedings of 6th International Workshop on Machine Learning in Medical Imaging*, vol. 9352, pp. 313–320, Munich, Germany, October 2015.

[21] S. Poria, E. Cambria, and A. Gelbukh, "Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 17–21, Lisbon, Portugal, 2015.

[22] F. Aiolli and M. Donini, "EasyMKL: a scalable multiple kernel learning algorithm," *Neurocomputing*, vol. 169, pp. 215–224, 2015.

[23] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "SimpleMKL," *Journal of Machine Learning Research*, vol. 9, pp. 2491–2521, 2008.

[24] X. Xu, I. W. Tsang, and D. Xu, "Soft margin multiple kernel learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 5, pp. 749–761, 2013.

[25] C. Cortes, M. Mohri, and A. Rostamizadeh, "L2 regularization for learning kernels," in *Proceedings of the Twenty-fifth Conference on Uncertainty in Artificial intelligence*, pp. 109–116, Montreal Quebec, Canada, June 2009.

[26] M. Kloft, U. Brefeld, S. Sonnenburg, and A. Zien, "lp-norm multiple kernel learning," *Journal of Machine Learning Research*, vol. 12, pp. 953–997, 2011.

[27] D. Wang, D. S. Yeung, and E. C. C. Tsang, "Weighted mahalanobis distance kernels for support vector machines," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1453–1462, 2007.

[28] M. Gonen and E. Alpaydin, "Multiple kernel learning algorithms," *Journal of Machine Learning Research*, vol. 12, pp. 2211–2268, 2011.

[29] Y. Gu, C. Wang, D. You, Y. Zhang, S. Wang, and Y. Zhang, "Representative multiple kernel learning for classification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7, pp. 2852–2865, 2012.

[30] L. Li, C. Sun, L. Lin, J. Li, and S. Jiang, "A dual-layer supervised Mahalanobis kernel for the classification of hyperspectral images," *Neurocomputing*, vol. 214, pp. 430–444, 2016.

[31] L. Li, C. Sun, L. Lin, J. Li, and S. Jiang, "A Mahalanobis metric learning-based polynomial kernel for classification of hyperspectral images," *Neural Computing and Applications*, vol. 29, no. 4, pp. 1103–1113, 2018.

[32] Y. Gu, Q. Wang, H. Wang, D. You, and Y. Zhang, "Multiple kernel learning via low-rank nonnegative matrix factorization for classification of hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2739–2751, 2015.

[33] J. E. Liu and F. P. An, "Image classification algorithm based on deep learning-kernel function," *Scientific Programming*, vol. 2020, Article ID 7607612, 14 pages, 2020.

[34] T. Wang, L. Zhang, and W. Hu, "Bridging deep and multiple kernel learning: a review," *Information Fusion*, vol. 67, pp. 3–13, 2021.

[35] T. Wang, H. Su, and J. Li, "DWS-MKL: depth-width-scaling multiple kernel learning for data classification," *Neurocomputing*, vol. 411, pp. 455–467, 2020.

[36] J. Liu and Y. Qiao, "Quasiconformal mapping kernel machine learning-based intelligent hyperspectral data classification for internet information retrieval," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8873366, 14 pages, 2020.

[37] X. Chen, X. Peng, R. Duan, and J. Li, "Deep kernel learning method for SAR image target recognition," *Review of Scientific Instruments*, vol. 88, no. 10, p. 104706, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6645961, 9 pages
https://doi.org/10.1155/2021/6645961

Research Article

An E-Commerce Economic Dynamic Data Evaluation Model Based on Multiuser Demand Constraints

Chenyuan Wang

School of Entrepreneurship, Yiwu Industrial & Commercial College, Yiwu, Zhejiang 322000, China

Correspondence should be addressed to Chenyuan Wang; wangchenyuan88@ywicc.edu.cn

Received 3 January 2021; Revised 29 January 2021; Accepted 5 February 2021; Published 18 February 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Chenyuan 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.

Forecasting the future earnings of listed companies based on multiuser constraints is the focus of investors, securities dealers, creditors, and management. Some empirical studies at home and abroad indicate that the financial reports issued by listed companies regularly contain information about future changes in earnings. On this basis, this article uses the Bayesian dynamic regression model to predict the changes in the future earnings of listed companies and compares the results with traditional analysis models. Through case analysis, it can be seen that the prediction effect of the Bayesian dynamic regression model is generally better than that of the traditional regression model. The Bayesian model can better predict the results, and through the prediction results, it can also establish an evolutionary game model of the industrial innovation replication dynamic system, which can assist enterprises in making profit decisions.

1. Introduction

The sustained profitability of a company in the future is a common concern among investors, securities firms, and creditors because only a company with sustained profitability can bring sustained dividend returns to investors. The most important and direct information that ordinary investors can get about the operating conditions of listed companies comes from the financial reports that listed companies publish regularly. Therefore, from the standpoint of ordinary investors, the following questions are worth studying. Do these financial reports contain information about the company's future profitability? If the answer is yes, can we find a better way to use this information to achieve better forecasting results [1]?

The results of scholars at home and abroad have answered the first question in a positive way. For example, we select 68 variables from the financial reports of listed companies in the United States and build a Logit regression model around the probability of earnings growth in the next year. We find that Pr is positively correlated with future earnings changes. Or nine of the basic signals commonly used by analysts are selected as predictors, and the linear regression model of cross-sectional data is used to examine their ability to predict future earnings. The analysis results also show that this prediction ability exists [2].

In China, there are also research studies on the usefulness of surplus figures in China's stock market and so on. It is generally believed that the company's current profit information is contained in the stock market price, but for long-term investors (holding more than one year of stock), it is important to look ahead. In addition to the winning figures, financial reports contain a large number of data describing the company's operating conditions. With sufficient credibility, these data should provide investors with information about the company's future profitability. This paper analyzes the annual reports of listed companies in the A-share market by using foreign scholars' methods and draws a similar conclusion: the information provided by the financial reports of listed companies can be used to predict the future profit changes of listed companies [3]. However, whether we can find a better way to use these information to obtain better prediction results is the problem to be studied in this paper. The ARMA model, exponential weighted model, and Bayesian dynamic regression model are commonly used in the economic field [4]. Compared with the mature stock markets of Western developed countries, China's stock market

(represented by the A-share market) is an emerging and developing capital market. Its main characteristics are as follows: short history, large changes (expanding scale, changing policies), and less information disclosure (except annual reports, the history of the Chinese stock market is very short, and there is no quarterly report so far) [5], not standard enough. Based on the above reasons, we believe that the ARMA model and the exponential weighted model will not have very good prediction results because both models require a long history and more stable data. The Bayesian dynamic regression model has the following advantages. First, the dynamic model assumes that the model parameters change with time, which is in line with the actual situation of the stock market. Secondly, based on the Bayesian view, the main feature of the dynamic model is to modify the distribution of model parameters according to all the past information (i.e., prior information) and the current data, obtain a posterior distribution, and then predict the future from the posterior distribution. As this correction is carried out over time, it can better simulate the stock market movement, so it is expected to get better forecasting results [6].

The rest of this paper is organized as follows. Section 2 discusses the Bayesian estimation model, followed by the empirical analysis of the Bayesian estimation model in Section 3. To the domestic city, a certain financial revenue forecast analysis is discussed in Section 4. Section 5 concludes the paper with a summary and future research directions.

2. Bias Model

The main formulas of the Bayesian estimation model are Formulas (1)–(3), Formula (1) is the observation equation, Formula (2) is the equation of state, and Formula (3) is the initial prior state. y_t is the observed value vector of the predicted variable at time t, F_t is the observed value matrix of the predicted variable at time t, t is the parameter vector of the linear regression model at time t, v_t is the normal zero mean error term, and v_t is the variance [7, 8]. Different from the traditional static linear regression model, the model parameter vector theta t is regarded as a random variable in the dynamic linear regression model, which changes with time. In the equation of state, it is assumed that theta t has a variation W_t with respect to theta $t-1$, and its distribution is normal, zero mean, and variance W_t. It is assumed that the observation and state error sequences ($\{V_t\}$ and $\{W_t\}$) are independent of each other and that the T variables are independent of each other at different times within each sequence. When $W_t = 0$, the state equation degenerates to an identity, and the dynamic linear regression model degenerates to a static linear regression model [9–11]:

$$y_t = F_t\theta_t + v_t, \quad v_t \sim N[0, V_t], \tag{1}$$

$$\theta_t = \theta_{t-1} + w_r, \quad w_r \sim N[0, W_t], \tag{2}$$

$$(\theta_0 \mid D_0) \sim N[m_0, C_0]. \tag{3}$$

The information set D_t at any time t is

$$D_t = \{y_t, F_t, D_{t-1}\}, \tag{4}$$

where D_{t-1} is the set of information at time $t-1$, and $\{Y_t, F_t\}$ is the current data. Firstly, the prior distribution D_{t-1} is obtained, then the predictive distribution Y_t/D_{t-1} is obtained, then the posterior distribution D_t is obtained, and finally, the predictive distribution Y_{t+1}/D_{t-1} is obtained. The above process continues over time [12].

Push the correction. When the observed variance V_t is known, the recurrence formula is as follows [13–15]:

The posterior distribution of $T-1$ moments:

$$(\text{theta}_{T-1}\, xD_{t-1}) \sim N[M_{T-1}, C_{t-1}]. \tag{5}$$

The a priori distribution of T moments:

$$(\text{theta}\, t \mid D_t - 1) \sim N[a_t, R_t]. \tag{6}$$

The $T-1$ moment prediction distribution:

$$(Y_T, D_{t-1}) \sim N[f_t, Q_t]. \tag{7}$$

The posterior distribution of T moments:

$$(\text{theta}\, t \mid D_t) \sim N[M_T, C_t], \tag{8}$$

where

$$\left\{\begin{array}{c} A_1 = M_T - 1 \\ R_t = C_t - 1 + W_T \\ F_t = F_t{}' a_t \\ Q_t = F_t{}' R_t F_t + V_T \\ E_t = Y_t - f_t \\ A_t = \dfrac{F_t R_t}{Q_t} \\ M_t = mT + 1 + A\, tet \\ C_t = R_t - A_t A_t{}' Q_t \end{array}\right\}, \tag{9}$$

where W_t is difficult to set, so R_t is usually determined by the discount method.

$$R_t - 1 = \text{delta}\, C_t - 11, \tag{10}$$

where delta is a discount factor, $0 < \text{delta} < 1$. That is, the prior accuracy is equal to the last posterior precision of the discount. By selecting different values, the best prediction model can be selected. When the observed variance V_t is unknown, a certain prior distribution should be assumed for V_t, and the model parameters are corrected recursively. The accurate distribution of the model parameters is t distribution [16].

Table 1: The forecasting variable.

Number	Variable	Formula
1	Stock	Inventory difference − difference in main business income
2	Accounts receivable	Data net accounts receivable − data main business income
3	Maori	Data main business revenue − data main business profit
4	Consumption and management costs	Data consumption and management expenses − data main business income
5	Allowance for bad debts	Data accounts receivable − data for bad debts
6	Real tax rate	Current pretax income ∗ (last year's effective tax rate − current effective tax rate)
7	Turnover of total assets	Current main business income ∗ 2/(fixed assets plus fixed assets in the previous year)
8	Asset liability ratio	Changes in pretax earnings ∗ (1 − effective tax rate in the previous year)
9	Turnover of fixed assets	Log (total assets)
10	Data PTE	Inventory difference − difference in main business income

3. Empirical Results

3.1. Prediction Models, Samples, and Variables. The Bayesian dynamic regression prediction model we use is described in the previous section [17]:

$$Y_T = (Y_{1T}, \cdots, y_{nt})'. \tag{11}$$

Subscript 1 to n is the company number, and t is the year [17]:

$$Y_{it} = \text{delta EPS } it + 1 = \text{EPS } T + 1 - \text{EPS } T, \tag{12}$$

where Y_{it} is the observed value matrix of the forecasting variable at time t and theta t is the regression coefficient vector. We should not only directly examine the prediction effect of the Bayesian dynamic model but also compare it with that of the static regression model. The basic data and literature we used are from the annual financial reports of listed companies in Shanghai and Shenzhen A-share markets from 1992 to 1999. In order to predict $T + 1$ years of income, we need to use data from $T - 2$ to $T + 2$ years [18]. The predictor variables are shown in Table 1.

Using the financial reported data for the past eight years, we can predict only four cycles, and the corresponding sample size is shown in Table 2 [19].

3.2. Analysis Results. When the Bayesian dynamic regression model is used to predict, choosing appropriate discount factors can achieve better prediction results. We try different values of delta in fitting and forecasting and find that the best result is when the value of delta equals 0.6 in most cases. We use the percentage of the actual value falling into the prediction interval to measure the prediction effect [20]. The confidence of the prediction interval is 85%, 90%, and 95%, respectively. Table 3 shows the accuracy of prediction at the time of delta = 0.6, from which we can see that the prediction effect of the Bayesian dynamic regression model is quite good. That is to say, the accuracy of the actual forecast is higher than the confidence in most cases; only in some cases is it equivalent to the confidence [21].

Table 2: The corresponding sample size (million).

Year	1995–1996	1996–1997	1997–1998	1998–1999
Shenzhen	129	207	345	365
Shanghai	132	231	356	383
Sum	262	435	642	744

Table 3: The accuracy of prediction at the time of delta = 0.6.

Year	1995–1996	1996–1997	1997–1998	1998–1999
Accuracy	98%	93%	93%	93%
95	98%	91%	91%	92%
90	98%	94%	93%	93%
85	98%	95%	95%	96%

Table 4: The results of the comparison.

Year	Index	Static regression model	Bias regression model
Predict 1995	MAD	18.4%	18.4%
1996	MSE	18.9%	18.9%
Predict 1996	MAD	20.3%	16.1%
1997	MSE	51.9%	7.4%
Predict 1997	MAD	21.5%	18.0%
1998	MSE	38.2%	10.9%
Predict 1998	MAD	19.9%	21.0%
1999	MSE	13.4%	19.3%

We use two evaluation criteria to compare the prediction results of the static regression model and the Bayesian dynamic model. One is the mean absolute error MAD = 6. The second is the mean square error MSE = 6. EN is the difference between the predicted value and the actual observed value. The smaller the value, the better the prediction effect. The results of the comparison are shown in Table 4 [22].

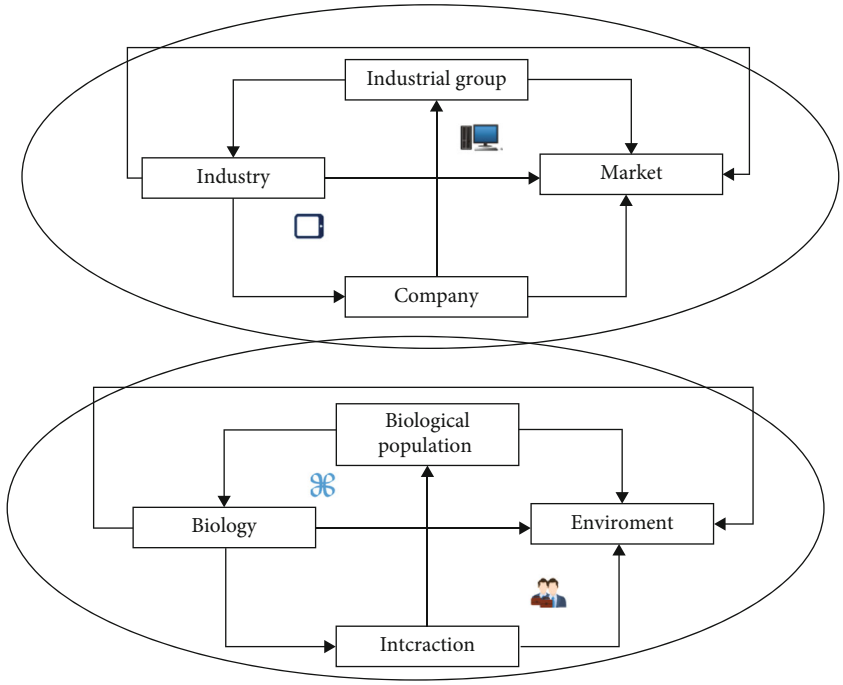


Figure 1: Certain degree of comparability between the industry and biological groups.

It can be seen that in the first forecast year, the prediction effect of the Bayesian dynamic regression model and the static regression model is equivalent. This is because, at the beginning, we have no prior information about the past, so the Bayesian dynamic regression model cannot provide better prediction results. In the next two years, the prediction effect of the Bayesian dynamic regression model is obviously better than that of the static regression model. But when the 1998 data are used to predict the 1999 profit, the prediction effect of the Bayesian dynamic regression model is not as good as that of the static regression model. The reason may be that the deterioration of the macroeconomic situation has led to a greater jump in the performance of listed companies compared with the past. Because the history of the Chinese stock market is very short, the accumulated data is very few, and our research funds and scale are limited, the current research results show the superiority of the Bayesian dynamic regression model in predicting the company's profitability to a certain extent but cannot fully show its full effectiveness [23].

4. Enterprise Innovation and Profit Constrained by Multiuser Demand

4.1. The Evolution of the Industrial Innovation Biological Population. In his 1974 monograph *The Economics of Industrial Innovation*, the British economist Freeman expounded the concept of industrial innovation systematically for the first time. Freeman believes that "industrial innovation" is "innovation" in the "industry." This definition is also the consensus of Western scholars and is basically consistent with the definition of industrial innovation in *The Handbook of Industrial Innovation*. Domestic industrial innovation research is a little late but develops rapidly. Based on the definition of industrial innovation in domestic and foreign works of literature, industrial innovation is defined as follows: industrial innovation refers to the process of technological innovation, product innovation, market innovation, or combination innovation carried out by a single enterprise or several large-scale enterprises jointly to become the common innovation of the whole industry through diffusion or through the enterprise.

The evolutionary game theory studies the repeated games of players in dynamic systems and the equilibrium state of the system. It is named after biological evolution, but it is widely used in economic and social fields. Building an innovative country has become an important strategic decision of our country. The industrial innovation system is a bridge linking the national innovation system and the enterprise innovation system and has important significance.

There are many similarities between the process of strategic choice of enterprises and the process of environmental adaptation of biological individuals. (1) The enterprise strategy is not fixed but is constantly changing in the process of interaction with the market. (2) The corporate strategy is genetically identical to biological genes. This kind of heredity shows that the strategy of enterprises is stable and sustainable in a period of time. (3) Enterprises, like organisms, are bounded rational individuals. Limited rationality of organisms is manifested in their choice behavior of seeking advantages and avoiding disadvantages and learning imitation behavior to a certain extent. The limited rationality of enterprises is rooted in the rapid change of the market environment. The completely correct and forever correct enterprise strategy is impossible to exist; the choice of the enterprise strategy is also a process of trial-and-error learning.

The industry is a collection of enterprises with similar attributes, and there is a certain degree of comparability

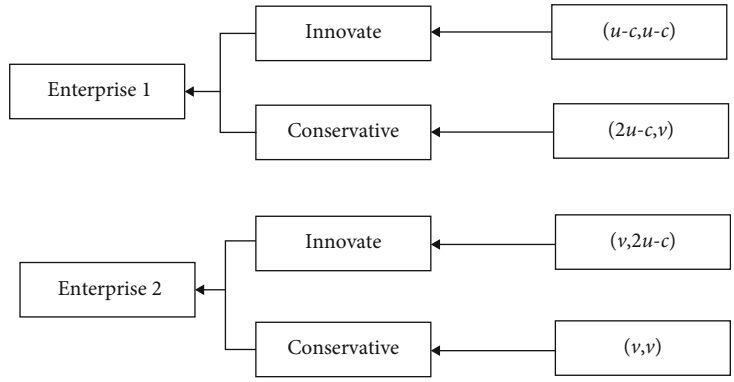


Figure 2: The matrix representation of the random pairing game.

between the industry and biological groups as shown in Figure 1. Therefore, the innovation process of the industry is similar to the evolution process of the biological population. This is the core idea of this article. The construction of the mathematical model and the analysis of the evolutionarily stable state are the core parts of this paper, which involve the evolutionary game theory, replication dynamic mechanism, dynamic system, and other theoretical knowledge.

4.2. Characteristics of Industrial Innovation under Demand Constraints. If there are demand constraints in the industry, the benefits of the enterprise innovation strategy will decrease with the increase of innovative strategy users. Taking product innovation as an example, suppose an enterprise launches a new product (product innovation), which can better meet the needs of consumers, and consumers rush to buy. The enterprise can set a higher price for the product and get excess profits. When other enterprises see that the product is profitable, they will turn to the production of the product so that the supply of the product substantially increased, and the demand for the product is limited; that is, there is a demand constraint; when the product supply exceeds demand, the price will fall, and the profits of enterprises will also be reduced. The greater the supply of products, the greater the price declines.

This phenomenon is very common in the mobile phone industry. The world's first mobile phone, Motorola's Motorola DynaTAC 8000X, was priced at $3,995. At such a high price, mobile phone manufacturers can get the excess profits brought by innovation. As the technology spreads or other companies innovate and acquire it, the number of mobile phone manufacturers increases, leading to a sharp drop in the price of mobile phones, resulting in a decrease in innovative profits. Innovations in products such as color-screen phones and camera phones have also experienced the same decline in profit margins.

4.3. The Basic Assumptions of the Model. The profit of the innovation strategy is greater than that of the conservative strategy. If the innovation strategy cannot do this, we can hardly call it innovation; gamers can only use the pure strategy; this assumption is made because the implementation of the hybrid strategy is often very complex; enterprises have the same size. This assumption is to facilitate research; the group size is large; that is, the number of players is n to infinity.

Suppose an industry is made up of n enterprises. In such an industry, 22 enterprises conduct repeated games with random pairing. Each enterprise can adopt two strategies (hereinafter referred to as strategies), that is, innovative management s1 and conservative management s2. If two enterprises adopt the conservative management strategy at the same time, they share the benefits of $2V(V > 0)$ equally. If two firms adopt an innovative strategy at the same time, they share the benefits of $2U(2U > 2V > 0)$ equally. At the same time, each firm must bear the cost of innovative operation (e.g., the cost of purchasing new equipment because of the original mechanical equipment eliminated by innovation). If an enterprise adopts the innovation strategy [8, 9], the other enterprise adopts the conservative strategy. The enterprise that adopts the innovation strategy obtains $2u - c$. The conservative enterprise still obtains v. The matrix representation of the random pairing game is shown in Figure 2.

It is assumed that the proportion of N in the total number of enterprises in the industry is X after adopting the innovation strategy s1. Because $N >$ infinity, X can be regarded as a continuous variable. Through calculation, we get

$$\frac{Dx}{dt} = x(\mathrm{PI}_1 - \mathrm{PI})(\mathrm{PI}). \tag{13}$$

To represent the average benefit of the strategy, the substitution ofPI_1, PI_2, and $\mathrm{PI}(1)$ is

$$\frac{Dx}{dt} = x(\mathrm{PI}_1 - \mathrm{PI})'' = x\left[''(2u - c - V) - ux\right]. \tag{14}$$

Formula (14) replicates the dynamic equation of $d_y n, i, e$. The point of $d_t d_x = 0$ in the dynamic equation, where $x = 0$, $1, 2u - c - vu$, and the evolutionarily stable state and the evolutionarily stable strategy of the "dynamic system of industrial innovation replication" are obtained.

4.4. Dynamic Equilibrium Analysis of Industrial Innovation Replication under Demand Constraints. It has been pointed out in the summary that when demand constraints exist, the benefits of adopting innovative strategies in industrial innovation will decrease with the increase of innovative adopters.

Under the analysis framework of Section 2, the mathematical expression of the requirement constraint $U(x)$: $(0 < x < 1)$ is a continuously decreasing function in the domain $[0, 1]$. By substituting $U(x)$ into the formula, the dynamic equation of replication under demand constraints is obtained:

$$\frac{Dx}{dt} = x(1-x) * [2U(x) - c - v] - U(x)\, x+ = F(x). \quad (15)$$

The fixed point of the equation is all x satisfying $F(x)$, which is expressed as

$$\begin{gathered} X = 0, 1, \\ X * x|2U(x) - c - v] - U(x)x = 0+. \end{gathered} \quad (16)$$

The following is the obtained function through analysis:

$$G(x) = [2U(x) - c - v] - U(x), \quad (17)$$

where x is continuously monotonically reduced in the interval $[0, 1]$, and the maximum and minimum values are

$$\left\{ \begin{array}{l} \text{Max } G(x) = 2U(0) - c - V \\ \text{Min } G(x) = U(1) - c - v \end{array} \right\}. \quad (18)$$

At the same time, according to the properties of continuous monotone functions, if

$$G(x) = [2U(x) - c - v] - U(x)x = 0, \quad (19)$$

there is a solution in the interval $[0, 1]$, and there is only one solution.

It is assumed that $X-$ is a solution of $[2U(x) - c - v] - U(x)x = 0$, and 0 is less than $x_{-1} = 1$, that is,

$$\begin{gathered} [2U(x-) - c - v] - U(x-)x- = 0, \\ \left\{ \begin{array}{l} \dfrac{Dx}{dt} = x(1-x) * [2U(x) - c - v] - U(x)x+ < 0 \\ \dfrac{Dx}{dt} = x(1-x) * [2U(x) - c - v] - U(x)x+ > 0 \end{array} \right\}. \end{gathered} \quad (20)$$

Therefore, we can get the phase diagram of the $x-$ point replication dynamic system, as shown in Figure 3.

It can be seen from Figure 3 that if the perturbation occurs to the right of $x-$, the system will automatically recover to the $x-$ point and vice versa. Therefore, $x-$ is a gradual stable state of replication of dynamic systems, that is, evolutionary stability. The evolutionary stability of the "industrial innovation replication dynamic system" under demand constraints is discussed below.

When $x-$ is $(0, 1)$, according to the above analysis, $x-$ $(0, 1)$ is the stable state of the replication dynamic system.

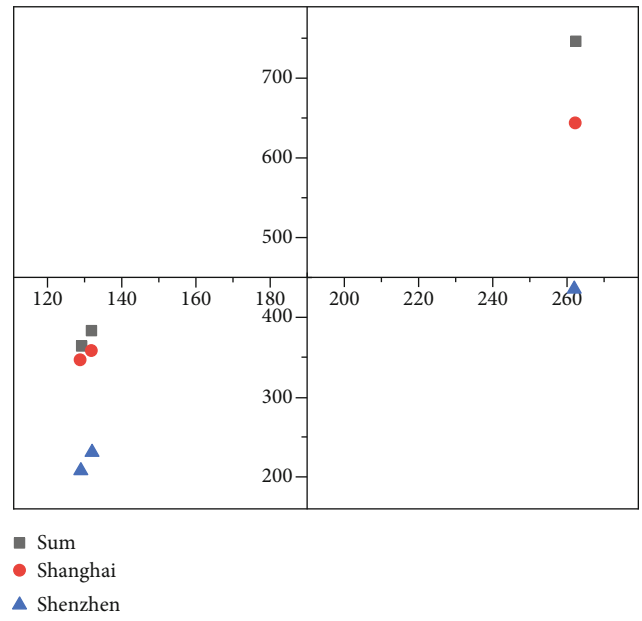


Figure 3: The phase diagram of the $x-$ point replication dynamic system.

When $x- = 0$ or 1 is based on the above analysis, $x- = 0$ or 1 replicates the steady state of the dynamic system.

There is no solution when

$$G(x) = [2U(x) - c - v] - U(x)x = 0 \text{ is in } [0, 1]. \quad (21)$$

The following two parts are discussed:

$$\left\{ \begin{array}{c} \max G(x) = 2U(0) - c - v < 0 \\ F(x) = x(1-x)G(x) \\ \dfrac{Dx}{dt} = x(1-x)G(0) = 0 \\ F(0+) = \lim x \text{ to } 0+ \\ F(x) = \lim x \text{ to } 0+ \\ X(1-x)[2U(0+) - c - v] < 0 \\ F(0-) = \lim x \text{ to } 0- \\ F(x) = \lim x \text{ to } 0- \\ X(1-x)[2U(0-) - c - v] > 0 \\ \text{Also } \max G(x) = 2U(0) - c - V \\ O \min G(x) = U(1) - c - v < 0 \\ \%F(1-) = \lim x \text{ to } 1- \\ F(x) = \lim x \text{ to } 1- \\ X(1-x)[U(1-) - c - v] < 0 \\ F(1+) = \lim x \text{ to } 1+ \\ F(x) = \lim x \text{ to } 1+ \\ X(1-x)[U(1+) - c - v] > 0 \end{array} \right\}. \quad (22)$$

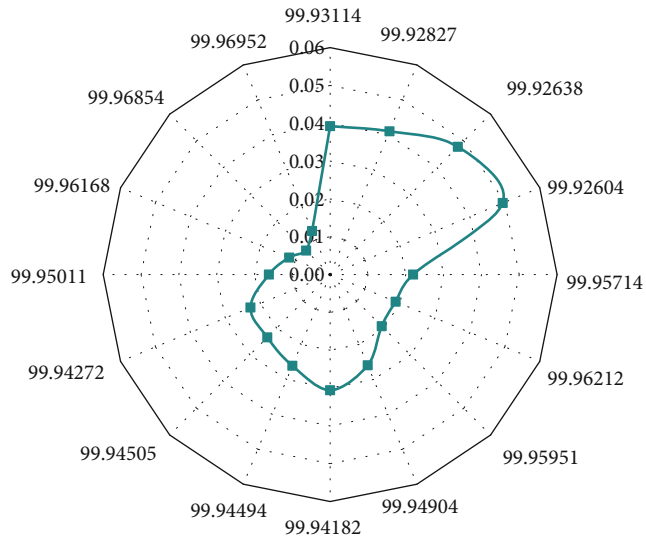


Figure 4: $x = 0$ is the equilibrium point for replication of dynamic systems.

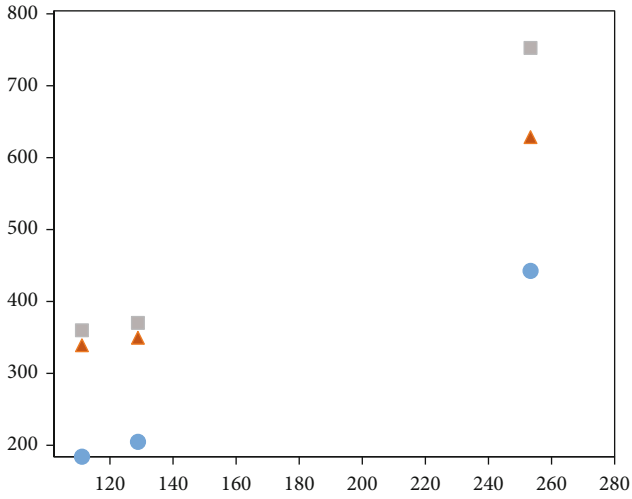


Figure 5: $x = 1$ point is the evolutionarily stable state of the replicated dynamic system.

The phase diagram of the evolutionary replication system is shown in Figure 4. Figure 4 shows that $x = 1$ is the equilibrium point for replication of dynamic systems.

The analysis procedure is similar to (1), get $F(1-) > 0$, $F(1+) < 0$, and $F(0-) < 0$. Therefore, the phase diagram of the replication system is shown in Figure 5. Figure 5 shows that the $x = 1$ point is the evolutionarily stable state of the replicated dynamic system [21–23].

In the process of simulation analysis, this paper takes a certain financial revenue of some domestic large- and medium-sized cities as the basic sample for analysis and research. Specific raw data are shown in Table 5.

It is clearly evident from Table 5 that with the passage of time, the fiscal revenue of this industry in major cities continues to increase, which also indicates that the user demand of this industry increases and the government needs to focus on this industry.

Table 5: The corresponding sample size (million).

Year	2005	2006	2007	2008
Shenzhen	136	215	345	369
Shanghai	132	231	349	378
Beijing	156	266	384	410
Guangzhou	145	236	365	385
Hangzhou	144	240	362	388
Nanjing	150	223	359	379

Table 6: Revenue forecast (million).

Year	2009	2010	2011	2012
Shenzhen	421	486	520	569
Shanghai	416	498	510	536
Beijing	409	536	540	596
Guangzhou	415	489	526	568
Hangzhou	411	485	496	529
Nanjing	408	480	526	544

The Bayesian estimation mentioned in this paper is adopted to forecast and analyze the fiscal revenue of big cities, and the results are shown in Table 6.

We compared the predicted data with the actual data from 2009 to 2012 and analyzed the error and got the error curve as shown in Figure 6.

It is clearly evident from Figure 6 that the predicted results of the Bayesian estimation model basically conform to the trend of data, which is an important data support for some conventional data prediction.

It is clearly evident from Figure 7 that the prediction results of the Bayesian estimation model are very good. The predicted results are in line with the actual results, and the errors are within the range of the theory.

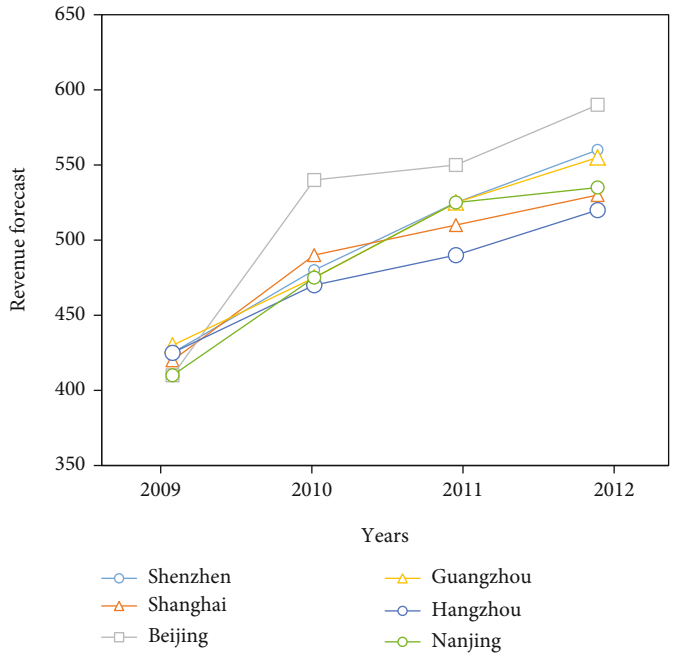


Figure 6: Revenue results.

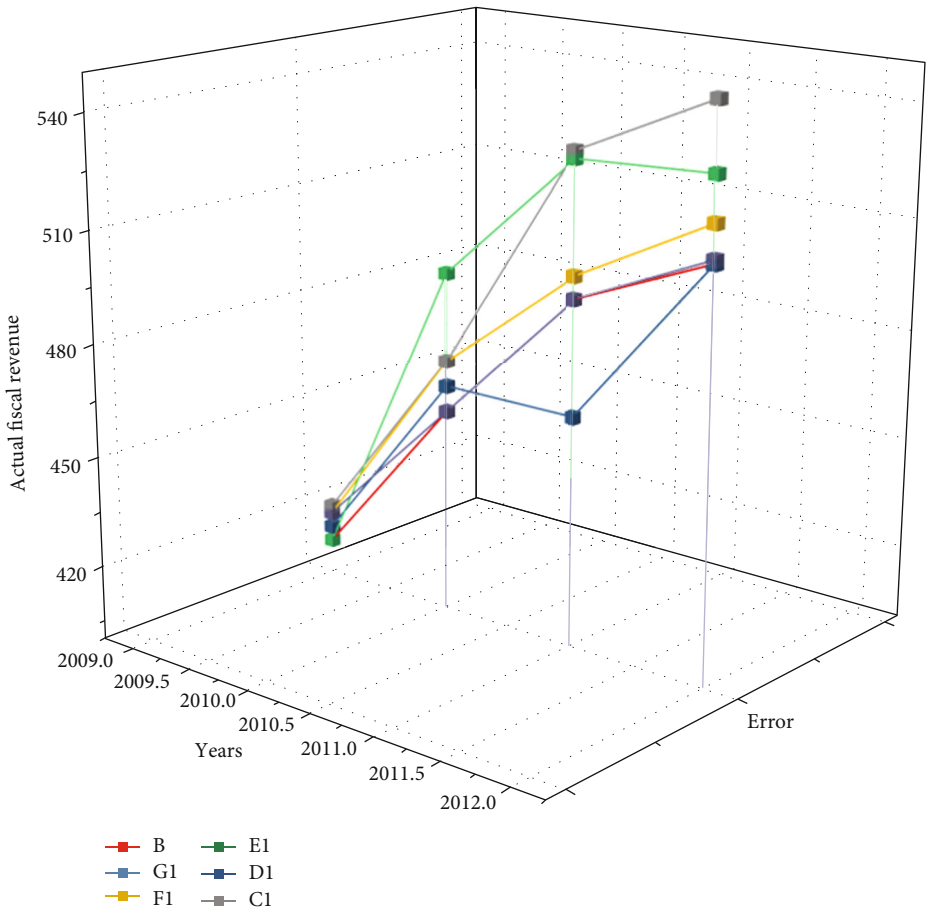


Figure 7: Comparison of error curves of financial revenue.

5. Conclusion

The forecast of the future earnings of listed companies is not only a forecast of the future development of the national economy but also a focus of investors. Through the Bayesian dynamic regression model, the information provided by the annual financial report of the listed company's A-share market is used to predict the company's future profit changes and achieve better prediction results. Compare the results with the results of the static regression model. The average absolute error and the mean square error are used as the evaluation criteria of the model. Through case analysis, it can be seen that the prediction effect of the Bayesian dynamic model is better than that of the static regression model. However, when the macroeconomic situation changes dramatically, the effect of using the Bayesian dynamic model to predict is worse than that of using the static regression model. The reason is not because of the Bayesian dynamic model itself but because we have not yet been able to introduce macroeconomic changes into the model.

Data Availability

We already included the analysis data in our manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M. Asai, "Heterogeneous asymmetric dynamic conditional correlation model with stock return and range," *Journal of Forecasting*, vol. 32, no. 5, pp. 469–480, 2013.

[2] H. M. Zhu, R. Li, and S. Li, "Modelling dynamic dependence between crude oil prices and Asia-Pacific stock market returns," *International Review of Economics & Finance*, vol. 29, no. 1, pp. 208–223, 2014.

[3] Y. Purwono, I. A. Ekaputra, and Z. A. Husodo, "Estimation of dynamic mixed hitting time model using characteristic function based moments," *Computational Economics*, vol. 51, no. 2, pp. 295–321, 2018.

[4] N. Lazzati and A. A. Menichini, "A dynamic model of firm valuation," *Social Science Electronic Publishing*, vol. 53, no. 3, pp. 499–531, 2014.

[5] S. J. Koopman and R. Lit, "A dynamic bivariate Poisson model for analysing and forecasting match results in the English Premier League," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 178, no. 1, pp. 167–186, 2015.

[6] Y. Liu, K. Sun, and Y. Liu, "A measurement-based power system model for dynamic response estimation and instability warning," *Electric Power Systems Research*, vol. 124, pp. 1–9, 2015.

[7] B. Wu, L. Zong, X. Yan, and C. Guedes Soares, "Incorporating evidential reasoning and TOPSIS into group decision-making under uncertainty for handling ship without command," *Ocean Engineering*, vol. 164, pp. 590–603, 2018.

[8] C. M. Kaul and V. Raman, "Analysis of a dynamic model for subfilter scalar dissipation rate in large eddy simulation based on the subfilter scalar variance transport equation," *Combustion Theory and Modelling*, vol. 17, no. 5, pp. 804–834, 2013.

[9] K. Barhoumi, O. Darné, and L. Ferrara, "Dynamic factor models: a review of the literature," *Working Papers*, vol. 2013, no. 2, pp. 73–107, 2014.

[10] A. Beloglazov and R. Buyya, "Managing overloaded hosts for dynamic consolidation of virtual machines in cloud data centers under quality of service constraints," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1366–1379, 2013.

[11] L. Guerrieri and M. Iacoviello, "OccBin: a toolkit for solving dynamic models with occasionally binding constraints easily," *Journal of Monetary Economics*, vol. 70, no. 1, pp. 22–38, 2015.

[12] Y. Yang, C. Zhang, and Z. Chen, "A method for state-of-charge estimation of Li-ion batteries based on multi-model switching strategy," *Energy Procedia*, vol. 75, pp. 2635–2640, 2015.

[13] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Transactions on Information and Systems*, vol. E102.D, no. 7, pp. 1374–1383, 2019.

[14] X. Zhou, J. Nakajima, and M. West, "Bayesian forecasting and portfolio decisions using dynamic dependent sparse factor models," *International Journal of Forecasting*, vol. 30, no. 4, pp. 963–980, 2014.

[15] A. Rubia and L. Sanchis-Marco, "On downside risk predictability through liquidity and trading activity: a dynamic quantile approach," *International Journal of Forecasting*, vol. 29, no. 1, pp. 202–219, 2013.

[16] V. Demiguel, F. J. Nogales, and R. Uppal, "Stock return serial dependence and out-of-sample portfolio performance," *Review of Financial Studies*, vol. 27, no. 4, pp. 1031–1073, 2014.

[17] W. K., & Automation, D. O, "Web information automatic extraction based on DOM tree and visual feature," *Computer Engineering*, vol. 39, no. 10, pp. 309–312, 2013.

[18] Z. Huang, G. Shan, J. Cheng, and J. Sun, "TRec: an efficient recommendation system for hunting passengers with deep neural networks," *Neural Computing and Applications*, vol. 31, Supplement 1, pp. 209–222, 2019.

[19] Z. Chen, Y. Zhang, C. Wu, and B. Ran, "Understanding individualization driving states via latent Dirichlet allocation model," *IEEE Intelligent Transportation Systems Magazine*, vol. 11, no. 2, pp. 41–53, 2019.

[20] Z. Huang, X. Xu, J. Ni, H. Zhu, and C. Wang, "Multimodal representation learning for recommendation in Internet of Things," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10675–10685, 2019.

[21] B. Wu, X. Yan, Y. Wang, and C. G. Soares, "An evidential reasoning-based CREAM to human reliability analysis in maritime accident process," *Risk Analysis*, vol. 37, no. 10, pp. 1936–1957, 2017.

[22] Y. Wang, E. Zio, X. Wei, D. Zhang, and B. Wu, "A resilience perspective on water transport systems: the case of Eastern Star," *International Journal of Disaster Risk Reduction*, vol. 33, pp. 343–354, 2019.

[23] Z. Huang, J. Tang, G. Shan, J. Ni, Y. Chen, and C. Wang, "An efficient passenger-hunting recommendation framework with multi-task deep learning," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7713–7721, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5565980, 8 pages
https://doi.org/10.1155/2021/5565980

Research Article

Research on Supply Chain Financial Risk Assessment Based on Blockchain and Fuzzy Neural Networks

Yijie Wang

School of Digital Commerce, Zhejiang Institute of Mechanical & Electrical Engineering, Zhejiang, Hangzhou 310053, China

Correspondence should be addressed to Yijie Wang; wangyijie@zime.edu.cn

Received 10 January 2021; Revised 25 January 2021; Accepted 2 February 2021; Published 17 February 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Yijie 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.

With the development of supply chain finance, the credit risk of small- and medium-sized financing enterprises from the perspective of supply chain finance has arisen. Risk management is one of the key tasks of the credit business of banks and other financial institutions, which runs through all aspects of the credit business before, during, and after the loan. This article combines blockchain and fuzzy neural network algorithms to study the credit risk of SME financing from the perspective of supply chain finance. This article builds a supply chain financial system through blockchain technology and integrates supply chain financial information into blocks. The fuzzy neural network algorithm is used for financial data processing and risk assessment, effectively solving and improving the risk processing level of the supply chain. Through further simulation, the application effect of blockchain and machine learning algorithms in the supply chain financial system was verified.

1. Introduction

The banking industry is one of the lifeline industries of the country's economy, and it plays a vital role in economic development, social stability, and people's livelihood. With the progress of the times, the banking industry is also facing new tasks and challenges [1, 2]. Against this background, supply chain finance is springing up and growing. Supply chain finance began in the 1970s and went through the bill discounting and trade finance phases.

Since the concept of supply chain financial risk was put forward, it has attracted the attention of many scholars [3, 4]. From the early stage of the research, scholars discussed the risks of different supply chain finance models, to the mid-term of the research, scholars conducted a comprehensive risk analysis of the entire supply chain finance through a macroperspective, and in recent years, scholars analyzed the risks of a single industry through an industry perspective [5]. Supply chain characteristics analyze the operating mechanism of the industry supply chain financial model and explore the industry supply chain financial risks. Berge defines the wait for defining supply chain finance from the perspective of SME loans [6, 7]. They believe that it is difficult to obtain loans due to the lack of good credit support for small- and medium-sized enterprises and proposed a new financing model; that is, large enterprises or financial institutions control transactions to finance small- and medium-sized enterprises that are difficult to finance [8, 9].

Aberdeen defines supply chain finance from the perspective of financial institutions such as commercial banks. He believes that the core of supply chain finance is composed of financial institutions, core enterprises, and information platforms [10, 11]. By focusing on financing and cost settlement in the supply chain, the cost of enterprises in the supply chain can be optimized and reduced [12]. Against this background, supply chain finance is springing up and growing. The base defines it as a financial institution that provides financial services to companies in the supply chain to help them manage logistics and information flow [13, 14]. Lamoureux and Evans proposed that supply chain finance is an innovation that combines financial services and technology. It strengthens logistic services through a service platform and reduces the financing costs of all participants in the supply chain [15]. Compared with traditional financial

models, supply chain finance has the advantages of low financing costs and efficient loans. Favored by financial institutions such as enterprises and banks, supply chain finance has not only improved the difficulty of financing for SMEs [16].

It also provides a new way for banks with operating difficulties to transform. However, with the development of supply chain finance, credit risk arises with the constraints of the quality of SMEs and supply chain finance management policies [17, 18]. Credit risk runs through all aspects of the credit business before, during, and after the loan. In the context of preventing financial risks, it is important to evaluate the credit risk of supply chain finance based on the perspective of commercial banks. With the development of supply chain finance, the credit risk of small and medium financing enterprises from the perspective of supply chain finance has arisen [19, 20]. Risk management runs through all aspects of the credit business before, during, and after the loan. By combining blockchain and fuzzy neural network algorithms, this paper studies the credit risk of financing for small- and medium-sized enterprises from the perspective of supply chain finance.

2. Blockchain Technology

In recent years, the research and industrial development of blockchain technology has attracted more and more attention from various countries. The blockchain has become an important competitive link in the deployment of the next generation of Internet information technology in various countries [21, 22]. As one of the most informatized countries in the world, China, we should seize the opportunity of this information industry revolution in time, increase capital and policy investment, cultivate blockchain talents in enterprises and universities, speed up the construction of blockchain standards, and grasp the right to speak and command heights in the development of the global blockchain [23]. The development of the blockchain is a process of point by line and side by side, from a single digital currency to the financial field, and then to a comprehensive range of industries, realizing "blockchain+" to empower the industry and promote the development of the industry [24]. Figure 1 shows the research on financial risk assessment of supply chain based on the blockchain.

At present, the blockchain is between 2.0 and 3.0. If you want to reach the 3.0 stage, it requires the popularization of the blockchain concept in various industries and the whole society and solve the constraints such as storage space [25]. Several core technical difficulties are in large-scale industrial applications. For the blockchain, it is composed of several blocks. The block is a data structure, which mainly includes two parts: intrablock structure and interblock structure [26]. The intrablock structure is mainly composed of data blocks, and the data blocks include two parts: block header and block body. The block header is mainly composed of the metadata of the block, which is mainly used to verify the block.

The composition of the block body mainly includes information such as version number, previous block hash value, random number, timestamp, Merkle root, and block hash value [27]. Blockchain technology is extracted from Bitcoin. From the data point of view, the blockchain is regarded as an unalterable distributed database that all nodes jointly maintain and safely store data. The blockchain has the characteristics of decentralization, tamper-proof, traceability, and maintainability [28]. The structure and technology of the data layer can protect the integrity of the block. The network layer is responsible for the transmission and verification of data between distributed nodes. The consensus layer combines the application layer with the appropriate scenario to complete the integration with the actual application. This scheme verifies the consistency of node data and completes the filtering of nodes to ensure consensus [29].

3. Supply Chain Financial Risk Assessment

The concept of supply chain finance originated in the 1980s, and it was Stemler and Securing's proposal of "supply chain finance" that first attracted academic attention. Berger and Udall and others put forward the idea of supply chain finance earlier. Hofmann made a theoretical definition of supply chain finance [30]. He believes that supply chain finance is a new interdisciplinary combining finance, logistics, and supply chain management. It is to plan, guide, and jointly create value for the members of the supply chain. Methods: The Aberdeen group believes that supply chain finance is a process of optimizing supply chain cost processes.

According to Michael Lamoureux's definition, supply chain finance is a process of systematically optimizing the availability and cost of funds in a corporate ecosystem dominated by core companies. Initially, enterprises solved the problem of shortage of funds by discounting bills. Before the bills expire, the enterprises endorsed the bills receivable to the bank for discounting [31]. The discounted amount is the remaining value after the maturity value of the bill minus the interest. By the 1990s, banks developed trade financing from bill discounting to grant credit to companies. This is in the field of import and export trade. Banks applied structured short-term financing tools based on prepayments, deposits, and accounts receivable in commodity transactions. Current assets such as loans are a new way for companies to issue credit. By the beginning of the 21st century, supply chain financial services have become popular. Banks have developed supply chain financial services based on core enterprises, using accounts receivable, inventory, and prepaid accounts [32]. Construction framework of the supply chain financial risk indicator system is shown in Figure 2.

The credit risk assessment of general commercial banks is based on the financial indicators of financing enterprises, and the credit risk assessment of small- and medium-sized enterprises in the financing process is comprehensively examined [33]. The same is true for chain finance, but the difference is that supply chain finance can be used as an indicator of credit risk evaluation in addition to the financing enterprise itself, and the conditions of the core enterprise related to the financing enterprise and the entire supply chain can be used as a reference indicator for credit risk evaluation.

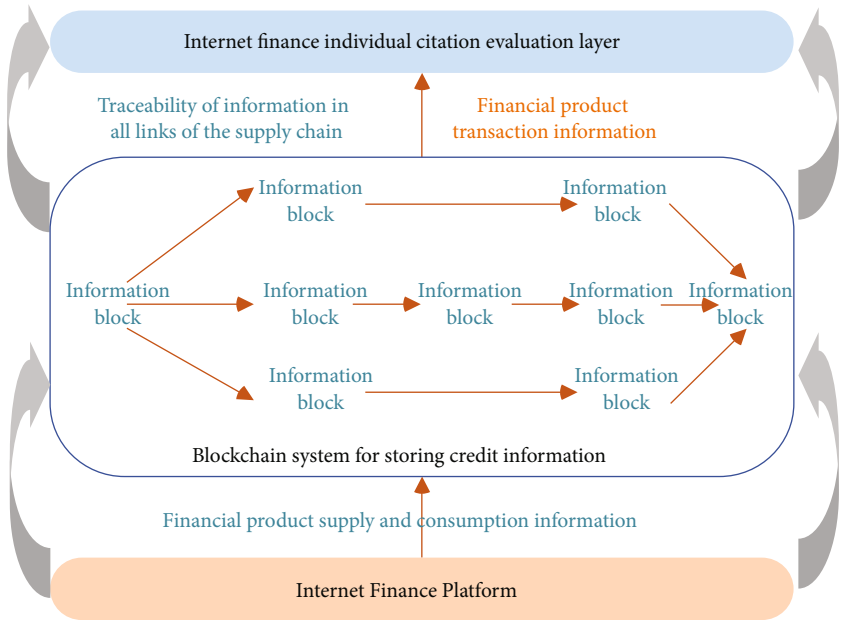


Figure 1: Research on the supply chain financial risk assessment based on the blockchain.

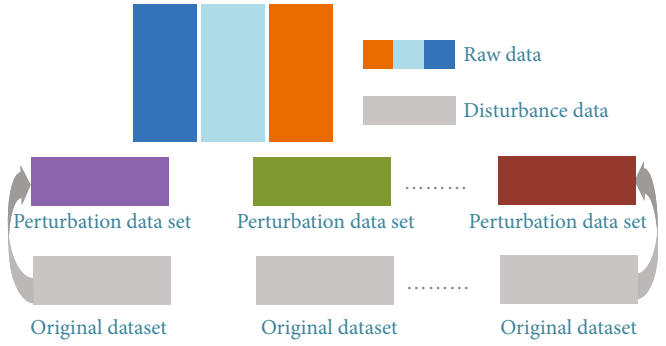


Figure 2: Construction framework of the supply chain financial risk indicator system.

4. Supply Chain Financial Risk Assessment Based on the Fuzzy Neural Network

Humans' understanding of objective things is often based on abstracting their characteristics first and then generalizing them. This leads to the connotation and extension of every concept, and its extension is the uncertainty of determining the concept. Based on this principle, Lazada proposed the fuzzy set theory and for the first time, also proposed the membership function representing the membership attribute, laying the foundation of the fuzzy theory.

The development of fuzzy theory has formed a rigorous mathematical science [34]. This article specifically introduces the basic theories of fuzzy sets, fuzzy matrices, and fuzzy inference systems in fuzzy theory. The traditional neural network structure and initial weights are generally determined based on empirical knowledge. Parameters such as network weights are obtained through offline training and then applied to the control system. It is obviously inappropriate to use a fuzzy neural network trained offline. In order to solve these problems, this paper proposes a fuzzy online adaptive neural network control method based on the fuzzy control and fuzzy neural network, which realizes the online determination and adaptive adjustment of the structure and parameters of the fuzzy neural network [35]. This leads to the connotation and extension of every concept, and its extension is the uncertainty of determining the concept. Based on this principle, Lazada proposed the fuzzy set theory and for the first time, also proposed the membership function representing the membership attribute, laying the foundation of the fuzzy theory. The fuzzy neural network has the advantages of fuzzy theory and neural network. It can not only complete the processing of fuzzy information. With the uncertainty of the characteristics, the neural network structure can be used to complete the learning and adaptation of the model.

Each basic unit of the fuzzy neural network is composed of fuzzy neurons. The basic ones included are the following types of fuzzy neurons. The function of the fuzzy neuron is to convert the determined value into a fuzzy output value.

$$yt = Ft\theta t + vt, vt \sim N[0, Vt]. \quad (1)$$

Fuzzy neurons convert fuzzy values into certain signals and output them. This type of neuron, which is glided with camphor, is formally similar to previous neurons, but has the opposite effect.

$$\theta t = \theta t - 1 + wr, wr \sim N[0, Wt]. \tag{2}$$

Because different samples have differences in the range of data, the neural network learning algorithm needs to determine the parameters of the network based on the data. The differences in these data will make the parameters of the neural network difficult to determine, and the accuracy cannot be guaranteed.

$$(\theta 0 \mid D0) \sim N[m0, C0]. \tag{3}$$

To solve this problem, this research normalizes all the fuzzy adaptive data to the range. The method of normalization is given by

$$Dt = \{yt, \mathrm{Ft}, Dt\text{-}1\}. \tag{4}$$

$$(\mathrm{theta}\, t \mid Dt - 1) \sim N[at, Rt]. \tag{5}$$

The adaptive adjustment of the neural network has a great advantage, and on the contrary, people are not clear about the logic.

$$(YT, Dt - 1) \sim N[\mathrm{ft}, Qt]. \tag{6}$$

The structure of the fuzzy neural network is extended from the RBF neural network, the extended structure does not exceed three layers, and its connection weight is not preset but determined by the input function. The dynamic fuzzy neural structure used in this article has five layers. The structure of four-input single-output is still adopted here.

$$X = 0, 1 \text{ and } X * x|2U(x) - c - v] - U(x)\, x = 0+, \tag{7}$$

$$G(x) = [2U(x) - c - v] - U(x). \tag{8}$$

Each node of this layer corresponds to the fuzzy rules of the fuzzy rule library, and the main function is to serve as the antecedent of matching fuzzy rules and then to solve the fitness of the corresponding rules. That is, the solution of ignition intensity is mainly obtained by combining each fuzzy neural node. The input and output of each node in this layer represent

$$\left\{ \begin{array}{l} \mathrm{Max}\ G(x) = 2U(0) - c - V \\ \mathrm{Min}\ G(x) = U(1) - c - v \end{array} \right\}, \tag{9}$$

$$[2U(x-) - c - v] - U(x-)\, x- = 0. \tag{10}$$

In the conjugate gradient method, the first search is specified as the negative gradient direction, and the second and subsequent search directions are all conjugated with the previous search direction. As one of the most informatized countries in the world, China, we should seize the opportunity of this information industry revolution in time, increase capital and policy investment, cultivate blockchain talents in enterprises and universities, speed up the construction of blockchain standards, and grasp the right to speak and command heights in the development of the global blockchain. The development of the blockchain is a process of point by line and side by side, from a single digital currency to the financial field and then to a comprehensive range of industries, realizing "blockchain+" to empower the industry and promote the development of the industry. In the fry method, after improving the FR method, the convergence speed of the fuzzy neural network algorithm is greatly improved. Perform a linear search along the new search direction to determine the best learning rate. Compared with the standard neural network algorithm, this method can usually converge more quickly.

5. Case Analysis of Supply Chain Financial Risk Assessment

Since the knowledge about classification contained in the training data is the only source of information for optimizing the classifier, sample data that contains good prior information about the problem to be processed has a direct impact on the final performance of the classifier and the test effect.

Therefore, it is necessary to achieve the two purposes of improving data separability and improving system processing efficiency through the data preprocessing process. Before conducting empirical analysis, due to the large number of indicators, it is necessary to eliminate indicators with higher correlation based on sample data. This article will use SPSS 25.0 to deal with the correlation between indicators in the risk assessment index system. In the 95% confidence interval, the indicators with the absolute value of the correlation coefficient greater than or equal to 0.7 are filtered and deleted. In addition, since the credit level of financing companies is relatively high, the score is 5 points, so removing this indicator is not included in the analysis. During the training process, the classification of the training sample set is shown in Figure 3.

First, randomly select the generated _5000 samples in the model and extract them into a combination of 3300 training sample sets and 1700 test sample sets. In order to eliminate the robustness of the model, this paper sets the model to perform 10 extractions in total, and the extraction becomes a combination of 10 different 3300 training sample sets and 1700 test sample sets for training. Distribution of the fuzzy neural network test results is shown in Figure 4.

Finally, this paper establishes a credit risk assessment model applicable to Internet supply chain finance. The training sample set consisting of 3300 samples first trains the fuzzy neural network model, and then the fuzzy neural network model analyzes and learns through the classification of the training sample set. Figure 5 shows a comparison of the classification accuracy of different index systems.

After that, it is compared with the experimental results of the traditional supply chain financial credit risk evaluation index system. In the traditional supply chain financial credit risk evaluation index system, there is no Internet-based

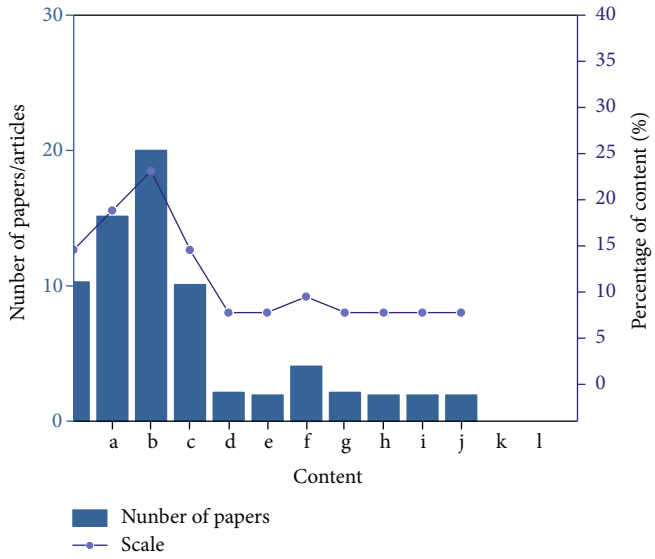


Figure 3: The classification of the fuzzy neural network training sample set.

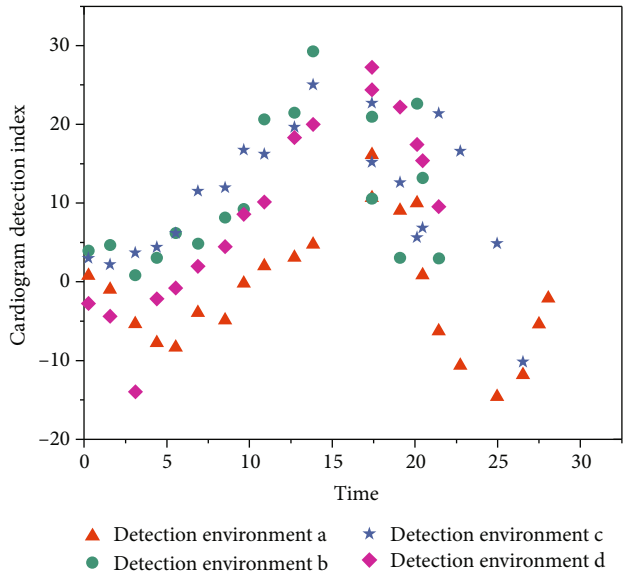


Figure 4: Distribution of fuzzy neural network test results.

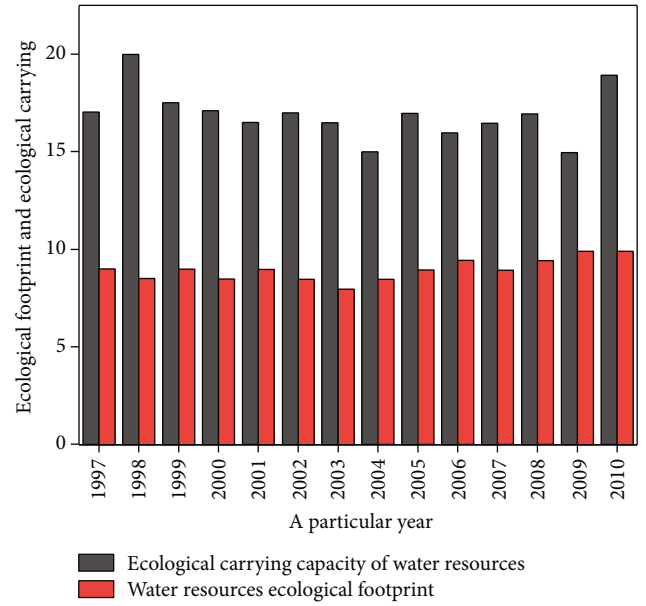


Figure 5: Comparison of classification accuracy rates for different index systems.

supply chain financial credit risk evaluation index system for the transaction situation. In the investigation, there are no two indicators of transaction volume and total profit margin of the supply chain. The first 24 three-level indicators constitute the traditional supply chain financial credit risk evaluation indicator system as independent variables. Still, use 1 to indicate that there are no overdue loans or accounts and 0 to indicate that there are overdue loans or accounts. Comparison of the effect of fuzzy neural algorithms in data reset is shown in Figure 6. Comparison of the effects of different methods in data reset is shown in Figure 7.

In order to eliminate the robustness of the model, this paper sets the model to perform 10 extractions in total, and the extraction becomes a combination of 10 different 3300 training sample sets and 1700 test sample sets for training. The training sample set composed of 3300 samples first trains the fuzzy neural network model, and the fuzzy neural network model analyzes and learns through the classification of the training sample set. After that, it is compared with the experimental results of the traditional supply chain financial credit risk evaluation index system. In the traditional supply chain financial credit risk evaluation index system, there is no Internet-based supply chain financial credit risk evaluation index system for the transaction situation. The discounted amount is the remaining value after the maturity value of the bill minus the interest. By the 1990s, banks developed trade financing from bill discounting to grant credit to companies. This is in the field of import and export trade. Banks applied structured short-term financing tools based on

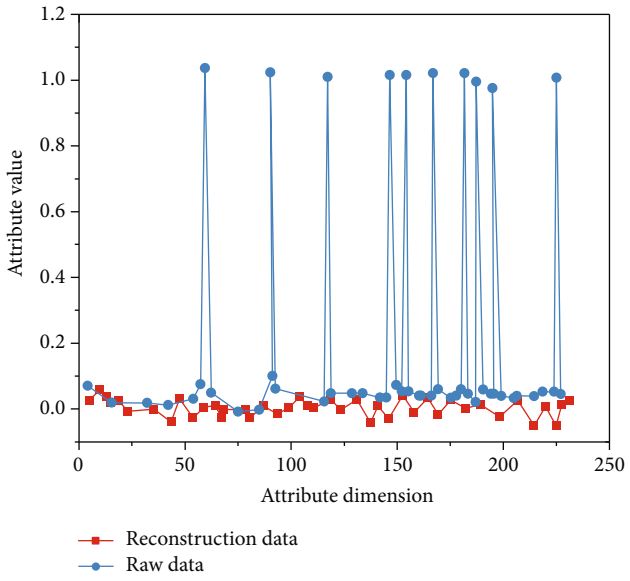


Figure 6: Comparison of the effect of fuzzy neural algorithms in data reset.

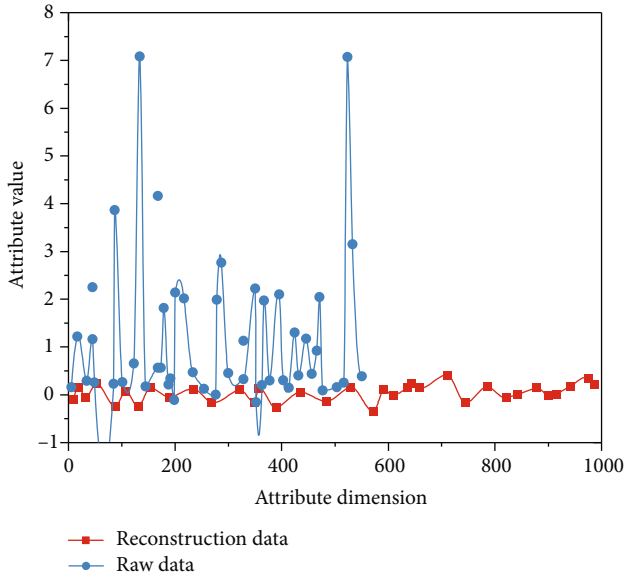


Figure 7: Comparison of the effects of different methods in data reset.

prepayments, deposits, and accounts receivable in commodity transactions. In the investigation, there are no two indicators of transaction volume and total profit margin of the supply chain.

6. Conclusion

The article first defines the related concepts of supply chain finance and supply chain financial credit risk and analyzes the generation mechanism of supply chain financial credit risk from a theoretical level. Secondly, from the level of development status, analyze the current status and existing problems of supply chain financial credit risk management and lay the foundation for building a supply chain financial credit risk assessment system, then focus on analyzing the sources and influencing factors of supply chain financial credit risks, and from financing enterprises, core enterprises, the overall operation of supply chain finance, and the macroeconomic environment to select indicators to build a supply chain finance credit risk assessment system. Finally, a modified fuzzy neural network model is used to conduct empirical analysis and research on credit risk.

Credit risk runs through all aspects of the credit business before, during, and after the loan. In the context of preventing financial risks, it is important to evaluate the credit risk of supply chain finance based on the perspective of commercial banks. The same is true for chain finance, but the difference is that supply chain finance can be used as an indicator of credit risk evaluation in addition to the financing enterprise itself, and the conditions of the core enterprise related to the financing enterprise and the entire supply chain can be used as a reference indicator for credit risk evaluation. With the development of supply chain finance, the credit risk of small and medium financing enterprises from the perspective of supply chain finance has arisen. Risk management is one of the key tasks of the credit business of banks and other financial institutions, which runs through all aspects of the credit business before, during, and after the loan. This article combines blockchain and fuzzy neural network algorithms to study the credit risk of SME financing from the perspective of supply chain finance.

Data Availability

The authors approve that data used to support the finding of this study are included in the article.

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 Soft Science Research Project of Zhejiang Province (2021C35072), Social Science Planning Project of Zhejiang Province (21NDJC311YBM), and Visiting Scholar Fund of Zhejiang Provincial Education Department (FX2019075).

References

[1] H. Zhou, H. Zhao, and Y. Zhang, "Nonlinear system modeling using self-organizing fuzzy neural networks for industrial applications," *Applied Intelligence*, vol. 50, no. 7, 2020.

[2] S. Zhao and Q. Zhu, "Remanufacturing supply chain coordination under the stochastic remanufacturability rate and the random demand," *Annals of Operations Research*, vol. 257, no. 1-2, pp. 661–695, 2017.

[3] Y. G. Jiang, Q. Dai, T. Mei, Y. Rui, and S. F. Chang, "Super fast event recognition in internet videos," *IEEE Transactions on Multimedia*, vol. 17, no. 8, pp. 1174–1186, 2015.

[4] M. Zhang, Y. K. Tse, J. Dai, and H. K. Chan, "Examining green supply chain management and financial performance: roles of social control and environmental dynamism," *IEEE Transactions on Engineering Management*, vol. 66, no. 1, pp. 20–34, 2017.

[5] Y. H. Kim and D. Henderson, "Financial benefits and risks of dependency in triadic supply chain relationships," *Journal of Operations Management*, vol. 36, pp. 115–129, 2017.

[6] Y. Sheng, F. L. Lewis, Z. Zeng, and T. Huang, "Lagrange stability and finite-time stabilization of fuzzy memristive neural networks with hybrid time-varying delays," *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 2959–2970, 2019.

[7] G. Yangyong and W. Juan, "Modeling of false information on microblog with block matching and fuzzy neural network," *International Journal of Modern Physics C*, vol. 12, p. 2150019, 2020.

[8] W. Xie and Q. Zhu, "Input-to-state stability of stochastic nonlinear fuzzy Cohen–Grossberg neural networks with the event-triggered control," *International Journal of Control*, vol. 93, no. 9, pp. 2043–2052, 2020.

[9] X. Wu, H. Han, Z. Liu, and J. Qiao, "Data-knowledge-based fuzzy neural network for nonlinear system identification," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 9, pp. 2209–2221, 2020.

[10] R. M. Wang, Y. N. Zhang, Y. Q. Chen, X. Chen, and L. Xi, "Fuzzy neural network-based chaos synchronization for a class of fractional-order chaotic systems: an adaptive sliding mode control approach," *Nonlinear Dynamics*, vol. 100, no. 2, 2020.

[11] K. M. Thompson and D. A. Kalkowska, "Potential future use, costs, and value of poliovirus vaccines," *Risk Analysis*, vol. 5, no. 9, 2020.

[12] M. Stachowski, A. Fiebig, and T. Rauber, "Autotuning based on frequency scaling toward energy efficiency of blockchain algorithms on graphics processing units," *Journal of Supercomputing*, vol. 77, no. 1, 2020.

[13] M. Shabani, "Blockchain-based platforms for genomic data sharing: a de-centralized approach in response to the governance problems?," *Journal of the American Medical Informatics Association*, vol. 26, 2018.

[14] J. Rubin, A. Ottosen, A. Ghazieh, J. Fournier-Caruana, A. K. Ntow, and A. R. Gonzalez, "Managing the planned cessation of a global supply market: lessons learned from the global cessation of the trivalent oral poliovirus vaccine market," *Journal of Infectious Diseases*, vol. 216, suppl_1, p. S40, 2017.

[15] L. V. Pavao, C. Pozo, C. B. B. Costa, M. A. S. S. Ravagnani, and L. Jimenez, "Financial risks management of heat exchanger networks under uncertain utility costs via multi-objective optimization," *Energy*, vol. 139, pp. 98–117, 2017.

[16] V. Nasir, S. Nourian, S. Avramidis, and J. Cool, "Stress wave evaluation for predicting the properties of thermally modified wood using neuro-fuzzy and neural network modeling," *Holzforschung*, vol. 73, no. 9, pp. 827–838, 2019.

[17] I. Monasterolo, S. Battiston, A. C. Janetos, and Z. Zheng, "Vulnerable yet relevant: the two dimensions of climate-related financial disclosure," *Climatic Change*, vol. 145, 2017.

[18] L. Liu, M. Du, and X. Ma, "Blockchain-based fair and secure electronic double auction protocol," *IEEE Intelligent Systems*, vol. 35, 2020.

[19] H. Li, C. Li, D. Ouyang, S. K. Nguang, and Z. He, "Observer-based dissipativity control for T-S fuzzy neural networks with distributed time-varying delays," *IEEE Transactions on Cybernetics*, vol. 10, no. 99, pp. 1–11, 2020.

[20] S. Kusi-Sarpong, H. Gupta, and J. Sarkis, "A supply chain sustainability innovation framework and evaluation methodology," *International Journal of Production Research*, vol. 57, no. 7, pp. 1990–2008, 2019.

[21] M. C. Kees, J. A. Bandoni, and M. S. Moreno, "An optimization model for managing the drug logistics process in a public hospital supply chain integrating physical and economic flows," *Industrial & Engineering Chemistry Research*, vol. 58, no. 9, pp. 3767–3781, 2019.

[22] H. Jun-Ho and S. Kyungryong, "Blockchain-based mobile fingerprint verification and automatic log-in platform for future computing," *Journal of Supercomputing*, vol. 75, pp. 3123–3139, 2018.

[23] M. W. D. Hanson-Heine and A. P. Ashmore, "Computational chemistry experiments performed directly on a blockchain virtual computer," *Chemical Science*, vol. 11, no. 18, 2020.

[24] B. Hamdaoui, M. Alkalbani, T. Znati, and A. Rayes, "Unleashing the power of participatory IoT with blockchains for increased safety and situation awareness of smart cities," *IEEE Network*, vol. 34, no. 2, pp. 202–209, 2020.

[25] D. A. Ghadge, S. K. Jena, S. Kamble, D. Misra, and M. K. Tiwari, "Impact of financial risk on supply chains: A manufacturer-supplier relational perspective," *International Journal of Production Research*, vol. 6, no. 3, 2020.

[26] L. E. Higham, C. Scott, K. Akehurst et al., "Effects of financial incentives and cessation of thinning on prevalence of Campylobacter: a longitudinal monitoring study on commercial broiler farms in the UK," *Veterinary Record*, vol. 183, p. 595, 2018.

[27] B. S. Das, K. Devi, J. R. Khuntia, and K. K. Khatua, "Discharge estimation in non-prismatic compound channel using adaptive neuro-fuzzy inference system," *Canadian Journal of Civil Engineering*, vol. 47, 2019.

[28] F. D'Amore, N. Sunny, D. Iruretagoyena, F. Bezzo, and N. Shah, "European supply chains for carbon capture, transport and sequestration, with uncertainties in geological storage capacity: Insights from economic optimisation," *Computers & Chemical Engineering*, vol. 129, article 106521, 2019.

[29] K. Christidis, D. Sikeridis, Y. Wang, and M. Devetsikiotis, "A framework for designing and evaluating realistic blockchain-based local energy markets," *Applied Energy*, vol. 281, 2021.

[30] S. G. Chen, F. J. Lin, C. H. Liang, and C. H. Liao, "Intelligent maximum power factor searching control using Recurrent Chebyshev fuzzy neural network current angle controller for SynRM drive System," *IEEE Transactions on Power Electronics*, vol. 36, no. 3, pp. 3496–3511, 2020.

[31] C.-H. Chen, "A cell probe-based method for vehicle speed estimation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E103.A, no. 1, pp. 265–267, 2020.

[32] J. Barrette, E. Thiffault, A. Achim, M. Junginger, D. Pothier, and L. De Grandpré, "A financial analysis of the potential of dead trees from the boreal forest of eastern Canada to serve

as feedstock for wood pellet export," *Applied Energy*, vol. 198, 2017.

[33] S. Azimi, M. M. Azhdary, and M. S. A. Hashemi, "Prediction of annual drinking water quality reduction based on groundwater resource index using the artificial neural network and fuzzy clustering," *Journal of Contaminant Hydrology*, vol. 220, pp. 6–17, 2019.

[34] M. C. Alison, "Duke fraud case highlights financial risks for universities," *Science*, vol. 353, pp. 977-978, 2018.

[35] A. Roehrs, C. A. da Costa, R. da Rosa Righi, V. F. da Silva, J. R. Goldim, and D. C. Schmidt, "Analyzing the performance of a blockchain-based personal health record implementation," *Journal of Biomedical Informatics*, vol. 92, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6682340, 12 pages
https://doi.org/10.1155/2021/6682340

Research Article

Energy Efficiency Optimization of Cognitive UAV-Assisted Edge Communication for Semantic Internet of Things

Yilong Gu,[1] **Yangchao Huang**,[2] **Hang Hu**,[2] **Weiting Gao**,[2] **and Yu Pan**[1]

[1]*Graduate School, Air Force Engineering University, Xi'an 710077, China*
[2]*College of Information and Navigation, Air Force Engineering University, Xi'an 710077, China*

Correspondence should be addressed to Hang Hu; xd_huhang@126.com

Received 17 October 2020; Revised 11 January 2021; Accepted 23 January 2021; Published 10 February 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Yilong Gu 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 consolidation of the Internet of Things (IoT), the unmanned aerial vehicle- (UAV-) based IoT has attracted much attention in recent years. In the IoT, cognitive UAV can not only overcome the problem of spectrum scarcity but also improve the communication quality of the edge nodes. However, due to the generation of massive and redundant IoT data, it is difficult to realize the mutual understanding between UAV and ground nodes. At the same time, the performance of the UAV is severely limited by its battery capacity. In order to form an autonomous and energy-efficient IoT system, we investigate semantically driven cognitive UAV networks to maximize the energy efficiency (EE). The semantic device model for cognitive UAV-assisted IoT communication is constructed. And the sensing time, the flight speed of UAV, and the coverage range of UAV communication are jointly optimized to maximize the EE. Then, an efficient alternative algorithm is proposed to solve the optimization problem. Finally, we provide computer simulations to validate the proposed algorithm. The performance of the joint optimization scheme based on the proposed algorithm is compared to some benchmark schemes. And the simulation results show that the proposed scheme can obtain the optimal system parameters and can significantly improve the EE.

1. Introduction

An unmanned aerial vehicle (UAV), which is developed originally for military purposes, has been widely applied to the civilian domain. With the miniaturization of equipment, the continuous reduction of manufacturing cost, and the continuous improvement of communication performance, the UAV-based wireless communication has attracted much attention. In the IoT, ultrahigh data rates and reliability are critical to a lot of user connections and network sensors. In order to improve the performance of IoT networks, a massive MIMO two-way relaying system has been studied [1]. The reliability of the communication systems can be improved by Simultaneous Wireless Information and Power Transfer (SWIPT). Furthermore, with the large-scale use of mobile IoT devices, such as wearable devices, driverless cars, and intelligent terminals, the problem of making the GBS cover more communication terminals needs to be solved urgently. The coordinated direct and relay transmission (CDRT) system is investigated to extend the communication coverage in [2]. In order to improve the spectrum efficiency, the uplink and downlink transmissions are designed jointly via network coding. Meanwhile, the advantages of the proposed scheme are verified by simulations. Meanwhile, the demand for the IoT based on UAV will be more urgent. UAV can improve the communication performance of the IoT nodes, especially in the network, whose coverage quality is poor, such as construction sites, disaster areas, highways, and narrow lanes. However, the performance of the UAV will be affected by its limited battery capacity. The method of improving the EE has become a significant direction.

The crucial drivers of IoT contain the daily addition of new devices with their own data models [3]. This leads to an enormous increase of the structured data in size and complexity. There is a demand to abstract the heterogeneity of devices so that their functions can be represented as virtual combination platforms [4–6]. A semantic IoT framework can accept heterogeneous models, which can support data sharing between devices. Thus, it needs an efficient method to update the metadata. The IoT directory sustaining semantic

description of IoT objects is proposed in [7]. Semantic description and semantic collaboration have become effective methods to realize information interaction and sharing between objects.

Usually, UAVs operate in an unlicensed spectrum. In the IoT, there are many wireless technologies, such as Wi-Fi, Bluetooth, and cellular network. The spectrum resources of the IoT become more and more scarce due to the explosive growth of wireless devices [8]. As is known to us all, a cognitive radio (CR) permits a secondary user (SU) to opportunistic access the licensed spectrum without interfering with the primary user (PU). Employing CR technology for the spectrum scarcity issue of IoT is a potential solution.

Under this background, a large number of researches on UAV-assisted IoT communication and cognitive UAV communication have been carried out. The problem of energy saving and consumption reduction of mobile IoT becomes more and more prominent. The joint optimal deployment of static ground nodes and UAVs is studied in Reference [9]. Due to the lack of dynamic research between UAVs and ground nodes, it is necessary to consider the UAV mobile relay system. Under the constraints of UAV mobility, the transmission power of the transmitter and UAV trajectory is optimized to maximize the throughput [10]. Moreover, the spectrum efficiency and energy efficiency are maximized by optimizing time allocation, flight speed, and trajectory of UAV [11]. The results show that the tradeoff between the maximum achievable spectrum efficiency and energy efficiency can be achieved by the design of the trajectory.

Aiming to improve the spectrum efficiency of IoT, it is necessary to study IoT-based CR; meanwhile, the devices in the IoT are able to sense the wireless environment [12]. The use of spectrum sharing technology in the IoT is also considered to be a mutually beneficial solution [13]. When the transmission rate of PU is lower than the requirement, a part of bandwidth is authorized to the SU to achieve the purpose of cooperative transmission. Compared with ground spectrum sensing technology, UAV spectrum sensing can obtain better sensing performance. Spectrum sensing performance can be improved by increasing the number of samples which is related to spectrum sensing time. However, improving the sensing performance may consume more sensing time and thus decrease the transmission time. It is proved that there is a sensing-throughput tradeoff in spectrum sensing of the CR, i.e., there exists an optimal sensing time that makes throughput of the CR achieve the maximal value [14]. In [15], the cooperative method based on decode-and-forward and physical-layer network coding is proposed to reduce the transmission delay of wireless sensor networks. Furthermore, the cooperative compressed spectrum sensing algorithm for UAV is proposed to improve sensing accuracy, which allows ground nodes to send compressed sensing information to UAV [16–18].

Semantic IoT devices with cognitive function can solve the problem of low spectrum efficiency by spectrum sensing [19]. UAV, as the air interface of the IoT, can effectively solve the problem of long-distance communication under the condition of energy limitation of IoT. However, there are few studies applying these two advantages of UAV to the semantic IoT. This paper considers that the CR-based UAV can opportunistically use the spectrum to provide information for the remote semantic IoT devices. In order to achieve the green and autonomous IoT, the EE of the cognitive UAV system is investigated.

In this paper, the maximization of semantic IoT's EE based on cognitive UAV is studied. On the basis of building a semantic device model for cognitive UAV to assist IoT communication, we jointly optimize the UAV sensing time, UAV flight speed, and UAV communication distance to maximize the EE. We solve this problem by dividing it into three subproblems. And an efficient alternative algorithm is proposed to obtain the optimal solution. Through computer simulations, we verify the proposed scheme. It is concluded that the proposed scheme can achieve better performance compared to the benchmark schemes and can obtain the maximum EE.

The rest of this paper is organized as follows. In Section 2, we consider a cognitive fixed-wing UAV with a circular trajectory. The UAV can opportunistically access the spectrum of GBS. And the semantic device model for cognitive UAV-assisted IoT communication is constructed, which is used to establish autonomous communication links between UAV and the ground nodes farther away from the ground base station (GBS). In Section 3, an algorithm is proposed to obtain the maximal EE of UAV cognitive communication. In Section 4, the algorithm is simulated. And the comparison shows that our joint optimization scheme can significantly improve the EE. Finally, conclusions are drawn in Section 5.

2. System Model

We consider a UAV-assisted edge IoT model consisting of a GBS, a UAV, and a group of ground nodes, such as car nodes, smartphone nodes, and computer nodes. As a relay node, the cognitive UAV is aimed at serving the ground nodes with low communication quality due to the long distance from the GBS. At the same time, the cognitive UAV can use the spectrum of GBS to communicate with the ground nodes and transmits its own messages when the spectrum of the GBS is idle [20]. Since the battery capacity of the UAV is usually limited, we aim to maximize the energy efficiency of UAV relay communication on the basis of improving spectrum efficiency in the IoT.

The UAV flies in a circle around the GBS which is located at the center of the circular flight path. We consider that the UAVs are equipped with a Global Positioning System (GPS); thus, their geographical coordinates can be obtained. The flying speed of UAV in a circular motion is v, and the radius of the UAV flight path is d_u. The advantage of using a circular trajectory is that the UAV can not only save energy consumption greatly [21] but also provide regular communication service. When the UAV flies to position A, B, or C, it can communicate with ground nodes in the circle A', B', or C' directly below it, which is shown in Figure 1.

The association rules between UAV and ground nodes can be determined by semantic technology. In order to support the autonomous scheduling of communication, the

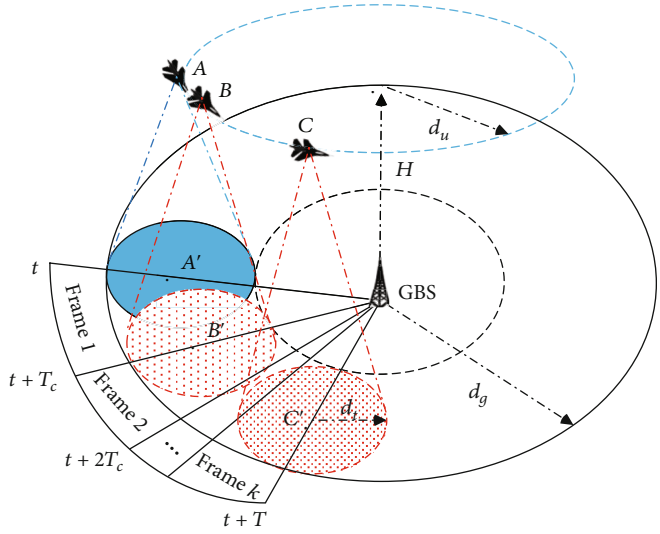


Figure 1: UAV-assisted edge communication system for IoT.

semantic representation model of IoT nodes is constructed from four aspects of basic equipment information, communication capability, communication state, and operation control, as shown in Figure 2.

The basic parameters are the semantic descriptions of its name, type, and technical parameters, where the location parameter represents the three-dimensional position of the node, which is needed for cognitive UAV to establish an assisted communication link. Application description describes the relevant parameter information of the node in different environments. And the communication mode, protocol, and communication coverage range of UAV are described by communication capability.

Let the task state indicate whether the node can establish the communication link. And the communication quality of IoT nodes covered by GBS is defined by coverage quality. The sensing state represents the information obtained by the UAV through environment sensing.

In order to enable UAV to adjust speed, sensing time, and communication coverage range in response to changing electromagnetic environment, we conducted a semantic model for operational control. The acquisition of ground nodes' state and control of communication link establishment are also described in the model of operating control.

Then, at a certain moment, the UAV can establish communication link with ground nodes in a circular region S_c with radius d_t. S_c is in the projection area below the UAV's flight position, as shown in Figure 3. d_t is determined by the UAV's communication range of the semantic model.

We define M as the total number of ground nodes in S_c, which are served by UAV at time t. Suppose that the ground nodes are randomly and evenly distributed in the circular area centered on PU with a radius of d_g, then we have $d_u = d_g - d_t$. The density of ground users is represented by ρ; hence, M can be expressed as

$$M = \lfloor \rho \pi {d_t}^2 \rfloor, \tag{1}$$

where $\lfloor x \rfloor$ denotes the largest integer no more than x.

The frame T_c is divided into the sensing slot and transmission slot. In the sensing slot, the UAV senses whether the GBS is active within time τ. If GBS is detected inactive, the UAV opportunistically uses the primary spectrum belonging to the GBS to send information to ground nodes within transmission slot $T_c - \tau$.

Then, two types of channels are considered: (i) the GBS to UAV channel and (ii) the UAV to ground node channel. These channels can be assumed to be line of sight (LOS). We assume that the UAV is flying horizontally with a fixed altitude K, which meets the minimum height required to avoid the actual terrain or buildings, so that frequent adjustment of flight altitude is not required. Thus, the impact of severe ground fading on the sensing and transmission performance can be ignored. The channel power gain of the UAV to a ground node follows the free-space path loss model, which can be expressed as

$$h_t = \frac{\beta_0}{K^2 + d(t)^2}, \tag{2}$$

where β_0 denotes the channel power at the reference distance of one meter. $d(t)$ denotes the horizontal distance from the UAV to the ground node at a certain time t. In addition, we suppose that the Doppler effect produced by the UAV's movement is offset.

In this paper, we consider that the bandwidth of the GBS has been allocated in advance. It is assumed that the total available bandwidth provided to UAV is W in Hz. To facilitate the analysis, we use $\rho \pi {d_t}^2$ as the number of ground nodes in S_c , and the bandwidth allocated to each ground node is

$$W' = \frac{W}{\rho \pi {d_t}^2}. \tag{3}$$

The transmission power of UAV to each ground node is defined as P_t. To guarantee the quality of the whole

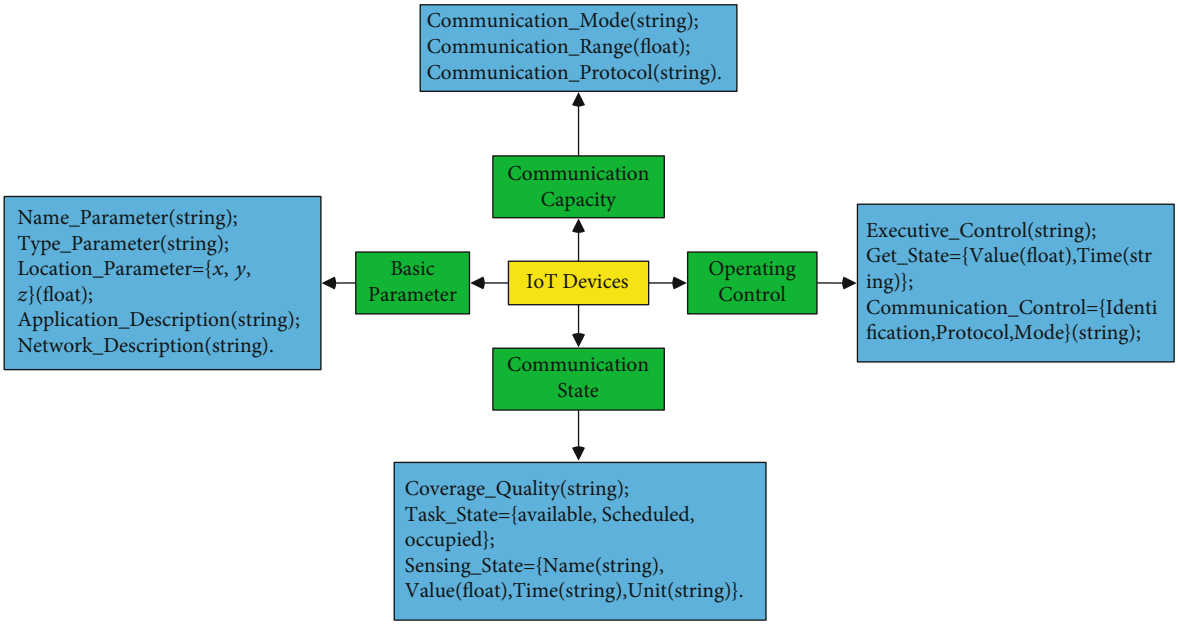


Figure 2: Semantic model for UAV-assisted communication in the IoT.

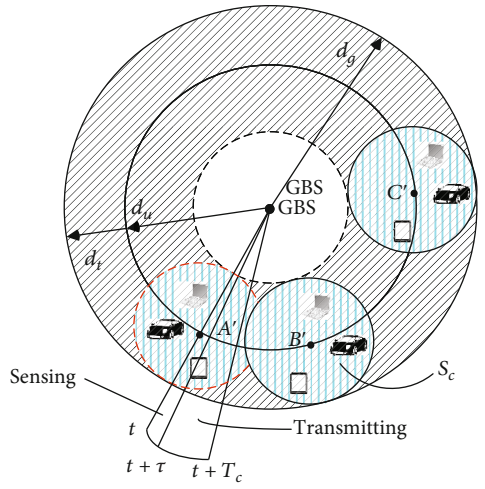


Figure 3: Spectrum sensing model.

communication network, we consider the worst case that the UAV communicates with the edge nodes of S_c.

Let C_0 denote the transmission rate from UAV to each ground node, which can be expressed as

$$C_0 = \log_2\left(1 + \frac{\beta_0 P_t}{\left(K^2 + {d_t}^2\right) W' \sigma^2}\right), \tag{4}$$

where σ^2 denotes the noise variance.

Next, the situation in which the UAV can use the frequency band of GBS is divided into the following two cases:

(i) When true status of the GBS is inactive and the sensing result of the UAV shows that it is inactive, the achievable throughput of the UAV's communication link is $((T_c - \tau)/T_c)C_0$

(ii) When true status of the GBS is active and the sensing result of the UAV shows that it is inactive, the achievable throughput of the UAV's communication link is $((T_c - \tau)/T_c)C_1$

Suppose that the primary signal is a binary phase shift keying (BPSK) signal, the noise is a real-valued Gaussian variable. And the energy detection method is used to detect the GBS's status. The signal to noise ratio (SNR) received by UAV is denoted as γ. The false alarming probability and the detection probability are denoted as P_f and P_d, respectively. And the sampling frequency is f_s. In the practice of low SNR, in order to protect the GBS from the interference of UAV, the detection probability of target $\bar{P}_d$ is required to be close to 1 but less than 1.

Then, for given target probabilities $\bar{P}_d$ and $\bar{P}_f$, P_f and P_d are given by [22]

$$\begin{aligned} P_f &= Q\left(\sqrt{2\gamma + 1}Q^{-1}\left(\bar{P}_d\right) + \sqrt{\tau f_s}\gamma\right), \\ P_d &= Q\left(\frac{1}{\sqrt{2\gamma + 1}}Q^{-1}\left(\bar{P}_f\right) + \sqrt{\tau f_s}\gamma\right), \end{aligned} \tag{5}$$

where $Q(x)$ is the distribution function of standard Gaussian, which is given by $Q(x) = (1/\sqrt{2\pi})\int_x^\infty \exp\left(-(t^2/2)\right)dt$.

When the frequency band W is given, $P(H_1)$ is the active probability of GBS, and $P(H_0)$ is the inactive probability of GBS. Therefore, the probability that the first case happens is $(1 - P_f)P(H_0)$, and the probability for the second case to happen is $(1 - P_d)P(H_1)$.

Hence, the effective throughput of the UAV communication link without active GBS is given as follows:

$$R_0 = \frac{T_c - \tau}{T_c} C_0 \left(1 - P_f\right) P(H_0). \tag{6}$$

When the GBS is inactive, the throughput is

$$R_1 = \frac{T_c - \tau}{T_c} C_1 (1 - P_d) P(H_1). \tag{7}$$

Then, the average throughput of the UAV communication link is

$$R = R_0 + R_1. \tag{8}$$

The following three cases will be considered when we analyze R in practice.

(i) When the UAV fails to sense that the GBS is active, there is interference between UAV and GBS; hence, the value of C_1 will be small, or the ground nodes may not decode the data transmitted from the UAV

(ii) The target detection probability is usually greater than 0.9 for the SNR of -20 dB, which is in IEEE 802.22 WRAN

(iii) The GBS's active probability $P(H_1)$ is small enough, less than 0.3 [21], so that it is economically advisable to explore the licensed spectrum band

Based on the above analysis, $C_1(1 - P_d)P(H_1) \ll C_0(1 - P_f)P(H_0)$; then, the average throughput can be approximated as

$$\tilde{R} \approx R_0 = C_0 P(H_0) \left(1 - \frac{\tau}{T_c}\right) \left(1 - Q\left(\alpha + \sqrt{\tau f_s}\gamma\right)\right), \tag{9}$$

where $\alpha = \sqrt{2\gamma + 1} Q^{-1}(\bar{P}_d)$.

Next, we analyze the energy consumption of UAV. The energy consumption of UAV is divided into two parts. One is the energy consumption associated with communication; the other is the energy required to maintain a fixed-wing drone flight. In practice, the power associated with communication is usually much smaller than the power needed to fly. The power needed to communicate is typically a few watts, but the UAV's propulsion power is usually hundreds of watts. Therefore, the power associated with UAV communication is ignored in this paper.

For constant speed circular flight, we have $\|v(t)\| = v$ and $a(t)^T v(t) = 0$. Particularly, $a(t)$ denotes the centripetal acceleration of UAV, whose direction must be perpendicular to the direction of velocity to ensure the constant speed. It can be seen from [23] that the power consumption of the UAV's steady circular flight is further derived:

$$\Xi = \left(c_1 + \frac{c_2}{g^2 {d_u}^2}\right) v^3 + \frac{c_2}{v}, \tag{10}$$

where g is the gravitational acceleration with normal value 9.8 m/s^2. c_1 and c_2 are determined by the weight of the aircraft, wing area, air density, and so on. This model conforms to the known classical aircraft power consumption model of aerodynamics theory.

According to formulas (9) and (10), the energy efficiency of IoT based on cognitive UAV can be expressed as

$$\eta = \frac{\tilde{R}}{\Xi} = \frac{P(H_0)(1 - \tau/T_c)\left(1 - Q\left(\alpha + \sqrt{\tau f_s}\gamma\right)\right) \log_2\left(1 + \beta_0 P_t \rho \pi {d_t}^2 / \left(K^2 + {d_t}^2\right) W\sigma^2\right)}{\left(c_1 + c_2 / g^2 \left(d_g - d_t\right)^2\right) v^3 + c_2 / v}. \tag{11}$$

It is a function with τ, d_t, v as variables.

3. Solutions of the Optimization Problem

Our goal is to maximize the energy efficiency of the IoT based on cognitive UAV relay communication by optimizing UAV flying speed v, the sensing slot duration τ, and the communication distance threshold d_t. This optimization problem can be formulated as

$$\begin{aligned} (\mathrm{OP}): \max_{\tau, v, d_t} & \eta, \\ 0 \le \tau \le T_c, & \\ 0 \le v \le v_{\max}, & \\ 0 \le d_t \le d_g. & \end{aligned} \tag{12}$$

In the objective function, τ determines the molecular part and v influences the denominator part. Then, we can take into account that optimization of objective function can be solved by dividing the problem OP into three subproblems.

3.1. Optimization of v with Fixed τ and d_t. The first subproblem OP_1 can be worked out by solving the optimal v to minimize the energy consumption of UAV with fixed τ and d_t:

$$\begin{aligned} (\mathrm{OP}_1): \min_{v} & P, \\ 0 \le v \le v_{\max}. & \end{aligned} \tag{13}$$

In order to obtain the optimal flight speed of UAV, we take its first derivative of the Ξ to analyze the monotonicity:

$$\frac{\partial \Xi}{\partial v} = 3v^2\left(c_1 + \frac{c_2}{g^2 d_u^2}\right) - \frac{c_2}{v^2}. \tag{14}$$

Then, we have $\lim_{v\to\infty} \partial\Xi/\partial v = +\infty$ and $\lim_{v\to 0}\partial\Xi/\partial v = -\infty$. Next, the second derivative of Ξ with respect in v is given as follows:

$$\frac{\partial^2 \Xi}{\partial v^2} = 6v\left(c_1 + \frac{c_2}{g^2 d_u^2}\right) + \frac{2c_2}{v^3} > 0. \tag{15}$$

As the result shown in equation (15), we can know that Ξ is convex in v. By solving equation (14) equal to 0, we have

$$v_0 = \sqrt[4]{\frac{c_2}{3c_1 + 3c_2/g^2\left(d_g - d_t\right)^2}}. \tag{16}$$

Obviously, the function Ξ decreases monotonically in the interval $(0, v_0)$ and increases monotonically when v is larger than v_0. Define the optimal v as v^*. Then, if $v_0 < v_{\max}$, we have $v^* = v_0$, and if $v_0 > v_{\max}$, we have $v^* = v_{\max}$. Hence, the minimum energy consumption of UAV is taken as $v = v^*$, when τ and d_t are fixed.

3.2. Optimization of τ with Fixed v and d_t. The second subproblem is to achieve the maximal average throughput when v and d_t are fixed:

$$\begin{aligned} (\mathrm{OP}_2):\ &\max_{\tau} \tilde{R}, \\ &0 \le \tau \le T_c. \end{aligned} \tag{17}$$

Therefore, we can obtain that the average throughput is a function of τ. In order to prove that the optimal τ exists, we take the first derivative of $\tilde{R}$:

$$\frac{\partial \tilde{R}}{\partial \tau} = C_0 P(H_0)\left[\frac{\gamma\sqrt{f_s}(1-\tau/T_c)}{2\sqrt{2\pi\tau}}\ \exp\left(-\frac{\left(\alpha + \sqrt{\tau f_s}\gamma\right)^2}{2}\right) - \frac{1}{T_c} + \frac{1}{T_c}Q\left(\alpha + \sqrt{\tau f_s}\gamma\right)\right]. \tag{18}$$

Then,

$$\lim_{\tau\to 0}\frac{\partial \tilde{R}}{\partial \tau} = +\infty. \tag{19}$$

Since $Q(x)$ is decreasing with upper boundary 1, $Q(\alpha + \sqrt{T_c f_s}\,\Upsilon) - 1 < 0$, we have

$$\lim_{\tau\to T_c}\frac{\partial \tilde{R}}{\partial \tau} = \frac{1}{T_c}\left(Q\left(\alpha + \sqrt{T_c f_s}\gamma\right) - 1\right)C_0 P(H_0) < 0, \tag{20}$$

and the second derivative of $\tilde{R}$ is as follows:

$$\begin{aligned} \frac{\partial^2 \tilde{R}}{\partial \tau^2}\frac{1}{C_0 P(H_0)} = {} & \frac{-\gamma\sqrt{f_s}/T_c + -\gamma\sqrt{f_s}\left(\alpha + \gamma\sqrt{\tau f_s}\right)/2\sqrt{\tau}}{8\pi\tau} 2\sqrt{2\pi\tau}\,\exp \\ & \cdot\left(-\frac{\left(\alpha + \gamma\sqrt{\tau f_s}\right)^2}{2}\right) - \gamma\left(1 - \frac{\tau}{T_c}\right)\sqrt{\frac{2\pi f_s}{\tau}}\exp \\ & \cdot\left(-\frac{\left(\alpha + \gamma\sqrt{\tau f_s}\right)^2}{2}\right) - \frac{1}{2T_c}\sqrt{\frac{f_s}{2\pi\tau}}\exp\left(-\frac{\gamma^2 f_s \tau}{2}\right) < 0. \end{aligned} \tag{21}$$

Equation (21) shows that $\partial\tilde{R}/\partial\tau$ is monotonic decline. Equations (19) and (20) show that $\tilde{R}(\tau)$ increases when the sensing time approaches 0, and $\tilde{R}(\tau)$ decreases when the sensing time approaches T_c. Therefore, define the optimal τ as τ^*, and it is reliable that τ^* exists in the interval $(0, T_c)$. For any given d_t and v, the Golden section method is used to obtain the optimal sensing time τ^*.

3.3. Optimization of d_t with Fixed τ and v. The last subproblem OP_3 is given by maximizing the energy efficiency with fixed τ and v:

$$\begin{aligned} (\mathrm{OP}_3):\ &\max_{d_t} \eta, \\ &0 \le d_t \le d_g. \end{aligned} \tag{22}$$

Notice that we need to adjust the value of d_t to achieve the maximum energy efficiency by tradeoff between maximum throughput and minimum energy consumption.

For any fixed radius d_t, the optimal speed v^* is obtained by solving P_1. And we notice that d_t and τ are two independent parameters. Thus, the multivariate function is reduced to a function with d_t as a single variable. We have

$$\frac{\partial \eta}{\partial d_t} = \frac{P(H_0)(1-\tau/T_c)\left(1 - Q\left(\alpha + \sqrt{\tau f_s}\gamma\right)\right)}{\Xi^2} \cdot\left[\frac{2K^2\Xi}{\ln 2 d_t\left(K^2 + d_t^2\right)} - \frac{C_0 2c_2 v^3}{g^2\left(d_g - d_t\right)^3}\right]. \tag{23}$$

Define

$$G = \frac{2K^2\Xi}{d_t\left(K^2 + d_t^2\right)\ln 2} - \frac{C_0 2c_2 v^3}{g^2\left(d_g - d_t\right)^3}, \tag{24}$$

when d_t is close to 0, Ξ is close to a positive constant and C_0 is close to 0. Then, we have

$$\lim_{d_t\to 0} G = +\infty. \tag{25}$$

1: Input
$\rho, W, K, P_t, T_c, \sigma, v_{\max}$.
2: Let $d_t = d_g$ and $d_t{}' = d_g - \Delta$.(Δ is a number small enough).
3: According to algorithm 2 to obtain τ^*.
4: Use d_t and $d_t{}'$ obtain the v_0 and $v_0{}'$ according to equation (16).
5: If $v_0 < v_{\max}$
$v^* = v_0$
else
$v^* = v_{\max}$.
6: If $\eta(d_t, v^*, \tau^*) < \eta'(d_t{}', v^{*'}, \tau^{*'})$,let $d_t = d_t{}', d_t{}' = d_t - \Delta$.
7: Repeat step 3 and 4 calculate until $\eta(d_t, v^*, \tau^*) > \eta'(d_t{}', v^{*'}, \tau^{*'})$.
8: Record $d_t{}^* = d_t$ and $\eta^* = \eta$.
9: Output
$(d_t{}^*, v^*, \tau^*, \eta^*)$

Algorithm 1: Algorithm for energy efficiency maximization problem.

1: Initialize interval $(0, T_c)$ as (a_1, b_1), tolerance level μ and the iterative number N.
2: Calculate the result of $\tilde{R}$ according to the equation (9).
3:$k = 1, \lambda_1 = a_1 + 0.382(b_1 - a_1), \mu_1 = a_1 + 0.618(b_1 - a_1)$.
4: While $b_k - a_k > \mu$ or $k < N$
if $\tilde{R}(\lambda_k) < \tilde{R}(\mu_k)$
$a_{k+1} = \lambda_k, b_{k+1} = b_k$
$\lambda_{k+1} = a_{k+1} + 0.382(b_{k+1} - a_{k+1})$
else
$a_{k+1} = a_k, b_{k+1} = u_k$
$\mu_{k+1} = a_{k+1} + 0.618(b_{k+1} - a_{k+1})$
$k = k + 1$
end while
$\tau^* = (b_k + a_k)/2$
5: Update d_t to satisfy the constraints (14), repeat 3, 4 and 5 steps to get τ^*.

Algorithm 2: Subalgorithm for Algorithm 1.

Let

$$G_1 = \frac{2K^2\left(c_1 v^3 + c_2 v\right)}{d_t\left(K^2 + d_t{}^2\right)\ln 2},$$
$$G_2 = \frac{2c_2 v^3\left[K^2\left(d_g - d_t\right) - C_0 d_t\left(K^2 + d_t{}^2\right)\ln 2\right]}{d_t\left(K^2 + d_t{}^2\right) g^2\left(d_g - d_t\right)^3 \ln 2}; \quad (26)$$

equation (24) can be written as

$$G = G_1 + G_2. \quad (27)$$

Since $\lim_{d_t \to d_g} G_2 = -\infty$ and $\lim_{d_t \to d_g} G_1 = 2H^2(c_1 v^3 + c_2 v)/d_g$ $(K^2 + d_g{}^2)\ln 2$, thus

$$\lim_{d_t \to d_g} G = -\infty. \quad (28)$$

Obviously, $\Xi^2 > 0$. Therefore, $\lim_{d_t \to 0} \partial\eta/\partial d_t = +\infty$ and $\lim_{d_t \to d_g} \partial\eta/\partial d_t = -\infty$. It is shown that η increases when d_t is small and decreases when d_t approaches d_g. By calculating the limit of η, we have $\lim_{d_t \to 0} \eta = 0$ and $\lim_{d_t \to d_g} \eta = 0$. Hence, there is a maximum point of $\eta(d_t)$ within interval $(0, d_g)$. Define this optimal d_t as $d_t{}^*$, which can be efficiently found numerically by iterative search. And the maximum energy efficiency is obtained when $d_t = d_t{}^*$. After obtaining

Table 1: Parameters in numerical analysis.

Parameter	Value	Parameter	Value
$P(H_0)$	0.8	$\rho\ (1/\mathrm{m}^2)$	0.001
d_g (m)	2000	σ^2 (dBm)	-150
T_c (s)	1	K (m)	200
P_t (w)	1	β_0 (dB)	-50
W (MHz)	6	c_1	9.26×10^{-4}
c_2	2250	γ (dB)	-20
$g\ (\mathrm{m/s}^2)$	9.8		

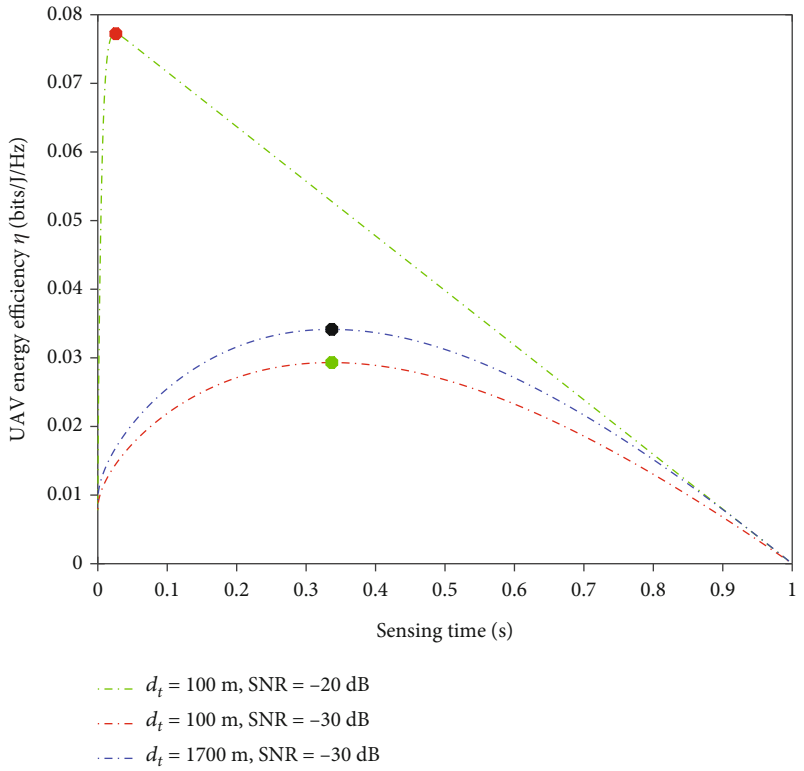


Figure 4: Energy efficiency vs. sensing time: $T_c = 1$ s.

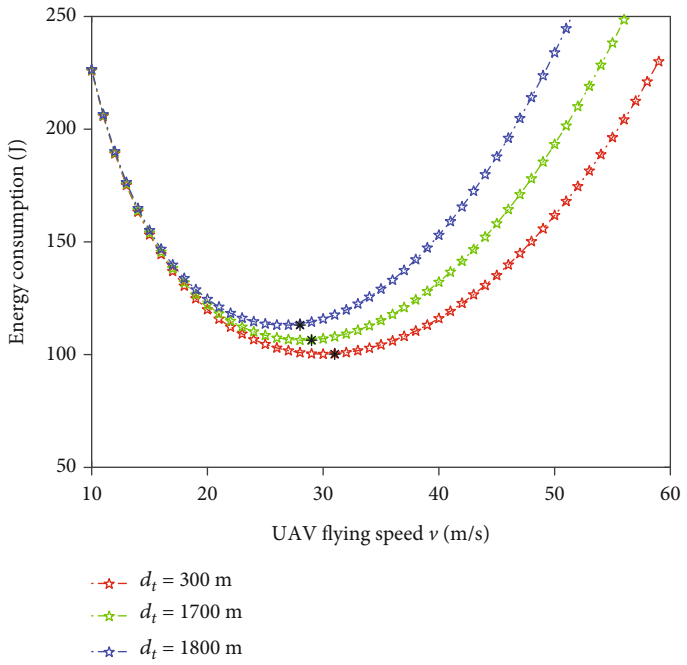


Figure 5: UAV energy consumption vs. UAV speed under different d_t.

the optimal sensing time and optimal speed, exhaustive search is used to obtain optimal d_t. Then, the algorithm to solve the problem OP is shown as follows.

For each given d_t, the computation complexity of Algorithm 2 is β, where $\beta = \lceil \log_2 T_c/\mu \rceil$ is the number of iterations that the Golden section method takes to terminate and $\lceil (\cdot) \rceil$ represents the smallest integer not less than $(\cdot)$ [24]. The computation complexity to find optimal d_t is θ, where $\theta = d_g/\Delta$. Therefore, the overall complexity of Algorithm 1 is $\beta\theta$.

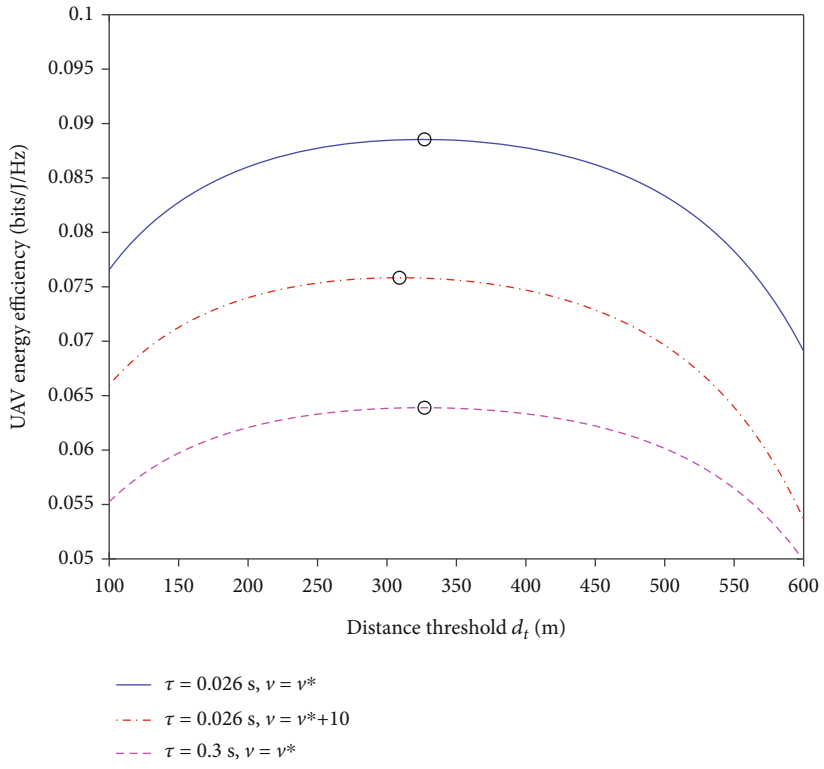


Figure 6: UAV energy efficiency vs. distance threshold.

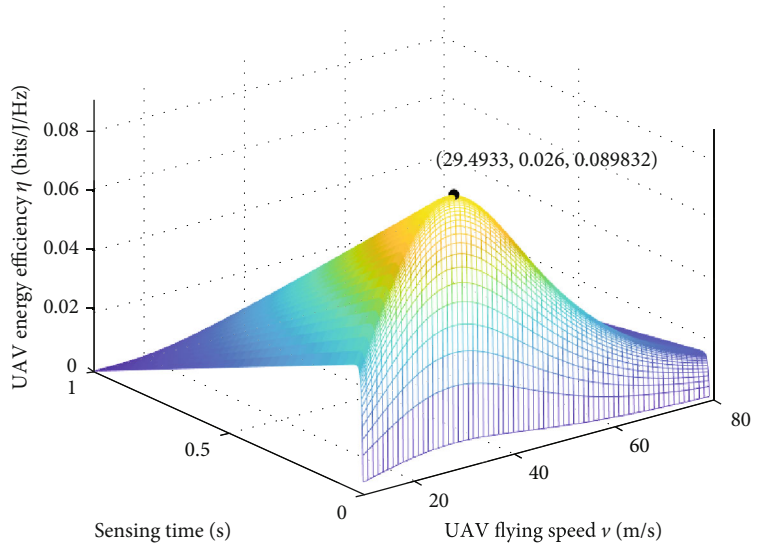


Figure 7: Three-dimensional map of η as a function of v and τ: $d_t = 300$ m.

4. Numerical Results

In this section, we verify the performance of our proposed design through simulations. We set the speed of the UAV between 10 and 80. Choose $\bar{P}_d = 0.9$, which is in IEEE802.22 WRAN. Suppose that the central frequency of ground base station is 2.4 GHz. The specific setting of remaining parameters is shown in Table 1.

Figure 4 shows that the energy efficiency varies with the sensing time τ. We can intuitively see that when d_t is set to a fixed value, such as $d_t = 100$ and $d_t = 1700$, there is an optimal τ to maximize the EE η. It can be seen that the value of d_t can determine the value of η, which means η can be affected by the semantic description of UAV's communication coverage range. Moreover, in the case of the same d_t, τ^* under the condition of low receiving SNR is longer than τ^* when receiving SNR is high, which means that it needs longer sensing time to improve the detection probability. Therefore, it is effective to set $\tau = \tau^*$ in T_c to improve the EE.

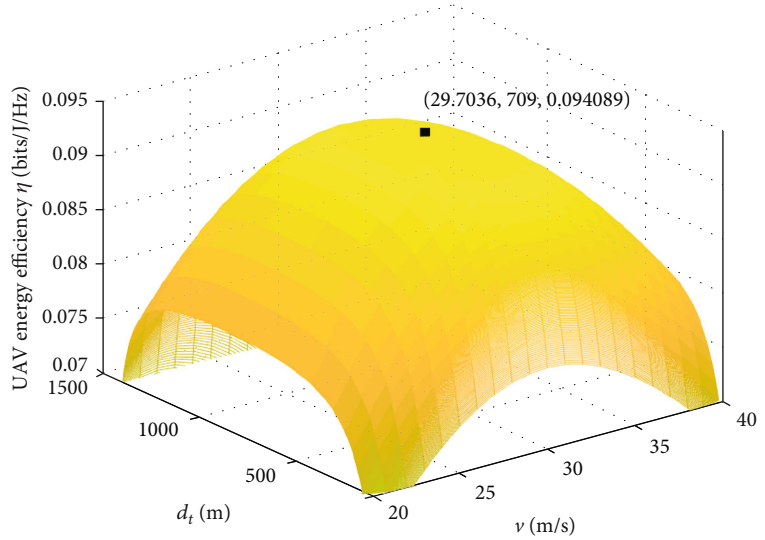


Figure 8: Three-dimensional map of η as a function of v and d_t.

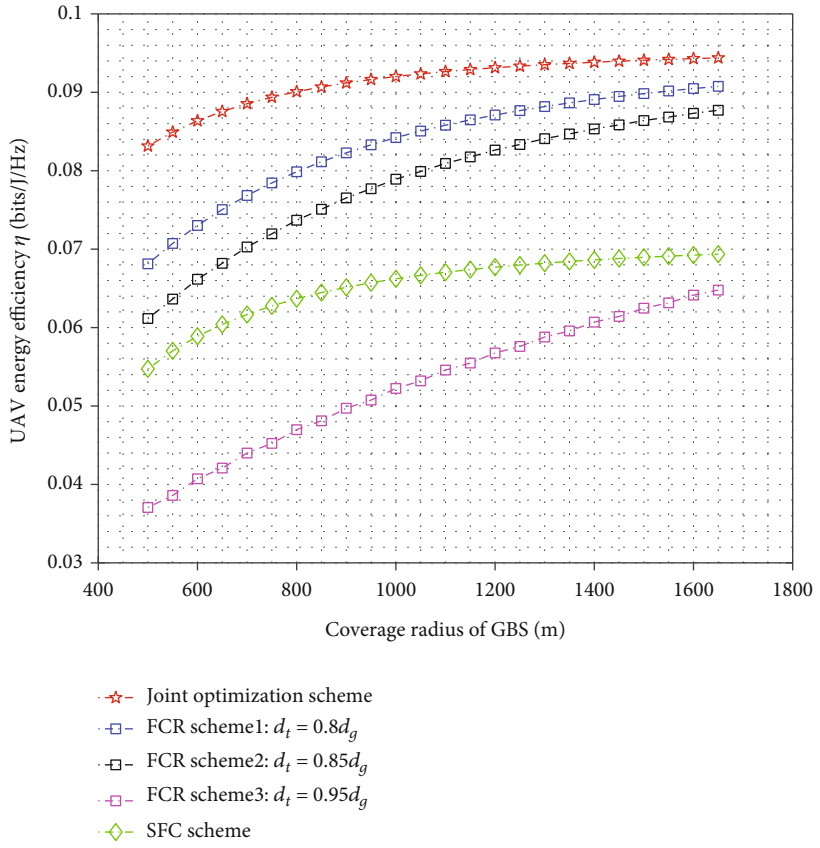


Figure 9: Comparison to FCR scheme and SFC scheme.

Figure 5 plots how energy consumption changes as a function of the UAV flying speed, which are under three different d_t. It shows that the energy consumption drops sharply when the speed is between 10 and 20 and increases sharply when the speed is greater than 40. Obviously with the increase of speed, the time of UAV flying over the same distance will be reduced, so it is reasonable to reduce the energy consumption of UAV. However, the increase of UAV speed is bound to be accompanied by the increase of energy consumption. Hence, it is necessary to set a suitable value of v.

From Figure 5, we also notice that the energy consumption increases with the increase of d_t. From Figure 4, we get that d_t also affects throughput; aiming to maximize EE, d_t

must be set reasonably. Therefore, we simulate the EE curves with d_t, as shown in Figure 6. It can be seen that the maximum η can be achieved if the d_t is set properly, when v and τ are fixed. As can be seen from Figure 2, τ^* is equal to 0.026 when the received SNR is -20 dB. It is obvious from Figure 6 that if the sensing time is equal, η is better when $v = v^*$. And if the flying speed is equal, η is better when $\tau = \tau^*$.

When communication distance threshold $d_t = 300$ m, Figure 7 shows the effects of τ and v on η. It indicates that there exists an optimal speed $v^* \approx 29\,(\mathrm{m/s})$ and $\tau^* = 0.026\,(\mathrm{s})$, which maximizes η and reaches $\eta \approx 0.08\,(\mathrm{bits/J/Hz})$. At the same time, Figure 7 also shows that there is an optimal point to maximize η; it means that there is an optimal cognitive communication scheme for UAV to obtain higher energy efficiency in the IoT. In addition, we can observe from Figures 7 and 8 that a slice parallel to the x-axis shows how the EE changes with v; the slice parallel to the y-axis in Figures 7 and 8, respectively, shows that EE varies with τ and d_t, which are consistent with the above analysis results.

Figure 9 shows a comparison between the proposed scheme and other two schemes. In the fixed communication range (FCR) scheme, the communication radius of UAV is a fixed value which can achieve higher throughput. The shortest flight cycle (SFC) scheme requires the UAV to fly at maximum speed. In this scheme, the time of energy consumption is reduced as much as possible. It can be seen that the energy efficiency decreases more and more sharply as d_t approaches the value of d_g in the FCR scheme. From Section 3, it is concluded that the average throughput will increase with the increase of d_t. Thus, for the same coverage radius of GBS, the joint optimization scheme obtains lower throughput. However, the energy efficiency of the proposed scheme is much better than FCR and SFC schemes. The result also indicates that the optimal energy efficiency can be obtained by applying the joint optimization scheme to the IoT with different parameters.

5. Conclusions

Based on the consideration of improving spectrum efficiency in semantic IoT, this paper mainly studies the process of UAV opportunity using the spectrum of GBS to communicate with edge nodes. A semantic device model for cognitive UAV-assisted IoT communication is constructed to ensure the automation of the IoT system. Aiming to solve the problem of UAV's limited energy supply, we propose a joint optimization algorithm to maximize the communication energy efficiency of UAV under certain constraints. In order to understand the model better, the experimental simulations are carried out. The simulation results are consistent with the theoretical analysis. It is also shown that the maximum energy efficiency for cognitive UAV in IoT can be obtained by the proposed algorithm. At the same time, the UAV's communication coverage range d_t of the semantic model will greatly influence the energy efficiency. And the performance of the joint optimization scheme is better than the fixed trajectory flight scheme with higher throughput.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

There is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under grant 61901509, in part by the Natural Science Foundation of Shaanxi Province of China under grant no. 2018JM6098, in part by the Research Foundation for Talented Scholars of Xijing University under grant no. XJ17B06, and in part by the President Foundation of Air Force Engineering University under grant XZJK2019033.

References

[1] J. Wang, G. Wang, B. Li, H. Yang, and A. Schmeink, "Massive MIMO two-way relaying dystems with SWIPT in IoT networks," *IEEE Internset ofsss Things Journal*, 2020.

[2] Y. Xu, B. Li, N. Zhao, Y. Chen, and X. Wang, "Coordinated direct and relay transmission with NOMA and network coding in Nakagami-m fading channels," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 207–222, 2021.

[3] A. Palavalli, D. Karri, and S. Pasupuleti, "Semantic Internet of Things," in *2016 IEEE Tenth International Conference on Semantic Computing*, pp. 91–95, Laguna Hills, CA, USA, Febuary 2016.

[4] J. Kiljander, A. D'elia, F. Morandi et al., "Semantic interoperability architecture for pervasive computing and Internet of Things," *IEEE Access*, vol. 2, pp. 856–873, 2014.

[5] G. Xiao, J. Guo, L. D. Xu, and Z. Gong, "User interoperability with heterogeneous IoT devices through transformation," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1486–1496, 2014.

[6] J. Takalo-Mattila, J. Kiljander, F. Pramudianto, and E. Ferrera, "Architecture for mixed criticality resource management in Internet of Things," in *Proceedings of 2014 TRON Symposium (TRONSHOW)*, pp. 1–9, Tokyo, Japan, December 2014.

[7] S. Chun, S. Seo, B. Oh, and K. Lee, "Semantic description, discovery and integration for the Internet of Things," in *Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing*, pp. 272–275, Anaheim, CA, USA, Febuary 2015.

[8] B. Li, C. Chen, R. Zhang, H. Jiang, and X. Guo, "The energy-efficient UAV-based BS coverage in air-to-ground communications," in *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 578–581, Sheffield, UK, July 2018.

[9] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Optimal transport theory for power-efficient deployment of unmanned aerial vehicles," in *2016 IEEE International Conference on Communications*, pp. 1–6, Kuala Lumpur, Malaysia, May 2016.

[10] Y. Zeng, R. Zhang, and T. J. Lim, "Throughput maximization for UAV-enabled mobile relaying systems," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 4983–4996, 2016.

[11] J. Zhang, Y. Zeng, and R. Zhang, "Spectrum and energy efficiency maximization in UAV-enabled mobile relaying," in *2017 IEEE International Conference on Communications*, pp. 1–6, Paris, France, May 2017.

[12] A. Athar, M. H. Rehmani, and A. Rachedi, "When cognitive radio meets the Internet of Things?," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 469–474, Paphos, Cyprus, September 2016.

[13] B. Zhang and K. Chen, "Selective spectrum leasing in internet of things via Nash bargaining solutions," in *2012 IEEE 2nd International Conference on Cloud Computing and Intelligence Systems*, pp. 841–845, Hangzhou, China, October-November 2012.

[14] X. Liu, M. Guan, X. Zhang, and H. Ding, "Spectrum sensing optimization in an UAV-based cognitive radio," *IEEE Access*, vol. 6, pp. 44002–44009, 2018.

[15] B. Li, J. Yang, H. Yang, G. Liu, R. Ma, and X. Peng, "Decode-and-forward cooperative transmission in wireless sensor networks based on physical-layer network coding," *Wireless Networks*, 2019.

[16] N. Zhao, F. Cheng, F. R. Yu et al., "Caching UAV assisted secure transmission in hyper-dense networks based on interference alignment," *IEEE Transactions on Communications*, vol. 66, no. 5, pp. 2281–2294, 2018.

[17] F. Cheng, S. Zhang, Z. Li et al., "UAV trajectory optimization for data offloading at the edge of multiple cells," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6732–6736, 2018.

[18] H. C. Chen, H. Kung, D. Vlah, D. Hague, M. Muccio, and B. Poland, "Collaborative compressive spectrum sensing in a UAV environment," in *2011 - MILCOM 2011 Military Communications Conference*, Baltimore, MD, USA, November 2011.

[19] P. Rawat, K. D. Singh, and J. M. Bonnin, "Cognitive radio for M2M and internet of things: a survey," *Computer Communications*, vol. 94, pp. 1–29, 2016.

[20] Y. Gao, H. Tang, B. Li, and X. Yuan, "Robust trajectory and power control for cognitive UAV secrecy communication," *IEEE Access*, vol. 8, no. 6, pp. 49338–49352, 2020.

[21] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, 2008.

[22] H. Hu, Y. Huang, X. Da, L. Ni, and Y. Pan, "Optimization of energy management for UAV-enabled cognitive radio," *IEEE Wireless Communications Letters*, vol. 9, no. 9, pp. 1505–1508, 2020.

[23] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747–3760, 2017.

[24] S. Boyd and L. Vandenberghe, *Convex Optimization*, Carmbridge University Press, 2013.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6651145, 8 pages
https://doi.org/10.1155/2021/6651145

Research Article

Design of English Diagnostic Practice Sentence Repetition Recognition System Based on Matching Tree and Edge Computing

Tanping Xi

Foreign Language School, Shanxi Xueqian Normal University, Xi'an 710100, China

Correspondence should be addressed to Tanping Xi; 40001@snsy.edu.cn

Received 25 December 2020; Revised 10 January 2021; Accepted 18 January 2021; Published 9 February 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Tanping 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.

English reading ability is an important indicator to measure learners' English ability. However, because reading ability cannot be directly observed, people usually take tests to judge the reading ability of learners. Therefore, it is very necessary to design a reasonable diagnostic practice sentence repetition recognition system to analyze and test, inform learners of the advantages and disadvantages in reading, and give corresponding countermeasures. In order to properly solve the problem of repeated recognition of English diagnostic sentences, we have developed a new recognition system combined with matching tree and edge computing technology. First, the matching tree algorithm is used for the repetitive diagnosis of English sentences. The algorithm has achieved good results in the repetitive diagnosis and matching. Secondly, an English diagnostic practice sentence repetition recognition system architecture is built through edge computing algorithms, which improves the efficiency of the English diagnostic system. Finally, through the simulation test of the English diagnostic practice system, the applicability of the established repeated recognition model is verified.

1. Introduction

Testing is an essential part of language teaching. According to different purposes, language tests can be divided into diagnostic tests, proficiency tests, ability tests, placement tests, etc. Compared with other types of tests, the relationship between diagnostic tests and teaching is closer [1, 2]. If properly designed, it can provide teachers with detailed feedback information, which is of great significance to language teachers and learners [3, 4]. However, due to the existence of test-oriented education, the development of diagnostic tests has not attracted enough attention [5]. Especially because of the existence of test-oriented education, the development of diagnostic testing has been neglected, so its development is particularly important. In addition, listening ability has always been an important part of language ability [6]. With the development of computers and the Internet, in various public tests at home and abroad, many listening tests are trying to use online listening tests. Traditional listening tests are gradually being replaced by online listening tests [7, 8].

At present, most college students lack the ability to self-study under the influence of self-study, traditional classroom teaching, and test-oriented education mode. And the dependence on teachers is still very strong. The contradiction between the increasing number of students and the shortage of teaching resources makes students urgently need to improve their self-learning ability [9, 10]. The improvement of computer network technology and teaching facilities has injected new vitality into college English teaching and learning. With its unique novelty, interactivity, and visibility, it attracts all students in the university and provides unprecedented advantages [11, 12]. It is a computer-based Internet technology requirement (trial version) proposed in the college English course, student-centered autonomous learning, focusing on general English ability, especially listening and speaking ability, emphasizing the importance of formative evaluation develop in the direction of learning personality and self-direction, teaching students to establish a dominant position in the process [13, 14]. At present, many universities in China have used broadband transmission technology to establish campus networks, thereby providing favorable conditions for universities to train foreign language talents. How will powerful media or the Internet serve language teaching on the Internet? One of the possible methods is online

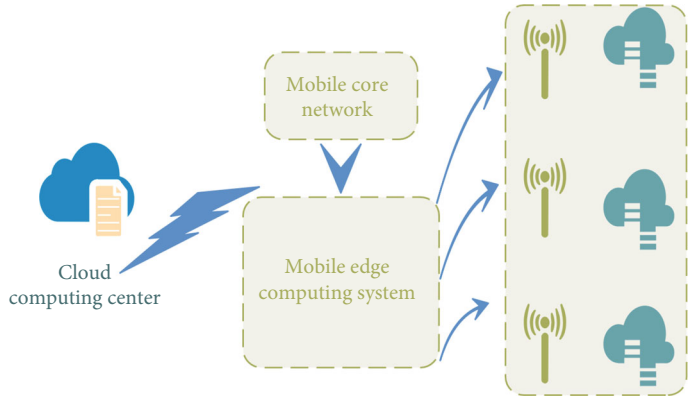


Figure 1: Schematic diagram of the value chain model of mobile edge computing.

testing. The online English test supplements the traditional language test, which is also a developmental language test of the traditional language. The organic combination of the two will communicate with each other in modern EFL teaching. Language testing is an important aspect of teaching in the learning process [15].

Previous studies mostly focused on analyzing the problems of learners in reading and did not put forward the advantages of learners in reading. At the same time, there are very few studies on these problems that give learners corresponding feedback. Therefore, it is very necessary to design a reasonable diagnostic practice sentence repetition recognition system to analyze and test, inform learners of the advantages and disadvantages in reading, and give corresponding countermeasures. In order to properly solve the problem of repeated recognition of English diagnostic sentences, we have developed a new recognition system combined with matching tree and edge computing technology. The matching tree algorithm and edge computing technology will be used in the model establishment and architecture design of the English diagnostic practice system. The improvement of computer network technology and teaching facilities has injected new vitality into college English teaching and learning. With its unique novelty, interactivity, and visibility, it attracts all students in the university and provides unprecedented advantages.

2. Overview of Related Technologies

2.1. Edge Computing. Traditional cloud computing can solve the core computing tasks of the network. However, with the continuous increase of data in the cloud, the transmission delay, and calculation delay of the cloud will be challenged [16, 17]. The fundamental problem of cloud computing is the higher service delay and intermittent connection between mobile devices and cloud servers, which may not be able to meet the real-time services of different emerging applications, such as augmented reality and online traffic monitoring systems [18, 19]. Therefore, with the advancement of information technology, data communication, sensing, processing, and control are transforming the transportation system from a traditional technology-driven system to a more powerful data-driven intelligent transportation system. Edge computing is a computing architecture that allows execution at the edge of the network [20]. It needs to sink the work of server computing to the edge of the network. Generally, there are three roles for devices in edge computing: terminal devices, edge servers, and cloud servers. Terminal devices are generally devices with network communication capabilities at the edge of the network. For example, for individual users, the terminal devices are mobile phones, laptops, and tablets. A schematic diagram of the value chain model of mobile edge computing is shown in Figure 1.

Mobile edge computing helps to realize the edge cache function and improve the current low efficiency of mobile content distribution. Based on the MEC server, wireless analysis applications can be deployed at the edge of the wireless access network to provide content servers with wireless environment information (such as real-time throughput of wireless links) [21, 22]. Using this type of information, the content server can perform more reasonable control over content transmission and content encoding (for example, online video) and match the transmission content with the wireless link capacity.

2.2. English Diagnostic Exercises. Since word recognition and language comprehension capabilities make important meaning independent of their contribution to reading ability, based on this research, the speech efficiency theory (VET) proposed by Perfetti in 1985 was adopted [23]. Perfetti's model emphasizes the importance of low-level accuracy and rapid word recognition and high-level language comprehension, which are part of improving reading progress. It explains that Automatism Theory (AT) reading is a complex process, in which simultaneous word recognition and obsessive-compulsive disorder can only occur when they show comprehensive cognitive needs. The document will not exceed the reader's available resources [24]. Interaction between English listening and speaking-source is shown in Figure 2.

Knowledge is divided into two categories: language knowledge and other related knowledge (social knowledge, personal experience, and background) and psychological activities of listening comprehension [25]. If a learner tries to improve his level, foreign language listening comprehension must go through three stages, namely voice recognition,

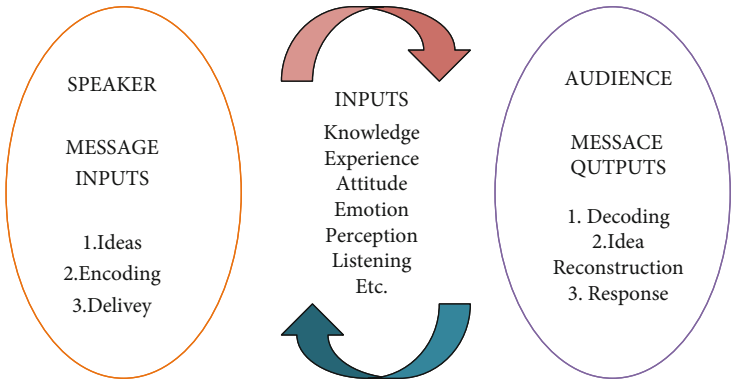


Figure 2: Interaction between English listening and speaking-source.

sentence comprehension, and paragraph comprehension. Rivers believes that foreign language listening comprehension includes two levels of speech activity recognition and selection. It has a guiding guide for understanding listening comprehension and development. Used in listening comprehension testing and teaching. For beginners' listening comprehension, test more materials related to sound recognition. For intermediate or advanced personnel, choose the sentence in the listening test [26].

Only when readers recognize words quickly and automatically without using a lot of attention resources, they can make full use of most of their attention resources to achieve successful reading ability. Although VET also focuses on the automation of decoding, this is different from AT [27]. VET expands the concept of the underlying decision-making process. They provide interesting reprinters of the relationship between word recognition and reading comprehension with different perspectives. According to Tracy and Morrow, VET is based on three assumptions: the learner's reading of a printed text is regarded as their final hearing of the text; the time the reader spends reading an isolated word indicates the reader's understanding of the word, and the reader's decoding ability is the main source of determining its word recognition.

3. English Diagnostic Practice Repetitive Sentence Recognition Based on Matching Tree

The diagnostic evaluation module of the current system is the early design of the system, and does not consider aspects of learning status, knowledge point correlation analysis, question type correlation analysis, four-level score prediction, etc. [28].

The learning guidance for learners is very limited, and it cannot be personalized study suggestions. Therefore, this chapter first analyzes and designs the diagnostic evaluation model and then preprocesses the data collected when the system is used. The three dimensions of learning status evaluation, question type association analysis, and college English four-level score prediction are modeled separately. Finally, after combining these submodels, a relatively complete and reliable diagnostic evaluation model was obtained, and related verification was carried out [29].

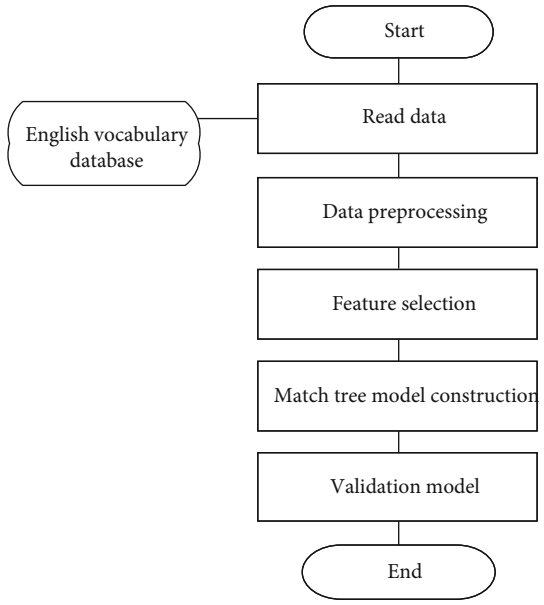


Figure 3: Recognition of repetitive sentences in English diagnostic exercises.

The matching tree algorithm first preprocesses all subscription collections into a tree. After the message arrives, it traverses the established tree to find matches between all messages and events. Its core is the matching tree. Figure 3 shows a tree with two subscriptions S1 and S2. The edge represents the test result, while the nonleaf node represents a test. The subscription conditions that need to be matched are associated with the leaf nodes.

$$f = f(x). \tag{1}$$

Constraints:

$$\begin{aligned} &x_j' \le x_j \le x_j' \quad (j = 1, 2, \cdots, n), \\ &g_j \le g_j' \quad (j = 1, 2, \cdots, n), \\ &h_j' \le h_j(x) \quad (j = 1, 2, \cdots, n), \\ &w_j' \le w_j(x) \le w_j' \quad (j = 1, 2, \cdots, n). \end{aligned} \tag{2}$$

X represents the test edge, which represents the subscription that can be reached through X, where the edge does not care about the test result. Such subscription is a predicate that contains leaf nodes.

$$Q(x,q)=\frac{f}{f_0}+\sum_j p_x(x_j)+q\left[\sum_j p_g\left(g_j\right)+\sum_j p_h(h_j)+\sum_j p_w(w_j)+\right], \tag{3}$$

in which p_x, p_g, p_h, and p_w are a penalty factor for constrained design variables and state variables. By using the gradient method for unconstrained optimization problems, the iterative formula is expressed as follows [30].

$$x^{(j+1)}=x^{(j)}+s_j d^{(j)}. \tag{4}$$

When algorithms are matched, in order to be able to quickly respond to query requests, we usually build indexes. Through the index, when responding to query requests, the system can quickly find the needs of users. The index establishes a mapping relationship between information and content storage location information. The information index is divided into the forward index and inverted index [31, 32].

$$\begin{aligned}\left|f^{(j)}-f^{(j-1)}\right|&\le\tau,\\ \left|f^{(j)}-f^{(b)}\right|&\le\tau.\end{aligned} \tag{5}$$

The forward index is expressed in the form of "document-keyword," and the position information of each keyword in the document is recorded in the table.

$$D(v)=\sum C\left(v_i,v_j\right),\ \text{where } v_i,v_j\in V. \tag{6}$$

When searching, scan the words in each document in turn according to the document number, and output all matching keywords. The establishment process and organization of the positive index are relatively simple and easy to maintain. However, all matching conditions must be scanned when querying, which results in a long response time and low efficiency.

$$D(v_0)\le D(v_1)\le D(v_2)\le\cdots\le D(v_n). \tag{7}$$

The design idea of the inverted index is the catalog, and we can quickly find the specific location of the content through the catalog. The structure of the directory is <key, address>. In order to be able to match quickly, we set a keyword whose structure is the directory of <keyword, addres>, which we call an inverted index.

$$D(v_{n+1})=\min\left\{D(v_n)+C(v_i,w)\right\},\ \text{where } v_i\in U, w\in V-U. \tag{8}$$

In the inverted index, the word in the document is used as the "directory" keyword, and the document number where the word appears is used as the storage address, that is, <word, key>. Generally, a word will appear in multiple documents or one or more times in a document. Therefore, the address in the "directory" is a list of document numbers. Each item in the list contains information about all the locations where the word appears in the document.

4. Case Verification of Repetitive Sentence Recognition in English Diagnostic Exercises

Based on the fast-matching algorithm of multilevel constraint search tree, this paper designs the matching mechanism of pub/subsystem based on inverted index. The fast-matching algorithm of multilayer constraint search tree reduces repeated matching through constraint coverage when one event matches multiple subscription conditions. However, there are still many repeated matching problems when multiple events match multiple subscription conditions. In order to solve this problem, by introducing an inverted index based on a multilayer constrained search tree, an efficient matching mechanism suitable for the parallel interaction of large-scale publishers and subscribers is constructed. The inverted index is used to analyze the coverage relationship between subscription conditions and events, which reduces the number of matching subscription conditions and events and further improves matching efficiency.

4.1. Technical Performance Test. At present, due to the existence of test-oriented education, the development of diagnostic tests has not attracted enough attention. Especially because of the existence of test-oriented education, the development of diagnostic testing has been neglected, so its development is particularly important. In addition, the listening ability has always been an important part of language ability. With the development of computers and the Internet, in various public tests at home and abroad, many listening tests are trying to use online listening tests, and traditional listening tests are gradually being replaced by online listening tests.

In Figure 4, the inverted index and the multilevel constraint search tree fast matching algorithm are applied to a large-scale subscription condition matching model. The specific construction algorithm of the model is as follows: (C1) Statistical analysis of the subscription conditions in the system, using the inverted index. The algorithm constructs an inverted index structure based on "constraint-subscription," which is aimed at quickly tracing back to the subscription conditions through constraints. (C2) Statistically analyze the events in the system and use the inverted index algorithm to construct an inverted index structure based on "attribute-event," aimed at quickly tracing back to events through attributes.

4.2. Comparison of Matching Tree Algorithm and Other Algorithms. Figure 5 shows the relationship between time delay and energy loss. Different θ has different proportions of time delay and energy loss. As the value of θ continues to increase, the proportion of time delay in the main problem

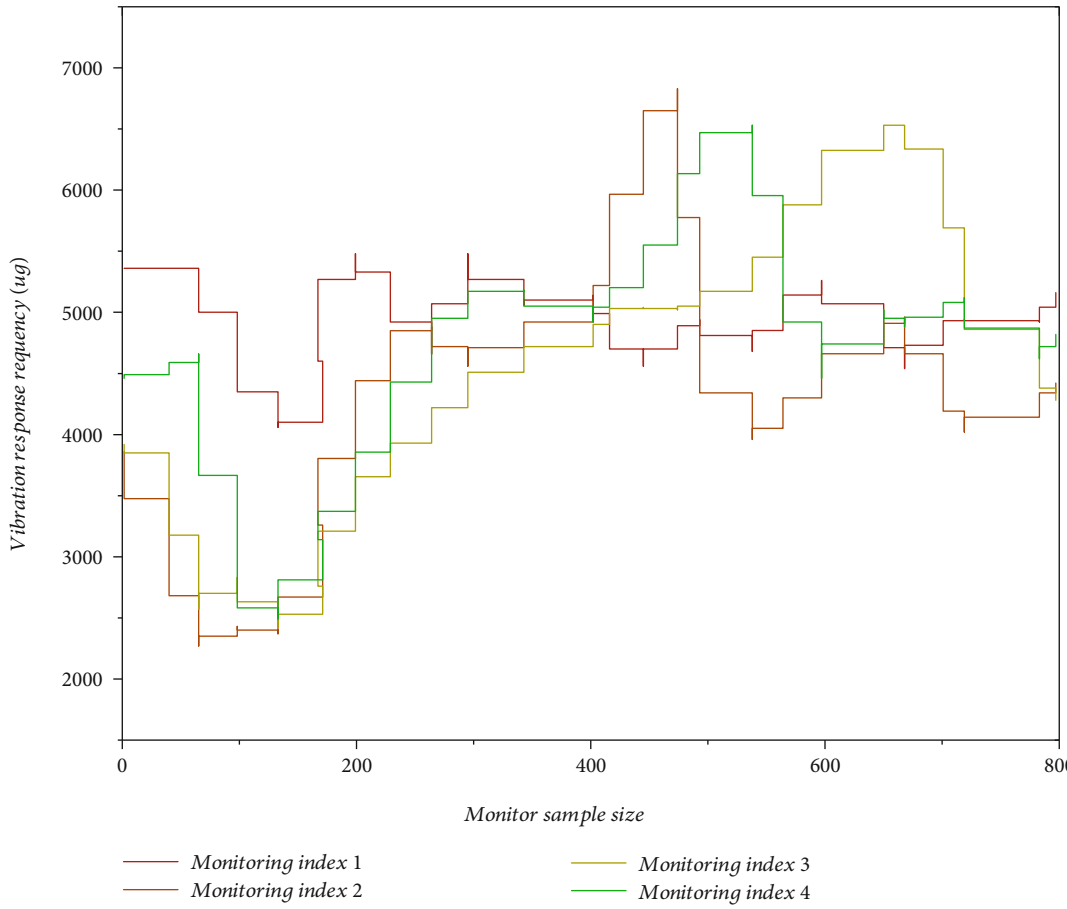


Figure 4: English diagnostic exercise matching model test based on an inverted index.

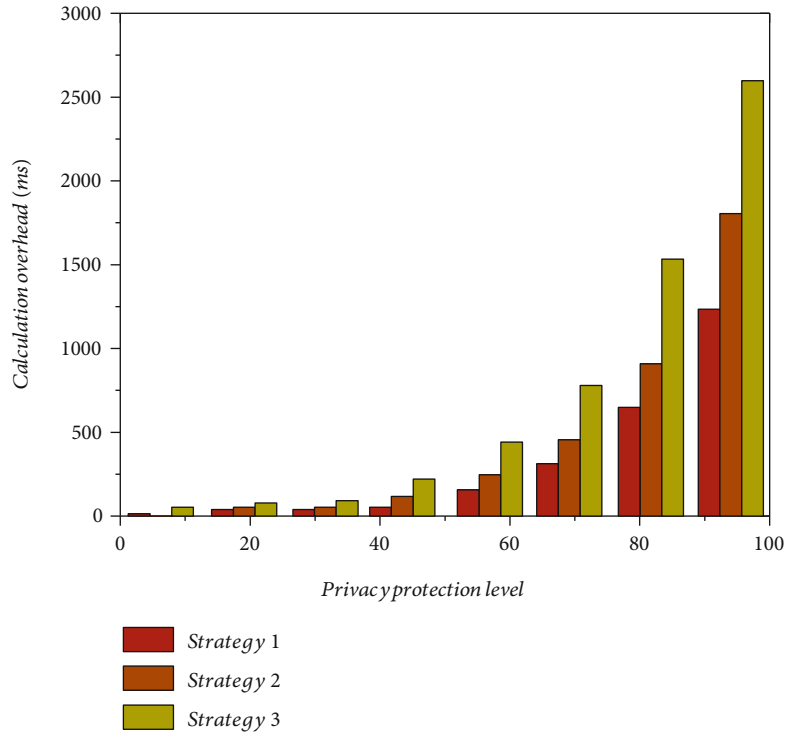


Figure 5: Comparison of the delay of matching tree algorithm with other algorithms.

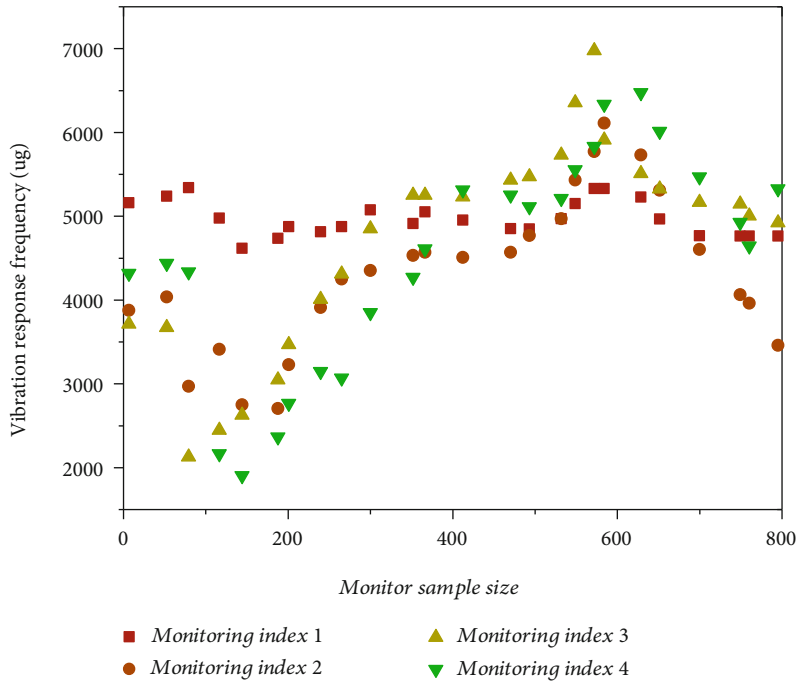


Figure 6: Comparison of edge computing power consumption for English diagnostic exercises.

continues to increase, and the proportion of energy loss continues to decrease.

When $\theta = 0$, it means that the time delay problem is not considered in the unloading process, so the value of the objective function depends largely on the energy loss value, which is an energy-optimal unloading algorithm. When $0 = 1$, it means that we do not consider the energy loss during the unloading process and focus on the delay problem, but it will still produce energy loss to the terminal equipment. The objective function value not only depends on the delay but also considers the energy loss. Therefore, it can be found from Figure 5 that energy loss and time delay are inevitable during the unloading process.

Figure 6 shows the relationship between the power consumption comparison of edge computing in English diagnostic exercises. This paper studies the optimal task offloading strategy of delay and energy consumption and deploys an effective priority task scheduling algorithm on the edge server, targeting vehicles. For the complex and diverse messages generated by the terminal, we adopt a message priority division algorithm, so that emergency messages of different priority levels can be processed in different ways. Therefore, the emergency messages of each priority level have different processing delays.

5. Conclusion

The results of the research prove that compared with the traditional diagnostic online hearing test, it can play a timelier diagnostic role, so that it can quickly and accurately feedback test information. Teachers can adjust teaching in time based on feedback information to improve teaching quality. Despite its limitations, the online diagnostic hearing test system is still a test method with great development and application prospects. English reading ability is an important indicator to measure learners' English ability. However, because reading ability cannot be directly observed, people usually take tests to judge the reading ability of learners. Therefore, it is very necessary to design a reasonable diagnostic practice sentence repetition recognition system to analyze and test, inform learners of the advantages and disadvantages in reading, and give corresponding countermeasures. The matching tree algorithm and edge computing technology will be used in the model establishment and architecture design of the English diagnostic practice system. The improvement of computer network technology and teaching facilities has injected new vitality into college English teaching and learning. With its unique novelty, interactivity, and visibility, it attracts all students in the university and provides unprecedented advantages. In the future, we will collect more experimental data to optimize the matching tree algorithm based on the diagnostic evaluation model, and improve the repeated recognition of English diagnostic contact sentences.

Data Availability

The authors approve that data used to support the finding of this study are included in the article.

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. Zhou, H. Liao, B. Gu, K. M. S. Huq, S. Mumtaz, and J. Rodriguez, "Robust mobile crowd sensing: when deep learning meets edge computing," *IEEE Network*, vol. 32, no. 4, pp. 54–60, 2018.

[2] M. Beller and E. Or, "The crossroads between lifelong learning and information technology a challenge facing leading universities," *Journal of Computer-Mediated Communication*, vol. 4, no. 2, 2010.

[3] T. M. Sobh and A. Tibrewal, "Parametric optimization of some critical operating system functions–an alternative approach to the study of operating systems design," *Journal of Stem Education Innovations & Research*, vol. 7, no. 3, pp. 354–369, 2006.

[4] J. Y. Kim, "A study of perceptional typologies on computer based assessment (CBA): instructor and student perspectives," *Journal of Educational Technology & Society*, vol. 18, no. 2, pp. 80–96, 2015.

[5] M. Wlodarczyk, "Clifford algebras meet tree decompositions," *Algorithmica*, vol. 81, no. 2, pp. 497–518, 2019.

[6] R. Wang, J. Yan, D. Wu, H. Wang, and Q. Yang, "Knowledge-centric edge computing based on virtualized D2D communication systems," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 32–38, 2018.

[7] K. Wang, H. Yin, W. Quan, and G. Min, "Enabling collaborative edge computing for software defined vehicular networks," *IEEE Network*, vol. 32, no. 5, pp. 112–117, 2018.

[8] D. A. J. P. van de Sande, A. Hoogeveen, J. Hoogsteen, and H. M. C. Kemps, "The diagnostic accuracy of exercise electrocardiography in asymptomatic recreational and competitive athletes," *Scandinavian Journal of Medicine & Science in Sports*, vol. 26, no. 2, pp. 214–220, 2016.

[9] A. J. Simpson, S. M. Drake, and S. J. Fowler, "Exercise-induced bronchoconstriction: a survey of diagnostic practice in secondary care across the United Kingdom," *Allergy*, vol. 75, no. 8, pp. 2130–2132, 2020.

[10] T. Ramiadantsoa, M. A. Stegner, J. W. Williams, and A. R. Ives, "The potential role of intrinsic processes in generating abrupt and quasi-synchronous tree declines during the Holocene," *Ecology*, vol. 100, no. 2, p. e02579, 2019.

[11] J.-Y. Feng, Y.-T. Chang, H.-Y. Chang, W. S. Erdley, C.-H. Lin, and Y.-J. Chang, "Systematic review of effectiveness of situated e-learning on medical and nursing education," *Worldviews on Evidence-Based Nursing*, vol. 10, no. 3, pp. 174–183, 2013.

[12] C. Prost-Squarcioni, F. Caux, E. Schmidt et al., "International Bullous Diseases Group: consensus on diagnostic criteria for epidermolysis bullosa acquisita," *British Journal of Dermatology*, vol. 179, no. 1, pp. 30–41, 2018.

[13] C. Pedlar, M. Brown, J. Otto et al., "Temporal sequence of athlete's heart regression during prescribed exercise detraining: diagnostic implications," *Journal of the American College of Cardiology*, vol. 69, no. 11, p. 1414, 2017.

[14] H. Liu, Y. Zhang, and T. Yang, "Blockchain-enabled security in electric vehicles cloud and edge computing," *IEEE Network*, vol. 32, no. 3, pp. 78–83, 2018.

[15] S. D. Bush, N. J. Pelaez, J. A. Rudd et al., "On hiring science faculty with education specialties for your science (not education) department," *CBE—Life Sciences Education*, vol. 5, no. 4, pp. 297–305, 2006.

[16] A. Lindsay and J. T. Costello, "Realising the potential of urine and saliva as diagnostic tools in sport and exercise medicine," *Sports Medicine*, vol. 47, no. 1, pp. 11–31, 2017.

[17] K. Kaur, S. Garg, G. S. Aujla, N. Kumar, J. J. P. C. Rodrigues, and M. Guizani, "Edge computing in the industrial internet of things environment: software-defined-networks-based edge-cloud interplay," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 44–51, 2018.

[18] K. Krüger, T. Reichel, and C. Zeilinger, "Role of heat shock proteins 70/90 in exercise physiology and exercise immunology and their diagnostic potential in sports," *Journal of Applied Physiology*, vol. 126, no. 4, pp. 916–927, 2019.

[19] N. N. Jiang, L. P. Wen, H. Li, and J. Yin, "A new diagnostic criteria of wheat-dependent, exercise-induced anaphylaxis in China," *Chinese Medical Journal*, vol. 131, no. 17, pp. 2049–2054, 2018.

[20] P. Jacobsson, W. D. Hamilton, G. Cook et al., "Refining the Hallstatt Plateau: short-term1 4C variability and small scale offsets in 50 consecutive single tree-rings from Southwest Scotland Dendro-dated to 510–460 BC," *Radiocarbon*, vol. 60, no. 1, pp. 219–237, 2018.

[21] B. Hu and T. Chomiak, "Wearable technological platform for multidomain diagnostic and exercise interventions in Parkinson's disease," *International Review of Neurobiology*, vol. 147, pp. 75–93, 2019.

[22] J. Y. Hogrel, J. B. E. Janssen, I. Ledoux et al., "The diagnostic value of hyperammonaemia induced by the non-ischaemic forearm exercise test," *Journal of Clinical Pathology*, vol. 70, no. 10, pp. 896–898, 2017.

[23] L. He, K. Ota, and M. Dong, "Learning IoT in edge: deep learning for the internet of things with edge computing," *IEEE Network*, vol. 32, no. 1, pp. 96–101, 2018.

[24] P. K. Gupta and H. Godara, "Bradycardia induced PVCS versus exercise induced PVCS: a diagnostic dilemma," *Journal of the American College of Cardiology*, vol. 73, no. 9, p. 2457, 2019.

[25] H. Guo and J. Liu, "Collaborative computation offloading for multiaccess edge computing over fiber–wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4514–4526, 2018.

[26] J. Cunha, Ú. Martins, D. Gomes, J. Matos, J. Moreira, and C. Aguiar-Branco, "P-45 Conservative treatment of traumatic extensor carpi ulnaris instability in a tennis player: case report," *British Journal of Sports Medicine*, vol. 50, Supplement 1, pp. A56.2–A5A57, 2016.

[27] X. Chen, Q. Shi, L. Yang, and J. Xu, "ThriftyEdge: resource-efficient edge computing for intelligent IoT applications," *IEEE Network*, vol. 32, no. 1, pp. 61–65, 2018.

[28] V. A. Chavez, C. A. Gilligan, and F. Bosch, "Variability in commercial demand for tree saplings affects the probability of introducing exotic forest diseases," *Journal of Applied Ecology*, vol. 56, no. 1, pp. 180–189, 2018.

[29] B. Brede, K. Calders, A. Lau et al., "Non-destructive tree volume estimation through quantitative structure modelling: comparing UAV laser scanning with terrestrial LIDAR," *Remote Sensing of Environment*, vol. 233, no. January, article 111355, 2019.

[30] T. Bharucha, A. Rutherford, S. Skeoch et al., "Diagnostic yield of FDG-PET/CT in fever of unknown origin: a systematic review, meta-analysis, and Delphi exercise," *Clinical Radiology*, vol. 72, no. 9, pp. 764–771, 2017.

[31] P. C. Balfour, J. A. Gonzalez, P. W. Shaw et al., "High-frequency QRS analysis to supplement ST evaluation in exercise stress electrocardiography: incremental diagnostic accuracy and net reclassification," *Journal of Nuclear Cardiology*, vol. 27, no. 6, pp. 2063–2075, 2020.

[32] Q.-H. Zhao, L. Wang, B. Pudasaini et al., "Cardiopulmonary exercise testing improves diagnostic specificity in patients with echocardiography-suspected pulmonary hypertension," *Clinical Cardiology*, vol. 40, no. 2, pp. 95–101, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6625184, 10 pages
https://doi.org/10.1155/2021/6625184

Research Article

Integrating Sensor Ontologies with Global and Local Alignment Extractions

Xingsi Xue,[1,2,3] **Xiaojing Wu**,[2] **Chao Jiang**,[2] **Guojun Mao**,[1] **and Hai Zhu**[4]

[1]*Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou, Fujian 350118, China*
[2]*Intelligent Information Processing Research Center, Fujian University of Technology, Fuzhou, Fujian 350118, China*
[3]*Guangxi Key Laboratory of Automatic Detecting Technology and Instruments, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China*
[4]*School of Network Engineering, Zhoukou Normal University, Zhoukou, Henan 466001, China*

Correspondence should be addressed to Xingsi Xue; jack8375@gmail.com

Received 21 December 2020; Revised 10 January 2021; Accepted 17 January 2021; Published 8 February 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Xingsi Xue 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 order to enhance the communication between sensor networks in the Internet of things (IoT), it is indispensable to establish the semantic connections between sensor ontologies in this field. For this purpose, this paper proposes an up-and-coming sensor ontology integrating technique, which uses debate mechanism (DM) to extract the sensor ontology alignment from various alignments determined by different matchers. In particular, we use the correctness factor of each matcher to determine a correspondence's global factor, and utilize the support strength and disprove strength in the debating process to calculate its local factor. Through comprehensively considering these two factors, the judgment factor of an entity mapping can be obtained, which is further applied in extracting the final sensor ontology alignment. This work makes use of the bibliographic track provided by the Ontology Alignment Evaluation Initiative (OAEI) and five real sensor ontologies in the experiment to assess the performance of our method. The comparing results with the most advanced ontology matching techniques show the robustness and effectiveness of our approach.

1. Introduction

With the wide range of applications and distribution of wireless communication systems, especially about the fifth-generation (5G) networks, the update speed technology in the Internet of things (IoT) is advancing by leaps and bounds [1, 2]. The popularity of IoT is increasing with artificial intelligence technology, which covers a wide range of applications [3–13]. As crucial components of IoT, sensor networks include connected sensors and associated devices, which deploy plenty of heterogeneous sensing nodes for capturing environmental data. On the one hand, due to their key features, i.e., autonomy, ease of deployment, and self-configurable and effective energy management, the sensor network has become a very active research field where an assortment of systems has been developed [14–17]. On the other hand, it cannot be ignored that the scarcity of integration and communication among these networks always separates critical data streams and exacerbates the existing problem of too much data and lack of knowledge [18]. For the purpose of solving this problem, the Semantic Web (SW) technology is used to date sensor networks to annotate sensor data with spatial, temporal, and topic semantic metadata. The communication and fusion between Semantic Web technology and sensor network give birth to the Semantic Sensor Web (SSW), which is dedicated to providing semantics for sensor perception, enabling advanced applications such as situational awareness of transformers to have interoperability and advanced analysis capabilities. The sensor ontology is combined with semantic annotation to integrate and share conceptual models, enhance the semantics of sensor data, and use different sensor ontologies to annotate

sensor data. Thus, it is essential to set up strong links between semantically concerning sensor information, which is the so-called sensor ontology matching [19].

Sensor ontology matching is aimed at determining the correspondences between heterogeneous entities that exist in two sensor ontologies, and the output sensor mapping set is defined as sensor ontology alignment [20]. Many works of experts, scholars, and institutions have been done to enhance the quality of ontology alignment. One of the most authoritative organizations is the Ontology Alignment Evaluation Initiative (OAEI), which is aimed at comparing ontology matching systems on precisely defined test cases [21]. These test cases can be composed of ontologies of varying complexity (from simple thesauri to expressive OWL ontologies) and use different evaluation modalities (e.g., blind evaluation, open evaluation, or consensus). The results of the OAEI show that there are still challenges in improving the quality of the matching work. More precisely, one of the critical issues in ontology matching is obtaining high-quality ontology comparisons. Thus, the way to pick, combine, and adjust distinct ontology matchers containing a specific matching technique has become a significant challenge [22]. Although plenty of techniques about machine learning have been proposed to determine the optimal aggregating weights for the matchers [23–29], the catch is that they pay little attention to the effects engendered by each entity mapping's preferences on different matchers, which decreases the quality of ontology alignment [30]. To overcome this drawback, this work proposes a novel sensor ontology alignment extracting method based on debating mechanism (DM) [31]. In our proposal, a semantic sensor network integration technique is used with the ontology alignment extraction technology. For the purpose of improving alignment's quality, we introduce the correctness factor of each matcher to determine a correspondence's global factor, and utilize the support strength and disprove strength in the debating process to calculate its local factor. Through comprehensively considering these two factors, the confidence of the judgment factor of an entity mapping is improved.

The rest of the paper is arranged as follows: Section 2 introduces the related concepts of sensor ontology and ontology alignment extraction; Section 3 describes in details the DM-based ontology alignment extracting method with global and local factors; Section 4 externalizes the experimental results and makes the corresponding analysis; Section 5 draws the conclusion.

2. Sensor Ontology and Ontology Alignment Extraction

2.1. Sensor Ontology. An ontology is considered to be the solution to data heterogeneity on the web, which formally specifies the domain concepts and their relationships to give the definition of domain knowledge [29, 32]. With the sensor ontology becoming a powerful tool for integration and dealing with heterogeneous, more and more sensor ontologies have appeared. The Semantic Sensor Network (SSN) (https://www.w3.org/TR/vocab-ssn) ontology was proposed by the W3C Semantic Sensor Network Incubator group (SSN-XG) as an OWL 2 ontology to shape sensors and observations, which describe sensors from various aspects, i.e., capabilities, measurement processes, and observations. OSSN (https://www.w3.org/2005/Incubator/ssn/wiki/SSN#Sensor) is an ontology developed by SSN-XG in 2005. The IoT-Lite ontology (http://www.w3.org/Submission/2015/SUBM-iot-lite-20151126) outlines the resources, entities, and services of the Internet of things (IoT), which is a lightweight ontology as well as an instance of the SSN ontology. The Sensor, Observation, Sample, and Actuator (SOSA) (http://www.w3.org/ns/sosa) ontology provides a lightweight kernel for SSN, which is designed to broad target audience and application areas which has access to ontologies. What is more, SOSA acts as a minimal interoperability fall-back level, i.e., it defines those common classes and properties for which data can be safely exchanged across all uses of SSN, its modules, and SOSA. SensorOntology2009 (https://www.w3.org/2005/Incubator/ssn/wiki/Sensor Ontology 2009) is an ontology developed by Michael Compton, Holger Neuhaus, and Nguyen Tran from CSIRO (Australia). It has been used as the initial version of the Semantic Sensor Network ontology.

A sensor ontology is defined as $O = (C, P, I)$, in which C, P, and I separately represent class set, property set, and instance set, respectively. A sensor ontology alignment A is an entity correspondence set where the entity correspondence inside is defined as a 4-tuples $\text{corr} = (e, e', n, \text{relation})$, e and e' are the entities of two ontologies, respectively, $n \in [0, 1]$ represents a similarity value between e and e', while relation is the equivalence relation. The process of determining the entity correspondence between sensor ontologies is called sensor ontology matching.

Figure 1 shows an example of two sensor ontologies and their alignment. As depicted in Figure 1, the symbol of a rectangle denotes the entity; the line with two arrows suggests the mapping between two entities. The ontology matcher takes two entities as input and outputs a real number that reflects the similarity in $[0, 1]$. The matching process cannot guarantee the comparative quality obtained by a single matcher, so multiple matchers usually need to work together to increase evidence of potential matches or discrepancies. All entity correspondences' similarity values obtained by an ontology matcher are restored in the corresponding similarity matrix, where the similarity values higher than the threshold are regarded as a correct mapping by this matcher.

2.2. Ontology Alignment Extraction. Various ontology matchers employ different information or features between two ontologies to measure the similarity values between two entities. Each ontology matcher has a particular similarity measure, divided into three categories, i.e., syntactic-based measures, linguistic-based measures, and structure-based measures [33]. To be specific, syntax-based similarity measures calculate two strings' similarity values according to their syntax information [34–36]. Linguistic-based similarity measure under the use of a background knowledge base, i.e., WordNet [37] and the work done by structure-based similarity measure is under the use of two concepts' parent-child concepts defined in their ontologies to do the work.

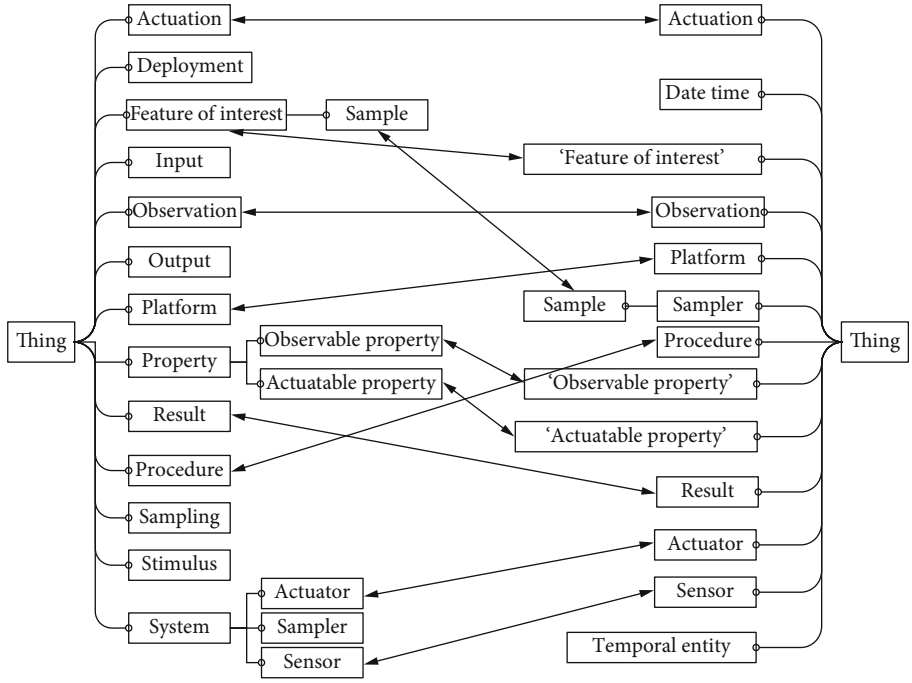


Figure 1: An example of sensor ontology matching.

Although these relevant methods have been extensively researched and taken into practice, it is still a challenge to satisfy the need for an effective ontology matcher on some specific matching task. With the emergence of various ontology matchers, ontology matching systems usually apply multiple matchers to enhance the alignment's quality [25, 38–40]. To better aggregate various ontology matchers, this work proposes a novel ontology alignment extracting method, whose framework is shown in Figure 2.

As shown in Figure 2, our approach is divided into two steps, i.e., various ontology alignment determination and DM-based alignment extraction. Preprocessing step generally includes ontology analysis and conversion ontology format. In this work, an argument ar is defined as follows:

$$ar = \{c, n, v, h\}, \tag{1}$$

where $c = (e, e')$ is a correspondence and v ($v \in N$) represents the artificially preset matcher number, and h ($h \in \{0, 1\}$) is the measure factor of similarity value determined as follows:

$$h = \begin{cases} 1, & \text{if } n \geq \delta, \\ 0, & \text{if } n < \delta, \end{cases} \tag{2}$$

where δ ($\delta \in [0, 1]$) is the threshold of similarity value. In particular, when $h = 1$, the matcher accepts c,otherwise rejects. Supposing k is the number of matchers, c might belong to one of the five categories C_i, $i = 1, \cdots, 5$, which are, respectively, defined as follows:

(i) If c is accepted by k matchers, $c \in C_1$

(ii) If c is rejected by k matchers, $c \in C_5$

When $k \geq 2$ and k is an even number:

(iii) If c is accepted by more than half matchers, $c \in C_2$

(iv) If c is rejected by more than half matchers, $c \in C_4$

(v) Otherwise, $c \in C_3$

When $k \geq 3$ and k is an odd number:

(vi) If c is accepted by more than $(k + 1)/2$ matchers, $c \in C_2$

(vii) If c is rejected by more than $(k + 1)/2$ matchers, $c \in C_2$

(viii) Otherwise, $c \in C_3$

Since $c \in C_1$ is accepted by all the matchers as a correct correspondence, and $c \in C_5$ is rejected by all the matchers, the correspondences of types C_1 and C_5 are directly regarded as correct or wrong correspondences in the extraction process, which do not participate in the argumentation. Besides, correspondences belong to C_2, C_3, and C_4 should participate in the argumentation process.

Given two arguments be set as $a = \{c_1, n_1, v_1, h_1\}$ and $b = \{c_2, n_2, v_2, h_2\}$. There are four relationships between b and a, i.e., unite, attack, supporting, and disproval. Among them, union relationship is marked as $U(b, a)$, attack relationship is $A(b, a)$, and support and disprove are $S(b, a)$ and $D(b, a)$, respectively. To be specific, the four relationships are defined as the following descriptions:

(i) When $c_1 = c_2, v_1 \neq v_2, h_1 = h_2$, b is united with a, which is denoted as $U(b, a)$

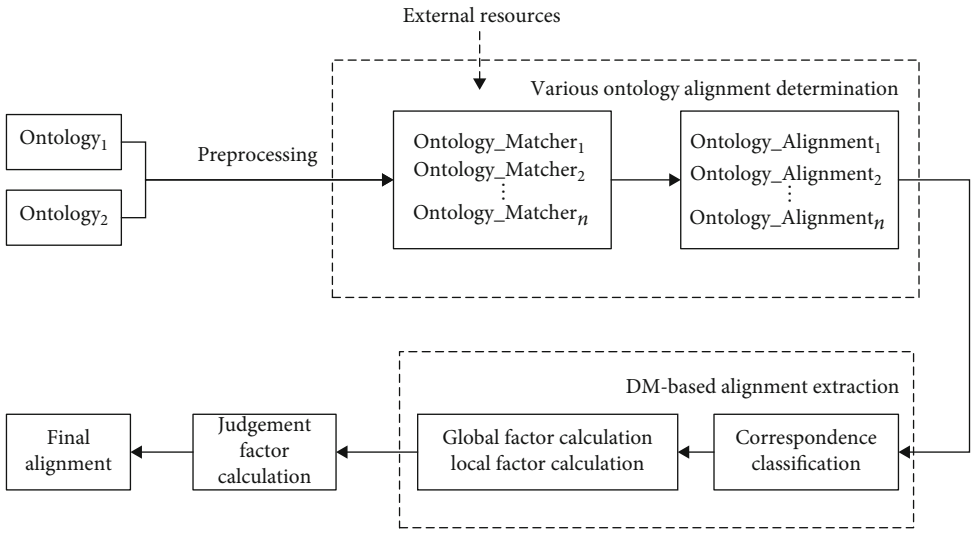


Figure 2: The framework of sensor ontology alignment extraction.

(ii) When $c_1 = c_2$, $v_1 \neq v_2$, $h_1 \neq h_2$, b attacks a, which is denoted as $A(b, a)$

(iii) When $c_2 = C_2$ or $c_2 = C_3$, $v_1 = v_2$, $n_1 > n_2$, $h_1 = h_2 = 1$, or when $c_2 = C_4$ or C_3, $n_1 < n_2$, $h_1 = h_2 = 0$, b supports a, which is denoted as $S(b, a)$

(iv) When $v_1 = v_2$, $c_1 = C_i$, $c_2 = C_j$, $i > j(i, j \in (2, 3, 4))$, $n_1 > n_2$, $h_1 = 1$, $h_2 = 0$, or when $v_1 = v_2$, $c_1 = C_i$, $c_2 = C_j$, $i < j(i, j \in (2, 3, 4))$, $n_1 < n_2$, $h_1 = 0$, $h_2 = 1$, b disproves a, which is denoted as $D(b, a)$

3. Debate Mechanism-Based Ontology Alignment Extraction

DM's argumentation is defined as a 7-tuple: $\{ar, \text{strength}, U, A, S, D, M\}$, where U, A, S, and D are the relationships mentioned above, $M = \{m_1, m_2, \cdots, m_n\}$ is the set of matchers which include n matchers, ar is the argument attached to correspondence c, and strength stands for the strength value with respect to c in terms of a matcher m_i, which is defined as follows:

$$\text{Strength}_c^{m_i} = \frac{\sum_{ar \in AR} |\{x \mid x \in AR \wedge S(x, ar)\}| - \sum_{ar \in AR} |\{x \mid x \in AR \wedge D(x, ar)\}|}{\sum_{ar \in AR} |\{x \mid x \in AR \wedge S(x, ar)\}| + \sum_{ar \in AR} |\{x \mid x \in AR \wedge D(x, ar)\}|}. \tag{3}$$

In order to extract the sensor ontology alignment from various alignments determined by different matchers, we define a correspondence's global factor as the correctness factor of each matcher to consider their similarity from different aspects, and local factor as the support strength and disprove strength in the debating process to consider the possibility of being a correct mapping from the perspective of other correspondences. In this work, a sensor entity correspondence c in each matcher is regarded as an argument made by it, and we need to calculate its judgment factor to determine whether it can be extracted into the final alignment. Therefore, it is critical to enhance the confidence of an entity mapping's judgment factor. To this end, we use the correctness factor of each matcher to determine a correspondence's global factor and utilize the support strength and disprove strength in the debating process to calculate its local factor. After that, the judgment factor of an entity mapping can be obtained by comprehensively considering these two factors.

To be specific, assuming $r_c \in \{0, 1\}$ is the judgment factor of c. The following debating process determines its value:

Step 1. It is obvious that r_c is 1 (or 0) when c belongs to C_1 (or C_5), and we can delete similarity values in the corresponding row and column of c from the similarity matrices.

Step 2. Calculate the correctness factor of the matcher m_i as follows:

$$\sigma_{m_i} = \frac{\sum_{m_i} |\{c \mid c \in (C_1, C_5)\}|}{\sum_{m_i} |c|}, \tag{4}$$

whose molecular calculates the number of correspondences belong to C_1 (or C_5) and denominator of the number of correspondences in total.

Step 3. In each matcher, the debating process is launched according to the relationships "support" and "disprove":

(1) For the C_2 type of argument, the matchers that support it account for the majority, so it is calculated whether the supporting party could defeat the disproving party successfully. Below we explain the situation above. Assuming that for c, the three matchers m_1, m_2, and m_3 support it but m_4 disprove it, the support strength Ss of matcher m_i is defined as follows:

$$Ss_t^{m_i} = \frac{\sum_{x \in AR} n_x - \sum_{y \in AR} n_y}{\sum_{x \in AR} n_x + \sum_{y \in AR} n_y}, \tag{5}$$

where argument $x = \{c, n_x, v_x, h_x\}$, argument $y = \{c, n_y, v_y, h_y\}$, $S(x,t)$, $D(y,t)$, $v_x = v_y$, and where $\sum_{x\in AR} n_x$ calculates the sum of similarity value of the arguments supporting argument $t = \{c, n_x, v_x, h_x\}$ and the same to $\sum_{y\in AR} n_y$.

It is determined that the mapping can be established between e and e' when $Ss_t^{m_1} > Ss_t^{m_4}$, $Ss_t^{m_2} > Ss_t^{m_4}$, and $Ss_t^{m_3} > Ss_t^{m_4}$, and r_c is set to 1. Otherwise, c is transformed into C_3, performing Step 4.

(2) For the C_4 type of argument, the matchers that disprove it account for the majority, so it is calculated whether the attack from the disproving party to the supporting party is successful. Below we explain the situation above. Assuming that for c, the three matchers m_1, m_2, and m_3 disprove it but m_4 support, the disprove strength Ds of matcher m_i is defined as follows:

$$Ds_t^{m_i} = \frac{\sum_{x\in AR} n_x - \sum_{y\in AR} n_y}{\sum_{x\in AR} n_x + \sum_{y\in AR} n_y}, \tag{6}$$

where argument $x = \{c, n_x, v_x, h_x\}$, argument $y = \{c, n_y, v_y, h_y\}$, $D(x,t)$, $S(y,t)$, $v_x = v_y$, and where $\sum_{x\in AR} n_x$ calculates the sum of similarity value of the arguments disproving argument $t = \{c, n_x, v_x, h_x\}$ and the same to $\sum_{y\in AR} n_y$. It is determined that the mapping cannot be established between e_i and e_j' when $Ds_t^{m_1} > Ds_t^{m_4}$, $Ds_t^{m_2} > Ds_t^{m_4}$, and $Ds_t^{m_3} > Ds_t^{m_4}$, and $r_c = 0$. Otherwise, c is transformed into C_3, performing Step 4.

Step 4. For the C_3 type of argument, the matchers with opposite opinions are almost evenly matched. It is necessary to calculate which of the two parties attacked each other successfully and calculate the strength value of each matcher. The attack with the higher the strength value is successful. Assuming that for c, the two matchers m_1 and m_2 disprove it but m_3 and m_4 support it. $r_c = 0$ if $\delta_{m_1} \bullet \text{Strength}_c^{m_1} + \delta_{m_2} \bullet \text{Strength}_c^{m_2} > \delta_{m_3} \bullet \text{Strength}_c^{m_3} + \delta_{m_4} \bullet \text{Strength}_c^{m_4}$. Otherwise, $r_c = 1$.

Step 5. Select c with $r_c = 1$ and determine the correspondence set based on them.

4. Experiment

4.1. Experimental Configuration. In the experiment, we use the bibliographic track from Ontology Alignment Evaluation Initiative (OAEI) to test the robustness of our approach through statistical comparison with the most advanced ontology matching techniques, as well as five sensor ontologies to validate the effectiveness of our proposal. We use four popular ontology matchers, which are, respectively, based on Levenshtein distance [41], Jaro-Winkler distance [42], similarity flooding (SF) [43], and WordNet similarity [44]. In particular, Levenshtein distance and Jaro-Winkler distance are terminology-based measures, WordNet is a semantic-based measure, and SF uses a versatile graph matching algorithm. The similarity threshold is empirically set as 0.85, which can ensure the average highest quality of alignment in all testing cases. Typically, the alignment is assessed in terms of two measures, commonly known as precision and recall [45], which are, respectively, defined as follows:

Table 1: Brief description on OAEI's bibliographic track.

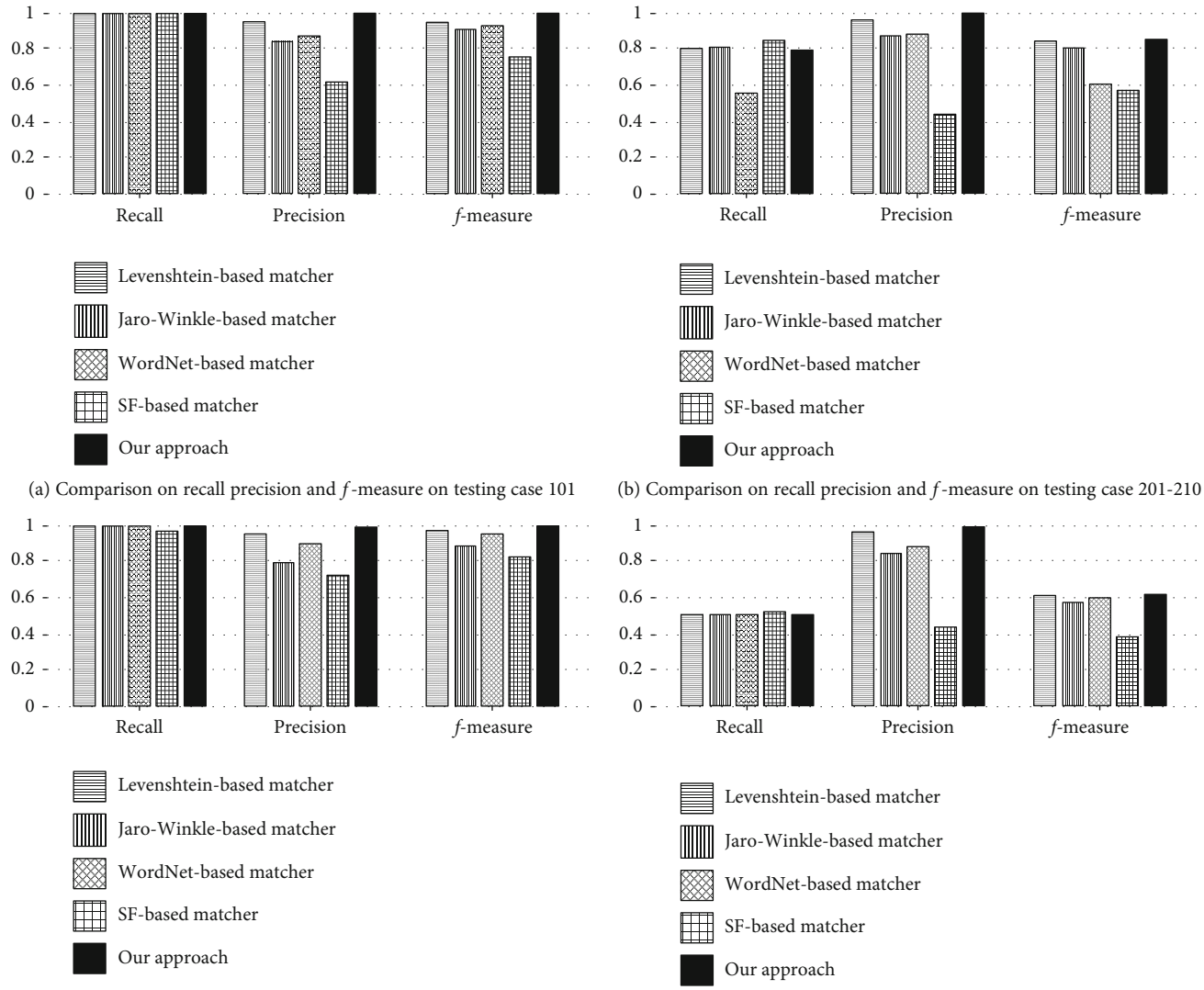
Testing case	Description
101	Identical ontologies
201–202	Ontologies varying in terminology and semantic characteristics
221–237	Ontologies varying in structure characteristics
248–261	Ontologies varying in terminology, semantics, and structure characteristics

$$\begin{aligned} \text{Recall} &= \frac{|R \bigcap A|}{|R|}, \\ \text{Precision} &= \frac{|R \bigcap A|}{|A|}, \end{aligned} \tag{7}$$

where A is the obtained alignment and R is the reference alignment. recall and precision are designed to examine the validity of the proposal. They use a set of standard answers as a reference to calculate the relative ratio as shown in the formula, which is about the result of the proposal and the reference. In our application, recall is used to measure the ability of the proposal to find the correct matching pair, and precision is used to measure the correct rate of the proposal to find the matching pair. In particular, recall = 1 means that all correct matching pairs have been found, and precision = 1 means that all matching pairs found are correct matching pairs. To trade off these two metrics, we further use the f-measure, which is the harmonic mean of recall and precision [46]:

$$f\text{-measure} = \frac{2\,\text{precision} - \text{recall}}{\text{recall} + \text{precision}}. \tag{8}$$

The bibliographic track (http://oaei.ontologymatching.org/2016/results/benchmarks) provided by the Ontology Alignment Evaluation Initiative (OAEI) is applied to test the robustness of our approach, and five real sensor ontologies are used to test our approach's effectiveness. Table 1 provides an abstract of OAEI's bibliographic track, where each testing case comprises of two to-be-mapped ontologies and one reference alignment for assessing the effectiveness of the ontology matcher. Figures 3(a)–3(d), respectively, compare our approach with four single similarity measures in terms of recall, precision, and f-measure. And relevant data is given by the Ontology Alignment Evaluation Initiative (OAEI) (http://oaei.ontologymatching.org/2016/). Figure 4, respectively, compares our approach with OAEI's participants [21], i.e., AML, edna, and LogMapLt in terms of f-measure. Table 2 describes our proposal's results on five sets of real



(a) Comparison on recall precision and f-measure on testing case 101

(b) Comparison on recall precision and f-measure on testing case 201-210

(c) Comparison on recall precision and f-measure on testing case 221-237

(d) Comparison on recall precision and f-measure on testing case 248-266

Figure 3: Comparison with four single similarity measures in terms of recall, precision, and f-measure.

sensor ontology matching tasks and compares them with other OAEI competitors in SOSA and SN matching tasks.

As shown in Figure 3, comparing with four single similarity measures, our approach's f-measure is highest in all testing cases presented. We pick testing case 101, which contains strictly identical ontologies as a representative one of the testing cases about two identical ontologies in OAEI's benchmark. As Figure 3(a) indicates that our proposal achieves good results, which proves the superiority of it, therefore, our proposal has perfect accuracy in establishing the correct matching pair. DM adopts local factors on matchers in order to fully consider the contexts of the entities and external resources, aiming to improve the recall value. Besides, the mechanism fully considers the use of global factors for the alignment of different matchers to improve the precision value. In more detail, in testing cases 201-210 presented in Figure 3(b), which describe two ontologies with different terminology and semantic characteristics in each case, case 201 contains ontologies without entity names and 202 without entity. In these cases, our proposal's performance is significantly better than the other methods, indicating that we have done a good job due to the consideration of support and disprove strength. In Figure 3(c), testing cases 221-247 are about ontologies with different structural characteristics. And results illustrate the effectiveness of our proposals and edna. In addition, our proposal is superior to edna in testing cases 221-223, which are about the ontologies with no specialization, with a flattened hierarchy, and the ontology with an expended hierarchy. A similar situation is shown in Figure 3(d), which compares the performance on testing cases about ontologies with different terminology, semantics, and structure characteristics. Our method considers the contextual information based on the ontology concept structure in calculating the strength value of the correspondence so that it can effectively overcome the problems of literal heterogeneity and information asymmetry when dealing with testing cases 201-261. In particular, our proposal has a certain competence with the most OAEI competitors in the aspect of recall, i.e., in case 101-247, where our approach can reach a level of 1.0 in recall value, indicating that the ability to find reference alignments is strong. To

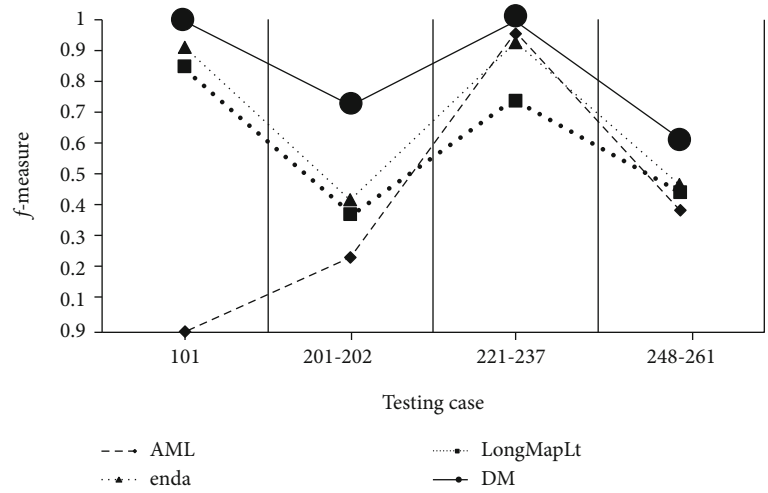


Figure 4: Comparison with OAEI's participants in terms of f-measure.

Table 2: Comparison on two pairs of real sensor ontologies with four matchers.

Matching task	Ontology quality measure	SF-based matcher	Jaro-Winkler-based matcher	WordNet-based matcher	Levenshtein-based matcher	Our approach
SOSA-OSSN	Recall	0.50	1.00	1.00	1.00	1.00
	Precision	0.20	1.00	0.67	1.00	1.00
	f-measure	0.29	1.00	0.80	1.00	1.00
SOSA-SN	Recall	1.00	1.00	1.00	1.00	1.00
	Precision	0.07	0.75	0.33	0.75	1.00
	f-measure	0.13	0.86	0.50	0.86	1.00
SSN-IoT	Recall	1.00	1.00	1.0	1.00	1.00
	Precision	0.01	1.00	0.33	1.00	1.00
	f-measure	0.03	1.00	0.50	1.00	1.00
SSN-OSSN	Recall	0.35	1.00	0.97	1.00	0.97
	Precision	0.06	0.94	0.80	1.00	1.00
	f-measure	0.11	0.97	0.88	1.00	0.98
SSN-SN	Recall	0.56	1.00	1.00	1.00	1.00
	Precision	0.02	0.90	0.52	1.00	1.00
	f-measure	0.04	0.95	0.70	1.00	1.00

improve the recall value and precision value, we introduce different ontology matching techniques to mine more potential correspondences and excavate the relationship between different arguments by support strength and disprove strength to explore more correct correspondences. It is worth learning that since the hypothetical goal of SF is to accurately obtain a matching element for each element in the source ontology, this method can get a higher recall rate.

Since our proposal takes into account each entity mapping's preference on different matchers, the alignment quality is relatively high. Figure 4 shows the result of the comparison with OAEI's participants in terms of f-measure, which indicates the certain comparability of our proposal in the field of ontology matching system. To be specific, the f-measure value generated by our proposal is better than these OAEI's participants owing to the fact that we integrate the advantages of different basic similarity measures through DM so as to comprehensively consider the matching problem from different perspectives in the ontology matching process. To conclude, our proposal is able to improve the result's precision value significantly and, meanwhile, ensure a high recall value, which makes it outperforms other competitors on various matching tasks.

4.2. Real Sensor Ontology. The five sensor ontologies used are SSN, SOSA, IoT, SN, and OSSN, which are described in Table 3, and more details on them are presented in Section 2.1.

When matching the real sensor ontology, it can be seen from Table 2 that our method can obtain almost the same result as golden alignment. When matching the new SSN

Table 3: Main features of sensor ontologies.

Ontology	Number of entities	Website
Semantic Sensor Network (SSN) ontology	55	https://www.w3.org/ns/ssn/
Sensor, Observation, Sample, and Actuator (SOSA) ontology	42	http://www.w3.org/ns/sosa/
IoT-Lite (IoT) ontology	40	https://www.w3.org/Submission/2015/SUBM-iot-lite-20151126
SensorOntology2009 (SN) ontology	152	https://www.w3.org/2005/Incubator/ssn/wiki/SensorOntology2009
Original Semantic Sensor Network (OSSN) ontology	107	https://www.w3.org/2005/Incubator/ssn/wiki/SSN#Sensor

and OSSN, our method gets a recall rate of 0.97. This is because we assume that the alignment base is one-to-one. That is to say, a concept in an ontology can only be mapped to another one in another ontology, and vice versa. However, in some practical tasks, the cardinality may be one-to-many. For example, the concept "stimultium" in the new SSN is mapped to the concept "stimultium" in OSSN, which is the same as another assertion, "Sensor Input." In our method, it will no longer participate in the filtering process if the "Stimulus" in the new SSN is first mapped to the "Stimulus" in the original SSN in terms of ID (although the "Stimulus" annotation in the new SSN is the same as the "Sensor Input" annotation in the OSSN), which reduces the recall value of our method. In our proposal, we make the three metrics to reach the level of perfection.

5. Conclusion

Since artificial intelligence has greatly changed the data link form, IoT technologies are making significant progress these years, where ontology matching plays an advance role [47–51]. As far as the quality of the ontology alignment is concerned, the final alignment is composed of mappings. Therefore, it is particularly important to extract correct mappings from the alignments produced by different matchers and to filter out incorrect mappings. Different from the previous ontology meta-matching technology used by Xue et al. [52, 53], the ontology alignment extraction technology using DM considers the matching opinions of each basic matcher to each mapping, so as to output the final alignment. The sensor ontology extracting approach is aimed at finding high-quality alignment from various sensor ontology alignments, which can be used to bridge the semantic gap between heterogeneous sensor ontologies and integrate the knowledge defined inside. To enhance the quality of alignment, we propose a novel sensor ontology extracting method, which uses the debating mechanism. During various ontology matcher's debating processes, we propose to calculate both entity correspondence's global factor and local factor to determine its judgment factor, which is able to enhance its confidence. Although the experimental results show that our approach is able to decide on high-quality sensor ontology alignments effectively, there is still room for improvement. In the next work, we will try to describe the similarity between matching pairs by applying a new method and improve the extraction method to make it more suitable for the characteristics of the sensor ontology.

Data Availability

The data used to support this study can be found in http://oaei.ontologymatching.org.

Conflicts of Interest

The authors declare that they have no conflicts of interest in the work.

Acknowledgments

This work is supported by the Natural Science Foundation of Fujian Province (No. 2020J01875), the Guangxi Key Laboratory of Automatic Detecting Technology and Instruments (No. YQ20206), the Scientific Research Foundation of Fujian University of Technology (No. GY-Z17162), and the National Natural Science Foundation of China (Nos. 61773415, 61801527, and 61103143).

References

[1] R. Zhao, X. Wang, J. Xia, and L. Fan, "Deep reinforcement learning based mobile edge computing for intelligent Internet of things," *Physical Communication*, vol. 43, p. 101184, 2020.

[2] M. Ye, Y. Wang, C. Dai, and X. Wang, "A hybrid genetic algorithm for minimum exposure path problem of wireless sensor network based on a numerical functional extreme model," *IEEE Transactions on Cybernetics*, vol. 65, no. 10, pp. 8644–8657, 2016.

[3] T. Mavroeidakos and V. Chaldeakis, "Threat landscape of next generation IoT-enabled smart grids," in *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pp. 116–127, Springer, Cham, 2020.

[4] S. De Dutta and R. Prasad, "Security for smart grid in 5G and beyond networks," *Wireless Personal Communications*, vol. 106, no. 1, pp. 261–273, 2019.

[5] A. I. Ali, S. Z. Partal, S. Kepke, and H. P. Partal, "ZigBee and LoRa based wireless sensors for smart environment and IoT applications," in *2019 1st Global Power, Energy and Communication Conference (GPECOM)*, pp. 19–23, Nevsehir, Turkey, Turkey, June 2019.

[6] S. A. A. Elmustafa and E. Y. Mujtaba, "Internet of things in smart environment: concept, applications, challenges, and future directions," *World Scientific News*, vol. 134, no. 1, pp. 1–51, 2019.

[7] A. Hameed and A. Leivadeas, "IoT traffic multi-classification using network and statistical features in a smart environment," in *2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Pisa, Italy, Italy, September 2020.

[8] D. Glaroudis, A. Iossifides, and P. Chatzimisios, "Survey, comparison and research challenges of IoT application protocols for smart farming," *Computer Networks*, vol. 168, p. 107037, 2020.

[9] G. Ramya and N. S. Lohitha, "The role and advancements of IoT technology in smart farming for agribusiness," *Journal of the Gujarat Research Society*, vol. 21, no. 16, pp. 2404–2413, 2019.

[10] T. M. Fernández-Caramés and P. Fraga-Lamas, "A review on the use of blockchain for the Internet of things," *IEEE Access*, vol. 6, pp. 32979–33001, 2018.

[11] H. Chen, "A cell probe-based method for vehicle speed estimation," *IEICE Transactions on Fundamentals of Electronics. Communications and Computer Sciences*, vol. 103, no. 1, pp. 265–267, 2020.

[12] C. H. Chen, F. J. Hwang, and H. Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Transactions on Information and Systems*, vol. 102, no. 7, pp. 1374–1383, 2019.

[13] C. H. Chen, "An arrival time prediction method for bus system," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4231-4232, 2018.

[14] J. Hill and D. Culler, *A Wireless Embedded Sensor Architecture for System Level Optimization. UC Berkeley Technical Report*, University of California, Berkeley, CA, USA, 2001.

[15] J. Rabaey, J. Ammer, J. L. Da Silva, and D. Patel, "PicoRadio: ad-hoc wireless networking of ubiquitous low-energy sensor/-monitor nodes," in *Proceedings IEEE Computer Society Workshop on VLSI 2000. System Design for a System-on-Chip Era*, pp. 9–12, Orlando, FL, USA, USA, April 2000.

[16] J. M. Kahn, R. H. Katz, and K. S. J. Pister, "Next century challenges: mobile networking for "smart dust"," in *Proceeding of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pp. 271–278, Seattle, WA, USA, August 1999.

[17] G. Asada, M. Dong, T. S. Lin et al., "Wireless integrated network sensors: low power systems on a chip," in *Proceeding of the 24th European Solid-State Circuits Conference (ESSCIRC 1998)*, pp. 9–16, Hague, Netherlands, 1998.

[18] A. Sheth, C. Henson, and S. S. Sahoo, "Semantic sensor web," *IEEE Internet Computing*, vol. 12, no. 4, pp. 78–83, 2008.

[19] X. Xue, "A compact firefly algorithm for matching biomedical ontologies," *Knowledge and Information Systems.*, vol. 62, no. 7, pp. 2855–2871, 2020.

[20] X. Xue and J. Chen, "Optimizing sensor ontology alignment through compact co-firefly algorithm," *Sensors*, vol. 20, no. 7, p. 2056, 2020.

[21] M. Achichi, M. Cheatham, Z. Dragisic et al., "Results of the ontology alignment evaluation initiative 2016," in *11th International Workshop on Ontology Matching (OM 2016)*, pp. 73–129, Kobe, Japan, 2016.

[22] P. Shvaiko and J. Euzenat, "Ontology matching: state of the art and future challenges," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 1, pp. 158–176, 2013.

[23] G. Alexandru-Lucian and A. Iftene, "Using a genetic algorithm for optimizing the similarity aggregation step in the process of ontology alignment," in *9th Roedunet International Conference*, pp. 43898–43907, Sibiu, Romania, 2010.

[24] X. Xue and J. Liu, "A compact brain storm algorithm for matching ontologies," *IEEE Access*, vol. 8, pp. 551–556, 2020.

[25] X. Xue and J. Liu, "Collaborative ontology matching based on compact interactive evolutionary algorithm," *Knowledge-Based Systems*, vol. 137, pp. 94–103, 2017.

[26] X. Xue and X. Yao, "Efficient user involvement in semi-automatic ontology matching," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2018, pp. 1–11, 2018.

[27] X. Xue, J. Chen, and X. Yao, "Ontology alignment based on instance using NSGA-II," *Journal of Information Science*, vol. 41, no. 1, pp. 58–70, 2014.

[28] X. Xue, J. Chen, and J. S. Pan, *Evolutionary Algorithm Based Ontology Matching Technique*, Science Press, Beijing, 2018.

[29] X. Xue and Y. Wang, "Using memetic algorithm for instance coreference resolution," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 580–591, 2016.

[30] X. Xue and A. Ren, "An evolutionary algorithm based ontology alignment extracting technology," *Journal of Network Intelligence*, vol. 2, no. 2, pp. 205–212, 2017.

[31] P. M. Dung, "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games," *Artificial Intelligence*, vol. 77, no. 2, pp. 321–357, 1995.

[32] X. Xue and Y. Wang, "Optimizing ontology alignments through a memetic algorithm using both MatchFmeasure and unanimous improvement ratio," *Artificial Intelligence*, vol. 223, pp. 65–81, 2015.

[33] J. Euzenat and P. Shvaiko, "Classifications of ontology matching techniques," in *Ontology Matching*, pp. 73–84, Springer, 2013.

[34] W. E. Winkler, "The state of record linkage and current research problems," in *Statistical Research Division*, Bureau of the Census, 1999.

[35] J. Oliva, J. I. Serrano, M. D. del Castillo, and Á. Iglesias, "SyMSS: a syntax-based measure for short-text semantic similarity," *Data & Knowledge Engineering*, vol. 70, no. 4, pp. 390–405, 2011.

[36] V. Mascardi, A. Locoro, and P. Rosso, "Automatic ontology matching via upper ontologies: a systematic evaluation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 5, pp. 609–623, 2009.

[37] T. Pedersen, S. Patwardhan, and J. Michelizzi, "WordNet: similarity-measuring the relatedness of concepts," *AAAI*, vol. 4, pp. 25–29, 2004.

[38] J. Gao, B. Zhang, and X. Chen, "A WordNet-based semantic similarity measurement combining edge-counting and information content theory," *Engineering Applications of Artificial Intelligence*, vol. 39, pp. 80–88, 2015.

[39] G. Pirrro and N. Seco, "Design, implementation and evaluation of a new semantic similarity metric combining features and intrinsic information content," in *On the Move to Meaningful Internet Systems: OTM 2008*, pp. 1271–1288, Springer, 2008.

[40] L. Meng, R. Huang, and J. Gu, "An effective algorithm for semantic similarity metric of word pairs," *International*

Journal of Multimedia and Ubiquitous Engineering, vol. 8, no. 2, pp. 1–12, 2013.

[41] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions and reversals," *Soviet Physics-Doklady*, vol. 10, no. 8, p. 707, 1996.

[42] W. W. Cohen, P. Ravikumar, and S. E. Fienberg, "A comparison of string distance metrics for name-matching tasks," in *Proceedings of the IJCAI-2003 Workshop on Information Integration on the Web (IIWeb-03)*, pp. 73–78, Acapulco, Mexico, 2013.

[43] S. Melnik, H. Garcia-Molina, and E. Rahm, "Similarity flooding: a versatile graph matching algorithm and its application to schema matching," in *Proceedings of 18th International Conference on Data Engineering*, pp. 117–128, Los Alamitos, CA, USA, February 2002.

[44] G. A. Miller, "WordNet," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[45] X. Xue and X. Yao, "Interactive ontology matching based on partial reference alignment," *Applied Soft Computing*, vol. 72, pp. 355–370, 2018.

[46] X. Xue and J. Chen, "Using compact evolutionary tabu search algorithm for matching sensor ontologies," *Swarm and Evolutionary Computation*, vol. 48, pp. 25–30, 2019.

[47] Z. G. du, J. Pan, S. Chu, H. J. Luo, and P. Hu, "Quasi-affine transformation evolutionary algorithm with communication schemes for application of RSSI in wireless sensor networks," *IEEE Access*, vol. 8, pp. 8583–8594, 2020.

[48] S. Chu, T. Dao, and J. Pan, "Identifying correctness data scheme for aggregating data in cluster heads of wireless sensor network based on naive Bayes classification," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, 2020.

[49] H.-C. Huang, S.-C. Chu, J.-S. Pan, C.-Y. Huang, and B.-Y. Liao, "Tabu search based multi-watermarks embedding algorithm with multiple description coding," *Information Sciences*, vol. 32, pp. 3379–3396, 2011.

[50] J. Pan, P. Hu, and S. Chu, "Novel parallel heterogeneous meta-heuristic and its communication strategies for the prediction of wind power," *Processes*, vol. 7, no. 11, p. 854, 2019.

[51] P. Song, J. Pan, and S. Chu, "A parallel compact cuckoo search algorithm for three-dimensional path planning," *Applied Soft Computing*, vol. 94, article 106443, 2020.

[52] X. Xue and J. Pan, "A compact co-evolutionary algorithm for sensor ontology meta-matching," *Knowledge and Information Systems*, vol. 56, no. 2, pp. 335–353, 2018.

[53] X. Xue, H. Yang, J. Zhang, J. Zhang, and D. Chen, "An automatic biomedical ontology meta-matching technique," *Journal of Network Intelligence*, vol. 4, no. 3, pp. 109–113, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6680701, 20 pages
https://doi.org/10.1155/2021/6680701

Research Article

Face Image Publication Based on Differential Privacy

Chao Liu,[1,2] **Jing Yang**,[1] **Weinan Zhao**,[2] **Yining Zhang**,[3] **Jingyou Li**,[1,2] **and Chunmiao Mu**[2]

[1]*Harbin Engineering University, Harbin 150001, China*
[2]*Qiqihar University, Qiqihar 161000, China*
[3]*DaQing Vocational College, DaQing 163000, China*

Correspondence should be addressed to Jing Yang; 00819@qqhru.edu.cn

Received 9 October 2020; Revised 28 November 2020; Accepted 10 December 2020; Published 7 January 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Chao Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As an information carrier, face images contain abundant sensitive information. Due to its natural weak privacy, direct publishing may divulge privacy. Anonymization Technology and Data Encryption Technology are limited by the background knowledge and attack means of attackers, which cannot completely content the needs of face image privacy protection. Therefore, this paper proposes a face image publishing SWP (sliding window publication) algorithm, which satisfies the differential privacy. Firstly, the SWP translates the image gray matrix into a one-dimensional ordered data stream by using image segmentation technology. The purpose of this step is to transform the image privacy protection problem into the data stream privacy protection problem. Then, the sliding window model is used to model the data flow. By comparing the similarity of data in adjacent sliding windows, the privacy budget is dynamically allocated, and Laplace noise is added. In SWP, the data in the sliding window comes from the image. To present the image features contained in the data more comprehensively and use the privacy budget more reasonably, this paper proposes a fusion similarity measurement EM (exact mechanism) mechanism and a dynamic privacy budget allocation DA (dynamic allocation) mechanism. Also, for further improving the usability of human face images and reducing the impact of noise, a sort-SWP algorithm based on the SWP method is proposed in the paper. Through the analysis, it can be seen that ordered input can further improve the usability of the SWP algorithm, but direct sorting of data will destroy the ε-differential privacy. Therefore, this paper proposes a sorting method-*SAS* method, which satisfies the ε-differential privacy; SAS obtain an initial sort by using an exponential mechanism firstly. And then an approximate correct sort is obtained by using the Annealing algorithm to optimize the initial sort. Compared with LAP algorithm and SWP algorithm, the average accuracy rate of sort-SWP algorithm in ORL, Yale is increased by 56.63% and 21.55%, the recall rate is increased by 6.85% and 3.32%, and F1-sroce is improved by 55.62% and 16.55%.

1. Introduction

With the rapid development of information technology and multimedia technology, it is easier to obtain and share face digital images. Users can publish photos of their mobile phones or digital cameras to social networking platforms (such as Twitter, LinkedIn, WeChat) or other channels. Relevant statistics show that the number of face photos shared by users on major social networking platforms worldwide exceeds 3.2 billion every day. Also, there are numerous face image data derived from video. However, these digital images usually contain a wealth of personally sensitive information. If this information was collected and analyzed by a third party with ulterior motives, it may cause personal privacy disclosure and other unexpected losses.

Privacy is an emotional word, which implies different meanings to different people. According to the definition of International Organization Standardization (ISO), privacy refers to the characteristics that can distinguish individuals or groups from other individuals and groups. Different countries have different legal definitions of privacy, and different objects (individuals, enterprises, governments, etc.) define

the scope of privacy differently. For digital images, sensitive information can be a specific person or object in an image, a face, or fingerprint; the embedded information (photo location information and creation time, etc.); or an area that the image owner is concerned about. How to publish and analyze without disclosing sensitive information is the main purpose of privacy protection.

The early research uses Anonymization Technology or Data Encryption Technology to solve the privacy protection of face image. Anonymization Technology refers to cover up real data with methods of hidden or fuzzy. It generally adopts anonymous operations such as suppression [1], generalization [2], analysis [3], slicing [4], and separation [5]. Especially, k-anonymity [6] is a classic representative algorithm. K-anonymity proposed that the sensitive information covered by the data should at least be indistinguishable from other $k-1$ data. Because of its shortcomings and shortcomings, K-anonymity extends the l-diversity method which ensures that each equivalent class contains at least l different sensitive attribute values [7], the t-closeness method to improve the global distribution of sensitive attributes [8], the m-variance method for dynamic relational data [9], and the HD composition method [10]. References [11, 12] use an anonymization mechanism to propose the k-same method. This method anonymizes the published digital image so that the probability of the attacker to identify the user identity through the published digital image again is less than $1/K$. However, the main drawback of the traditional anonymization mechanism is too many assumptions about the attacker's background knowledge and attack model, but those assumptions are not completely successful in reality.

Data Encryption Technology is another important research direction of image information security, and its representative methods include secure multiparty computing [13], homomorphism encryption [14], and classification algorithm [15]. References [16–19] prevent the invasion of the third party by controlling the user communication protocol. References [20–23] propose using pixel replacement or pixel value substitution to encrypt image content. References [24–28] encrypt the image by changing the transform coefficient of the image in the frequency domain. Similar to the problems of anonymization technology, data encryption technology will also make corresponding assumptions for attacks and then design the corresponding encryption algorithm based on these assumptions. But this kind of encryption method will fall into the cycle of "new encryption methods are constantly proposed but constantly broken." Also, the encrypted image is not open.

Dwork first proposed differential privacy [29] in 2006, which disturbs sensitive data by adding noise to the output. Differential privacy can hide the influence of a single record. That means whether the record is in or not in the dataset, the output probability of the same result will not change significantly. The attacker's ability to further reasoning is limited. Therefore, differential privacy is better than other privacy protection technologies by not making any assumptions about the background knowledge of any potential attacker. Besides, differential privacy is further studied in a series of papers of Dwork [30–34], and its implementation mechanism is proposed in [35, 36]. McSherry pointed out that some differential privacy algorithms for complex privacy problems satisfy two combinatorial properties: sequence composability and juxtaposition and combination [37]. In recent years, differential privacy is mainly used in data publishing, including histogram publishing [38–43], graph data publishing [44–48], data mining [49–51], data stream publishing [52], and spatial data publishing [53]. Due to the complexity of image data, researchers are still in the exploratory stage to use differential privacy technology to protect sensitive information in images.

The real field matrix is a common representation of the image. Any pixel in the image can be mapped to a numerical relative position in a 2D matrix. It is the most direct method to add Laplace noise to all the values in the matrix. Although this method can satisfy ε-differential privacy, it will cause excessive distortion and low usability of the disturbing image. Fourier transform and Wavelet transform are commonly compression techniques in image processing. In reference [54], an image compression method based on a discrete Fourier transform is proposed. This method adds the corresponding Laplace noise to the compressed image. Although the noise error is reduced, the reconstruction error is introduced in the image compression process. To reduce the impact of noise on the original image and improve the usability of published images, this paper proposes a differential privacy protection method for image publishing. It is inspired by the noncorrelation of the values in the image matrix; this paper tries to use image segmentation technology to transform the image gray matrix into 1D ordered data stream and then use the sliding window model to model the data flow. By comparing the similarity of data in adjacent sliding windows, the privacy budget is dynamically allocated which is used to solve the problem of image privacy protection. This method not only satisfies the differential privacy but also has high usability of the published image.

2. Background

2.1. Differential Privacy. Dalenius raised a problem with statistical databases: by accessing the database, no one should be able to get any information about a person [55]. However, due to background knowledge, absolute privacy protection is not possible. Differential privacy sidesteps this issue and turns to relative privacy protection. Any potential privacy breach will be limited to a small multiplier. To be in attention is that serious leaks may occur, but it is not because of whether a particular piece of data exists in the database.

As a carrier of information, a digital image is usually stored and transmitted by a 3D matrix (i.e., a color image can be expressed as R, G, and B, three 2D matrices). To facilitate the data processing, it tries to make normalization treatment to the 3D image matrix and then get the corresponding 2D image gray matrix. Image x can be expressed as a 2D matrix X_{mn}, with m is the number of rows, and N is the number of columns of the matrix.

Formula (1) gives the specific calculation method.

$$X_{mn} = R_{mn} \times 0.299 + G_{mn} \times 0.587 + B_{mn} \times 0.114. \quad (1)$$

Before giving the formal definition of differential privacy, the definition of the neighborhood is given first by combining with X.

Definition 1. Given an image X, the image gray matrix X_{mn} is obtained after normalization, so

$$X|X_{mn} = \begin{bmatrix} x_{11}, x_{12}, & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1}, x_{m1}, & \cdots & x_{mn} \end{bmatrix}; x_{ij} \quad (2)$$

represents the gray value of the corresponding element in the matrix X_{mn}. If there is an X', and there is only one element difference between X and X', that is, $||X - X'| = x_{ij}(1 \le i \le m, 1 \le j \le n)$, then X and X' are said to be adjacent to each other.

Definition 2. A random algorithm M for image data publishing is supposed. Range (M) is the output range of M. if any output s of algorithm M on two two-dimensional matrices X and X' which are adjacent to each other satisfies Equation (3), then algorithm M satisfies ε-differential privacy.

$$\Pr[M(X) \in S] \le \exp(\varepsilon) \times \Pr\left[M\left(X'\right) \in S\right]. \quad (3)$$

In Equation (3), ε is usually a small positive number, which is used to weigh the relationship between privacy and precision. Relatively, if the ε is small, the privacy is higher, and the accuracy is lower, and vice versa. In general, the users select ε by executing a certain privacy policy. Besides, if an algorithm satisfies the ε-differential privacy when the neighbor database differs by one record, it satisfies the $k\varepsilon$-differential privacy when the neighbor database differs at most by K records.

To achieve differential privacy, a certain amount of random noise needs to be added to the query results. Intuitively, its magnitude should cover the maximum impact of a single record on the output. Therefore, the noise level is closely related to the global sensitivity of the corresponding query function.

Definition 3. Q is supposed to any query function and $Q : D \longrightarrow R^d$, the sensitivity of Q is expressed as

$$\Delta Q = \max_{X,X'} \left\| Q(X) - Q\left(X'\right) \right\|_\rho. \quad (4)$$

Laplace mechanism is the most common noise-adding mechanism. For achieving differential privacy, the noise generated by Laplace distribution (the noise distribution satisfies Laplace probability density function $pdf(x \mid \lambda) = (1/2\lambda) e^{-|x|/\lambda}$, $\lambda = \Delta Q/\varepsilon$) is added to disturb the real output.

Theorem 4. *Laplace mechanism: suppose Q is a query sequence of length D. random algorithm m, which takes database as input and outputs the following vectors. It will satisfy ε-differential privacy.*

$$M(D) = Q(D) + < Lap_1\left(\frac{\Delta Q}{\varepsilon}, \cdots, Lap_d\left(\frac{\Delta Q}{\varepsilon} > . \quad (5)\right.\right.$$

What should be noted here is that $\mathrm{Lap}_i(\Delta Q/\varepsilon)$ $(1 \le i \le d)$ is an independent Laplace noise. The magnitude of the noise is proportional to ΔQ and inversely proportional to ε.

Theorem 5. *Exponential mechanism: for any sampling method m under the exponential mechanism, M satisfies ε-differential privacy if it satisfies Equation (6).*

$$M(H, H_i) = \left\{ H_i : \mid \Pr[H_i \in \mathscr{H}] \propto \exp\left(\frac{\varepsilon u(H, H_i)}{2\Delta u}\right) \right\}. \quad (6)$$

The exponential mechanism mainly deals with the nonnumerical output of the sampling algorithm. In the mechanism, $u(H, H_i)$ is the scoring function, Δu is the global sensitivity of scoring function $u(H, H_i)$, and H is the output domain of the algorithm. According to formula (6), the higher the scoring function of H_i, the greater probability of output is selected.

2.2. Data Flow and Sliding Window Model. Differential privacy can ensure that the operation of inserting or deleting a record in a database will not affect the output of any query, thus ensuring that each record's deletion or joining the database will not pose a threat to its privacy. To define differential privacy on a data stream, it is necessary to give the nearest neighbor relationship between two data streams.

Definition 6. For the data stream D and D', if there is at most one record difference between them, then D and D' are neighbors to each other.

Definition 7. For the data stream D and D', they are adjacent to each other. A privacy algorithm A is assumed, if the result s of A on D and $d\wedge$ satisfies Equation (7), then algorithm A satisfies ε-differential privacy.

$$\Pr[A(D) \in S] \le \exp(\varepsilon) \times \Pr\left[A\left(D'\right) \in S\right]. \quad (7)$$

At any time, the sliding window model only needs to consider and process the most recently arrived N data. It can better reflect the characteristics that the importance of data in the data stream gradually decreases with time. Therefore, the sliding window model is usually used in data stream processing. The sliding window model is used to model a data stream D with the length of N. All sliding windows use the fixed size of $|W|$. The value of $|\omega|$ is equal to the number of data contained in the sliding window. The sliding amplitude s is the distance that the current sliding window moves forward compared with the previous one. Generally, the sliding amplitude is 1, which can be adjusted according to the

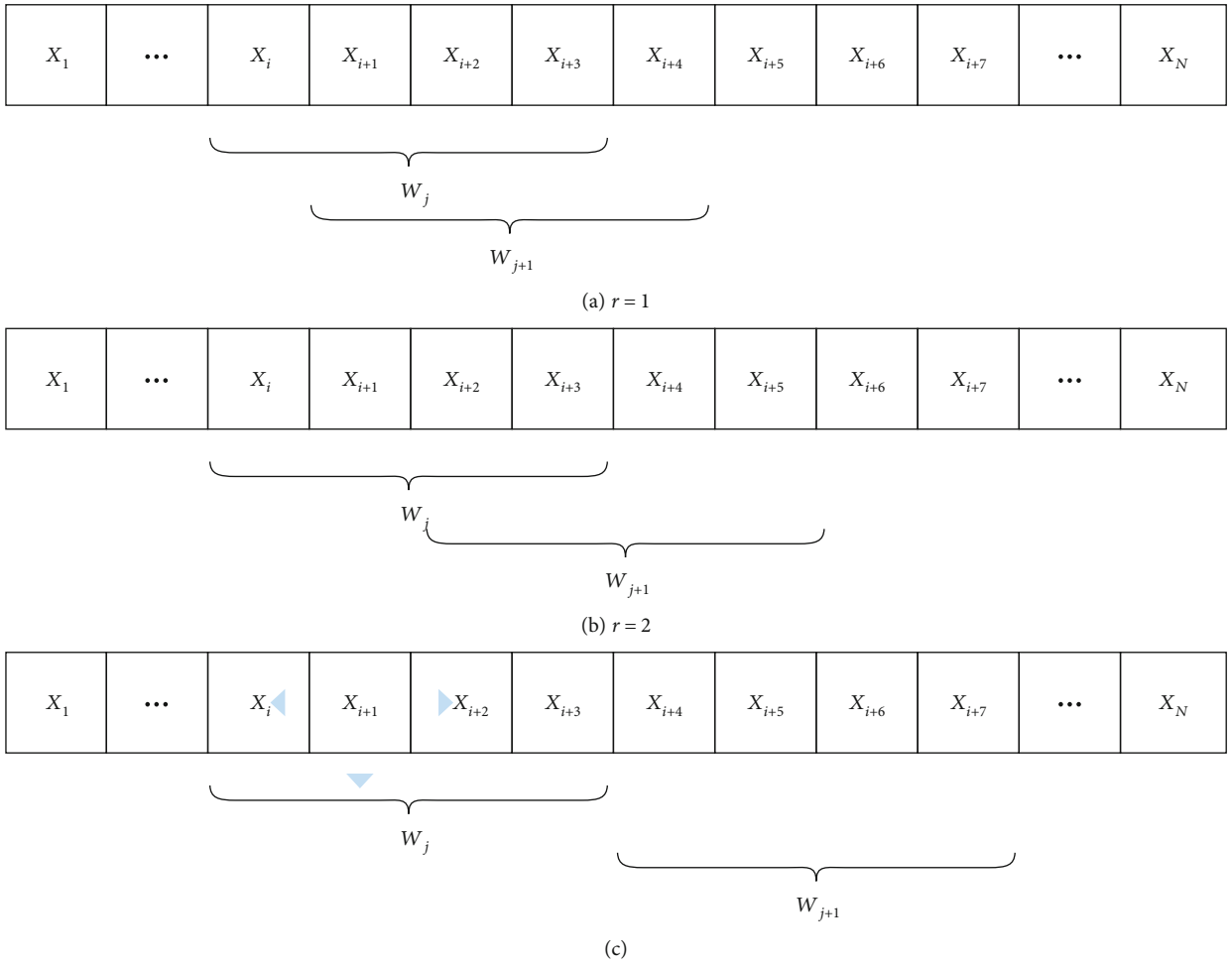


Figure 1: Flow sliding window model.

actual demand. However, to ensure the continuity of the sliding window in the data stream, it is necessary to ensure $1 \le r \le |W|$. The data stream $D = \{x_1, x_2, x_3, \cdots, x_N\}$ and two adjacent sliding windows W_j and W_{j+1}. Assumed $|W| = |W| = 4$ 且 $r = \{1, 2, 4\}$, Figure 1 shows the schematic diagram of adjacent sliding windows under different sliding amplitudes.

3. Methods

Most of the existing differential privacy methods for image publishing use image compression technology to transfer X_{mn}; then, Laplace noise is added to the transformed data. After that, the disturbing data is restored to obtain $X_{mn}{}'$ which includes noise. However, there are two kinds of errors in the process of obtaining $X_{mn}{}'$. One is the noise error LE $(X_{mn}{}')$ which is caused by the Laplace mechanism. The other is the reconstruction error RE $(X_{mn}{}')$ in the process of transformation. As a result, the total error Error $(X_{mn}{}')$ of "X" can be shown.

$$\text{Error}\left(X_{mn}{}'\right) = \text{LE}\left(X_{mn}{}'\right) + \text{RE}\left(X_{mn}{}'\right). \tag{8}$$

The process of converting an image into an image gray matrix can be understood as storing the gray value in each position of the corresponding two-dimensional matrix. As shown in Figure 2, X_{mn} is supposed to give each value a two-dimensional number to represent its position. This paper proposes to reconstruct the original two-dimensional number into a one-dimensional number, transform the image gray matrix into a data stream in the order of the one-dimensional number, and then, use the sliding window model to construct the data stream.

Because the image gray matrix does not have the overall mathematical meaning, namely, the numerical of X_{mn} is no correlation. It can be seen that the transformation of the 2D matrix into a 1D data stream does not destroy the distribution of the original image. Be in attention, the whole conversion process is reversible and lossless, so RE $(X_{m\times n}{}')$ is avoided. In the research of this paper, the total error of $X_{mn}{}'$ is expressed as

$$\text{Error}\left(X_{mn}{}'\right) = LE\left(X_{mn}{}'\right). \tag{9}$$

In this paper, the imaging differential privacy publishing method uses a sliding window model to construct the data

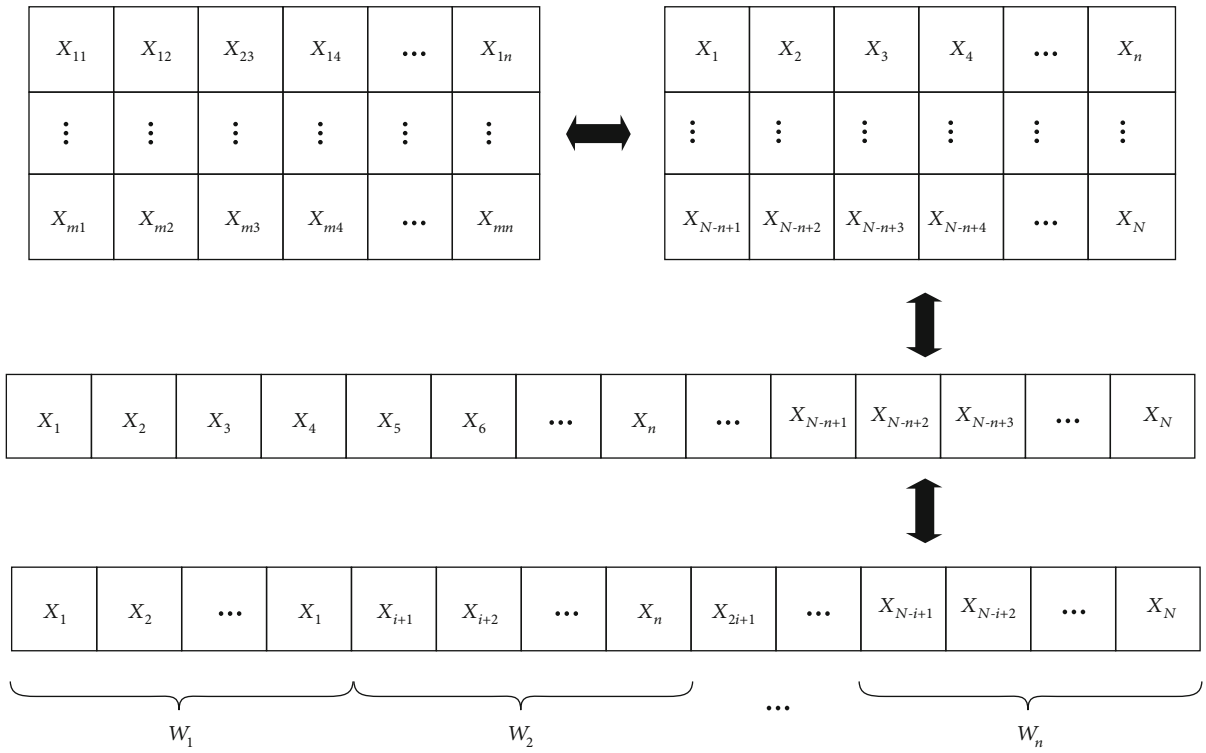


Figure 2: Matrix transform data flow.

stream. Assuming that there is a sliding window W_{j+1} at the moment of $j+1$, and set $r = |W|$. Whether the data in the current sliding window is noisy and published, it needs to be judged according to the preset threshold value. If W_{j+1} and the latest published W'_j (W'_j has been added noise by Laplace mechanism) is less similar than the default threshold, W'_j replace W_{j+1} to be released; otherwise, W_{j+1} is released as W'_{j+1} after it is allocated an appropriate privacy budget and adds noise. In the process of method implementation, the following three principles should be considered:

(1) The data samples contained in any two adjacent sliding windows are all from the image gray matrix. Comparing the similarity between the two samples by numerical difference alone cannot show all the features of the image. Therefore, the measurement method of samples similarity needs to be considered and designed from the image features

(2) Using a sliding window model to construct data flow needs to consider the allocation of the privacy budget. For any W'_j, the smaller privacy budget is allocated, the greater noise value is added, and the higher degree of privacy protection has, but this practice also caused the lower availability of W'_j. Therefore, it is necessary to establish a reasonable privacy budget allocation mechanism on the premise of satisfying ε-differential privacy

(3) The method should satisfy ε-differential privacy and improve the usability of the noisy image

3.1. LAP Algorithm. In this paper, the Laplace algorithm is proposed based on the Laplace mechanism. The method does not change the original data, but directly uses Laplace noise to disturbed the value of X_{mn} and then release $X_{mn}{}'$.

Given the gray matrix X_{mn} of an image, Laplace noise is added to each numerical matrix to obtain $X_{mn}{}' = X_{mn} + \sum_{i=1}^{m\times n} \mathrm{lap}(\Delta Q/\varepsilon)$. According to Theorem 4, the LAP algorithm satisfies the ε-differential privacy, and $\varepsilon = \varepsilon_{\mathrm{sum}}/mn$ is the privacy budget that is allocated to each number of the image gray matrix. $\varepsilon_{\mathrm{sum}}$ is the total privacy budget, and ε is the privacy budget allocated to each numerical. To explain the problem more simply, the size of ε will not be specified under the same conditions.

This paper will continue to describe the error size of the lap algorithm. Since there are only Laplace noises in the algorithm, the sum of squares of errors is obtained as:

$$\begin{aligned}\mathrm{Error}\left(\mathrm{X_{m\times n}}'\right) &= E\left(\sum_{i=1}^{m\times n}\left(x'_i - x_i\right)^2\right)\\ &= E\left(\sum_{i=1}^{m\times n}\left(x_i - x_i + \mathrm{lap}\left(\frac{\Delta Q}{\varepsilon}\right)\right)^2\right)\\ &= 2\mathrm{mn}\left(\frac{\Delta Q}{\varepsilon}\right)^2.\end{aligned} \tag{10}$$

In formula (10), the sensitivity is $\Delta Q = x_{\max} - x_{\min}$. Because the value of X_{mn} comes from the calculation result of the formula (1), so $\Delta Q \le 255$ exists. Besides, it can be seen from the above formula that the main factor determining the

error size in the lap algorithm is the selected image size ($m \times n$). When the selected image size is too large, using the lap algorithm will produce a large amount of noise, which will lead to the image with low usability after adding noise. To improve usability, this paper proposes an image privacy-preserving publishing method based on a sliding window.

3.2. The Measurement Method of Similarity. The image privacy protection publishing method proposed in this paper needs to rely on an accurate similarity measurement method. Different from an ordinary matrix, an image gray matrix not only has mathematical meaning but also includes a color feature, texture feature, and spatial distribution feature of the image itself. The commonly used similarity measurement methods include Manhattan distance, Chebyshev distance, and cosine distance. Due to the existing methods cannot analyze and compare samples from the perspective of images, the results obtained are available, but it will affect the usability of privacy protected images. Therefore, an accurate image data similarity measurement method—EM (exact mechanism)—is proposed.

Suppose W_x and W_y are two adjacent samples obtained by using a sliding window model after the image gray matrix is transformed into a data stream, where $W_x = \{X_1, X_2, X_3, \cdots, X_n\}$, $W_y = \{Y_1, Y_2, Y_3, \cdots, Y_n\}$. Euclidean distance is often used to compare the similarity of data. The formula is:

$$d(X, Y) = \sqrt[2]{\sum_{i=1}^{n} (X_i - Y_i)^2}. \tag{11}$$

Since the data in the sample comes from pixel values, the Euclidean distance is used to measure the similarity of image data, and the results can reflect the difference of color features between different samples. According to formula (11), the similarity formula of the two samples is

$$\theta = d(W_x, W_y). \tag{12}$$

There is an example to illustrate the limitation of θ obtained from Equation (12): there are three sample W_x, W_y, and W_z. Equation (12) is used to calculate θ_1 of W_x and W_y and θ_2 of W_x and W_z. If $\theta_1 < \theta_2$, W_x and W_y can be determined more similar. But actually, it is not like this. Different color feature, texture feature, is not based on pixel feature. It needs statistical calculation in the region containing multiple pixels. In pattern matching, this regional feature has great advantages and will not fail to match due to local bias. Suppose $W_x = \{0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}$, $W_y = \{80,80,80,80,80,80,80,80,80,80\}$, and $W_z = \{0, 0, 0, 0, 0, 0, 0, 0, 0, 255\}$, there is a large global deviation between W_x and W_y, and a minimum local deviation between W_x and W_z (only one data difference between different samples). Equation (12) does not consider the influence of texture features on image matching results.

Jaccard similarity index is used to measure the similarity between two sets. It is defined as the number of elements in the intersection of two sets divided by the number of elements in the union, the formula is

$$\rho(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}. \tag{13}$$

For considering the influence of texture features, the Jaccard similarity index of two samples is used as the optimal value to add to the calculation formula of θ, the formula is

$$\theta = \frac{d(W_x, W_y)}{\rho(W_x, W_y)} + \sigma. \tag{14}$$

In the formula, $d(W_x, W_y)$ is the Euclidean distance between two samples. $\rho(W_x, W_y)$ is the Jaccard similarity index between the two samples. σ appears as a correction value to ensuring the denominator is not 0. Besides, to reduce the influence of $\rho(W_x, W_y)$ to the calculation results of formula (14), the value range of σ should be controlled between 0.8 and 1.

Besides, hamming distance is often used to reflect the difference between two spatial vectors with the same structure and size. Hamming distance is a judgment method by comparing the values of the corresponding positions of two space vectors X and Y in turn. If the value of X_i and Y_i is the same, the result is 0. If it is different, it is 1. The result is to compare the values of all positions of the two space vectors in turn and then accumulate the result 1. The value obtained is the Hamming distance of the two space vectors. The formula of hamming distance is

$$\mu(X, Y) = \sum_{i=1}^{n} X_i \oplus Y_i. \tag{15}$$

The hamming distance can effectively measure the difference of spatial distribution characteristics between two samples, but its calculation result is a positive integer accumulated by the number 1, which cannot be directly used in the calculation of θ. Besides, the calculation results of Hamming distance are limited by the sample size, and the Hamming distance of different size samples is not comparable. To solve this problem, a method combining Hamming distance with a perceptual hash algorithm [55] is proposed to add a disturbance value φ to formula (14). It further enhances the accuracy of θ. The values of φ are as follows:

$$\begin{cases} \dfrac{\theta'(\rho(W_x, W_y) + \sigma)}{d(W_x, W_y)}, \dfrac{\theta'(\rho(W_x, W_y) + \sigma)}{d(W_x, W_y)}, \mu(W_x, W_y) < 0.078n, \\ 1, 0.078n \le \mu(W_x, W_y) \le 0.156n, \\ \dfrac{1}{\lg(n - \mu(W_x, W_y))}, \mu(W_x, W_y) > 0.156n. \end{cases} \tag{16}$$

In the above formula, n is the size of the sample, and θ' is the minimum value given in advance. If $\theta > \theta'$, the two

samples are not similar. If $\theta \le \theta'$, the two samples are similar. To sum up, the EM mechanism can be expressed as

$$\theta = \varphi\left(\frac{d(X, Y)}{\rho(X, Y)} + \sigma\right). \tag{17}$$

3.3. SWP Algorithm. Different from the lap algorithm, SWP (sliding window publication) uses a sliding window model to construct data flow. Through continuous moving of sliding window, it calculates $\theta(2 \le i \le k)$ between W_i and W'_{i-1} by EM mechanism. If $\theta \le \theta', W'_{i-1}$ replaces W_i to release W'_{i-1}; if $\theta > \theta'$, the DA mechanism is used to allocate an appropriate privacy budget to get W'_i, then release W'_i. The implementation process of the SWP algorithm is as follows.

Note that in the SWP algorithm, we use the iterative method to take a dynamic allocation privacy budget mechanism—DA (dynamic allocation) mechanism which can be used for a limited data stream. Different from the commonly used dichotomy privacy budget allocation mechanism, the DA mechanism solves the problem of too fast consumption of privacy budget in a data stream and reduces the impact of noise on the original data to a certain extent.

3.3.1. Dynamic Allocation. In the DA mechanism, the privacy budget for each consumption is assumed to be $\varepsilon_i = \{\varepsilon_1, \varepsilon_2, \varepsilon_3, \cdots, \varepsilon_m\}$, and the number of sliding windows affected by each consumption of privacy budget is $s_i = \{s_1, s_2, s_3, \cdots, s_m\}$, there is $s_1 + s_2 + s_3 + \cdots + s_m = k$, and $s_1 = 1$, K is the total number of sliding windows. Besides, $\varepsilon_i^{\text{left}}$ is the remaining privacy budget after i status is completed, then there is

$$\begin{aligned}
\varepsilon_1 &= \frac{\varepsilon}{k}, \quad \varepsilon_1^{\text{left}} = \frac{\varepsilon - \varepsilon}{k},\\
\varepsilon_2 &= \frac{\varepsilon_1^{\text{left}}}{(k - s_1)}, \quad \varepsilon_2^{\text{left}} = \varepsilon_1^{\text{left}} - \varepsilon_2,\\
\varepsilon_3 &= \frac{\varepsilon_2^{\text{left}}}{(k - s_1 - s_2)}, \quad \varepsilon_3^{\text{left}} = \varepsilon_2^{\text{left}} - \varepsilon_3,\\
\varepsilon_4 &= \frac{\varepsilon_3^{\text{left}}}{(k - s_1 - s_2 - s_3)}, \quad \varepsilon_4^{\text{left}} = \varepsilon_3^{\text{left}} - \varepsilon_4,\\
m-1.\varepsilon_{m-1} &= \frac{\varepsilon_{m-2}^{\text{left}}}{(k - s_1 - s_2 \cdots - s_{m-1})}, \quad \varepsilon_{m-1}^{\text{left}} = \varepsilon_{m-2}^{\text{left}} - \varepsilon_{m-1},\\
m.\varepsilon_m &= \frac{\varepsilon_{m-1}^{\text{left}}}{(k - s_1 - s_2 \cdots - s_m)}, \quad \varepsilon_m^{\text{left}} = \varepsilon_{m-1}^{\text{left}} - \varepsilon_m.
\end{aligned} \tag{18}$$

We can see from the above formula of DA that under any state, ε_i, and $\varepsilon_{i\text{left}}$ the general formula of ε_i and $\varepsilon_{i\text{left}}$ is as follows

$$\varepsilon_i = \frac{\varepsilon_{i-1}^{\text{left}}}{\left(k - \sum_{j=1}^{i-1} s_j\right)}, \tag{19}$$

$$\varepsilon_i^{\text{left}} = \varepsilon_{i-1}^{\text{left}} - \varepsilon_i. \tag{20}$$

As shown in Figure 3, in the implementation process of the DA mechanism, W_1 is added Laplace noise to generate W'_1 and publish it; the privacy budget ε/k is consumed with W'_1 and then publish it. Using EM mechanism to calculate the θ between W'_1 and W_2, if the condition $(\theta <= \theta')$ is satisfied, then W_2 is replaced by W'_1 and released. The above operation is repeated until finding all W_i that meet the conditions. When W_2, W_3, and W_4 are the suitable conditions, all of them are replaced with W'_1. Because the replacement operation does not consume the privacy budget, the total privacy budget consumed by the four sliding windows is ε/k. If θ between W'_1 and W_2 does not satisfy the condition $(\theta > \theta')$, W_5 adds Laplace noise to generate W'_5 and publish it; the privacy budget which cost by W_5 is $\varepsilon(k-1)/k(k-4)$..

Theorem 8. *In SWP, the privacy budget consumed will not exceed ε, that is $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \cdots + \varepsilon_m \le \varepsilon$.*

Certificate: according to formula (17), when $k - \sum_{j=1}^{m-1} s_j = 1$, $\varepsilon_{m-1}^{left} = \varepsilon_m$, then $\varepsilon_m^{left} = 0$. In this state, it is expressed as

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \cdots + \varepsilon_m = \varepsilon, \tag{21}$$

where $k - \sum_{j=1}^{m-1} s_j > 1$, $\varepsilon_{m-1}^{left}/\varepsilon_m > 1$, then $\varepsilon_m^{left} > 0$. In this state, the formula is

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \cdots + \varepsilon_m < \varepsilon. \tag{22}$$

The proof is complete.

Property 1 sequence combination property: D is assumed as a privacy data set, $A_1, A_2, A_3, \cdots, A_n$ is n random algorithms, and $A_i(1 \le i \le n)$ satisfies ε-differential privacy with $\varepsilon = \sum_{i=1}^{n} \varepsilon_i$.

Theorem 9. *SWP method satisfies ε-differential privacy*

Certificate: It can be seen from Theorem 4:

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \cdots + \varepsilon_m \le \varepsilon. \tag{23}$$

In the formula, $\varepsilon_i(1 \le i \le m)$ is the budget privacy which is consumed in each state. According to the sequence combination property of difference privacy (property 1), the SWP method satisfies ε-differential privacy. The proof is complete.

Theorem 10. *The error of the SWP method is not greater than the lap algorithm, it is:*

$$Error(SWP) \le Error(Lap). \tag{24}$$

Certificate: the DA used in the SWP method consumes a privacy budget of $\varepsilon_i = \{\varepsilon_1, \varepsilon_2, \varepsilon_3, \cdots, \varepsilon_m\}$. The number of sliding windows affected by each consumption of privacy budget is $s_i = \{s_1, s_2, s_3, \cdots, s_m\}$, and there is $s_1 + s_2 + s_3 + \cdots + s_m = k$. The Laplace mechanism adopted by the lap algorithm consumes a privacy budget ε/k and $\sum_{i=1}^{k} s'_i = k$ $(s'_1 = s'_2 = s'_3 = \cdots = s'_k = 1)$

input : sliding window $W = \{W_1, W_2, W_3, W_4, \cdots, W_k\}$, privacy budget ε, θ'
output : 1. $s = 0$;
2. $W_1' = W_1 + \text{lap}(\varepsilon/k)$;
3. $\varepsilon_{\text{left}} = \varepsilon(k-1)/k$;
4. for $i = 2 + s$ to k
5. use EM mechanism to calculate θ between W_{1+s}' and W_{2+s};
6. If $\theta <= \theta'$
7. $W_i = W_{1+s}'$;
8. $s = s + 1$;
9. else
10. $s = s + 1$;
11. $W_i' = W_i + \text{lap}(\varepsilon_{\text{left}}/(k-s))$;
12. $\varepsilon_{\text{left}} = \varepsilon_{\text{left}} - \varepsilon_{\text{left}}/(k-s)$;
13. $W_{1+s}' = W_i'$;
14. end if
15. end for
16. return $W' = \{W_1', W_2', W_3', W_4', \cdots, W_k'\}$.

Algorithm 1: SWP.

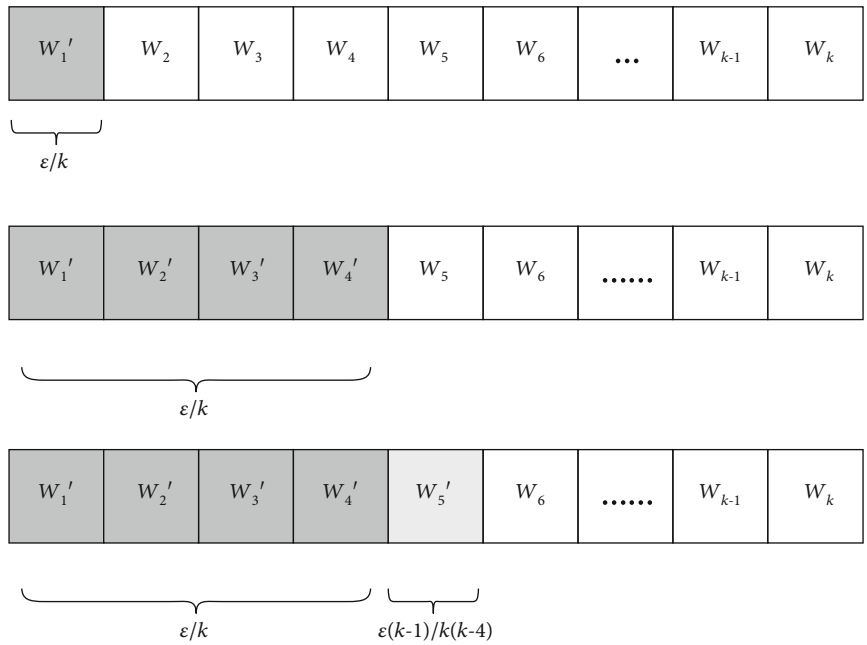


Figure 3: Private budget allocation.

First of all, the maximum error caused by the DA mechanism is calculated. In this case, it knows $s_1 = s_2 = s_3 = \cdots s_m = 1$, means $m = k$, so it is

$$\begin{aligned}\text{Error(SWP)}_{\max} &= \left(\text{lap}\left(\frac{\Delta Q}{\varepsilon_1}\right) + \text{lap}\left(\frac{\Delta Q}{\varepsilon_2}\right) + \text{lap}\left(\frac{\Delta Q}{\varepsilon_3}\right) + \cdots \text{lap}\left(\frac{\Delta Q}{\varepsilon_m}\right)\right)_{\max} \\ &= \sum_{i=1}^{k} \text{lap}\left(\frac{\Delta Q}{\varepsilon_i}\right) = k \times \text{lap}\left(\frac{\Delta Q}{(\varepsilon/k)}\right) = \text{Error(Lap)}.\end{aligned} \tag{25}$$

For other cases, it can be judged according to formula (20). After the completion of state i, the remaining privacy budget of DA is $\varepsilon_i^{\text{left}}$, the remaining privacy budget of the Laplace mechanism is $\acute{\varepsilon}_i^{\text{left}}$. The DA allows $s_i > 1$ and assume the existence of $s_1 = s_2 = s_3 = \cdots s_{i-1} = 1 (1 < i < m)$. From this, it can get

$$\begin{aligned}\varepsilon_i^{\text{left}} - \acute{\varepsilon}_i^{\text{left}} &= \left(\varepsilon - \varepsilon \times \frac{k - (i - 1 + s_i)}{k}\right) - \left(\varepsilon - \varepsilon \times \frac{k-i}{k}\right) \\ &= \varepsilon \times \left(\frac{k-i}{k} - \frac{k-(i-1+s_i)}{k}\right) = \varepsilon \times \frac{s_i - 1}{k}.\end{aligned} \tag{26}$$

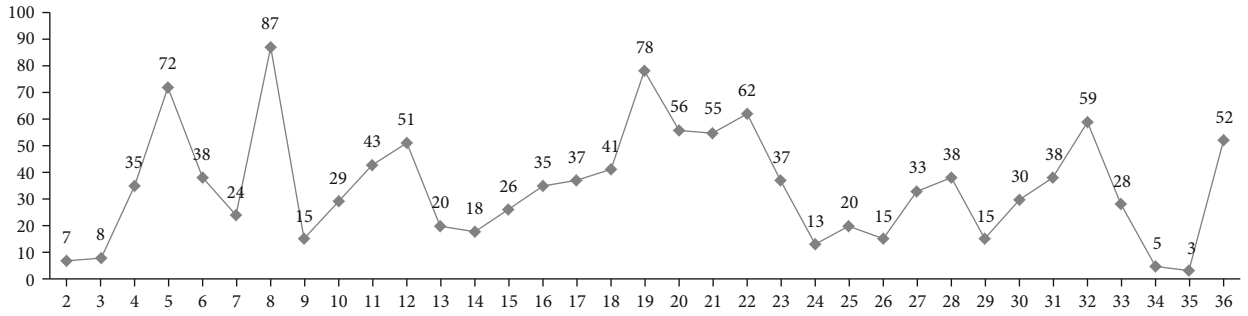


Figure 4: Original sequence.

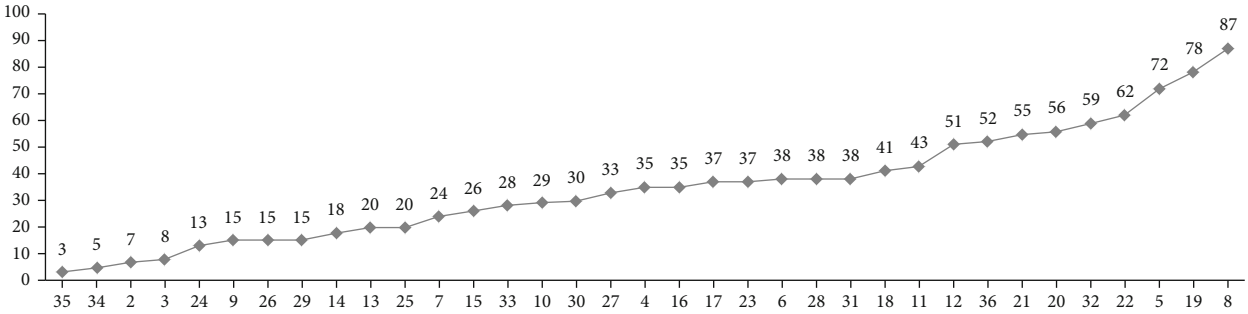


Figure 5: Sort sequence.

Due to $s_i > 1$, so $\varepsilon_i^{\text{left}} - \varepsilon_i^{\text{left}} > 0$, then $(\varepsilon_{i+1}, \varepsilon_{i+2}, \varepsilon_{i+3}, \cdots, \varepsilon_m) > \varepsilon/k$, it is shown that Error(SWP) < Error(Lap).The proof is complete.

3.4. Sort-SWP Algorithm. In the SWP algorithm, the sliding window is set to move forward only in one-dimensional space, so the impact range s_i of W_i' is completely determined by θ'. The verification method is not complicated. Assuming that there is an image x which use the SWP method to protect ε-privacy protection of X (assuming $k = 36$), and θ_i between W_1' and every $W_i (2 \le i \le 36)$ is calculated. Figure 4 shows the distribution of specific data (abscissa is i, the ordinate is θ). When $\theta' = 10$, W_2 and W_3 will be replaced by W_1' to publish. From this, we can see that $s_1 = 2$.

However, by looking at all θ_i in Figure 3, it is found when $\theta' = 10$, it is not only W_2 and W_3 satisfy the condition $\theta \le \theta'$ but also including W_{34} and W_{35}. According to the proof process of Theorem 10, under the same conditions, the larger the value of s_i is, the smaller the $\text{Error}(X')$ is. If the SWP method is executed after sorting θ, assumed $\theta' = 10$, $s_1 = 4$ can be obtained. The sorted result is shown in Figure 5.

However, the direct sorting of data will destroy the ε -differential privacy [56]. Reference [57] proposed to add independent Laplace noise to each subdataset and then use the added noise value to sort. But this method will cause inaccurate sorting results. Reference [58] proposed to use the repeated index mechanism to sample the data and then to sort the data according to the order of sampling. However, the accuracy of the ranking results is too dependent on the allocation of the privacy budget. Therefore, an approximate sorting method—SAS (Simulated Annealing-Sort)—is proposed to satisfy ε-differential privacy. In the SAS algorithm, the sampling technique in exponential mechanism is used to get the initial sorting results, and then Simulated Annealing Algorithm is used to optimize the initial sorting results, so more accurate sorting results can be obtained.

The Simulated Annealing Algorithm comes from the principle of solid annealing, and its core is the choice of solution. Simulated Annealing Algorithm is also a greedy algorithm. Random factors are introduced in its implementation process, which makes the algorithm accept a solution that is worse than the current solution with a certain probability. The annealing process in thermodynamics can be summarized as the state that the body with variable temperature drops slowly and reaches the lowest energy among molecules. Suppose there are n discrete states in the thermodynamic system S, and the energy of the current state is E_{old}, the energy of the next state is E_{new}. According to the Metropolis criterion, the probability P of state transition at temperature T $(1 \le \text{old}, \text{new} \le n)$ is expressed as follows:

$$p = \begin{cases} 1 & E_{\text{new}} < E_{\text{old}}, \\ \exp\left(-\dfrac{E_{\text{new}} - E_{\text{old}}}{\gamma T}\right) & E_{\text{new}} \ge E_{\text{old}}. \end{cases} \tag{27}$$

According to Equation (27), the receiving state is transferred when $E_{\text{new}} < E_{\text{old}}$. It will accept the state transition with probability P when $E_{\text{new}} \ge E_{\text{old}}$. As a positive number which is less than 1, γ is used to control the change rate of temperature T by an iterative method, which also shows that

Input : sliding window $W = \{W_1, W_2, W_3, W_4, \cdots, W_k\}$, Privacy budget ε_1 and ε_2, initial temperature T, Cooling coefficient γ, Termination temperature $T_{\min}$
Output : Approximation of $\tilde{W}$
1. $W_1' = W_1 + \text{lap}(\varepsilon_2)$;
2. Using *EM* mechanism to calculate the similarity between W_1' and all other sliding windows, get $\theta = \{\theta_1, \theta_2, \theta_3, \cdots, \theta_{k-1}\}$;
3. $z = 0$;
4. *while* $(k - z > 1)$
5. The scoring function of exponential mechanism ΔQ is $1/\theta + \sigma$;
6. An unlabeled $\theta_j j = 1 \ldots k - 1$ is selected with probability P $\propto \exp(\varepsilon_1 \Delta Q / 2\Delta u)$ by using the exponential mechanism
7. marked θ_j which is currently selected
8. $z = z + 1$;
9. $\widetilde{\theta}_z = \theta_j$;
10. $S_z = j$;
11. while $(T > T_{\min})$
12. Random select $\widetilde{\theta}_i i = 1 \cdots z - 1$;
13. $dE = \widetilde{\theta}_i - \widetilde{\theta}_{i+1}$;
14. *if*$(dE > = 0)$
15. Change the position of $\widetilde{\theta}_i$ and $\widetilde{\theta}_{i+1}$;
16. Update sorting of data in $\widetilde{\theta}$;
17. Change the position of S_i *and* S_{i+1};
18. *else*
19. *if*$(\exp(dE/T) > random(0, 1))$
20. Change the position of $\widetilde{\theta}_i$ and $\widetilde{\theta}_{i+1}$;
21. Update sorting of data in $\widetilde{\theta}$;
22. Change the position of S_i and S_{i+1};
23. *end if*
24. T = T $*$ γ;
25. Return the sliding window $\tilde{W}$ ={W_1, W_M, …} after sorting, initial position M = {$S_1 + 1, \cdots, S_z + 1$}.

Algorithm 2: SAS.

the higher the temperature is, the greater the value of P is. Conversely, the lower the temperature is, the smaller the value of P will be.

Because the Simulated Annealing Algorithm is probability-based, the solution obtained is not necessarily correct. It is not contrary to the core idea of differential privacy. Based on the above analysis, the implementation details of the SAS algorithm are given.

In the SAS algorithm, the first line is to distribute W_1 to get W_1'. In line 2, using EM mechanism calculates the similarity of W_1' with all other sliding windows and stored in θ. Lines 4-7 use the exponential mechanism to rank θ from small to large. The fifth row is the scoring function of the index mechanism. In the SAS algorithm, we want those sliding window which is more similar to W_1' extracted with a higher probability. σ is a small positive number to ensure that the denominator of the scoring function is not 0. Besides, because deleting or adding a record only affects one count of θ, $\Delta u = 1$. In line 9, the sorted results are stored in $\widetilde{\theta}$. Line 10 is to store the sequence number of the position before sorting θ into S. In line 11-24, the annealing algorithm is used to optimize the sorting results. In line 12-17, $\widetilde{\theta}_i$ is randomly selected in $\widetilde{\theta}$, and the position of $\widetilde{\theta}_i$ and $\widetilde{\theta}_{i+1}$ is changed. Assumed energy change from it is $dE = \widetilde{\theta}_i - \widetilde{\theta}_{i+1}$. If $dE > = 0$, the changed $\widetilde{\theta}$ and S are obtained after accepting the adjustment of this position, and the changed $\widetilde{\theta}$ and S are regarded as the starting point of the next change. Lines 18-23, if $de < 0$, it is accepted as the change with a certain probability and make the corresponding adjustment. Note that when $dE < 0$, $\exp(dE/T)$ ranges from 0 to 1. In line 24, the temperature T is adjusted by using γ (the value of γ is generally set as a positive number slightly less than 1). Since it is an annealing process, it can be seen from formula (12) that the value of T will decrease with the increase of iteration times, and this change will cause the acceptance probability to decrease.

Based on the above contents, a sort-SWP method which is an optimization method based on sorting is proposed. The implementation process of the algorithm is shown as follows.

The sort SWP method consists of three main steps: first, the SAS algorithm is used to sort W to get $\tilde{W}$, where ε_1 is used for exponential mechanism, and ε_2 is used in the Laplace mechanism. To ensure the consistency of the sort SWP method, $\varepsilon_2 = \varepsilon_3/k$ is set. Secondly, the SWP algorithm is used to add noise to $\widetilde{W}$, then gets $\tilde{W}'$; the third (lines 5-11), because the distribution of data in $\tilde{W}'$ destroys the structure of the original image, the position number stored in M is used to restore $\tilde{W}'$ to W'.

Theorem 11. *Sort SWP method satisfies ε-differential privacy*

Input : sliding window $W = \{W_1, W_2, W_3, W_4, \cdots, W_k\}$, Privacy budget ε, θ'
Output : Satisfying differential privacy W'
1. $\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$;
2. $\tilde{W} = \mathrm{SAS}(W, \varepsilon_1, \varepsilon_2)$;
3. $\tilde{W}' = \mathrm{SWP}(\tilde{W}, \varepsilon_3)$;
4. for $i = 1$ to k
5. if $i = 1$
6. $W_1' = \tilde{W}_1'$;
7. Else
8. $m = M_{i-1}$;
9. $W_m' = \tilde{W}_i'$;
10. endif
11. end for
12. return $W' = \{W_1', W_2', W_3', W_4', \cdots, W_k'\}$.

Algorithm 3: Sort-SWP.

Certificate: in the sort SWP method, the total privacy budget is divided into three parts, and there exists $\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$. *Inside,* ε_1 *used exponential machine to select the sliding window* W *and the probability is proportional to* $\exp(\varepsilon_1 \Delta Q / 2\Delta u)$, *it shows as*

$$\Pr[M(X, \Delta Q) = W] = \frac{\exp((\varepsilon_1 \Delta Q(X, W))/2\Delta u)}{\sum_{W' \in O} \exp\left(\left(\varepsilon_1 \Delta Q\left(X, W'\right)\right)/2\Delta u\right)}. \tag{28}$$

Given X and its adjacent X', for any value of W ($W \in O$), it can be seen from formula (28)

$$\begin{aligned}
&\frac{\Pr[M(X, \Delta Q) = W]}{\Pr\left[M\left(X', \Delta Q\right) = W\right]} \\
&= \frac{\exp(\varepsilon_1 \Delta Q(X, W)/2\Delta u) / \sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X, W'\right)/2\Delta u\right)}{\exp\left(\varepsilon_1 \Delta Q\left(X', W\right)/2\Delta u\right) / \sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X', W'\right)/2\Delta u\right)} \\
&= \left(\frac{\exp(\varepsilon_1 \Delta Q(X, W)/2\Delta u)}{\exp\left(\varepsilon_1 \Delta Q\left(X', W\right)/2\Delta u\right)}\right) \times \left(\frac{\sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X', W'\right)/2\Delta u\right)}{\sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X, W'\right)/2\Delta u\right)}\right) \\
&= \exp\left(\frac{\varepsilon_1\left(\varepsilon_1 \Delta Q(X, W) - \varepsilon_1 \Delta Q\left(X', W\right)\right)}{2\Delta u}\right) \\
&\quad \times \left(\frac{\sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X', W'\right)/2\Delta u\right)}{\sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X, W'\right)/2\Delta u\right)}\right) \le \exp\left(\frac{\varepsilon_1}{2}\right) \\
&\quad \times \left(\frac{\sum_{W' \in O} \exp(\varepsilon_1/2) \times \exp\left(\varepsilon_1 \Delta Q\left(X, W'\right)/2\Delta u\right)}{\sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X, W'\right)/2\Delta u\right)}\right) \\
&\le \exp\left(\frac{\varepsilon_1}{2}\right) \times \exp\left(\frac{\varepsilon_1}{2}\right) \times \left(\frac{\sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X, W'\right)/2\Delta u\right)}{\sum_{W' \in O} \exp\left(\varepsilon_1 \Delta Q\left(X, W'\right)/\varepsilon_1 \Delta Q\left(X, W'\right)\right)}\right) \\
&= \exp(\varepsilon_1).
\end{aligned} \tag{29}$$

It can be seen that the process of selecting W by using the exponential mechanism satisfies ε_1-differential privacy in Sort-SWP. Besides, according to Theorem 4, the Sort-SWP method satisfies ε_2-differential privacy. According to the proof process of Theorem 9, it can be concluded that the Sort-SWP method satisfies ε_3-differential privacy. Therefore, the whole process of the Sort-SWP method satisfies the ε -differential privacy. The proof is complete.

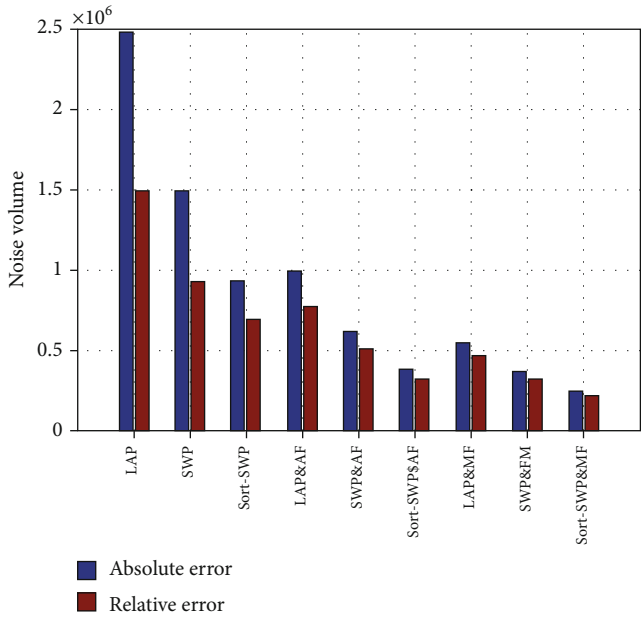
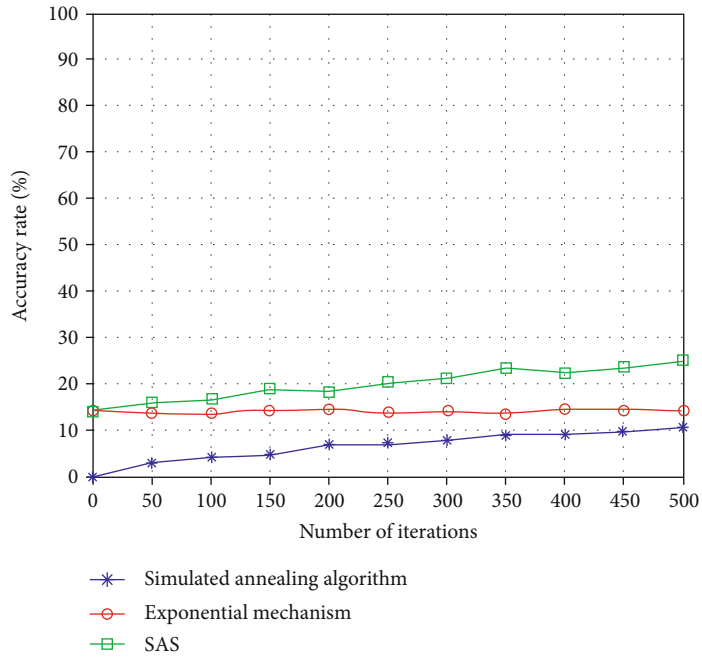


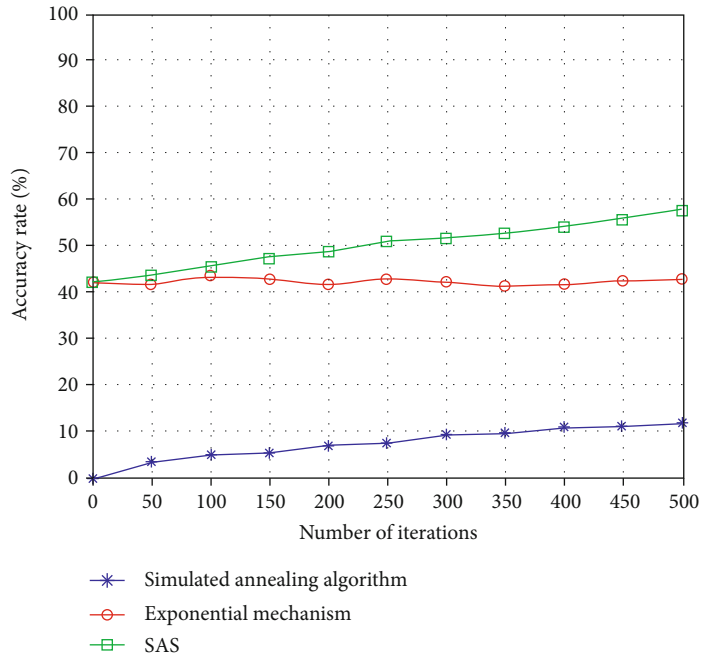
Figure 6: Noise statistics.

4. Experiment and Result Analysis

4.1. Sorting Method Experiment. Direct sorting will destroy the ε-differential privacy. Therefore, this paper adopts an approximate sorting method—SAS algorithm—which satisfies the ε-differential privacy. During the test, a one-dimensional array which is containing 57 random numbers is used as the experimental data set, and three sorting methods, namely, simulated annealing algorithm, exponential mechanism, and *SAS* algorithm, are selected for comparison. Besides, to verify the impact of the privacy budget on

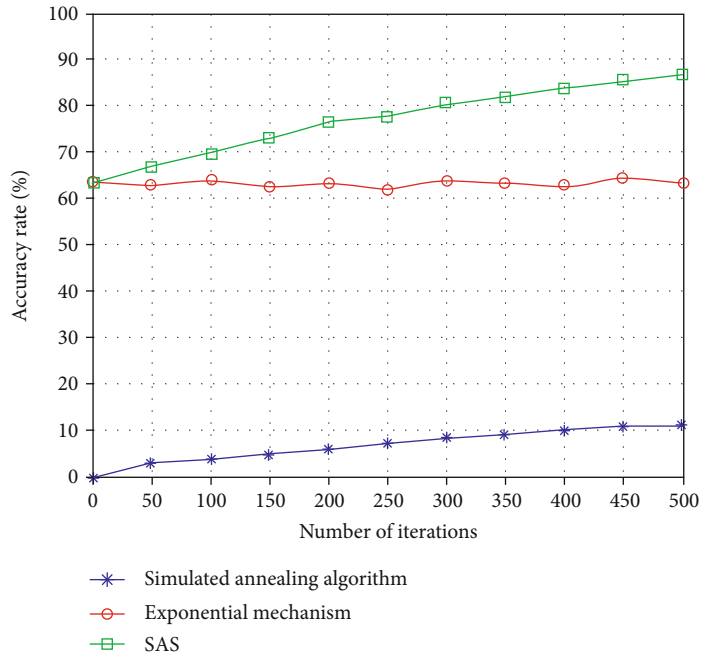


(a) $\varepsilon = 0.4$

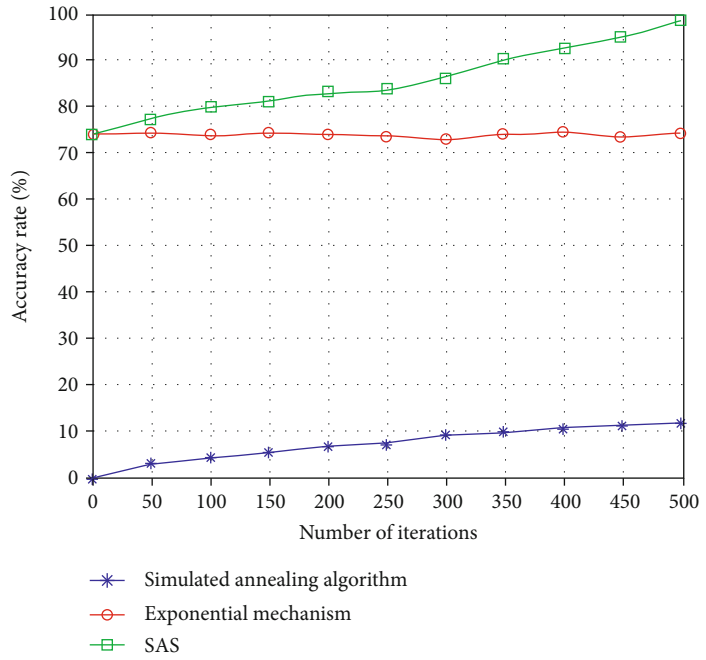


(b) $\varepsilon = 0.5$

Figure 7: Continued.



(c) $\varepsilon = \ln 2$



(d) $\varepsilon = 1$

Figure 7: Sorting result test.

different ranking results, ε is taken as 0.4, 0.5, ln2, and 1, respectively. The results are shown in Figure 6.

In Figure 7, the abscissa represents the number of times the algorithm is executed, and the ordinate represents the accuracy of the sorting. The following three conclusions can be drawn from the experimental results: (1) the sorting results of the simulated annealing algorithm are only related to the execution times of the algorithm and not relevant to the value of ε; (2) the ranking results of the exponential mechanism are only related to the value of ε, and the higher the value of ε is, the higher the sorting accuracy is. It is independent of the execution times of the algorithm; (3) the sorting results of the SAS algorithm are related to the value of ε and also related to the execution times of the algorithm, and they are proportional.

4.2. Simulation Experiment. In different kinds of images, the sensitive information of face photos is the most representative. Therefore, to verify the usability of the algorithm, it selects a front face photo $(190 * 160)$ to complete the

(a) Original image (b) LAP

(c) SWP (d) Sort-SWP

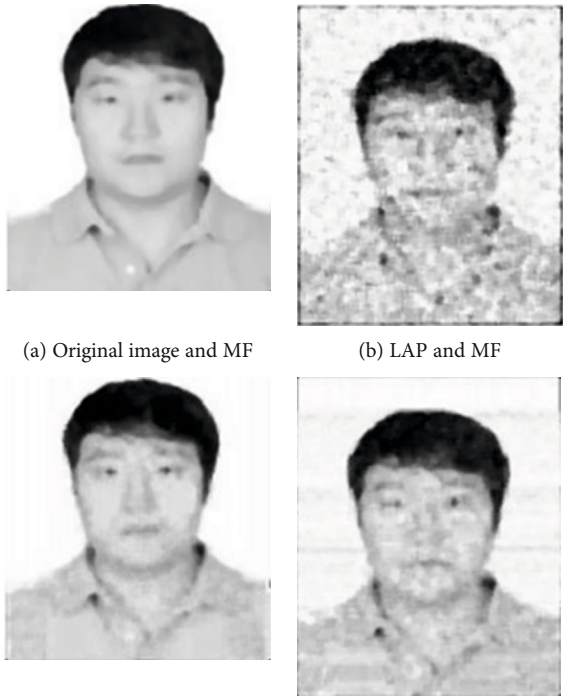
Figure 8: Experimental results.

(a) Original image and AF (b) LAP and AF

(c) SWP and AF (d) Sort-SWP and AF

Figure 9: Experimental results graph and mean filtering.

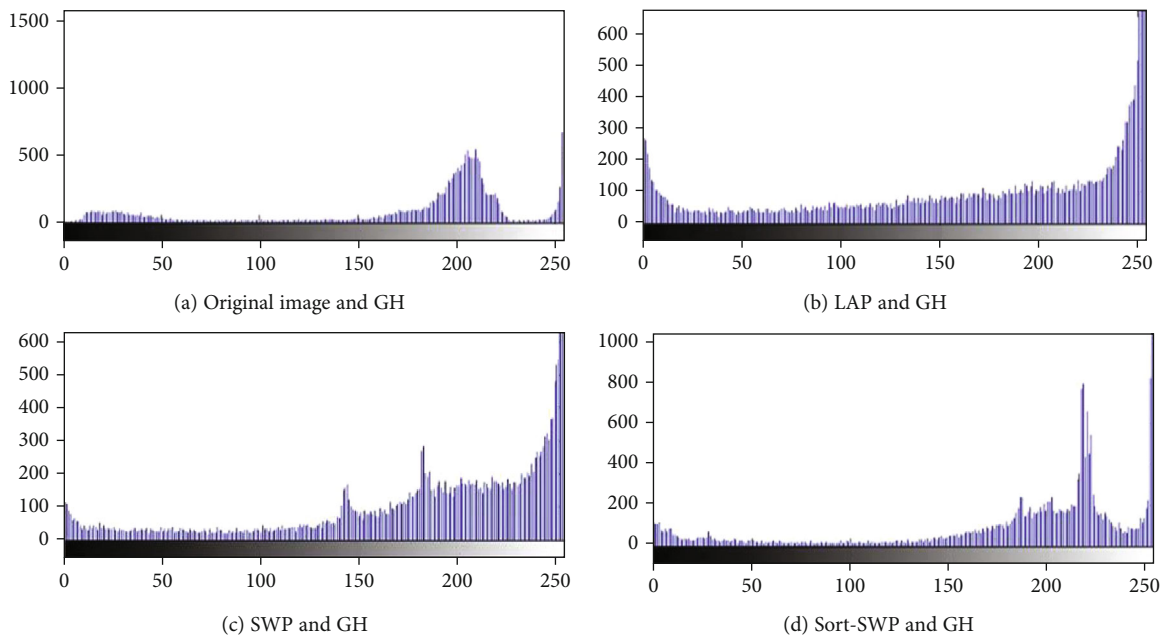
corresponding experiment in this paper. The experimental results of different algorithms under the same conditions are shown in Figure 8. Figure 8(a) is the original image; Figure 8(b) is the resulting graph obtained by directly adding Laplace noise to the original image; Figure 8(c) is the resulting graph obtained by using the SWP algorithm; Figure 8(d) is the resulting graph obtained by using sort SWP algorithm. Compared with the LAP algorithm, the



(a) Original image and MF (b) LAP and MF

(c) SWP and MF (d) Sort-SWP and MF

Figure 10: Experimental result graph and median filter.



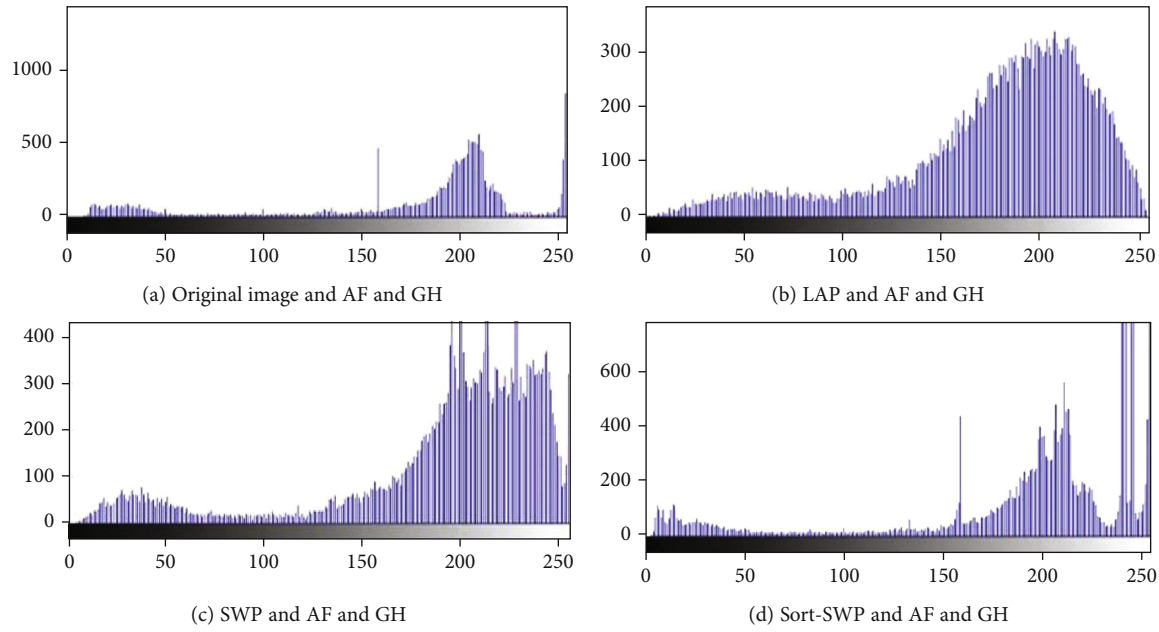
(a) Original image and GH (b) LAP and GH

(c) SWP and GH (d) Sort-SWP and GH

Figure 11: Gray histogram of experimental image.

SWP algorithm and sort SWP algorithm have better display effect.

The core of this paper is to use the Laplace mechanism in differential privacy protection to add noise to the image and to protect the sensitive information in the image. Therefore, to destroy the effect of privacy protection, it is a common method to eliminate or reduce the impact of image noise. For the methods of eliminating image noise, the representative methods are average filtering or median filtering. In Figures 9 and 10, the mean filter and median filter are used to interfere with the experimental results in Figure 8. The purpose of this is to evaluate the degree of privacy protection and the usability of the protected images after noise elimination.



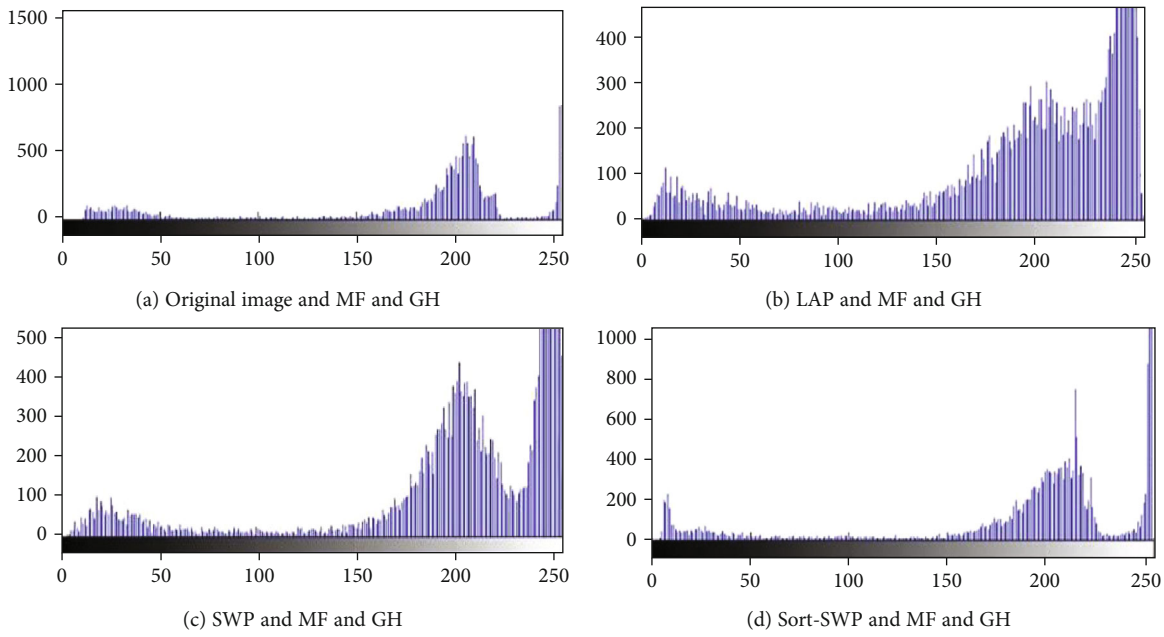
(a) Original image and AF and GH

(b) LAP and AF and GH

(c) SWP and AF and GH

(d) Sort-SWP and AF and GH

Figure 12: Experimental image gray histogram and average filtering.



(a) Original image and MF and GH

(b) LAP and MF and GH

(c) SWP and MF and GH

(d) Sort-SWP and MF and GH

Figure 13: Experimental image gray histogram and median filter.

Now, we have given a preliminary conclusion that the SWP algorithm and sort SWP algorithm are better than the LAF method. However, this conclusion is only obtained by observing the experimental results, so it may be one-sided or inaccurate. To further verify whether the conclusion is correct, the previous 12 images are transformed into a gray histogram for further comparison. Gray histogram reflects the relationship between the frequency of gray level pixels and gray level in an image, although the gray histogram cannot reflect the specific distribution of image pixels but reflect the impact of Laplace noise on the image as a statistical result.

It can be seen from Figures 11–13, under the same conditions, the histogram distribution of the SWP algorithm and lap algorithm is closer, but it is quite different from the distribution of pixels in the original image. The distribution of pixels in the histogram of the result of the sort SWP algorithm is consistent with the original image. Therefore, it can be considered that the noise interference caused by the sort SWP algorithm is the least.

In the papers related to differential privacy protection, the noise size is an important evaluation index to measure the advantages and disadvantages of the algorithm. Under

Table 1: ORL and precision.

Precision	LAP	SWP	Sort-SWP	LAP and AF	SWP and AF	Sort-SWP and AF	LAP and MF	SWP and MF	Sort-SWP and MF
$\varepsilon = 1$	6.52%	39.29%	53.86%	9.27%	47.09%	60.55%	14.95%	49.01%	65.39%
$\varepsilon = 2$	8.83%	43.45%	69.26%	15.86%	56.39%	72.06%	25.37%	57.61%	75.41%
$\varepsilon = 3$	11.37%	48.29%	71.66%	22.00%	63.39%	79.27%	36.61%	72.18%	83.68%
$\varepsilon = 4$	13.96%	55.18%	75.19%	27.71%	70.43%	82.51%	44.36%	80.97%	87.55%
$\varepsilon = 5$	16.67%	62.72%	82.83%	33.08%	76.36%	86.19%	51.98%	88.70%	91.33%

Table 2: ORL and recall.

Recall	LAP	SWP	Sort-SWP	LAP and AF	SWP and AF	Sort-SWP and AF	LAP and MF	SWP and MF	Sort-SWP and MF
$\varepsilon = 1$	71.32%	77.96%	83.58%	72.35%	81.63%	84.57%	80.12%	83.47%	88.14%
$\varepsilon = 2$	73.56%	75.68%	81.83%	75.52%	81.26%	83.67%	78.68%	79.85%	80.33%
$\varepsilon = 3$	75.79%	83.81%	85.25%	77.53%	84.36%	85.23%	81.18%	86.63%	84.54%
$\varepsilon = 4$	72.32%	74.22%	78.58%	73.52%	78.59%	82.18%	75.36%	81.23%	87.10%
$\varepsilon = 5$	77.91%	76.23%	80.12%	78.07%	80.11%	82.51%	80.12%	81.15%	83.94%

Table 3: ORL and F1-score.

F1-sroce	LAP	SWP	Sort-SWP	LAP and AF	SWP and AF	Sort-SWP and AF	LAP and MF	SWP and MF	Sort-SWP and MF
$\varepsilon = 1$	11.95%	52.25%	65.51%	16.43%	59.73%	70.57%	25.20%	61.76%	75.08%
$\varepsilon = 2$	15.77%	55.21%	75.02%	26.21%	66.58%	77.43%	38.37%	66.93%	77.79%
$\varepsilon = 3$	19.77%	61.27%	77.87%	34.27%	72.39%	82.14%	50.46%	78.75%	84.11%
$\varepsilon = 4$	23.40%	63.30%	76.85%	40.25%	74.29%	82.34%	55.85%	81.10%	87.32%
$\varepsilon = 5$	27.46%	68.82%	81.45%	46.47%	78.19%	84.31%	63.05%	84.76%	87.48%

Table 4: YALE and precision.

Precision	LAP	SWP	Sort-SWP	LAP and AF	SWP and AF	Sort-SWP and AF	LAP and MF	SWP and MF	Sort-SWP and MF
$\varepsilon = 1$	4.12%	27.33%	44.52%	7.89%	41.25%	53.91%	10.18%	44.74%	56.10%
$\varepsilon = 2$	6.77%	31.56%	57.38%	12.51%	48.98%	60.47%	18.23%	53.82%	64.58%
$\varepsilon = 3$	9.85%	39.41%	64.57%	18.01%	57.24%	69.13%	29.76%	59.71%	71.35%
$\varepsilon = 4$	11.26%	46.55%	72.31%	23.80%	65.50%	75.25%	37.61%	66.07%	78.69%
$\varepsilon = 5$	13.01%	59.32%	77.01%	27.12%	72.05%	83.27%	46.07%	75.15%	86.38%

the same conditions, less noise represents higher availability. Of course, as a special information carrier, image noise is only one of the evaluation indicators.

If the pixel value of the gray-scale image is regarded as a value, the value after adding Laplace noise will change with the change of ε. The smaller the ε is, the greater the noise is, the higher the degree of privacy protection is, and there is no limit to the value. However, in the experiment, the range of the pixel value of the gray image is limited between 0 and 255, so even if the obtained value is beyond this range under the influence of noise, it will be normalized. That means all values greater than 255 will be changed to 255, and all values less than 0 will be changed to 0. Therefore, the error of the image after privacy protection is divided into absolute error and relative error. For example, in an image x, there is a pixel $x_{mn} = 200$. If $x'_{mn} = 280$ after adding noise, the absolute error value is 80, and the relative error value is 55.

In Figure 6, the absolute error and relative error value of the experimental result graph under the influence of different algorithms are given. The conclusion is that the error caused by the three algorithms is in line with the previous expectations, and compared with the LAP algorithm and SWP algorithm, the sort-SWP algorithm will cause the smallest error. Also, the median filter is slightly better than the average filter for Laplace noise.

4.3. Result Analysis. To verify the feasibility of the algorithm, the ORL face database, and YALE face database are used as

Table 5: YALE and recall.

Recall	LAP	SWP	Sort-SWP	LAP and AF	SWP and AF	Sort-SWP and AF	LAP and MF	SWP and MF	Sort-SWP and MF
$\varepsilon = 1$	77.34%	82.51%	84.18%	79.58%	84.83%	83.05%	73.62%	82.49%	85.26%
$\varepsilon = 2$	78.12%	78.74%	82.47%	77.41%	80.24%	84.18%	76.14%	80.14%	86.14%
$\varepsilon = 3$	80.34%	83.56%	87.51%	78.87%	81.07%	86.91%	79.35%	81.61%	85.36%
$\varepsilon = 4$	79.24%	81.54%	80.63%	76.15%	77.69%	85.43%	80.45%	78.34%	84.81%
$\varepsilon = 5$	75.39%	82.37%	85.66%	78.56%	82.48%	88.75%	79.51%	81.95%	87.57%

Table 6: YALE and F1-score.

F1-sroce	LAP	SWP	Sort-SWP	LAP and AF	SWP and AF	Sort-SWP and AF	LAP and MF	SWP and MF	Sort-SWP and MF
$\varepsilon = 1$	7.82%	41.06%	58.24%	14.36%	55.51%	65.38%	17.89%	58.01%	67.67%
$\varepsilon = 2$	12.46%	45.06%	67.67%	21.54%	60.83%	70.38%	29.42%	64.39%	73.82%
$\varepsilon = 3$	17.55%	53.56%	74.31%	29.32%	67.10%	77.01%	43.29%	68.96%	77.73%
$\varepsilon = 4$	19.72%	59.27%	76.24%	36.27%	71.08%	80.02%	51.26%	71.68%	81.64%
$\varepsilon = 5$	22.19%	68.97%	81.11%	40.32%	76.91%	85.92%	58.34%	78.40%	86.97%

experimental data sets, and the experimental environment is Intel® Corei9-9900K CPU@ 3.60 GHz,32 G memories, GTX 21080TI GPU, Windows 10 operating system. A face recognition method based on an improved AlexNet convolution neural network is adopted in the experiment. Compared with the AlexNet model, this method has a simpler network structure and fewer parameters, which can save a lot of model training time and then make a fast prediction. When using the ORL face database for the experiment, 5 images of each person are used as a training set, and other images are used as a test set. The sliding window length is 56, $k = 184$. When using the YALE face database for experiments, random 5 images of each person are also used as the training set and the rest as the test set. The sliding window length is 160, $k = 486$. Privacy budget ε is 1, 2, 3, 4, and 5. The test items are accuracy rate, recall rate, and F1 score. The experimental results are shown in Tables 1–6.

5. Conclusions

To solve the privacy protection problem of face image publishing, this paper proposes to use the Laplace mechanism of differential privacy to add noise in the image, so that this should protect the sensitive information in the face image. Compared with directly adding Laplace noise to the image, the SWP algorithm and sort SWP algorithm can effectively reduce the impact of noise on the protected image. Especially sort SWP algorithm, which uses the SAS algorithm to sort and then execute the SWP algorithm, is proposed to further improve the usability of the protected face image. From the results of several test experiments, compared with the lap algorithm (ORL, Yale), the absolute error of SWP algorithm decreased by 40.81%, relative error decreased by 38.26%, accuracy increased by 35.01%, recall rate increased by 3.53%, and F1-sroce increased by 39.07%; absolute error of sort-SWP algorithm decreased by 64.12%, relative error decreased by 51.70%, accuracy increased by 56.63%, recall rate increased by 6.85%, and F1-sroce increased by 55.62%. Besides, in the test of SAS algorithm, compared with the exponential mechanism ranking method (the accuracy of simulated annealing algorithm is only related to the number of iterations), the average ranking accuracy of SAS algorithm is improved by 17.63%.

The differential privacy protection method proposed in this paper is implemented by global noise adding. However, most of the sensitive information contained in a face image exists in some specific areas (such as eyes, mouth, nose, and other organs or feature points). The next research direction will be devoted to accurately finding the region of sensitive information and adding noise for those local regions. At the same time, the research will try to further reduce the impact of noise and improve the usability of face image after privacy protection.

Data Availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

Conflicts of Interest

The authors declare that they have no competing interests.

Acknowledgments

The authors thank the project of the National Natural Science Foundation of China (Nos. 61672179, 61370083, and 61402126), Natural Science Foundation of Heilongjiang Province (No. F2015030), and Fundamental Research Funds in Heilongjiang Provincial Education Department (Nos. 135109247, 135109243, 12541872, and 135209239).

References

[1] K. Wang, B. C. M. Fung, and P. S. Yu, "Handicapping attackers confidence; an alternative to *k*-anonymization," *Knowledge and Information Systems*, vol. 11, no. 3, pp. 345–368, 2007.

[2] B. C. M. Fung, K. Wang, and P. S. Yu, "Anonymizing classification data for privacy preservation," *IEEE Trans on Knowledge and Data Engineering*, vol. 19, no. 5, pp. 711–725, 2007.

[3] X. Xiao and T. Y. Anatomy, "Simple and effective privacy preservation," in *Proc of the 32nd IntConf on Very Large Data Bases*, pp. 139–150, New York, 2006.

[4] T. Li, N. Li, J. Zhang, and I. Molloy, "Slicing: a new approach for privacy preserving data publishing," *IEEE Trans on Knowledge and Data Engineering*, vol. 24, no. 3, pp. 561–574, 2012.

[5] M. Terrovitis, N. Mamoulis, J. Liagouris, and S. Skiadopoulos, "Privacy preservation by diSASsociation," *Proceedings of the VLDB Endowment*, vol. 5, no. 10, pp. 944–955, 2012.

[6] L. Sweeney, "k-Anonymity: a model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2012.

[7] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam, "l-Diversity: privacy beyond k-anonymity," in *22nd International Conference on Data Engineering (ICDE'06)*, pp. 1–12, Atlanta, GA, USA, 2006.

[8] N. Li, T. Li, and S. Venkatasubramanian, "t-Closeness: privacy beyond k-anonymity and l-diversity," in *2007 IEEE 23rd International Conference on Data Engineering*, pp. 106–115, Istanbul, Turkey, April 2007.

[9] X. K. Xiao and Y. F. Tao, "Towards privacy-preserving republication of dynamic datasets," in *IEEE 51th Annual Symposium on Foundations of Computer Science*, pp. 619–642, Las Vegas, NV,USA, 2010.

[10] J. Xu, W. Wang, J. Pei, X. Wang, B. Shi, and A. W.-C. Fu, "Utility-based anonymization using local recoding," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '06*, pp. 785–790, Philadelphia, PA, USA, 2006.

[11] J. Zhang, G. Cormode, C. M. Procopiuc, D. Srivastava, and X. Xiao, "Private release of graph statistics using ladder functions," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data - SIGMOD '15*, pp. 731–745, Melbourne, Victoria, Australia, 2015.

[12] Y. Chen, A. Machanavajjhala, M. Hay, and G. Miklau, "PeGaSus: data-adaptive differentially private stream processing," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pp. 1375–1388, Dallas, TX, USA, October 2017.

[13] O. Goldreich, S. Micali, and A. Wigderson, "How to play any mental game," in *Proceedings of the nineteenth annual ACM conference on Theory of computing - STOC '87*, pp. 218–229, New York, USA, 1987.

[14] C. Gentry, "Fullyhomomorphic encryption using ideal lattices," *Stochastics*, vol. 9, p. 169, 2009.

[15] C. Liu, J. Yang, and J. Wu, "Web intrusion detection system combined with feature analysis and SVM optimization," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, Article ID 33, 2020.

[16] S. Ramezanian, T. Meskanen, M. Naderpour, V. Junnila, and V. Niemi, "Private membership test protocol with low communication complexity," *Digital Communications and Networks*, vol. 6, no. 3, pp. 321–332, 2020.

[17] V. Shukla, A. Chaturvedi, and N. Srivastava, "A secure stop and wait communication protocol for disturbed networks," *Wireless Personal Communications*, vol. 110, no. 2, pp. 861–872, 2020.

[18] C. Thammarat and W. Kurutach, "A lightweight and secure NFC-base mobile payment protocol ensuring fair exchange based on a hybrid encryption algorithm with formal verification," *International Journal of Communication Systems*, vol. 32, no. 12, article e3991, 2019.

[19] M. Anbarasan, S. Prakash, A. Antonidoss, and M. Anand, "Improved encryption protocol for secure communication in trusted MANETs against denial of service attacks," *Multimedia Tools and Applications*, vol. 79, no. 13-14, pp. 8929–8949, 2020.

[20] B. Liu, B. Zhan, C. Zhang, and L. Yang, "Research on visual control system of inverted pendulum based on pixel displacement," *Journal of Physics: Conference Series*, vol. 1550, article 062006, 2020.

[21] PixArt Imaging, "Patent issued for high accuracy displacement detection system with offset pixel array (USPTO 10, 609, 314)," pp. 8821–8835, 2020.

[22] A. Grigoriev, E. Danilova, V. Trusov, M. Miheev, and M. Uhanova, "Modelling of measurement error for vibrational displacement based on the blurring analysis of a round mark image," *Applied Computer Systems*, vol. 23, no. 1, pp. 69–74, 2018.

[23] G. Ye and X. Huang, "An image encryption algorithm based on autoblocking and electrocardiography," *IEEE Multimedia*, vol. 23, no. 2, pp. 64–71, 2016.

[24] L. Yuan, P. Korshunov, and T. Ebrahimi, "Privacy-preserving photo sharing based on a secure JPEG," in *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 185–190, Hong Kong, China, April 2015.

[25] Z. Moghaddasi, H. A. Jalab, and R. M. Noor, "Image splicing forgery detection based on low-dimensional singular value decomposition of discrete cosine transform coefficients," *Neural Computing and Applications*, vol. 31, no. 11, pp. 7867–7877, 2019.

[26] L. Yuan and T. Ebahimi, "Image privacy protection with secure JPEG transmorphing," *IEEE Transactions on Signal Processing*, vol. 11, no. 9, pp. 1031–1038, 2017.

[27] M. Sundararajan, M. Veerappan, and S. Anbazhagan, "Partial image encryption based on using discrete cosine transform coefficients and lightweight stream algorithm," *Journal of Computational and Theoretical Nanoscience*, vol. 16, no. 4, pp. 1573–1576, 2019.

[28] L. He, Y. Wang, and Z. Xiang, "Wavelet frame-based image restoration using sparsity, nonlocal, and support prior of frame coefficients," *The Visual Computer*, vol. 35, no. 2, pp. 151–174, 2019.

[29] C. Dwork, "Differential privacy," in *Proceedings of the 33rd International Colloquium on Automata Languages and Programming*, pp. 1–12, Berlin, 2006.

[30] C. Dwork, "Differential privacy: a survey of results," in *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation*, pp. 1–19, Xi'an, China, 2008.

[31] C. Dwork, "The differential privacy frontier (extended abstract)," in *Proceedings of the 6th Theory of Cryptography Conference*, pp. 496–502, San Francisco.CA.UAS., 2009.

[32] C. Dwork and J. Lei, "Differential privacy and robust statistics," in *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pp. 371–380, Bethesda, MD, USA, 2009.

[33] C. Dwork, "Differential privacy in new settings," in *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, Austin, Texas, USA, January 2010.

[34] C. Dwork, "The promise of differential privacy: a tutorial on algorithmic techniques," in *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, Palm Springs, CA, USA, October 2011.

[35] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proceedings of the 3rd Theory of Cryptography Conference*, pp. 265–284, New York, NY, USA, 2006.

[36] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pp. 94–103, Providence, Rhode Island, USA, October 2007.

[37] F. McSherry, "Privacy integrated queries," *Communications of the ACM*, vol. 53, no. 9, pp. 89–97, 2010.

[38] J. Xu, Z. Zhang, X. Xiao, Y. Yang, and G. Yu, "Differentially private histogram publication," in *2012 IEEE 28th International Conference on Data Engineering*, pp. 32–43, Washington, DC, USA, April 2012.

[39] X. Liu and S. Li, *Histogram Publishing Method Based on Differential Privacy*, 2018.

[40] M. Hay, V. Rastogi, G. Miklau, and D. Suciu, "Boosting the accuracy of differentially-private queries through consistency," CoRR, 2009, https://arxiv.org/abs/0904.0942.

[41] C. Piao, Y. Shi, J. Yan, C. Zhang, and L. Liu, "Privacy-preserving governmental data publishing: a fog-computing-based differential privacy approach," *Future Generation Computer Systems*, vol. 90, pp. 158–174, 2019.

[42] H. Li, J. Cui, X. Meng, and J. Ma, "IHP: improving the utility in differential private histogram publication," *Distributed and Parallel Databases*, vol. 37, no. 4, pp. 721–750, 2019.

[43] X. Xiao, G. Wang, and J. Gehrke, "Differential privacy via wavelet transforms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 8, pp. 1200–1214, 2011.

[44] M. Hay, C. Li, G. Miklau, and D. Jensen, "Accurate estimation of the degree distribution of private networks," in *2009 Ninth IEEE International Conference on Data Mining*, pp. 169–178, Miami, FL, USA, December 2009.

[45] J. Xia, W. Huang, Z. Ma, X. Dai, and L. He, "Gradient-based differential privacy optimizer for deep learning model using collaborative training mode," in *2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT)*, Dalian, China, October 2019.

[46] J. Liu, "Security and privacy problems and countermeasures of internet of things applications," in *Proceedings of the 2017 5th International Conference on Mechatronics, Materials, Chemistry and Computer Engineering (ICMMCCE 2017)*, Edmonton, AB, Canada, 2017.

[47] C. Task and C. Clifton, "A guide to differential privacy theory in social network analysis," in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 411–417, Istanbul, Turkey, August 2012.

[48] R. Chen, B. C. M. Fung, P. S. Yu, and B. C. Desai, "Correlated network data publication via differential privacy," *The VLDB Journal*, vol. 23, no. 4, pp. 653–676, 2014.

[49] W. Jinqiu, Q. Gang, and K. Pengbin, "Emerging 5G multicarrier chaotic sequence spread spectrum technology for underwater acoustic communication," *Complexity*, vol. 2018, Article ID 3790529, 7 pages, 2018.

[50] C.-H. Chen, F. Song, F.-J. Hwang, and L. Wu, "A probability density function generator based on neural networks," *Physica A: Statistical Mechanics and its Applications*, vol. 541, article 123344, 2020.

[51] J. Zhang, Z. Zhang, X. Xiao et al., "Functional mechanism: regression analysis under differential privacy," in *Proceedings of the 38th Conference of Very Large Database*, pp. 1364–1375, Istanbul, Turkey, 2012.

[52] J. Cao, Q. Xiao, G. Ghinita, N. Li, E. Bertino, and K. L. Tan, "Efficient and accurate strategies for differentially-private sliding window queries," in *Proceedings of the 16th International Conference on Extending Database Technology*, pp. 191–202, New York,USA, 2013.

[53] W. Qardaji, W. Yang, and N. Li, "Differentially private grids for geospatial data," in *2013 IEEE 29th International Conference on Data Engineering (ICDE)*, pp. 32-33, Brisbane, Australia, April 2013.

[54] X. J. Zhang, X. J. Zhang, C. C. Fu, and X. F. Meng, "Facial image publication with differential privacy," *Journal of Image and Graphics*, vol. 23, no. 9, pp. 1305–1315, 2018.

[55] T. Dalenius, "Towards a methodology for statistical disclosure control," *StatistikTidskrift*, vol. 15, article 429–222, 1977.

[56] G. Kellaris and S. Papadopoulos, "Practical differential privacy via grouping and smoothing," *Proceedings of the Very Large Database Endowment*, vol. 6, no. 5, pp. 301–312, 2013.

[57] X. Zhang, R. Chen, J. Xu, X. Meng, and Y. Xie, "Towards accurate histogram publication under differential privacy," in *Proceedings of the 2014 SIAM International Conference on Data Mining*, pp. 587–595, Philadelphia, PA, April 2014.

[58] X. Li, J. Yang, Z. Sun, and J. Zhang, "Differentially private release of the distribution of clustering coefficients across communities," *Security and Communication Networks*, vol. 2019, Article ID 2518714, 9 pages, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8841210, 12 pages
https://doi.org/10.1155/2020/8841210

Research Article

Analysis and Optimization of Asymmetric Wireless Power Transfer in Concrete

Zijian Tian,[1] Dairong Liu,[1] Wei Chen,[1,2] Wenqing Wang,[3] Shan Pan,[1] and Ni Guo[1]

[1]*School of Mechanical Electronic & Information Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China*
[2]*School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, Jiangsu 221000, China*
[3]*Beijing Polytechnic College, Beijing 100042, China*

Correspondence should be addressed to Wei Chen; chenwdavior@163.com

Received 31 August 2020; Revised 2 November 2020; Accepted 25 November 2020; Published 10 December 2020

Academic Editor: Chi-Hua Chen

Copyright © 2020 Zijian Tian 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 the Internet of things (IoT) technology, the application of IoT has been expanded greatly, and the disadvantages of the traditional battery power supply have become increasingly prominent. The power supply mode limits the development of the concrete structural health monitoring network. And the application of magnetic resonance coupled wireless power transfer technology can solve the problem of power supply to sensors embedded in concrete. The corrected transmission efficiency considering the concrete conductivity is proposed which establishes the relationship between the electromagnetic field and the circuit model. And the field-circuit coupled model of asymmetric wireless power transfer system in concrete is developed. The effects of radial offset and axial dislocation on the transmission efficiency at different concrete conductivity are further analyzed. The relationship between the resonant frequency and the transmission efficiency in different concrete conductivity is analyzed, and an optimization scheme is proposed to improve the transmission efficiency. Finally, the experimental setups are established, and the theoretical analysis is verified. The conclusions cannot only break through the bottleneck of the scale of the concrete structural health monitoring network but also further releases the application potential of IoT.

1. Introduction

In recent years, Internet of things (IoT) technology has been applied in many fields of our life. With the development of application, different IoT specifications and semantic mobile computing technologies have emerged [1–3]. The stable power supply of each part is the basis for realizing the service interoperability between different platforms. Lithium batteries are widely used in the field of wearable devices and industrial IoT, and the mainstream development trend is to reduce the loss current by simulating the design technology and power system technology to extend the system operation time and standby time. But the current power solutions cannot fundamentally solve the problem of system interruption caused by battery charging and replacement. In the field of Internet of vehicles, the power system needs to adopt a combined integrated power management solution with the increasing number of onboard intelligent modules to achieve stronger system power performance. But the transmission speed of information still poses a great technical challenge to the power solution.

The concrete structural health monitoring network is capable to realize the damage identification and the safety judgment for buildings, bridges, and tunnels through kinds of sensors embedded in concrete that collects real-time data and upload the data to the information processing center. Continuous and stable operation of system can be realized by introducing contactless power transfer technology, which can avoid the corrosion of concrete to damage power supply cable of sensors. And the cost of battery replacement can also be reduced for the battery-powered sensors embedded in concrete. This technology can be extended to the whole field of the IoT which can increase the coverage of IoT and reduce the production cost of products. It can also provide a new

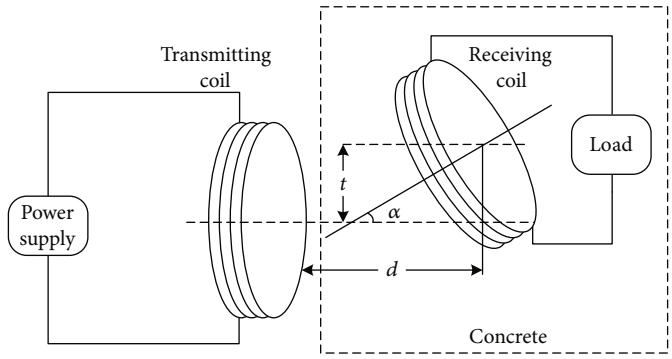


Figure 1: Asymmetric wireless power transfer system in concrete.

power solution for the IoT ecosystem. The landmark of contactless power transfer from concept to application was the proposal of magnetic resonance-coupled wireless power transfer which achieved efficient midrange power transfer [4]. Since then, scientists began to explore and research this emerging field and had considerable achievements [5–8]. The application of technology can further promote the development of the concrete structural health monitoring network.

A high-frequency rectangular patch antenna was proposed to demonstrate the feasibility of wireless power supply to sensors embedded in concrete in [9]. The problem of wireless power supply to sensors embedded in concrete was solved by the application of wireless power transfer technology in [10]. And the experimental setup was built to verify the transmission performance. Steel bars in concrete were proposed as the medium of low-frequency magnetic induction to realize long-distance wireless power transfer in [11]. The wireless power transfer system that the resonant frequency was fixed at 60 Hz was built in [12]. And the efficiency formula including copper core losses in concrete was derived, and the effect of the shape of magnet pole pieces on the transmission performance was analyzed. The effect of reinforced concrete widely used in metro construction on the transmission performance was analyzed in [13]. And the unnecessary loss including dielectric loss, the hysteresis loss, and the eddy current loss was derived through analytical formula and finite element analysis. The result showed the eddy current loss was the main reason of performance degradation. A novel multicoil resonator was designed which was capable to solve the problem of low transmission efficiency in [14]. And the interference of reinforcement array was canceled through adding the four middle coils to a certain extent. The transmission efficiency was compared between the system in dry concrete and free space in [15]. And the output power was measured in different concrete medium when the input power was constant. The models of the low power near-field system in dry and wet concrete were established in [16]. And the DC current and output power gradually increased with the drying of wet concrete. The wireless power transfer system in concrete was developed in [17]. And the concrete material mixed in a certain proportion was verified to improve the transmission performance. The effects of the height of transmitting coil, the depth of receiving coil, and the concrete humidity on the transmission performance were analyzed by the extended Debye model in [18]. The radial offset and axial dislocation between the coils are common asymmetrical conditions, which are important factors that cause the change of transmission performance. The formulas of self-inductance and mutual inductance of circular coil in near field wireless power transfer were given in [19]. And the formulas of the mutual inductance and the transmission efficiency in the case of misalignment also were derived, and the effect of radial offset and angular misalignment was studied. A mathematical model of wireless power transfer system with series-parallel compensation network was developed, which was capable to determine the circuit parameters under various misalignment conditions in [20]. And the mathematical model was used to realize the preliminary design of the system. The mutual inductance formulas in the case of radial offset, axial dislocation, and general misalignment between the coils were derived, and the relationship between the optimal transmission performance and the turns of the coils was defined in [21]. The change of transmission performance when the misalignment occurs was analyzed, and an optimization scheme was proposed to improve the transmission efficiency by changing the relative position between the coils in [22]. A novel coil structure was designed to realize stable transmission performance with the varying of angular misalignment in [23, 24]. A multivariable dynamic tuning method was proposed in [25]. And the maximum output power point was tracked through monitoring the size and phase of the inductance current.

In Section II, the eddy current equivalent resistance and corrected mutual inductance voltage is presented, and the corrected transmission efficiency formula is analyzed. In Section III, the asymmetric wireless power transfer in different concrete conductivity is analyzed, and an optimization scheme to improve the transmission efficiency is proposed. In Section IV, the experimental setups are established, and the theoretical analysis is verified. In Section V, the results are listed. In Section VI, the paper is summarized and the limitations are elaborated.

2. Theoretical Analysis

The schematic diagram of asymmetric wireless power system in concrete is shown in Figure 1. Here, t is the radial offset

distance, and α is the dislocation angle. And the receiving coil is sealed in concrete. The concrete medium has different conductivity due to the different humidity environment and manufacturing process. The conductivity of concrete medium is generally 0.4-5 S/m. There is sinusoidal alternating current in the transmitting loop and the receiving loop.

The sinusoidal alternating electromagnetic field generated by the sinusoidal alternation current is periodic field with the variation of time. And the complex Maxwell equations can omit the process of solving the time, reduce the complexity of solution, and realize the transformation from the time domain to frequency domain.

The complex Maxwell equations are as follows:

$$\begin{cases} \nabla \times H = \sigma E + j\omega\varepsilon E \\ \nabla \times E = -j\omega\mu H \\ \nabla \cdot \mu H = 0 \\ \nabla \cdot \varepsilon E = \rho, \end{cases} \tag{1}$$

where E is the electric field strength, H is the magnetic field strength, σ is the conductivity, μ is the permeability, and ε is the dielectric constant of the medium. According to (1) and (2), the constraint equation of electric field can be derived as follows:

$$\nabla \times (\nabla \times E) = \nabla(\nabla \cdot E) - \nabla^2 E, \tag{2}$$

$$\begin{cases} \nabla^2 E + \dfrac{1}{\sigma + j\omega\varepsilon} \nabla(\nabla \cdot \sigma E) = j\omega\mu(2\sigma + j\omega\varepsilon)E \\ \nabla \cdot E + \dfrac{1}{\sigma + j\omega\varepsilon} \nabla \cdot \sigma E = 0. \end{cases} \tag{3}$$

The electromagnetic field model of the coil is built in the cylindrical coordinate system, and the coordinates of any point in the field are (ρ, φ, z). The radius of the coil is r, and the effective value of the sinusoidal alternating current is I. The distance from the xOy plane is d, and the axial of the coil coincides with z-axis. The field is divided into two parts by the plane of the coil for ensuring that no external source in the solution area. And the boundary value problem of electric field is analyzed. Since the electric field is the eddy electric field parallel to the plane of the coil, the boundary value problem of electric field can be simplified as the boundary value problem of circumferential component of electric field.

According to (3), the constraint equation of electric field is simplified as follows:

$$\nabla^2 E_\phi - \frac{j\omega\mu(\sigma + j\omega\varepsilon) + 1}{\rho^2} E_\phi = 0. \tag{4}$$

Boundary conditions for the plane of the coil are expressed as follows:

$$\begin{aligned} E_{2\phi} - E_{1\phi} &= 0, \\ \frac{\partial E_{2\phi}}{\partial z} - \frac{\partial E_{1\phi}}{\partial z} &= j\omega\mu I \delta(\rho - r). \end{aligned} \tag{5}$$

Boundary condition for the infinite far place is expressed as follows:

$$\lim_{x \to \infty} eE_\phi = c_\phi, \tag{6}$$

where e is the distance from the coordinate origin to the point (ρ, φ, z) and c is the bounded constant vector independent of the coordinate system. By solving the ordinary differential equation, the electric field strength E is obtained from the above boundary conditions as follows:

$$E = -\frac{j\omega\mu r I}{2} \int_0^\infty \frac{\lambda J_1(\lambda r) J_1(\lambda \rho)}{\sqrt{\lambda^2 + j\omega\mu(\sigma + j\omega\varepsilon)}} e^{(z-d)\sqrt{\lambda^2 + j\omega\mu(\sigma + j\omega\varepsilon)}} d\lambda e_\varphi. \tag{7}$$

And the eddy current loss generated by the eddy electric field is as follows:

$$P_{Loss} = \int_v \sigma E^2 dv, \tag{8}$$

Where v is the solution area of eddy current loss between the coils.

According to (7) and (8), the eddy current loss is directly proportional to the square of the effective value of current. Therefore, the eddy current equivalent resistance is defined as the eddy current loss divided by the square of the effective value of current as follows:

$$R_{\text{Loss}} = \frac{P_{\text{Loss}}}{I^2}, \tag{9}$$

where R_{Loss} is a function of conductivity. The relationship between the electromagnetic field and the circuit model is established by the eddy current equivalent resistance which can simplify the analysis parameters and process. And it is easier to analyze the effect of concrete conductivity on the transmission efficiency.

The asymmetric coil model is shown in Figure 2. The radial offset and axial dislocation affect the mutual inductance between the coils and further reduce the transmission efficiency. In [22], the mutual inductance formula for radial offset and axial dislocation was derived as follows:

$$M = \frac{\mu_0 n_1 n_2}{4\pi} \int_0^{2\pi} \int_0^{2\pi} \frac{r_1 r_2 (\sin\theta \sin\varphi \cos\alpha + \cos\theta \cos\varphi) d\theta d\varphi}{R_{QN}}, \tag{10}$$

where the cylindrical coordinate system is adopted and the coordinates of the center of the receiving coil are $(0, t, d$

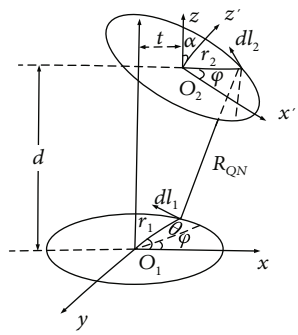


Figure 2: Asymmetric coil model.

). The mutual inductance voltage phase is 90° behind the excitation current phase in the air, but the mutual inductance voltage phase is no longer 90° behind the excitation current phase in concrete, and the effective value also has changes. The mutual inductance voltage formula in concrete is corrected, and the formula is as follows:

$$U^{12}_{\text{concrete}} = -f(\sigma)j\omega M I_{\text{concrete}} e^{-j\beta}, \tag{11}$$

where $f(\sigma)$ is a function of conductivity which is defined as the correction coefficient of the effective value and β is the difference of electric field strength phase between air and concrete. There are two differences between the circuit model and the corrected model: the eddy current equivalent resistance and corrected mutual inductance voltage as shown in Figure 3. U_S is the high-frequency voltage source. R_1 and R_2 are the equivalent resistance of the transmitting loop and the receiving loop. L_1 and L_2 are the self-inductance of the transmitting coil and the receiving coil. C_1 and C_2 are the equivalent compensation capacitance. R_{Loss1} and R_{Loss2} are the eddy current equivalent resistance. M is the mutual inductance between the coils. R_L is the equivalent load resistance.

According to Kirchhoff's law, the voltage equations can be expressed as follows:

$$\begin{aligned} U_1 &= I_1 R_1 + I_1 R_{\text{Loss1}} + j\omega L_1 I_1 + \frac{1}{j\omega C_1} I_1 - f(\sigma)j\omega M I_2 e^{-j\beta}, \\ -f(\sigma)j\omega M I_1 e^{-j\beta} &= j\omega L_2 I_2 + I_2 R_{\text{Loss2}} + I_2 R_2 + I_2 R_L + \frac{1}{j\omega C_2} I_2. \end{aligned} \tag{12}$$

In order to facilitate the calculation, assuming $R_1 = R_2 = R$, $C_1 = C_2 = C$, and $L_1 = L_2 = L$. And the current of the transmitting loop and the receiving loop is as follows:

$$\begin{aligned} I_1 &= \frac{(j\omega L + R + R_{\text{Loss2}} + R_L + (1/j\omega C))U_1}{(j\omega L + R + R_{\text{Loss2}} + R_L + (1/j\omega C))(j\omega L + R + R_{\text{Loss1}} + (1/j\omega C)) + f(\sigma)^2\omega^2 M^2 e^{-j2\beta}}, \\ I_2 &= \frac{f(\sigma)j\omega M e^{-j\beta} U_1}{(j\omega L + R + R_{\text{Loss2}} + R_L + (1/j\omega C))(j\omega L + R + R_{\text{Loss1}} + (1/j\omega C)) + f(\sigma)^2\omega^2 M^2 e^{-j2\beta}}. \end{aligned} \tag{13}$$

The transmission efficiency is determined as follows:

$$\eta = \frac{f(\sigma)^2\omega^2 M^2 e^{-j2\beta} R_L}{\left[(R + R_{\text{Loss1}})(R + R_L + R_{\text{Loss2}}) + f(\sigma)^2\omega^2 M^2 e^{-j2\beta}\right](R + R_L + R_{\text{Loss2}})}. \tag{14}$$

The transmission efficiency formula considering the concrete conductivity is derived which shows that the concrete conductivity is an important factor that causes the change of transmission performance. It is mainly reflected by the eddy current equivalent resistance and corrected mutual inductance voltage. Also, the radial offset and axial dislocation between the coils are common asymmetrical conditions in the specific engineering. Therefore, it is necessary to study the effect of radial offset and axial dislocation between the coils on the transmission efficiency at different concrete conductivity.

3. Establishment and Analysis of System Model

In this section, Maxwell and Simplorer are used to establish the model of asymmetric wireless power transfer system with the transmitting coil in the air and the receiving coil in the concrete. The three-dimensional system model established in Maxwell is shown in Figure 4. And the design parameters are shown in Table 1.

In order to study the effect of the conductivity of concrete medium, the dry concrete, wet concrete, and reinforced concrete is used as the transmission mediums. According to the normal standard parameters, the conductivity is set to 1 S/m, 3.8 S/m, and 4.9 S/m, respectively. The three-dimensional system model is imported into the external circuit that established in Simplorer. The stop time is 10 μs, the time step is 0.02 μs, and the circuit parameters are shown in Table 2.

The spatial distribution of magnetic flux density at different radial offset distance is shown in Figure 5. In Figure 5, the magnetic field is mainly distributed inside the coils and between the coils, indicating that the transmission of energy is achieved through the coupling of magnetic field. The magnetic flux density of the transmitting coil increases from 1.1232×10^{-4} T to 1.7974×10^{-4} T throughout the process, which shows that the energy of the transmitting coil cannot be effectively transmitted to the receiving coil with the increase of radial offset distance. In addition, the effect of concrete medium on the spatial distribution of magnetic flux

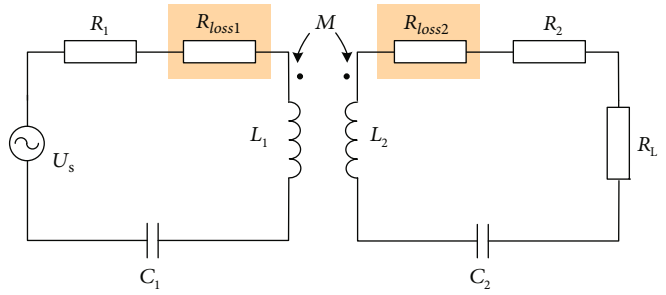


Figure 3: Corrected circuit mutual inductance model in concrete.

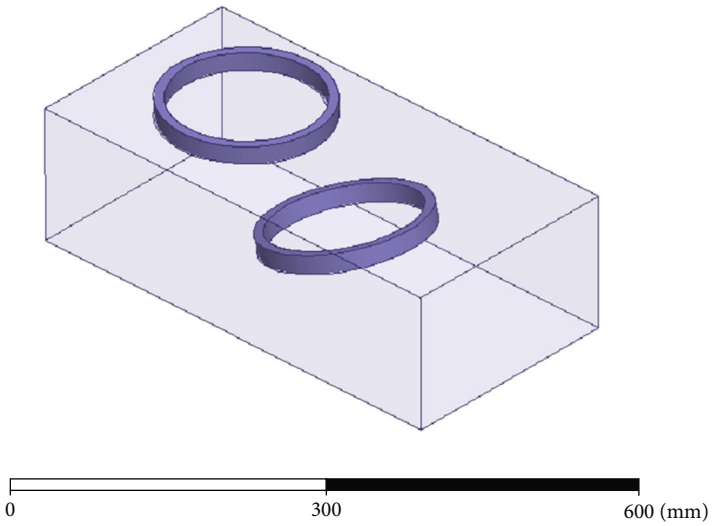


Figure 4: Three-dimensional system model.

Table 1: Design parameters.

	Value
Internal diameter	80 mm
External diameter	90 mm
Wire diameter	2 mm
Turns	5 turns
Axial distance	80 mm
Concrete size ($L \times W \times H$)	$500 \times 250 \times 150$ mm

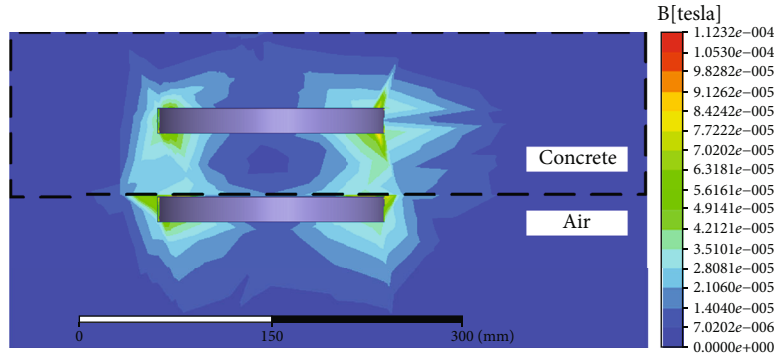
Table 2: Circuit parameters.

	Transmitting loop	Receiving loop
Resonant frequency	6 MHz	
Resistance	0.2 Ω	
Power supply voltage	18 V	—
Load	—	50 Ω
Self-inductance	72.19 μH	70.92 μH
Matching capacitance	9.74 pF	9.92 pF

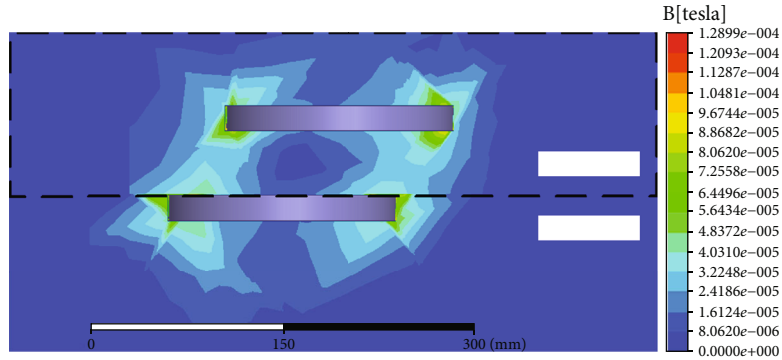
density always exists. The magnetic flux density is obviously weakened at the interface between concrete and air. And the weakening trend is more obvious with the increase of radial offset distance, which shows that the transmission of energy is hindered by concrete medium to a certain extent.

The simulation current can be obtained at different conditions through the Maxwell and Simplorer joint simulation method. The current simulation waveform when the radial offset distance $t = 0$ mm in dry concrete is shown in Figure 6.

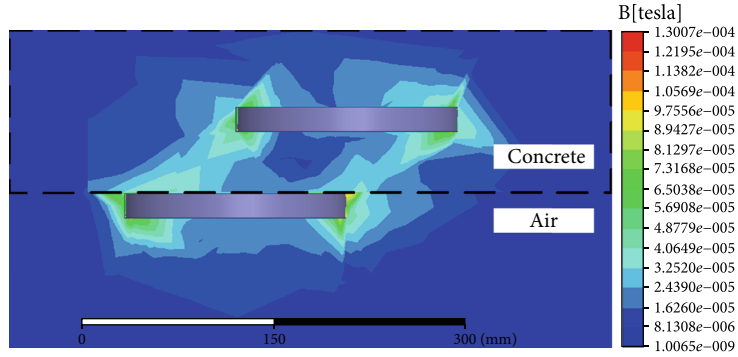
The effect of radial offset on the transmission efficiency in different concrete medium is shown in Figure 7. The transmission efficiency remains almost constant when the radial offset is in a small scope. With the increase of radial offset distance, the transmission efficiency decreases dramatically. The effect of concrete conductivity on the transmission efficiency is also significant. The transmission efficiency has overall increase with the continuous decrease of concrete conductivity. Compared with wet and reinforced concrete, the energy can be transmitted in dry concrete at a longer radial offset distance. And the trend of transmission efficiency is identical in different concrete conductivity with the increase of radial offset distance, which confirms the relationship between the eddy current loss and conductivity in the eddy current loss formula, which further proves that the eddy current loss is an important factor that causes the change of transmission performance.



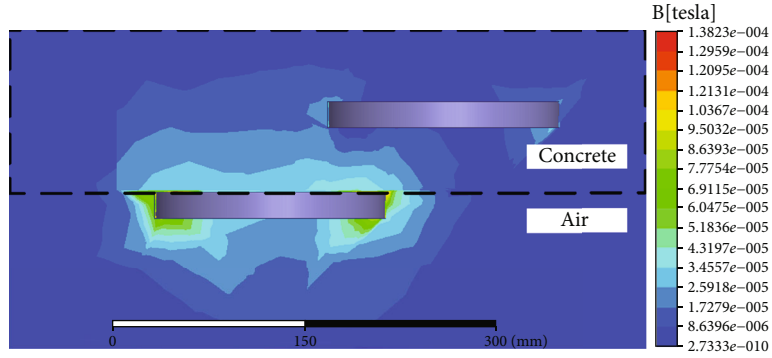
(a) $t = 0$ mm



(b) $t = 45$ mm

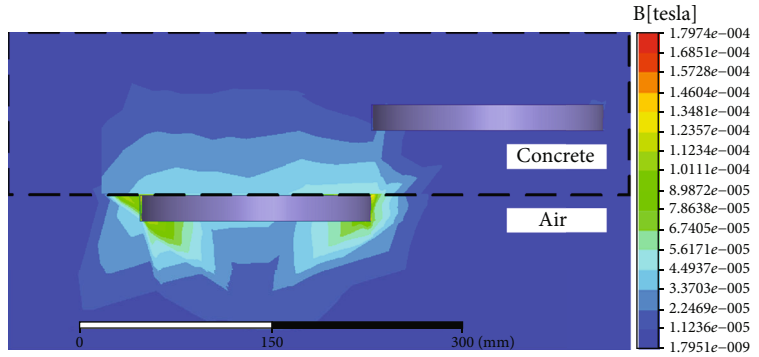


(c) $t = 90$ mm



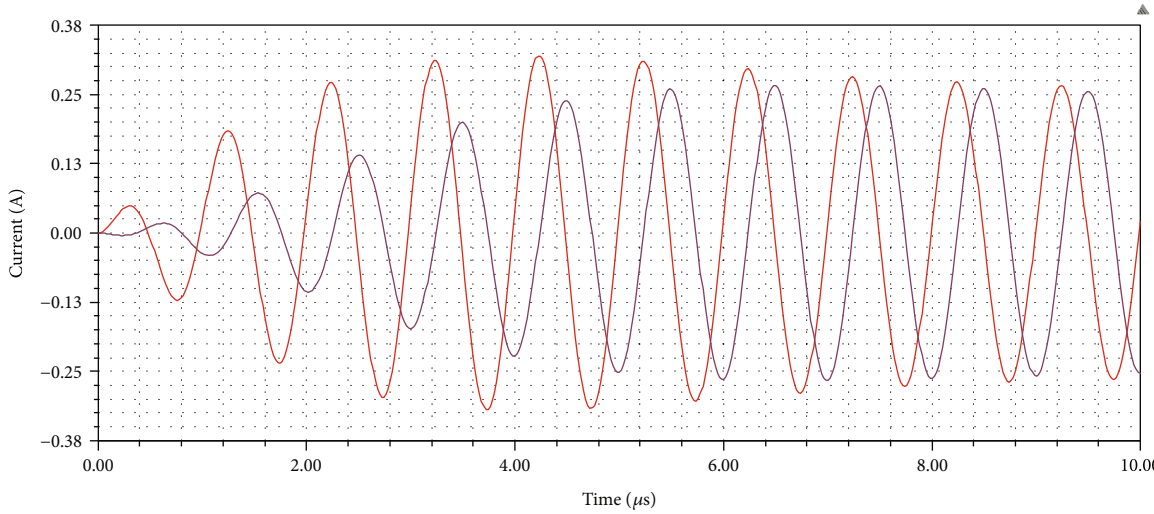
(d) $t = 135$ mm

Figure 5: Continued.



(e) $t = 180$ mm

Figure 5: The spatial distribution of magnetic flux density at different radial offset distance.



Curve Info	rms
Current (Phase_prim) Setup1 : Transient	0.1926
Current (Phase_sec) Setup1 : Transient	0.1829

Figure 6: Current simulation waveform of the two coils in dry concrete when $t = 0$ mm.

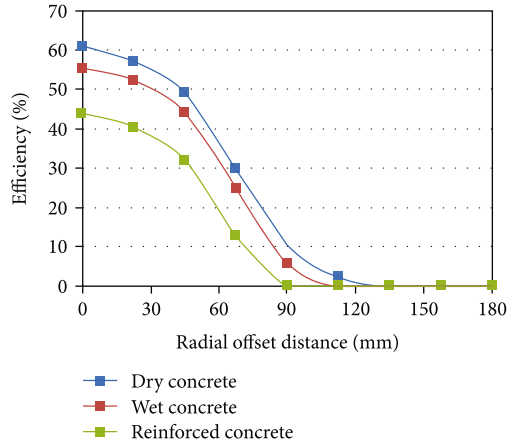


Figure 7: The effect of radial offset on the transmission efficiency in different concrete medium (dry concrete $\sigma = 1$ S/m, wet concrete $\sigma = 3.8$ S/m, and reinforced concrete $\sigma = 4.9$ S/m).

The spatial distribution of magnetic flux density at different dislocation angle is shown in Figure 8. In Figure 8, the magnetic field is mainly distributed inside the coils and between the coils, indicating that the transmission of energy is achieved through the coupling of magnetic field. The magnetic flux density of the whole system remains almost constant with the increase of dislocation angle α from 0° to 45° indicating that the transmission performance does not change much. In addition, the magnetic flux density is also weakened at the interface between concrete and air. Since the magnetic flux density remains almost constant, the weakening trend is not obvious with the increase of dislocation angle.

The effect of axial dislocation on the transmission efficiency in different concrete medium is shown in Figure 9. When the dislocation angle α varies from 0° to 45°, the transmission performance remains steady and the fluctuation of

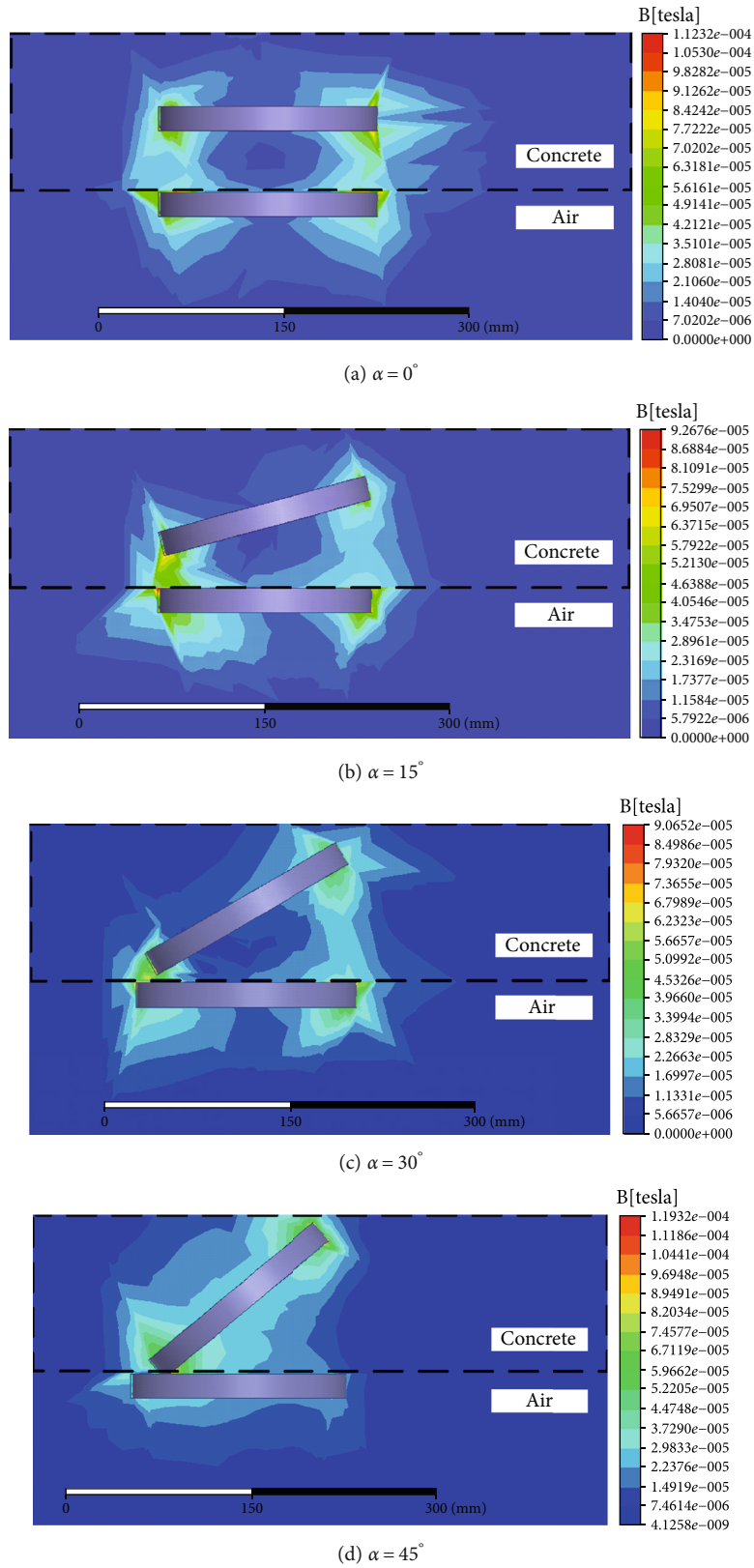


Figure 8: The spatial distribution of magnetic flux density at different dislocation angle.

the transmission efficiency is small. Figure 9 still shows that the eddy current loss is an important factor that causes the change of transmission performance.

Figure 10 shows that the effect of resonant frequency on the transmission efficiency in different concrete medium. The transmission efficiency drops sharply with the increase

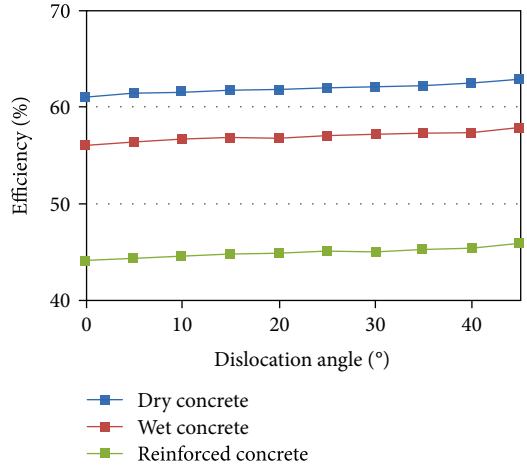


Figure 9: The effect of axial dislocation on the transmission efficiency in different concrete medium (dry concrete $\sigma = 1$ S/m, wet concrete $\sigma = 3.8$ S/m, and reinforced concrete $\sigma = 4.9$ S/m).

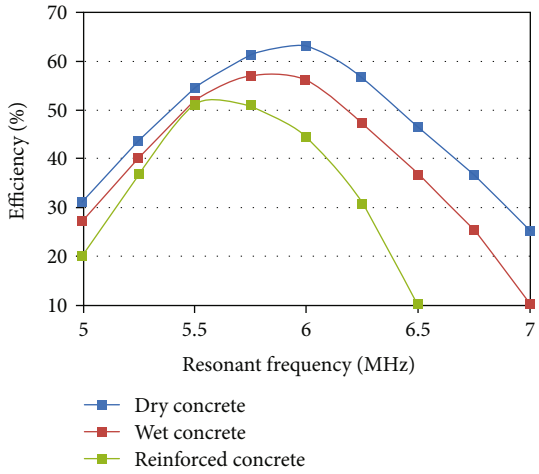


Figure 10: The effect of resonant frequency on the transmission efficiency in different concrete medium (dry concrete $\sigma = 1$ S/m, wet concrete $\sigma = 3.8$ S/m, and reinforced concrete $\sigma = 4.9$ S/m).

of conductivity at the same resonant frequency which further confirms the correctness of the conclusion analyzed from Figures 7 and 9. The optimal resonant frequency in dry concrete is approximately 5.93 MHz. The optimal resonant frequency in wet concrete is approximately 5.82 MHz. And the optimal resonant frequency in reinforced concrete is approximately 5.54 MHz. It shows that the optimal resonant frequency is gradually decreasing with the increase of conductivity. Therefore, the frequency shift is also an important factor that causes the change of transmission performance. According to the change rule of the transmission efficiency and the optimal resonant frequency in different concrete conductivity, an optimization scheme to improve the transmission efficiency is proposed. When the transmission efficiency is lower than the normal transmission standard in engineering due to the change of coil position or the difference of concrete conductivity, the transmission efficiency can be improved by properly reducing the resonant frequency to close the optimal resonant frequency. Figure 11 shows that the transmission efficiency can be improved in reinforced concrete by this optimization scheme. And compared with dry and wet concrete, the transmission efficiency has a great improvement in the reinforced concrete. Therefore, this optimization scheme is more effective for the system with high conductivity.

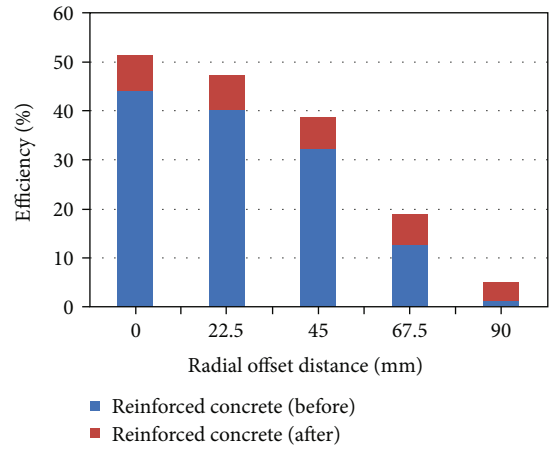


Figure 11: The effect of the optimization scheme in reinforced concrete (reinforced concrete $\sigma = 4.9$ S/m).

4. Experimental Validation

The experimental setups of wireless power transfer system are shown in Figure 12. The digital signal generator, the power amplifier, and the DC power supply constitute a high-frequency power supply which provides energy for magnetic resonance coupled wireless power transfer system. The impedance matching device ensures the optimal transmission performance before the experiments. The transmitting coil and the receiving coil are two identical coils, and the parameters of the coils are given as follows: wire diameter 2 mm, coil radius 90 mm, and 5 turns. To satisfy the experimental requirements, the receiving coil and load is sealed in the concrete. In order to study the effect of the conductivity of concrete medium, the dry concrete, wet concrete, and reinforced concrete are used as the transmission medium. The conductivity of the dry concrete, wet concrete, and reinforced concrete is measured as 1.31 S/m, 3.63 S/m, and 4.97 S/m. The load is a resistance box of the type EMC5307ALN, and the resistance value is 50 Ω.

The identical resonant frequency is a precondition for magnetic resonance coupled wireless power transfer, and the resonant frequency of the system is adjusted to 6 MHz by adding the corresponding compensation capacitance. The axial distance between the coils is always fixed at 80 mm during the experiment. Figure 13 shows that the effect of resonant frequency on the transmission efficiency in different concrete medium. The effects of the conductivity of concrete medium on the transmission efficiency from big to small are reinforced concrete, wet concrete, and dry concrete. And the optimal resonant frequency in different concrete

Figure 12: Experimental setups of wireless power transfer system in concrete.

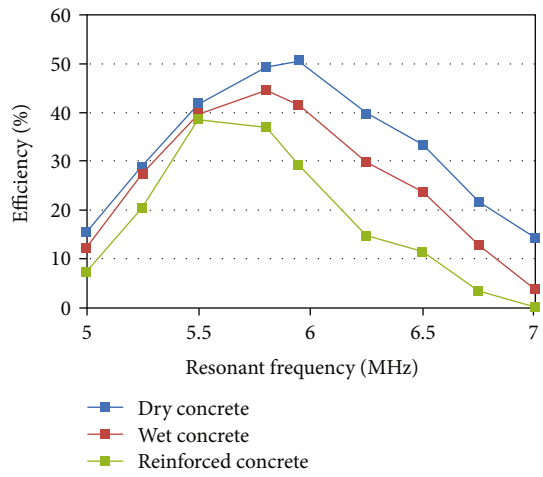


Figure 13: The effect of resonant frequency on the transmission efficiency in different concrete medium.

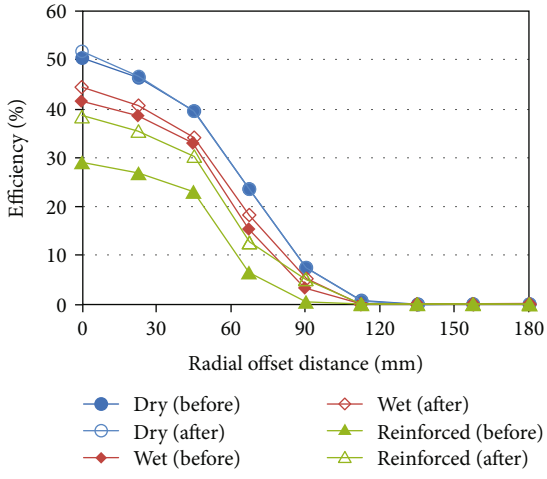


Figure 14: The effect of radial offset on the transmission efficiency in different concrete medium.

conductivity also changes. The optimal resonant frequency of dry concrete, wet concrete, and reinforced concrete is 5.95 MHz, 5.81 MHz, and 5.50 MHz, respectively. Since the power amplifier is used to adjust the resonant frequency directly in the experiment, it will lead to the loss of the transmission efficiency. However, it does not hinder the overall agreement with the simulation result.

In Figure 14, the transmission efficiency remains stable in a small scope and decreases rapidly with the gradual increase of radial offset distance. And the decreasing trend of transmission efficiency remains basically consistent with the increase of radial offset distance. According to the optimization scheme, the resonant frequency is adjusted to the optimal resonant frequency in different concrete conductivity, and the transmission efficiency is improved to a certain extent which has the most obvious effect on wireless power transfer in reinforced concrete. When the dislocation angle α varies from 0° to 45°, the transmission performance remains steady in Figure 15 which still proves the effectiveness of the proposed optimization scheme.

5. Results

The effects of radial offset and axial dislocation on the transmission performance in different concrete conductivity were analyzed. With the increase of radial offset distance, the transmission efficiency decreased dramatically. When the dislocation angle α varied from 0° to 45°, the fluctuation of the transmission efficiency was small. And the effect of concrete conductivity on the transmission efficiency was holistic. The transmission efficiency has overall decrease with the continuous increase of concrete conductivity at different radial offset and axial dislocation.

The relationship between the resonant frequency and the transmission efficiency in different concrete conductivity was analyzed; an optimization scheme was proposed. The optimal resonant frequency was gradually decreasing with the increase of conductivity. The frequency shift was also an important factor that caused the change of transmission performance. Thus, when the transmission efficiency was lower than the normal transmission standard in engineering due to the change of coil position or the difference of concrete conductivity, the transmission efficiency could be improved by properly reducing resonant frequency to close the optimal resonant frequency.

6. Discussion

This paper studied and optimized the asymmetric wireless power transfer system in concrete in the context of the concrete structural health monitoring network. Through the proposal of the eddy current equivalent resistance and corrected mutual inductance voltage, the influencing factors of transmission efficiency were simplified. The effects of radial

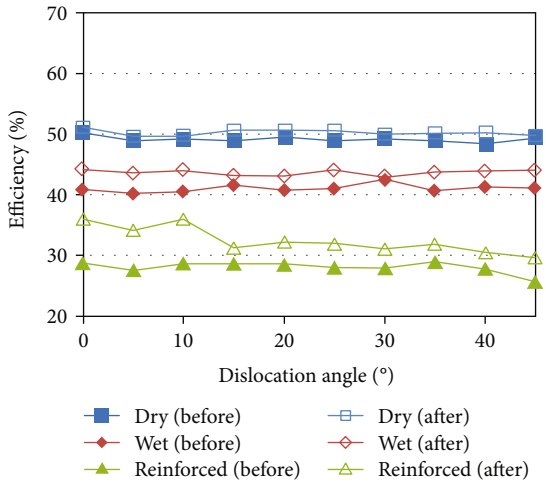


FIGURE 15: The effect of axial dislocation on the transmission efficiency in different concrete medium.

offset and axial dislocation on the transmission efficiency in different concrete conductivity were further analyzed by establishing the field-circuit coupled model. Through the relationship between the resonant frequency and the transmission efficiency in different concrete conductivity, and an optimization scheme was proposed to improve the transmission performance. Finally, the experimental setups were established, and the theoretical analysis was verified, which not only promoted the development of the concrete structural health monitoring network but also enriched the power solutions of IoT. With the continuous development of wireless power transfer technology, its deep integration with the IoT ecosystem will further reduce the production cost of products and improve the coverage of the IoT.

Although the correctness of the optimization scheme in concrete medium has been verified, it is uncertain whether the scheme is effective for the wireless power supply system of IoT under other transmission media. The relationship between the resonant frequency and the transmission efficiency in various transmission media will become the further research direction in the future.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was funded by the National Natural Science Foundation of China, grant number 52074305 and 51874300; National Natural Science Foundation of China and Shanxi Provincial People's Government Jointly Funded Project of China for Coal Base and Low Carbon, grant number U1510115; and the Open Research Fund of Key Laboratory of Wireless Sensor Network & Communication, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, grant numbers 20190902 and 20190913.

References

[1] X. Kong, Y. Xu, Z. Jiao, D. Dong, X. Yuan, and S. Li, "Fault location technology for power system based on information about the power internet of things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6682–6692, 2020.

[2] M. Choksi and M. A. Zaveri, "Multiobjective based resource allocation and scheduling for postdisaster management using IoT," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 6185806, 16 pages, 2019.

[3] Z. Safdar, S. Farid, M. Qadir, K. Asghar, J. Iqbal, and F. K. Hamdani, "A novel architecture for internet of things based E-health systems," *Journal of Medical Imaging and Health Informatics*, vol. 10, no. 10, pp. 2378–2388, 2020.

[4] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljacic, "Wireless power transfer via strongly coupled magnetic resonances," *Science*, vol. 317, no. 5834, pp. 83–86, 2007.

[5] L. Tan, K. E. I. Elnail, M. Ju, and X. Huang, "Comparative analysis and design of the shielding techniques in WPT systems for charging EVs," *Energies*, vol. 12, no. 11, p. 2115, 2019.

[6] V. Cirimele, F. Freschi, L. Giaccone, L. Pichon, and M. Repetto, "Human exposure assessment in dynamic inductive power transfer for automotive applications," *IEEE Transactions on Magnetics*, vol. 53, no. 6, pp. 1–4, 2017.

[7] F. Lu, H. Zhang, C. Zhu et al., "A Tightly Coupled inductive power transfer system for low-voltage and high-current charging of automatic guided vehicles," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 9, pp. 6867–6875, 2019.

[8] D. Kim, D. Jeong, J. Kim et al., "Design and implementation of a wireless charging-based cardiac monitoring system focused on temperature reduction and robust power transfer efficiency," *Energies*, vol. 13, no. 4, p. 1008, 2020.

[9] K. M. Z. Shams and M. Ali, "Wireless power transmission to a buried sensor in concrete," *IEEE Sensors Journal*, vol. 7, no. 12, pp. 1573–1577, 2007.

[10] J.-M. Kim, M. Han, H. J. Lim, S. Yang, and H. Sohn, "Operation of battery-less and wireless sensor using magnetic resonance based wireless power transfer through concrete," *Smart Structure and Systems*, vol. 17, no. 4, pp. 631–646, 2016.

[11] Z. Wang, O. Kypris, and A. Markham, "RePWR: wireless power transfer within reinforced concrete," in *Proceedings of the 4th International Workshop on Energy Harvesting and Energy-Neutral Sensing Systems*, pp. 1–6, California, CA, USA, 2016.

[12] H. Ishida and H. Furukawa, "Wireless power transmission through concrete using circuits resonating at utility frequency of 60 Hz," *IEEE Transactions on Power Electronics*, vol. 30, no. 3, pp. 1220–1229, 2015.

[13] S.-H. Lee, M.-Y. Kim, B.-S. Lee, and J. Lee, "Impact of rebar and concrete on power dissipation of wireless power transfer systems," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 1, pp. 276–287, 2020.

[14] Y. Jang, J.-K. Han, J.-I. Baek, G. Moon, J.-M. Kim, and H. Sohn, "Novel multi-coil resonator design for wireless power transfer through reinforced concrete structure with rebar

array," in *2017 IEEE 3rd International Future Energy Electronics Conference and ECCE Asia (IFEEC 2017 - ECCE Asia)*,, pp. 2238–2243, Kaohsiung, Taiwan, 2017.

[15] X. H. Jin, J. M. Caicedo, and M. Ali, "Near-field wireless power transfer to embedded smart sensor antennas in concrete," *Applied Computational Electromagnetics Society Journal*, vol. 30, no. 3, pp. 261–269, 2015.

[16] R. H. Bhuiyan, R. Islam, J. M. Caicedo, and M. Ali, "A study of 13.5-MHz coupled-loop wireless power transfer under concrete and near metal," *IEEE Sensors Journal*, vol. 18, no. 23, pp. 9848–9856, 2018.

[17] J.-M. Kim, M. Han, and H. Sohn, "Magnetic resonance-based wireless power transmission through concrete structures," *Journal of Electromagnetic Engineering and Science*, vol. 15, no. 2, pp. 104–110, 2015.

[18] O. Jonah and S. V. Georgakopoulos, "Wireless power transfer in concrete via strongly coupled magnetic resonance," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 3, pp. 1378–1384, 2013.

[19] S. R. Khan, S. K. Pavuluri, and M. P. Y. Desmulliez, "Accurate modeling of coil inductance for near-field wireless power transfer," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 9, pp. 4158–4169, 2018.

[20] B. K. Kushwaha, G. Rituraj, P. Kumar, and P. Bauer, "Mathematical model for the analysis of series-parallel compensated wireless power transfer system for different misalignments," *IET Circuits, Devices & Systems*, vol. 13, no. 7, pp. 970–978, 2019.

[21] F. Liu, Y. Yang, D. Jiang, X. Ruan, and X. Chen, "Modeling and optimization of magnetically coupled resonant wireless power transfer system with varying spatial scales," *IEEE Transactions on Power Electronics*, vol. 32, no. 4, pp. 3240–3250, 2017.

[22] P. Gao, Z. Tian, T. Pan, J. Wu, and W. Gui, "Transmission efficiency analysis and optimization of magnetically coupled resonant wireless power transfer system with misalignments," *AIP Advances*, vol. 8, no. 8, 2018.

[23] Z. Zhang and B. Zhang, "Angular-misalignment insensitive omnidirectional wireless power transfer," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 4, pp. 2755–2764, 2020.

[24] D. Liu and S. V. Georgakopoulos, "Cylindrical misalignment insensitive wireless power transfer systems," *IEEE Transactions on Power Electronics*, vol. 33, no. 11, pp. 9331–9343, 2018.

[25] L. Murliky, R. W. Porto, V. J. Brusamarello, F. R. de Sousa, and A. Triviño-Cabrera, "Active tuning of wireless power transfer system for compensating coil misalignment and variable load conditions," *AEU - International Journal of Electronics and Communications*, vol. 119, article 153166, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 6666228, 10 pages
https://doi.org/10.1155/2020/6666228

Research Article

Sensor Ontology Metamatching with Heterogeneity Measures

Xingsi Xue,[1,2,3,4,5] **Jiawei Lu**,[6] **Chengcai Jiang**,[7] **and Yikun Huang**[8]

[1]*Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou, Fujian 350118, China*
[2]*Intelligent Information Processing Research Center, Fujian University of Technology, Fuzhou, Fujian 350118, China*
[3]*Guangxi Key Laboratory of Automatic Detecting Technology and Instruments, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China*
[4]*Fujian Key Lab for Automotive Electronics and Electric Drive, Fujian University of Technology, Fuzhou, Fujian 350118, China*
[5]*Institute of Artificial Intelligence, Fujian University of Technology, Fuzhou, Fujian 350118, China*
[6]*School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou, Fujian 350118, China*
[7]*College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan 030024, China*
[8]*Concord University College, Fujian Normal University, Fuzhou 350117, China*

Correspondence should be addressed to Xingsi Xue; jack8375@gmail.com

Received 5 October 2020; Revised 2 November 2020; Accepted 12 November 2020; Published 25 November 2020

Academic Editor: Feng-Jang Hwang

Copyright © 2020 Xingsi Xue 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 heterogeneity problem among different sensor ontologies hinders the interaction of information. Ontology matching is an effective method to address this problem by determining the heterogeneous concept pairs. In the matching process, the similarity measure serves as the kernel technique, which calculates the similarity value of two concepts. Since none of the similarity measures can ensure its effectiveness in any context, usually, several measures are combined together to enhance the result's confidence. How to find suitable aggregating weights for various similarity measures, i.e., ontology metamatching problem, is an open challenge. This paper proposes a novel ontology metamatching approach to improve the sensor ontology alignment's quality, which utilizes the heterogeneity features on two ontologies to tune the aggregating weight set. In particular, three ontology heterogeneity measures are firstly proposed to, respectively, evaluate the heterogeneity values in terms of syntax, linguistics, and structure, and then, a semiautomatically learning approach is presented to construct the conversion functions that map any two ontologies' heterogeneity values to the weights for aggregating the similarity measures. To the best of our knowledge, this is the first time that heterogeneity features are proposed and used to solve the sensor ontology metamatching problem. The effectiveness of the proposal is verified by comparing with using state-of-the-art ontology matching techniques on Ontology Alignment Evaluation Initiative (OAEI)'s testing cases and two pairs of real sensor ontologies.

1. Introduction

A sensor network is composed of various sensors. In order to realize the information integration and communication between multiple sensor networks, a semantic sensor web is born, which is composed of the semantic web and the sensor network. In the semantic sensor web, sensor data is marked as metadata characterized by sensor ontology to increase interoperability [1–3]. Sensor ontology can reach a consensus on the meaning of concepts in sensor certain field and provide rich domain knowledge and semantic vocabulary for the interaction between application systems in the semantic sensor web. In recent years, there have been many sensor ontologies, such as SensorOntology2009 (https://www.w3.org/2005/Incubator/ssn/wiki/SensorOntology2009), IoT-Lite (https://www.w3.org/Submission/2015/SUBM-iot-lite-20151126/), original SSN (Semantic Sensor Network) (https://www.w3.org/2005/Incubator/ssn/wiki/SSN#Sensor), new SSN (https://www.w3.org/ns/ssn/), and SOSA (Sensor, Observation, Sample, and Actuator) (https://www.w3.org/ns/sosa/). These ontologies can represent the function, performance, and usage conditions of sensors, which can provide different data for different purposes and contexts [4, 5]. However, in the distributed and open environment, since the people who construct sensor ontologies might have diverse preferences, knowledge backgrounds, and styles of understanding knowledge, one concept could be defined with different names or granularities, and the context of a concept could also be different as well. These lead to the problem of ontology heterogeneity [6, 7]. For example, the existence of the concept "personal computer" in one ontology is defined as "PC" in another ontology, which results in syntax heterogeneity. To support the semantic interaction between intelligent systems based on sensor ontology, we need to determine the correspondence between their heterogeneous concepts in sensor ontologies, which is the so-called sensor ontology matching [8]. During the matching process, it is important to distinguish identical concepts by calculating their similarity value, and thus, similarity measure becomes the kernel technique that directly affects the alignment's quality. At present, similarity measures can be mainly divided into three categories, i.e., syntax-based measure, linguistics-based measure, and structure-based measure [9]. Although these methods have been extensively studied and used, it is still a challenge to find an effective similarity measure for a specific matching task. With the emergence of various similarity measures, it is a feasible way of using multiple similarity measures to overcome the limitations of a single one [10–12]. However, how to find suitable aggregating weights to combine them becomes another problem, i.e., ontology metamatching problem [13]. Traditional metamatching methods include Max, Min, and average approaches [14–16], which, respectively, take into account the maximum value, minimum value, and average value of the concepts measured by various similarity measures but ignore the heterogeneity features of concepts. This paper proposes a novel ontology metamatching approach to face this challenge, which uses the heterogeneity features of two ontologies to semiautomatically tune the weights for aggregating different similarity measures.

In particular, three ontology heterogeneity measures are firstly proposed to, respectively, evaluate the heterogeneity values in terms of syntax, linguistics, and structure heterogeneity features, and then, along with the aggregating weights of corresponding similarity measures which are given by experts, conversion functions that convert any two ontologies' heterogeneity values to the weights for aggregating the similarity measures are trained. The training process of conversion function is shown in Figure 1, where O_r and O_i represent the ith pair of ontologies for training (O_r is the reference ontology and O_i is a target ontology), and W_i, W_i', and W_i'' are the ith set of weights for aggregating the syntax-based, linguistics-based, and structure-based similarity measures, respectively, which are given by experts and have the best effectiveness to measure the ith pair of ontologies. The conversion function aims at converting the heterogeneity value into corresponding aggregating weight, such as converting the syntax heterogeneity to the aggregating weight for syntax-based similarity measure.

In the rest of the paper, Section 2 defines the related concepts about sensor ontology and the evaluation function; Section 3 describes in detail the ontology heterogeneity measure and semiautomatically learning method; Section 4 shows the experimental results; Section 5 summarizes the work and presents future work.

2. Preliminaries

2.1. Sensor Ontology and Entity Correspondence. Sensor ontology is a 3-tuple [17] $O = (C, P, I)$, where C is a nonempty set of classes, P is a nonempty set of properties, I is a set of instances (can be empty). In general, class, property (datatype property and object property), and instance are called entities.

The sensor ontology matching measure utilizes the information inside the sensor ontology and other knowledge bases outside the sensor ontology to determine the similar entities in two sensor ontologies [18]. The entity mapping set is called sensor ontology alignment, and each entity correspondence inside is defined as follows:

An entity correspondence is a 4-tuple [19] (e, e', n, r), where e and e' are two entities from the sensor ontologies, respectively; n represents the similarity value obtained by calculating e and e' through mathematical functions; and r is the semantic relationship between e and e', which is the equivalence relationship in this work [20].

2.2. Similarity Measure. A similarity measure is a function that takes as input two entities' information d and d' and returns their similarity value of $v = g(d, d') \in [0, 1]$. Next, we introduce three types of similarity measures in detail, i.e., syntax-based, linguistic-based, and structure-based.

2.2.1. Syntax-Based Similarity Measure. Syntax-based similarity measure utilizes the syntax information of two strings to calculate their similarity value. There are many similarity measures based on syntax, such as Jaro-Winkler distance [21], SMOA distance [22], and N-gram [23]. According to

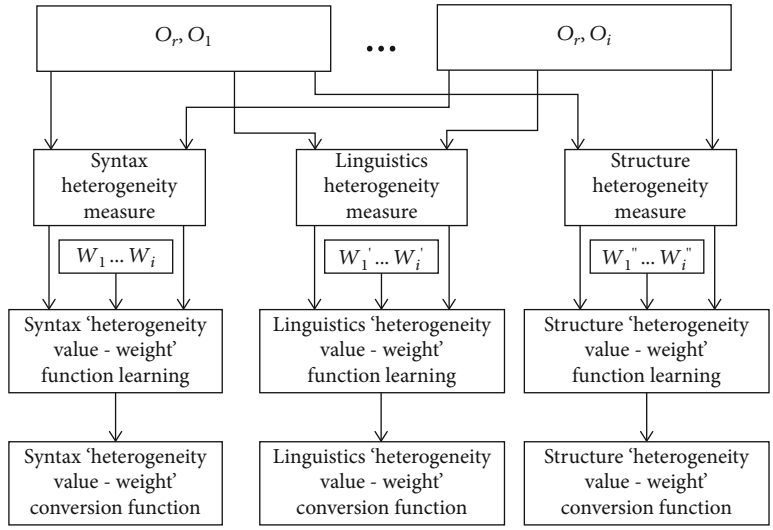


Figure 1: The construction progress of conversion function.

[23], N-gram similarity measure shows superior performance when concepts match, and especially when $N = 3$, the best effect is achieved. Therefore, this paper uses the N-gram as the representative of the syntax-based similarity measure. N-gram is defined as follows:

$$\text{N-gram}(s_1, s_2) = \frac{2 \times \text{comm}(s_1, s_2)}{N_{s_1} + N_{s_2}}, \tag{1}$$

where s_1 and s_2 are the strings to be matched, N_{s_1} and N_{s_2}, respectively, represent the number of substring in the two strings when $N = 3$, and $\text{comm}(s_1, s_2)$ represents the number of common substring in the two strings.

2.2.2. Linguistics-Based Similarity Measure. Linguistics-based similarity measure calculates two words' similarity value by using a background knowledge base, such as WordNet [24]. WordNet is an electronic language database which includes the set of synonyms of various vocabularies and defines the subordinate relation between different vocabularies. It is usually used to measure the similarity between concepts. The Wu and Palmer similarity measure used in this paper uses the depth of the two concepts in WordNet and the depth of their nearest common parent concept to calculate the similarity [25]. It considers the change of connection strength between concepts, and the result of measurement is more accurate. Given two words c_1 and c_2, the Wu and Palmer similarity measure is defined as follows:

$$\text{Sim}_{wp}(c_1, c_2) = \frac{2 \times \text{depth}(\text{LCA}(c_1, c_2))}{\text{depth}(c_1) + \text{depth}(c_2)}, \tag{2}$$

where $\text{LCA}(c_1, c_2)$ is the closest common parent concept about the concept c_1 and c_2 in the WordNet hierarchy. $\text{depth}(\text{LCA}(c_1, c_2))$, $\text{depth}(c_1)$, and $\text{depth}(c_2)$ represent the depth of the common parent concept, c_1 and c_2 in the hierarchy, respectively.

2.2.3. Structure-Based Similarity Measure. Structure-based similarity measure utilizes two concepts' parent-child concepts defined in their ontologies to calculate the similarity value. As a classic structure-based similarity measure, SimRank [26] works based on the assumption that if two concepts are related to similar concepts, then the two concepts are similar. The relationship graph between concepts is expressed as $G = (V, E)$, V as a node represents an object, and E as an edge connecting nodes represents the relationship between objects. The SimRank similarity measure is defined as follows:

$$s(a, b) = \begin{cases} 1, & a = b, \\ \dfrac{C}{|I(a)||I(b)|} \displaystyle\sum_{i=1}^{|I(a)|} \sum_{j=1}^{|I(b)|} s\big(I_i(a), I_j(b)\big), & I(a) \neq 0,\ I(b) \neq 0, \\ 0, & I(a) = 0,\ I(b) = 0, \end{cases} \tag{3}$$

where a and b are two concepts from two ontologies, respectively, $I_i(a)$ represents the ith concept that points to a in the relationship graph, and $|I(a)|$ represents the cardinality of the concepts pointing to a in the relationship graph. When $a = b$, the similarity value between concepts is 1; otherwise, the similarity value of a and b is the average value of the similarity value of all combinations of the concepts pointing to a and the concepts pointing to b. The constant C as the damping coefficient is generally assigned to 0.8. When the number of concepts pointing to a or b is 0, the similarity value between a and b is 0.

As a recursive algorithm, SimRank algorithm has an iterative process as follows. Let $R_k(a, b)$ be the similarity of (a, b) in the kth iteration process and initialized as

$$R_0 = (a, b) = \begin{cases} 1, & a = b, \\ 0, & a \neq b. \end{cases} \tag{4}$$

When $a = b$, the initial similarity value between concepts is 1; otherwise, the initial similarity value between concepts is 0, using the initialization result to start iteration:

$$R_{k+1}(a,b) = \frac{C}{|I(a)||I(b)|} \sum_{i=1}^{|I(a)|} \sum_{j=1}^{|I(b)|} R_k\left(I_i(a), I_j(b)\right). \quad (5)$$

When $a = b$, $R_{k+1}(a,b) = 1$. The similarity of each generation (a,b) will be updated according to the value of the previous generation. $R_{k+1}(a,b)$ converges with the increase of k. When $k \to \infty$, it can be considered that $R_k(a,b) = s(a,b)$. Usually, it has the best effect when $k \approx 5$.

2.3. Evaluation Metrics on Alignment's Quality. Traditionally, the quality of a sensor ontology alignment can be evaluated by recall, precision, and f-measure [27, 28]. The recall indicates the correct matching pairs among the total found matching pairs accounting for the reference matching pairs. When the recall is 1, it means that all correct matching pairs have been found. However, this does not mean that there are no wrong matching pairs among the found matching pairs, so the accuracy needs to be evaluated. The precision indicates the correct matching pairs among the total found matching pairs accounting for the found matching pairs. In order to comprehensively evaluate the matching quality, the weighted harmonic mean (i.e., f-measure) of the precision and recall is used. The precision, recall, and f-measure are defined as follows:

$$\begin{aligned} \text{Recall} &= \frac{|R \cap A|}{|R|}, \\ \text{Precision} &= \frac{|R \cap A|}{A}, \\ f\text{-measure} &= \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}}, \end{aligned} \quad (6)$$

where R is a reference alignment and A is an alignment.

3. Ontology Heterogeneity Measure and Conversion Function

In this work, we dedicate to semiautomatically tuning the aggregating weights for various similarity measures by taking into consideration two ontologies' heterogeneity features. Firstly, we construct an ontology hierarchy graph for each ontology, whose nodes are the ontology's concepts and the edge between two nodes is the relationship between the corresponding concepts [29]. Secondly, the concepts in each ontology are ranked in descending order according to their corresponding nodes' out-in-degrees [30] in the ontology hierarchy graph. After that, the first third concepts with the best ranked are selected as the representative concepts, which are further used to measure the heterogeneity value of two ontologies. On this basis, the representative datatype properties and object properties can also be determined. Then, given the similarity measures of their aggregating weights, we try to learn conversion functions that map any two ontologies' heterogeneity values to the weights for aggregating the corresponding similarity measures.

3.1. Ontology Heterogeneity Measure. To the best of our knowledge, this is the first time to quantify the degree of ontology heterogeneity features. To reduce the computational complexity, it is necessary to select the representative concepts from the ontology for the heterogeneity calculation. In this work, we first calculate the sum of each node's out-in-degrees, which reflects the close connection between this node and its surrounding nodes, as well as the importance of the nodes. Then, the first third concepts with the best ranked are selected as the representative concepts (i.e., classes). The datatype properties and object properties related to representative classes can be chosen as the representative ones.

Being inspired by traditional metrics such as recall and precision, the heterogeneity value in this work is calculated through the ratio of the heterogeneous part to the overall part. Supposing cardinality of the alignment in this paper is one-to-one mapping, E_1 and E_2 are two entity sets to be matched; the heterogeneity value is calculated as follows:

$$H = \frac{\min\{|E_1|, |E_2|\} - |E_{\text{mapped}}|}{\min\{|E_1|, |E_2|\}}, \quad (7)$$

where $|E_1|$ and $|E_2|$ are, respectively, E_1 and E_2's cardinalities, $|E_{\text{mapped}}|$ is the cardinality of the identical entities in the two entity sets. At present, there are three kinds of ontology heterogeneity, and the quantification of the heterogeneity of each of them should be carried out on the entities related to the heterogeneity. In different heterogeneity measures, the choice of entity sets E_1 and E_2 is different:

(1) Syntax Heterogeneity Measure. It is inspired by the syntax similarity measure N-gram, which directly measures the entities in the ontology. Then, the quantification of syntax heterogeneity feature will also focus on the entities in the ontology. E_1 and E_2, respectively, represent the representative entity sets in the two ontologies to be matched

(2) Linguistics Heterogeneity Measure. It is inspired by linguistics similarity measure Wu and Palmer, measuring the synonyms of the entities to be matched in WordNet, that the quantification of linguistic heterogeneity will also focus on the synonyms of the entities in the ontology in WordNet. E_1 and E_2, respectively, represent the set of representative entities in the two ontologies to be matched and their synonyms in WordNet

(3) Structure Heterogeneity Measure. The parent-child concepts of the entities to be matched in the ontology will be considered into the quantification of structural heterogeneity, and it is enlightened by the structural similarity measure SimRank, which measures the parent-child concepts of the entities to be matched in the ontology hierarchy graph. E_1 and E_2 , respectively, represent the set of representative

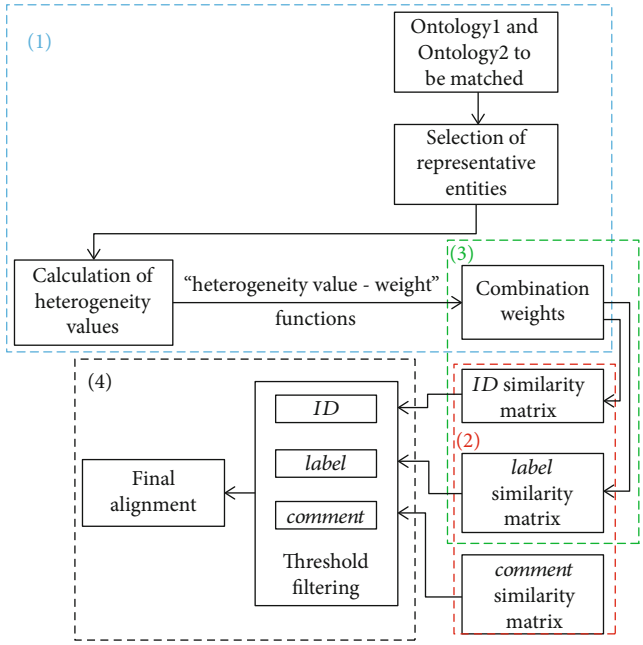


Figure 2: Ontology alignment determination.

entities in the two ontologies to be matched and their parent-child concepts in the ontology hierarchy

According to different heterogeneity features, different entities are selected as the input of the heterogeneity method (formula (7)), which can quantify the degree of different heterogeneity features more targeted.

3.2. Conversion Function Learning. The construction of the "heterogeneity value-weight" conversion function is aimed at exploring the functional relationship between the aggregating weights of various similarity measures and the heterogeneity features that cause the heterogeneity between ontologies, aiming to give more effective aggregating weights to various similarity measures. After determining three kinds of heterogeneity values between two ontologies, we need to further convert them to the aggregating weights for different similarity measures. So the conversion function is needed to construct. As shown in Figure 1, the learning data of the conversion function includes heterogeneity value and the corresponding aggregating weight (about N-gram, Wu and Palmer, or SimRank) given by experts.

This work uses the data analysis software Origin2020 (https://www.originlab.com) to construct the conversion function. The fitting algorithm behind the tool is least-squares. Compared to language interpolating polynomials and cubic spline interpolation, the least-squares algorithm is more suitable for forecasting. Furthermore, by observing that "heterogeneity value-weight" fitting data is an exponential function, and the least-squares algorithm is capable of fitting the exponential function.

Due to the different heterogeneity features of various entities in ontology, not all the three types of similarity measures are needed for each kind of entity.

(1) For class. The learning data for fitting can be defined as (X_i^j, Q_i^j), where $X_i (i = 1, 2, 3)$ represents a certain kind of heterogeneity value of the three, and Q_i is the aggregating weight of the similarity measure which belongs to the same kind, such as X_i is the syntax heterogeneity value and Q_i is the aggregating weight of the syntax-based similarity measure N-gram. $j = 1, 2, 3 \cdots$ represents the serial number of the data for fitting

(2) For properties (datatype property and object property). Due to the lack of structural features, SimRank is not suitable for them. So in (X_i^j, Q_i^j), $i = 1, 2$

Finally, seven conversion functions are obtained. The quality of the conversion functions will be evaluated with coefficient of determination (R-square) [31].

4. Ontology Metamatching

In this section, we describe the proposed ontology metamatching technique in detail. The entities in an ontology can be divided into three types, i.e., class, datatype property, and object property, and they all have such information as ID, label, and comment. Next, we take class as an example to illustrate the matching process: given a pair of ontologies to be matched, (1) select the representative classes from the two ontologies, calculate three kinds of heterogeneity values, and, respectively, substitute them into the corresponding "heterogeneity value-weight" functions to obtain the aggregating weights. In particular, the ith aggregating weights are denoted as W_i; (2) construct the similarity matrices on ID and label using different similarity measures, which are, respectively, denoted as M_i^{ID} and M_i^{label}. Since comment has no structural relationship and it is usually a sentence that SimRank and Wu and Palmer could not measure it effectively, we utilize N-gram to measure their similarity; (3) multiply the similarity matrix of ID and label with their corresponding aggregating weights and, respectively, add these weighted matrices to a comprehensive similarity matrix, i.e., $\sum_{i=1}^{3} M_i^{\mathrm{ID}} \times W_i$ and $\sum_{i=1}^{3} M_i^{\mathrm{label}} \times W_i$; (4) use a threshold to filter the comprehensive similarity matrix: first, filter the ID matching pairs; second, for those entities that do not match, filter them by their labels; third, for those entities that still do not match, filter by their comments; finally, put the mappings whose similarity values are above the threshold into the final alignment in each step. For the sake of clarity, we show the flowchart of the process of determining the final ontology alignment in Figure 2.

5. Experiment

5.1. Dataset and Experimental Configuration. In the experiment, we use the testing cases provided by the Ontology Alignment Evaluation Initiative (OAEI) (http://oaei.ontologymatching.org). A brief description on OAEI's testing cases is shown in Table 1. In particular, each case contains two ontologies to be matched and one reference alignment for evaluating the quality of the alignment. The

Table 1: The brief descriptions on OAEI's testing cases and real sensor ontologies.

Ontology		Brief description
OAEI	101-104	The ontology pairs to be matched in the cases are identical or have different constraints in OWL.
	201-210	The concept language features of the ontology pairs to be matched are different, but the structure of them is the same.
	221-247	The concept structure of the ontology pairs to be matched is different, but the language features of them are the same.
	248-262	The ontology pairs to be matched in the cases have different lexical, linguistic, and structure features.
Real sensor ontologies	Original SSN	Description of sensors and observations and related concepts.
	New SSN	Addressing changes in scope and audience, shortcomings of the original SSN, and new technical developments.
	SOSA	As a lightweight core for SSN.

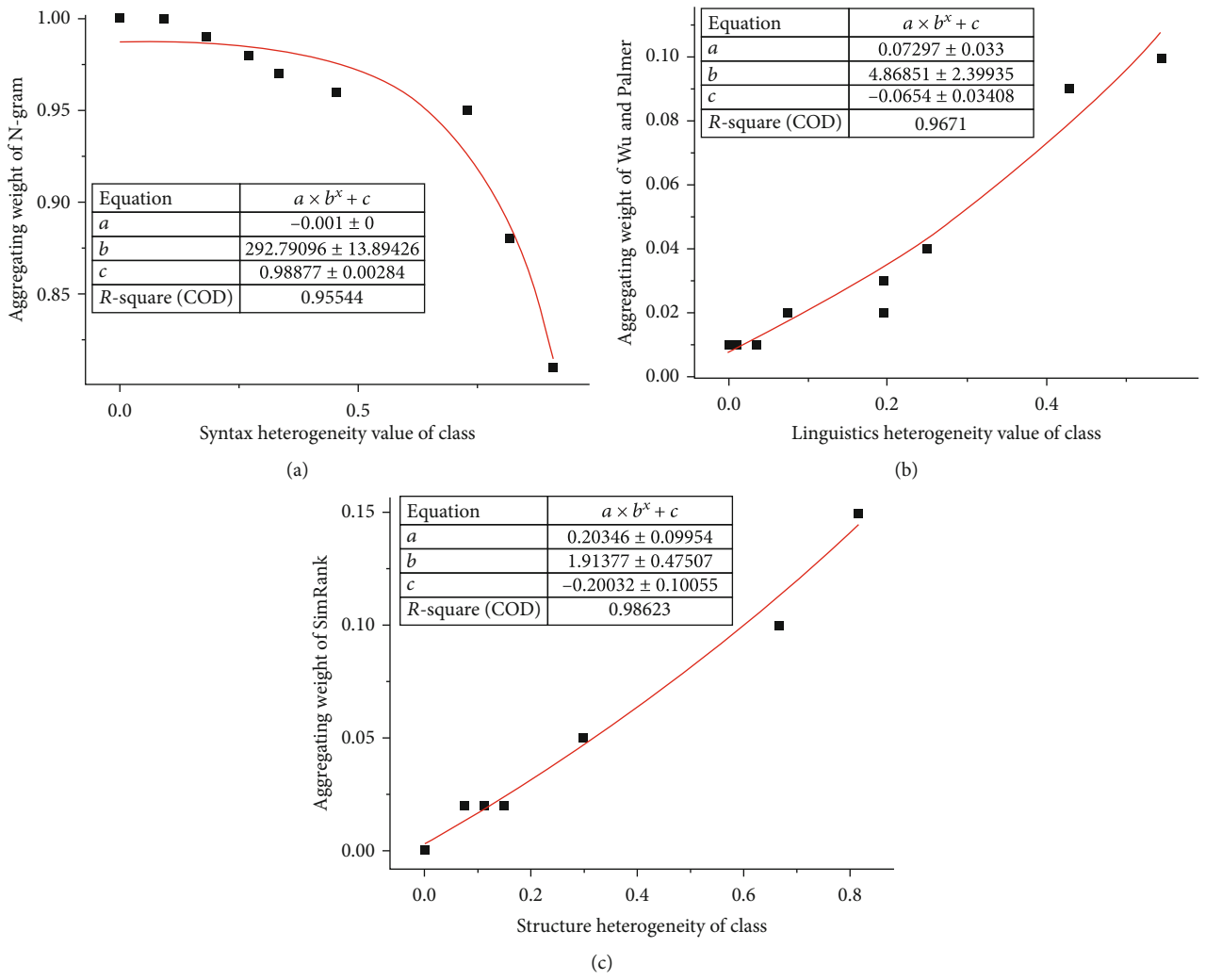


Figure 3: Conversion function of class. (a) "Heterogeneity value-weight" conversion function on syntax. (b) "Heterogeneity value-weight" conversion function on linguistics. (c) "Heterogeneity value-weight" conversion function on structure.

testing cases include another part of the benchmark above and two pairs of real sensor ontologies, i.e., new SSN vs. original SSN and new SSN vs. SOSA. Where the original SSN describes sensors and observations and related concepts, the new SSN improves the original SSN as follows: addressing changes in scope and audience, addressing shortcomings of

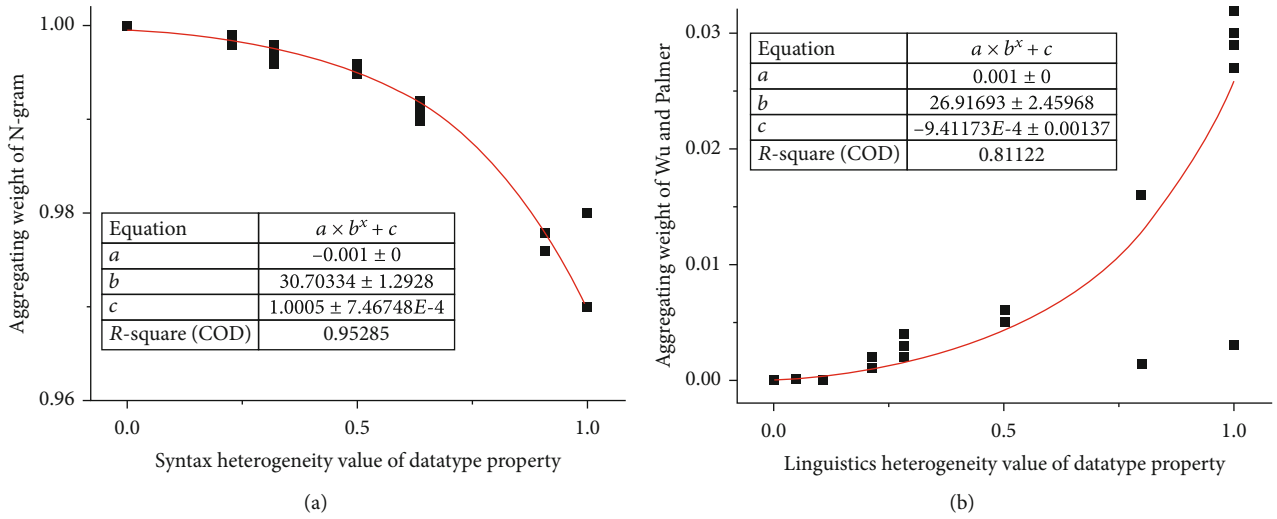


Figure 4: Conversion function of datatype property. (a) "Heterogeneity value-weight" conversion function on syntax. (b) "Heterogeneity value-weight" conversion function on linguistics.

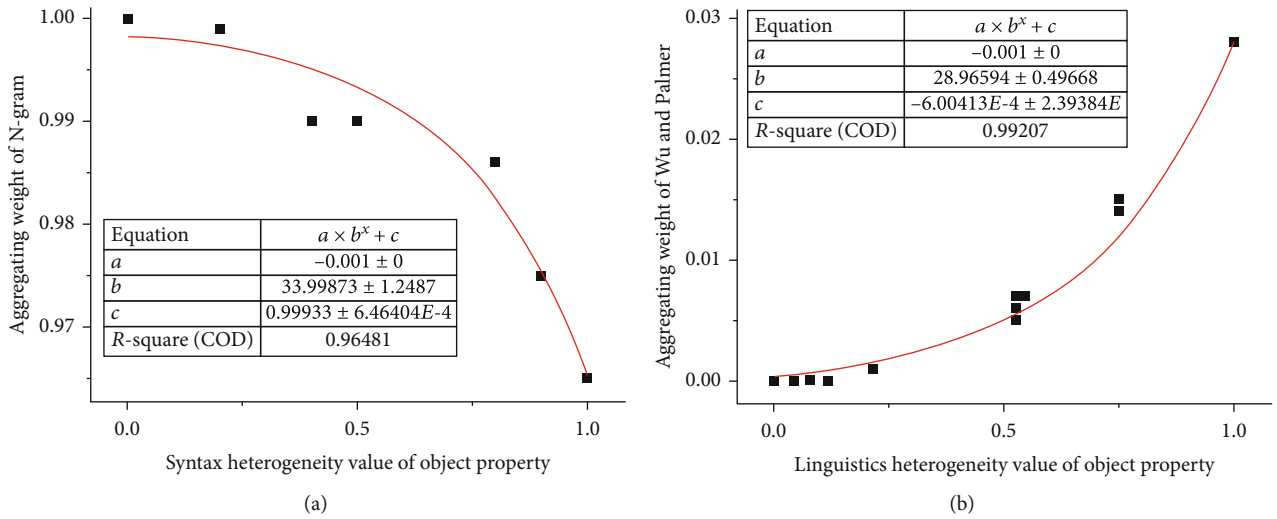


Figure 5: Conversion function of object property. (a) "Heterogeneity value-weight" conversion function on syntax. (b) "Heterogeneity value-weight" conversion function on linguistics.

the original SSN, and addressing technical developments; SOSA provides a lightweight core for SSN and aims at broadening the target audience and application areas that can make use of Semantic Web ontologies (https://www.w3.org/TR/vocab-ssn/). As classic sensor ontologies, SSN series sensor ontology is widely used in satellite imagery, large-scale scientific monitoring, industrial and household infrastructures, social sensing, citizen science, observation-driven ontology engineering, and the Web of Things. They greatly promote information exchange in various fields.

The relevant configuration of the experiment is as follows: the threshold of the similarity matrix is 0.68; the model of the conversion function is the exponential function $y(x) = a \times b^x + c$. In this experiment, through the benchmark, we compare our approach with OAEI's participants and three classic metamatching methods, i.e., maximum-based metamatching

Table 2: Comparison with OAEI's participants.

Matching system	Recall	Precision	f-measure
edna	0.68	0.50	0.56
AML	0.43	1.00	0.55
LogMap	0.65	0.92	0.72
LogMapLt	0.68	0.50	0.55
PhenoMF	0.02	0.02	0.02
PhenoMM	0.02	0.02	0.02
PhenoMP	0.01	0.02	0.02
XMap	0.59	0.96	0.68
LogMapBio	0.53	0.53	0.51
Our approach	0.74	0.98	0.82

Table 3: Comparison with the maximum-based metamatching approach.

Testing case	Recall		Precision		f-measure	
	Maximum-based approach	Our approach	Maximum-based approach	Our approach	Maximum-based approach	Our approach
101-104	1.00	1.00	1.00	1.00	1.00	1.00
201-210	0.73	0.73	0.83	0.95	0.75	0.80
221-247	0.98	1.00	0.97	0.99	0.98	1.00
248-262	0.56	0.58	0.90	0.98	0.68	0.71
101-262	0.73	0.74	0.92	0.98	0.80	0.82

Table 4: Comparison with the minimum-based metamatching approach.

Testing case	Recall		Precision		f-measure	
	Minimum-based approach	Our approach	Minimum-based approach	Our approach	Minimum-based approach	Our approach
101-104	0.98	1.00	0.97	1.00	0.97	1.00
201-210	0.60	0.73	0.89	0.95	0.66	0.80
221-247	0.96	1.00	0.96	0.99	0.96	1.00
248-262	0.16	0.58	0.49	0.98	0.24	0.71
101-262	0.49	0.74	0.70	0.98	0.54	0.82

Table 5: Comparison with the average-based metamatching approach.

Testing case	Recall		Precision		f-measure	
	Average-based approach	Our approach	Average-based approach	Our approach	Average-based approach	Our approach
101-104	0.98	1.00	0.97	1.00	0.97	1.00
201-210	0.60	0.73	0.89	0.95	0.66	0.80
221-247	0.96	1.00	0.96	0.99	0.96	1.00
248-262	0.16	0.58	0.49	0.98	0.24	0.71
101-262	0.49	0.74	0.70	0.98	0.54	0.82

approach, minimum-based metamatching approach, and average-based metamatching approach. Finally, we apply the method of this paper to two pairs of real sensor ontologies to verify the robustness.

5.2. Experimental Results and Analysis. As shown in Figures 3–5, seven conversion functions are achieved, among which class has three kinds of conversion functions, while due to the lack of structural features, datatype property and object property have two kinds of conversion functions, respectively. According to R-square, we can see that each conversion function is able to effectively reflect the mapping relation between heterogeneity value and aggregating weight. From the conversion functions, we can find that the aggregating weight of syntax-based similarity measure N-gram takes the largest percentage among the similarity measures for all three kinds of entities, while the others account for a small part. In addition, the aggregating weight of the syntax-based N-gram decreases exponentially as its heterogeneity value increases for all three kinds of entities; on the contrary, the aggregating weight of the linguistics-based similarity measure Wu and Palmer and the structure-based similarity measure SimRank increases exponentially as their heterogeneity value increases.

In Table 2, edna [32] is the baseline method, and AML [33], LogMap series [34], and XMap [35] are OAEI's participants. In terms of the comprehensive evaluation index f-measure, our method score is 0.82, which is much higher than other state-of-the-art methods. Since AML uses more than ten similarity measures, its precision is slightly higher than our method, but there will be suppression between multiple similarity measures, which yields low recall and high computational complexity on tuning these similarity measures and debugging the results. Our approach's recall is 0.74, which is higher than other state-of-the-art methods.

As shown in Tables 3–5 that in all 45 testing cases, our method achieves better results than other competitors in terms of 41 testing cases and has equal results with other competitors in terms of another 4 testing cases, which further shows the effectiveness of our method. For a more intuitive comparison, a bar graph of f-measure for the four methods is drawn in Figure 6.

Both in comparisons with the state-of-the-art methods in OAEI and three classic metamatching methods, our method achieves ideal results. This is because other methods ignore the effect of heterogeneity features, which makes them unable to identify more matching pairs effectively. The heterogeneity features into the process of obtaining aggregating weights

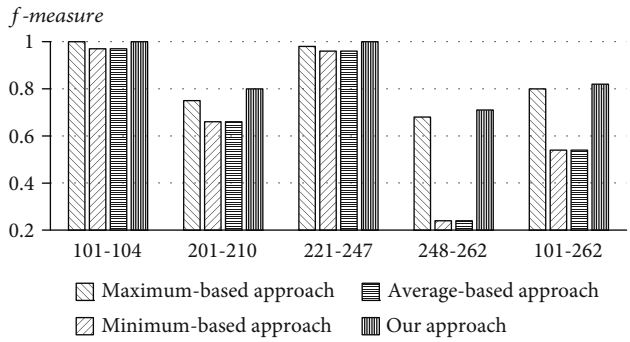


Figure 6: Comparison with the maximum-, minimum-, and average-based metamatching approach.

Table 6: Evaluation on two pairs of real sensor ontologies.

Testing case	Recall	Precision	f-measure
New SSN vs. original SSN	0.97	1	0.98
New SSN vs. SOSA	1	1	1

are first time taken into account in our method and find more matching pairs and realize effective matching.

When matching real sensor ontologies, new SSN is treated as the source ontology, and the other two real sensor ontologies are treated as the target ontologies. It can be seen from Table 6 that our approach is able to obtain results that are very close to the golden alignment. When matching the new SSN and the original SSN, the recall obtained by our approach is 0.97, which is due to that we suppose the alignment's cardinality is one to one, i.e., one concept in the ontology can only be mapped with one concept of another ontology and vice versa. However, the cardinality could be one to many in some practical tasks, e.g., the concept "Stimulus" in the new SSN is mapped with the concept "Stimulus" in the original SSN which is asserted the same as another class "Sensor Input." In our approach, according to the filtering sequence of step 4 in Figure 2, if "Stimulus" in the new SSN first is mapped with "Stimulus" in the original SSN by ID, it will no longer participate in the filtering process (although the comment of "Stimulus" in the new SSN is the same as the comment of "Sensor Input" in the original SSN), which decreases our approach's recall value.

6. Conclusion and Future Work

Addressing ontology metamatching problem requires the determination of suitable aggregating weights of several similarity measures for matching ontologies. Considering that the heterogeneity features of ontology have an important influence on the aggregating weight of similarity measures, this paper first proposes three heterogeneity measures to quantify the heterogeneity values of ontology heterogeneity features and then constructs the functional relationship between the heterogeneity feature and the aggregating weights of various similarity measures. The effectiveness of the proposal is verified by comparing with the state-of-the-art matching techniques in OAEI and other classic metamatching methods. The robustness of the method is proved by matching two pairs of real sensor ontologies.

In the future, we are interested in further improving this approach to deal with the matching task with the alignment's cardinality being many to many. We are also interested in addressing large-scale matching problem, where two sensor ontologies might own tens of thousands of concepts and instances. In addition, when the number of similarity measures is large, e.g., more than 50, the selection of suitable ones is the critical issue that we need to address. Finally, an efficient user interacting strategy should be used to further reduce the interacting time and the user's error rate.

Data Availability

The data used to support this study can be found in http://oaei.ontologymatching.org.

Conflicts of Interest

The authors declare that they have no conflicts of interest in the work.

Acknowledgments

This work is supported by the Natural Science Foundation of Fujian Province (No. 2020J01875), the Guangxi Key Laboratory of Automatic Detecting Technology and Instruments (No. YQ20206), the Scientific Research Foundation of Fujian University of Technology (No. GY-Z17162), the 2018 Program for Outstanding Young Scientific Researcher in Fujian Province University, the Fujian Province Undergraduate Universities Teaching Reform Research Project (No. FBJG20190156), the Science and Technology Planning Project of Fuzhou City (No. 2019-G-40), and the Foreign Cooperation Project in Fujian Province (No. 2019I0019).

References

[1] X. Xue and J. Chen, "Optimizing sensor ontology alignment through compact co-firefly algorithm," *Sensors*, vol. 20, no. 7, article 2056, 2020.

[2] X. Xue and J. Liu, "Collaborative ontology matching based on compact interactive evolutionary algorithm," *Knowledge-Based Systems*, vol. 137, pp. 94–103, 2017.

[3] X. Xue and J. Chen, "Using Compact Evolutionary Tabu Search algorithm for matching sensor ontologies," *Swarm and Evolutionary Computation*, vol. 48, pp. 25–30, 2019.

[4] X. Xue, J. Chen, and J. S. Pan, *Evolutionary Algorithm based Ontology Matching Technique*, Beijing: Science Press, 2018.

[5] X. Xue, J. Chen, and X. Yao, "Efficient user involvement in semiautomatic ontology matching," *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–11, 2018.

[6] X. Xue, H. Yang, J. Zhang, J. Zhang, and D. Chen, "An automatic biomedical ontology meta-matching technique," *Journal of Network Intelligence*, vol. 4, no. 3, pp. 109–113, 2019.

[7] X. Xue and A. Ren, "An evolutionary algorithm based ontology alignment extracting technology," *Journal of Network Intelligence*, vol. 2, no. 2, pp. 205–212, 2017.

[8] X. Xue and J. Chen, "A preference-based multi-objective evolutionary algorithm for semiautomatic sensor ontology matching," *International Journal of Swarm Intelligence Research*, vol. 9, no. 2, pp. 1–14, 2018.

[9] J. Euzenat and P. Shvaiko, "Classifications of ontology matching techniques," in *Ontology Matching*, pp. 73–84, Springer, Berlin, Heidelberg, 2013.

[10] J.-B. Gao, B.-W. Zhang, and X.-H. Chen, "A WordNet-based semantic similarity measurement combining edge-counting and information content theory," *Engineering Applications of Artificial Intelligence*, vol. 39, pp. 80–88, 2015.

[11] G. Pirró and N. Seco, "Design, implementation and evaluation of a new semantic similarity metric combining features and intrinsic information content," in *On the Move to Meaningful Internet Systems: OTM 2008*, pp. 1271–1288, Springer, Berlin, Heidelberg, 2008.

[12] L. Meng, R. Huang, and J. Gu, "An effective algorithm for semantic similarity metric of word pairs," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 8, no. 2, pp. 1–12, 2013.

[13] X. Xue and J.-S. Pan, "A compact co-evolutionary algorithm for sensor ontology meta-matching," *Knowledge and Information Systems*, vol. 56, no. 2, pp. 335–353, 2018.

[14] H.-H. Do and E. Rahm, "COMA — a system for flexible combination of schema matching approaches," *VLDB '02: Proceedings of the 28th International Conference on Very Large Databases*, pp. 610–621, 2002.

[15] J. Wang, Z. Ding, and C. Jiang, "GAOM: genetic algorithm based ontology matching," in *2006 IEEE Asia-Pacific Conference on Services Computing (APSCC'06)*, pp. 617–620, Guangzhou, Guangdong, China, December 2006.

[16] A. Doan, P. Domingos, and A. Halevy, "Learning to match the schemas of data sources: a multistrategy approach," *Machine Learning*, vol. 50, no. 3, pp. 279–301, 2003.

[17] X. Xue and Y. Wang, "Optimizing ontology alignments through a memetic algorithm using both MatchFmeasure and unanimous improvement ratio," *Artificial Intelligence*, vol. 223, pp. 65–81, 2015.

[18] X. Xue, "A compact firefly algorithm for matching biomedical ontologies," *Knowledge and Information Systems*, vol. 62, no. 7, pp. 2855–2871, 2020.

[19] X. Xue, J. Chen, and D. Chen, "Matching Biomedical Ontologies Through Compact Hybrid Evolutionary Algorithm," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 10, no. 1, pp. 110–117, 2019.

[20] X. Xue and J. Lu, "A compact brain storm algorithm for matching ontologies," *IEEE Access*, vol. 8, pp. 43898–43907, 2020.

[21] W. E. Winkler, *The State of Record Linkage and Current Research Problems*, Statistical Research Division, U.S. Bureau of the Census, 1999.

[22] X. Xue and Y. Wang, "Using memetic algorithm for instance coreference resolution," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 580–591, 2016.

[23] V. Mascardi, A. Locoro, and P. Rosso, "Automatic ontology matching via upper ontologies: a systematic evaluation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 5, pp. 609–623, 2010.

[24] G. A. Miller, "WordNet," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[25] Z. Wu and M. Palmer, "Verbs semantics and lexical selection," in *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, pp. 133–138, Las Cruces, New Mexico, 1994.

[26] J. Glen and J. Widom, "Sim Rank: a measure of structural-context similarity," in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '02*, pp. 538–543, Montreal, Canada, 2002.

[27] S.-C. Chuan, X. Xue, J.-S. Pan, and X. Wu, "Optimizing ontology alignment in vector space," *Journal of Internet Technology*, vol. 21, no. 1, pp. 15–22, 2020.

[28] C. J. V. Rijsberge, *Information Retrieval*, University of Glasgow, 1975.

[29] X. Xue and X. Yao, "Interactive ontology matching based on partial reference alignment," *Applied Soft Computing*, vol. 72, pp. 355–370, 2018.

[30] M. Barthelemy, "Betweenness centrality in large complex networks," *The European Physical Journal B - Condensed Matter*, vol. 38, no. 2, pp. 163–168, 2004.

[31] M. Awawdeh, T. Faisal, A. Bashir, and A. Sheikh, "Application of outlier detection using re-weighted least squares and R-squared for IoT extracted data," in *2019 Advances in Science and Engineering Technology International Conferences (ASET)*, pp. 1–6, Dubai, United Arab Emirates, March 2019.

[32] M. Achichi, M. Cheatham, Z. Dragisic et al., *Results of the Ontology Alignment Evaluation Initiative*, 2016.

[33] D. Faria, C. Pesquita, B. S. Balasubramani et al., "OAEI 2016 results of AML," in *11th International Workshop on Ontology Matching co-located with the 15th International Semantic Web Conference, CEUR Workshop Proceedings*, vol. 1766, Kobe, Japan, 2016.

[34] E. Jimenez-Ruiz, B. C. Grau, and V. V. Cross, "Logmap family participation in the OAEI 2016," in *Proceedings of the 11th International Workshop on Ontology Matching co-located with the 15th International Semantic Web Conference*, pp. 185–189, Kobe, Japan, 2016.

[35] D. E. Warith, M. T. Khadir, and S. B. Yahia, "XMap: results for OAEI 2016," in *11th International Workshop on Ontology Matching co-located with the 15th International Semantic Web Conference*, pp. 222–226, Kobe, Japan, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8870438, 10 pages
https://doi.org/10.1155/2020/8870438

Research Article

Blockchain-Based Internet of Vehicles Privacy Protection System

Tianhong Su,[1] **Sujie Shao**,[2] **Shaoyong Guo**,[2] **and Min Lei**[1]

[1]*Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, China*

Correspondence should be addressed to Sujie Shao; buptssj@bupt.edu.cn

Received 16 July 2020; Revised 7 August 2020; Accepted 21 August 2020; Published 7 September 2020

Academic Editor: Chi-Hua Chen

Copyright © 2020 Tianhong Su 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 development of wireless local area networks and intelligent transportation technologies, the Internet of Vehicles is considered to be an effective method to alleviate the severe situation of the current transportation system. The vehicles in the Internet of Vehicles system build the Vehicular Ad Hoc Networks through wireless communication technology and dynamically provide different services through the real-time driving information broadcast by the vehicles. Vehicle drivers can control the distance, planning the driving route, between vehicles according to the current traffic environment, which improves the overall safety and efficiency of the traffic system. Due to the particularity of the Internet of Vehicles system service, vehicles need to broadcast their location information frequently. Attackers can collect and analyze vehicle broadcast information to steal privacy and even directionally track the owner through the driving trajectory, bringing serious security risks. This paper proposes a blockchain-based privacy protection system for the Internet of Vehicles. The system combines the blockchain with the Internet of Vehicles system to design a safe and efficient two-way authentication and key agreement algorithm through encryption and signature algorithm, which also solves the central dependency problem of the traditional Internet of Vehicles system.

1. Introduction

With the improvement of living standards, traffic congestion becomes more and more serious with the increase of vehicles. With the emergence of 5th Generation, Bluetooth, sensor technology, cloud computing, big data, and other new technologies, the field of intelligent transportation has made great progress. In the Internet of Vehicles system, vehicles perceive the surrounding road condition environmental information through the on-board unit and communicate to roadside units and other vehicles in the network through the vehicle communication module so that the vehicle owner can obtain road condition information and navigation information, reducing traffic risks.

Because the Internet of Vehicles system is a complex system composed of vehicles, people, and other network facilities and it communicates through the wireless network, the system faces serious security problems. Therefore, how to protect user privacy while providing services is an urgent problem to be solved.

The concept of Vehicular Ad Hoc Networks (VANETS) originally originated from the Internet of Things and is the core technology of autonomous driving. The Internet of Vehicles system integrates advanced communication technology, and intervehicular communication, intravehicular communication, and vehicular mobile Internet are the three main communication components of the system [1].

Blockchain is essentially a decentralized and distributed database technology. It maintains a chain structure of data blocks among participating nodes, which is a continuously growing and immutable data record based on cryptography [2]. Each block can be logically divided into blockhead and block body; each block is concatenated with a hash value in the blockhead, and the transactions in each block are associated with the Merkle root. The blockchain will synchronize the transaction information to the entire network through the consensus mechanism, and each client stores the latest transaction information, forming a decentralized storage method. When some nodes fail, it will not affect the operation of the entire system [3].

In order to solve the problem of security and communication efficiency, this paper proposed the Blockchain-Based Internet of Vehicles Privacy Protection System. The system solves the problem of central failure caused by excessive access by managing the public key information of vehicles through a distributed mechanism. Due to the tamper-proof feature of blockchain, it can prevent the vehicle's public key information stored on the chain from being illegally tampered with. The system uses malicious behavior voting system to detect the message sent by the vehicle through the blockchain node, which prevents the vehicle from maliciously publishing illegal location information. The system uses blockchain and cryptographic algorithms to design efficient two-way authentication and key exchange strategies to reduce communication costs. At the same time, vehicles in the system can communicate anonymously with location service providers through blockchain, avoiding location service providers from stealing vehicle privacy. In summary, the Blockchain-Based Internet of Vehicles Privacy Protection System proposed in this paper is lower in cost, higher in security, and easier to maintain than the traditional Internet of Vehicles scheme.

The rest of the paper is organized as follows. Literature Review introduces existing techniques. Materials and Methods introduce the details of the system implementation. Security Analysis analyzes the security of the system. Results and Discussion present the performance results obtained for the system and provide a state-of-the-art comparison. Finally, Conclusions conclude the paper and outline some potential future work.

2. Literature Review

The traditional Internet of Vehicles system is based on PKI technology [4], which will generate a lot of communication costs. Due to the centralization of authentication nodes, the central node has heavy tasks, cannot be proxied, and is easy to be compromised, which brings the risk of leakage of user's sensitive information [5]. Salem et al. proposed a dynamic key distribution protocol based on PKI [6]. Vehicles do not need to store a large amount of key information but dynamically obtain keys from CA through RSU, which reduces the storage pressure on the vehicle side. When the vehicle certificate is revoked, the CA only needs to send a part of the revocation message without updating the entire Certificate Revocation List (CRL). Tan et al. proposed the SA-KMP algorithm and, based on the public key system, proposed a key agreement protocol based on the 3-dimension matrix to ensure secure communication between the vehicle and the RSU. In order to better protect the private data in the IoV system, Guo proposed a communication protocol based on pseudonyms [7]. The protocol combines homomorphic key agreement and digital signatures to manage and use pseudonyms to ensure the communication security and privacy protection of vehicles, but there is still the problem of central node dependence. Wang et al. proposed an anonymous certificate distribution mechanism for the Internet of Vehicles [8]. The vehicle obtains a temporary anonymous certificate from the passing RSU and uses it to broadcast messages. Therefore, there is no complicated certificate management problem, and the vehicle does not need to check the time-consuming CRL during the authentication process, which significantly improves authentication efficiency. Mei uses ring signatures and identity-based encryption technologies to authenticate communications between vehicles, but there is no experimental plan to analyze the complex network. Wei et al. proposed a fog-based privacy protection scheme that improves the security of a crowded vehicle network [9]. The schema can ensure that the fog user is anonymized during identity authentication between the fog user and the fog server. In order to reduce the central dependency problem, researchers combine blockchain technology with the Internet of Vehicles. The concept of blockchain was first proposed by Nakamoto in 2008 [10]. Lasla et al. proposed a blockchain-based lightweight authentication method to replace certificate authentication, using blockchain technology to track the certificate of each vehicle in a distributed immutable ledger, and proposed a completely distributed vehicle access mechanism [11]. Yang et al. proposed a blockchain-based vehicle reputation management system. In this system, the Bayesian inference model is used to evaluate adjacent vehicles and send rating information to the blockchain network which realized tamper-proof reputation information [12]. He proposed a key management mechanism of IoV based on blockchain technology [13], which improves the low efficiency of traditional key management schemes and enhances the security of keys.

In general, most of the relevant studies have flaws in communication efficiency, cost, or security. Combining the advantages and disadvantages of the abovementioned existing schema, this paper proposes a privacy protection system based on blockchain. Based on the open, self-organized, and fast-moving features of the Internet of Vehicles, the system uses the tamper-proof and distributed features of blockchain technology to design efficient two-way authentication and key agreement algorithm and an anonymous location service-providing algorithm which improves the security of the Internet of Vehicles system.

3. Materials and Methods

The architecture of the Internet of Vehicles system based on blockchain is shown in Figure 1. A large amount of real-time traffic information such as safety messages needs to be processed in the Internet of Vehicles. Due to the limited computing power of the vehicle, the system uses RSU as a blockchain network node. RSU has certain computing power and has routing and forwarding functions. It adopts a cloud computing strategy to forward the received data to the RSU server, which has powerful real-time computing power and enough storage space to provide necessary computing power for the blockchain system. The cloud server provides a high-speed data query interface for RSU. Cloud computing provides a guarantee for data storage and analysis in the Internet of Vehicles system [14]. The blockchain is used to manage the user's public key and other private information to prevent it from being illegally tampered with; the blockchain system can hold the relevant nodes accountable in the event of a

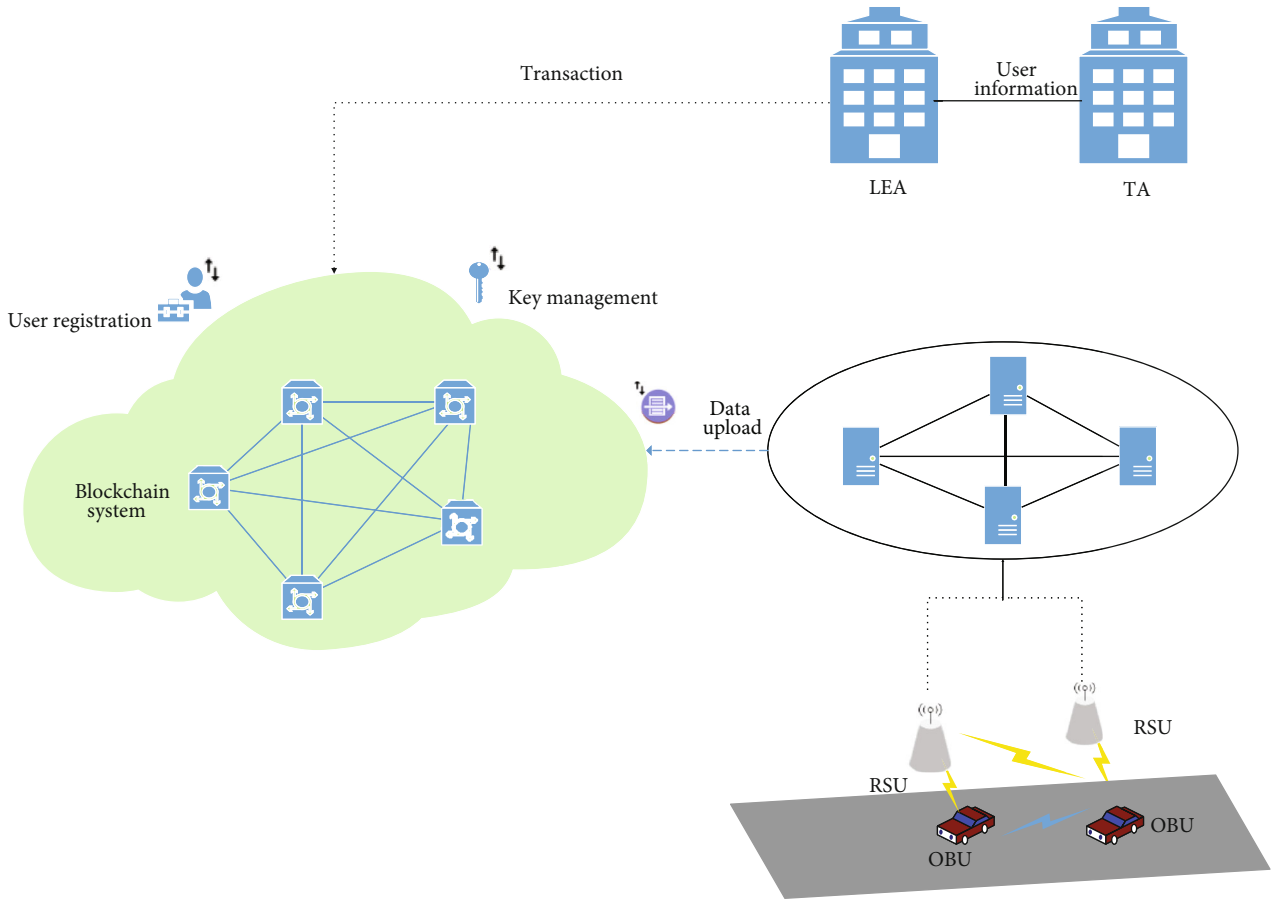


Figure 1: Blockchain-based IoV system.

dispute. Law Enforcement Agency (LEA) is also added to the system to review user information and TA-generated information, which is then signed and posted to the blockchain network (see Figure 1).

3.1. System Initialization. When the system is initialized, TA needs to create a blockchain account for each blockchain node, package the smart contract that implements the key management function into a transaction, and publish it to the blockchain network. TA records the contract address and interface information. When TA communicates with the blockchain node (RSU), it needs to provide the smart contract address to the node; only the correct contract address can trigger the contract to perform the predetermined function.

In order to realize the system's cryptosystem, TA needs to set the parameter $L = (p, a, b, G, n)$ for the system to determine the elliptic curve E in the finite field GF (p). TA needs to allocate public and private key pairs to the prearranged blockchain nodes (RSU) and upload the public key to the blockchain network.

3.2. Key Management. Key management is the core problem to be solved in the application system based on cryptography. In the traditional IoV system, the central authority has full authority to manage the binding of vehicle identity and public key, which requires huge communication and storage costs. As the number of vehicles increases in the system, it may cause the central server to crash. When the data in the central server is tampered and not discovered by the administrator in time, it will cause huge losses to the system.

The vehicles in the IoV system are traveling at high speed; vehicles need to make quick decisions. The central server cannot meet the communication delay problem due to the long communication distance [13]. The system proposed in this paper not only enhances the security of the key in the system but also reduces the communication delay.

3.2.1. User Registration. Before joining the Internet of Vehicles system, the vehicle needs to register with the local trusted agency. The detailed process is as follows:

(1) The owner sends the relevant personal information such as name, ID number, license plate number, vehicle type, and motor vehicle driving license to TA for review. The identity information needs to be strictly reviewed to ensure the legitimacy of the members in the system

(2) After TA reviews the user information, it generates a unique ID value for the user (UID), user public key (UPubK), user private key (UPriK), and validity

period VP (Valid Period) of the user public key information for the user

(3) TA submits the user information and the generated public key information to LEA for review. After checking the information, LEA encodes {UID, UPubK, VP} as hexadecimal data encapsulates it into a transaction and sends the transaction information to the blockchain network. Then, the blockchain node executes according to the logic defined by the registration function in the smart contract. After the contract is successfully executed, the transaction is successfully chained, and the blockchain system feeds back the successful execution result to TA

(4) After the TA obtains the execution result, it will package {UID, UPubK, UPriK, and VP} and send it to the vehicle user through the secure channel

3.2.2. Key Update. Regularly updating key information helps to enhance the security of the system. The key update is for keys whose validity period is about to expire. The user obtains a new set of public and private key pairs from the TA and continues to enjoy the services of the IoV system during the next validity period. The detailed process is as follows:

(1) The user packages the public key information and sends a key update request to the TA; the requested content includes the public key information before update like {UID, UPubK, and VP}

(2) TA reviews the user's request information and generates new public and private keys and validity period (newVP) information after verification, that is, {UID, newPubK, newPriK, and newVP}. TA encapsulates the information and sends an authentication request to LEA

(3) After receiving the verification request, LEA verifies the validity of the public key information, signs it after successful verification, packages it into a transaction, and sends it to the blockchain network

(4) After the blockchain receives the transaction information, the blockchain node executes the key update method in the smart contract. After the contract is successfully executed, the transaction is successfully chained and the blockchain system feeds back the successful execution result to TA

(5) After TA obtains the execution result, it packs the new user information and sends it to the vehicle user through the secure channel

3.2.3. Key Revocation. Key revocation is the process in which the blockchain system analyzes the behavior of vehicles to judge malicious users and revoke their public keys. The system designs a malicious behavior voting mechanism through the blockchain to record the malicious behaviors by using the smart contract. The voting results will be encapsulated into transactions and stored in the blockchain. Key revocation is aimed at finding out malicious users in the system. The detailed process is as follows:

(1) After receiving the distorted broadcast message from the malicious user, the blockchain node (RSU) records the vehicle ID and malicious behavior, packages it into a transaction, uploads it to the blockchain, and calls the voting record function in the smart contract. The function first queries the voting record Vote_v according to this id and then saves the record after adding one. The blockchain system will judge whether the current Vote_v exceeds the threshold THR_v of malicious behavior set by the system. If $\text{Vote}_v > \text{THR}_v$, the system will determine the vehicle as a malicious vehicle and encapsulate the vehicle ID, public key, and other identifying information to send a public key revocation request to TA

(2) After receiving the public key revocation request, TA checks the voting process. If the voting process is correct, the malicious user's information {UID, UPubK, and newVP} will be sent to LEA

(3) LEA reviews the revocation request, encapsulates a key revocation transaction, and sends it to the blockchain network

(4) After receiving the key revocation transaction, the blockchain executes the revocation function in the smart contract, stores the public key revocation transaction on the chain, marks the malicious user's public key as revoked, and sends a successful receipt to TA

The RSU can verify its location through the positioning system. The detailed process of the malicious behavior voting system based on smart contracts is as follows.

After RSU receives the vehicle message, it extracts the location information provided by the vehicle and judges whether the message is severely distorted based on its own location information.

$$\text{bol} = C_L(L_{\text{RSU}}, L_V)\#. \tag{1}$$

If bol is false, it means that the message is severely distorted; the system will drop the data packet and query the current malicious behavior of the vehicle through the smart contract and save the query result after adding one.

If bol is true, it means that RSU determines that this communication is normal.

Key revocation is of great significance for maintaining system security. The system needs to set reasonable thresholds to reduce system risks and ensure the normal operation of the system.

3.3. Communication Technology

3.3.1. Authentication. In order to ensure the legality of vehicles entering the system, when the vehicle enters the network and requests to communicate with other vehicles, it needs to be authenticated by RSU.

(1) Blockchain network node (RSU) broadcasts its beacon message, including RSU ID information (RID), RSU public key information (RPubK), time information (TS_{RSU}), which is used to verify whether the message has expired, and RSU public key proof (Mer_v), which is part of the Merkle tree in the block header. It can be used to verify the existence of the RSU public key

(2) After receiving the RSU broadcast message, the vehicle uses the on-board unit to calculate Mer_v to verify the correctness of the RSU public key. After the successful verification, the vehicle ID (UID) and public key information (UPubK) are generated by the hash function to maintain the integrity of the information. To ensure efficient operation of the system, this calculation process can be done offline or when the processor is idle

$$H_v = \mathrm{SHA}(\mathrm{UID}, \mathrm{UPubK}). \tag{2}$$

In order to ensure the authenticity of the message source, the vehicle uses its private key to sign the generated summary information. To enable RSU to verify the legitimacy of messages and avoid replay attacks, the vehicle needs to send clock information TS_v to the RSU node.

$$\begin{aligned} \mathrm{Sig}_v &= \mathrm{Sig}_{\mathrm{UPriK}}(H_v) = \mathrm{Sig}_{\mathrm{UPriK}}(\mathrm{SHA}(\mathrm{UID}, \mathrm{UPubK}))\#, \\ \mathrm{Msg}_v &= (\mathrm{UID}, \mathrm{UPubK}, \mathrm{TS}_v, H_v, \mathrm{Sig}_v)\#. \end{aligned} \tag{3}$$

(3) After receiving the Msg_v from the vehicle, the RSU checks its own clock information T_{RSU}. Check whether $T_{RSU} - TS_v < THR_t$ is established; if the time is within the preset range, the communication is allowed to continue. Otherwise, the data packet will be dropped. After verification, the system uses the same algorithm to calculate H_v'. Compare H_v' and H_v. If it matches, it can prove that the public key information has not been modified. If the matching fails, the message will be discarded

$$H_{v'} = \mathrm{SHA}(\mathrm{UID}, \mathrm{UPubK})\#. \tag{4}$$

After verifying the integrity, the RSU extracts the public key information (UPubK) queries whether it exists in the blockchain. If the public key exists, it can be used to verify the signature information. Otherwise, the packet is discarded. It is assumed that the user's private key will not be disclosed to other nodes; so when the signature verification is successful, the RSU can confirm the current communication vehicle legal.

$$\mathrm{Ver}_v = \mathrm{Ver}_{\mathrm{UPubK}}(H_v, \mathrm{Sig}_{\mathrm{UPriK}}(H_v))\#. \tag{5}$$

(4) In order to prevent malicious nodes from impersonating RSU for communication, vehicles also need to verify the authenticity of RSU nodes. RSU will assign an authentication status code (ASC) and an authentication validity period (AVP) to vehicles that have successfully authenticated. RSU sends RID, ASC, AVP, and timestamp information (T_{Rsu}') to the vehicle node with the private key signed

$$\begin{aligned} S &= \{\mathrm{RID}, \mathrm{ASC}, \mathrm{AVP}\}\#, \\ \mathrm{Sig}_{\mathrm{Rsu}} &= \mathrm{Sig}_{\mathrm{RPriK}}(S)\#. \end{aligned} \tag{6}$$

(5) After receiving the RSU information, the vehicle node first checks its own clock information T_v. Calculate $T_v - T_{Rsu}' < THR_t$ to verify whether the timestamp information is within the allowed range. After verification, use the RSU public key received previously to verify the RSU signature message. If the verification is successful, the vehicle saves ASC and AVP. Two-way authentication is completed

$$\mathrm{Ver}_v = \mathrm{Ver}_{\mathrm{UPubK}_v}(S, \mathrm{Sig}_{\mathrm{UPriK}}(S))\#. \tag{7}$$

After two-way authentication, the blockchain will record the authentication status and the validity period. If the vehicle enters the scope of other RSU management within the validity period, cross-domain certification can be quickly completed.

After the certified vehicle enters the communication range of other RSU and receives the RSU broadcast message, the vehicle needs to verify the authenticity of the RSU public key through public key proof in the broadcast message first. After successful verification, the vehicle sends ID information, ASC, and signature information to RSU; RSU determines whether the authentication status code and validity period are legal by querying the blockchain for the authentication information corresponding to the vehicle's public key. If so, RSU uses the private key to sign its own public key information to verify the validity of the RSU to the vehicle to complete the two-way authentication. If not, the RSU sends a reauthentication request to the vehicle.

If the vehicle is not within the communication range of the RSU, the vehicle needs to send a broadcast message requesting communication with the RSU to other vehicles within the communication range of the current vehicle. After receiving the message, vehicles within the RSU communication range forward the request message to the RSU. If the RSU verifies the public key information of the requested vehicle, the RSU broadcasts a beacon message, and the vehicles within the RSU communication range will try to forward the message to the requesting vehicle. The system provides

rewards for successfully forwarded vehicles. The rest of the steps are the same. If the vehicle does not have an RSU within its own communication range and there are no other vehicles, it cannot communicate.

3.3.2. Key Agreement. Since the symmetric cryptosystem has a faster encryption and decryption speed than the asymmetric cryptosystem, the system adopts symmetric cryptography in communication, and it is necessary for both parties to complete the key agreement before transmitting the information.

(1) The vehicle generates a random number A and Initial Key. The seed of the random number is dynamically generated from the identity information and the timestamp information to ensure the randomness of the number. Packages the message M_0 and uses the hash algorithm to generate a message digest

$$\begin{aligned} M_0 &= \{\text{RID, UID, } A\text{, Initial Key}\}\#, \\ H_1 &= \text{SHA}(M_0)\#. \end{aligned} \tag{8}$$

Pack the message M_0, summary result and timestamp information, encrypt it with the verified RSU public key, and send the ciphertext to RSU.

$$C_1 = E_{\text{RPubK}}(\text{RID, UID, } A\text{, Initial Key}, H_1, \text{TS}_v)\#. \tag{9}$$

(2) After receiving the message, RSU uses its private key to decrypt the message to obtain the plaintext content and then uses its own clock information T_{RSU} to check $T_{\text{RSU}} - \text{TS}_v < \text{THR}_t$,

$$M_1 = D_{\text{RPriK}}(C_1) = (\text{RID, UID, } A\text{, Initial Key}, H_1, \text{TS}_v)\#. \tag{10}$$

Verify message integrity. Use the received {RID, UID, A, and Initial Key} to calculate the message digest value $H_1{}'$; checks whether H_1 and $H_1{}'$ are equal; if so, it can be proved that the message has not been tampered with.

$$H_{1'} = \text{SHA}(\text{RID, UID, } A\text{, Initial Key})\#. \tag{11}$$

After verifying the integrity, the RSU extracts the random number A and generates a random number B. Using the random number A, B and Inital Key jointly generate Basic Key and its validity period. Store generated results. Finally, use Basic Key, random number A, random number B, RSU identification (RID), and vehicle identification (UID) to generate the final session key.

$$\begin{aligned} \text{Basic Key} &= \text{Generate}(\text{Initial Key}, A, B)\#, \\ \text{SK} &= \text{Generate}(\text{RID, UID, Basic Key}, A, B)\#. \end{aligned} \tag{12}$$

After generating the session key, the RSU uses the session key to generate the HMAC verification code.

$$\text{HMAC}_1 = \text{HMAC}_{\text{SK}}(\text{RID, Basic Key}, A, B)\#. \tag{13}$$

The RSU encapsulates the RID, A, B, Basic Key, summary information, and timestamp information, encrypts it with the vehicle public key, and sends it to the vehicle.

$$C_2 = E_{\text{UPubK}}\left(\text{RID}, A, B, \text{Basic Key}, B_{\text{key}}\text{VP}, \text{HMAC}_1, \text{TS}_{\text{RSU}}\right)\#. \tag{14}$$

(3) After receiving the message, the vehicle uses its private key to decrypt the message to obtain the plain text content. Check the vehicle clock information T_V, and check $T_V - \text{TS}_{\text{RSU}} < \text{THR}_t$ to ensure that the message was received within the threshold

$$M_2 = D_{\text{UPriK}}(C_2) = (\text{RID}, A, B, \text{Basic Key}, \text{HMAC}_1, \text{TS}_{\text{RSU}})\#. \tag{15}$$

The vehicle node extracts the Basic Key and uses the same algorithm to generate the final session key (SK$'$). Use SK$'$ to generate the verification code.

$$\text{HMAC}_{1'} = \text{HMAC}_{\text{SK}'}(\text{RID, Basic Key}, A, B)\#. \tag{16}$$

Compare $\text{HMAC}_{1'}$ with HMAC_1; if they are equal, it means that the key agreement is completed, store the Basic Key, and its validity period feedback the message that the key agreement is completed to the RSU. Subsequent communications can be encrypted using the session key SK.

(4) RSU packages the key agreement process into transactions and stores them in the blockchain for verification

In order to facilitate the communication parties to update the session key, when the communication parties update the key within the validity period of the Basic Key, they only need to exchange random numbers with each other and use the Basic Key and the new random number to directly generate a new session key. The nodes in the system will periodically check the saved Basic Key data and clear the invalid data.

3.4. Location-Based Service. The Internet of Vehicles system will provide many services based on location information. The traditional IoV system requires users to send personal information and real-time location information to the service provider during driving. It brings a lot of security risks.

In this system, assuming that the service provider believes in the blockchain, users only need to send their public key information and service request information to the service provider when registering for the service. To improve the efficiency of verification and ensure the real-time performance

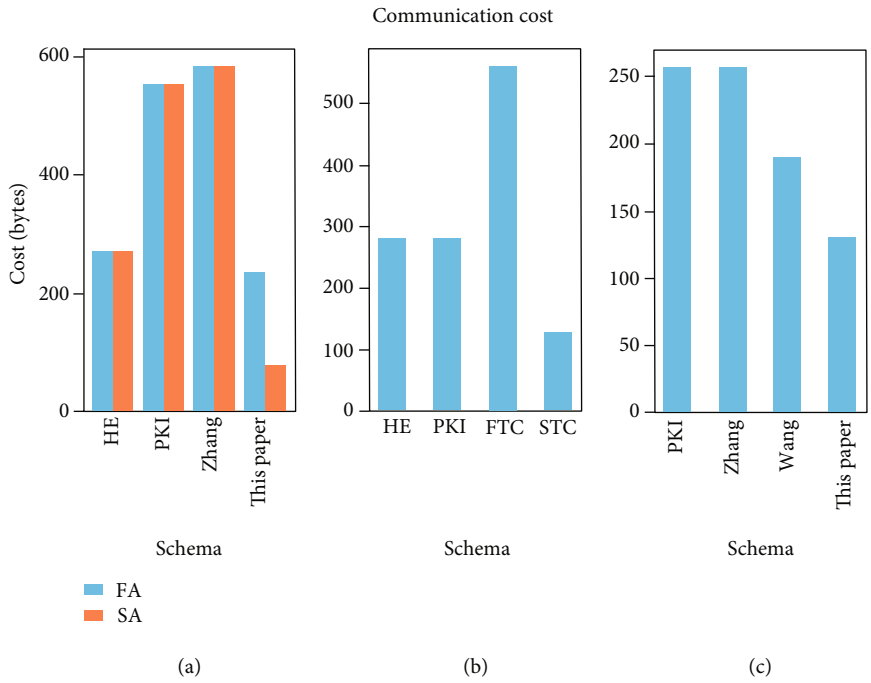


Figure 2: Communication cost. (a) Authentication cost in a different number of vehicles. (b) Key agreement communication cost in a different number of vehicles. (c) Message authentication communication cost in a different number of vehicles.

of zero-delay tolerant services, each service provider has a local cache of public key information in the blockchain. After receiving the message, the service provider will verify the authenticity of the public key from the blockchain cache; if the public key is true, record the user's public key information and feedback the user service application success information. The cache will be updated regularly. The public key can be used as the "pseudonym" of the user's identity information. Assuming that the TA will not disclose user registration information to other organizations, service providers cannot match the collected user location information with user identity information when providing services. When users apply for LBS while driving, they need to send the location information and service application information to the blockchain; the blockchain node verifies whether the location information is true. If it is true, the request information is signed and forwarded to the service provider. If not, the voting mechanism is used to record the malicious behavior of the user and the communication service is terminated.

4. Security Analysis

4.1. TA Security. TA has strong identity authentication. TA processes private information through salt value and hash function and stores private information such as keys in ciphertext. In the worst case, even if the private information is leaked, the attacker cannot know the plaintext information corresponding to the private information.

4.2. Public Key Security. The consensus algorithm adopted by the system is proof of work (Pow). When a malicious user wants to tamper with the user's public key information, it is necessary to control more than 51% of the computing power of the whole network. With the current computing power and the attacker's attack ability, such a large number of nodes cannot be controlled. And with the current computing power, the currently used 256-bit elliptic curve password cannot be cracked. In summary, the price that illegal users want to decipher the system exceeds its deciphering ability, so the public key in the system is safe.

4.3. Session Key Security. Since the session key is generated by the Initial Key and the random number negotiated by both parties through a key generation algorithm and the two parties need to exchange new random numbers when the key is updated, even if the randomness of the Initial Key cannot be guaranteed, the attacker cannot predict the random number generated in each round during the key agreement process, so the session key is safe.

4.4. Antieavesdropping Attack. Eavesdropping attack means that an attacker listens to the communication channel between the two parties to obtain sensitive data that is not encrypted by both parties. In the system proposed in this paper, the vehicle needs to complete the two-way authentication to confirm the user's legitimacy after entering the RSU communication range, and at the same time, the key agreement needs to be conducted after the authentication. All subsequent communications are encrypted and transmitted by using the session key to ensure that the information is not eavesdropped by unauthorized users.

4.5. Antireplay Attacks. Replay attacks refer to malicious nodes republishing legitimate data packets in the previous system to deceive the trust of legitimate vehicles. After receiving the message, the node first verifies whether the difference

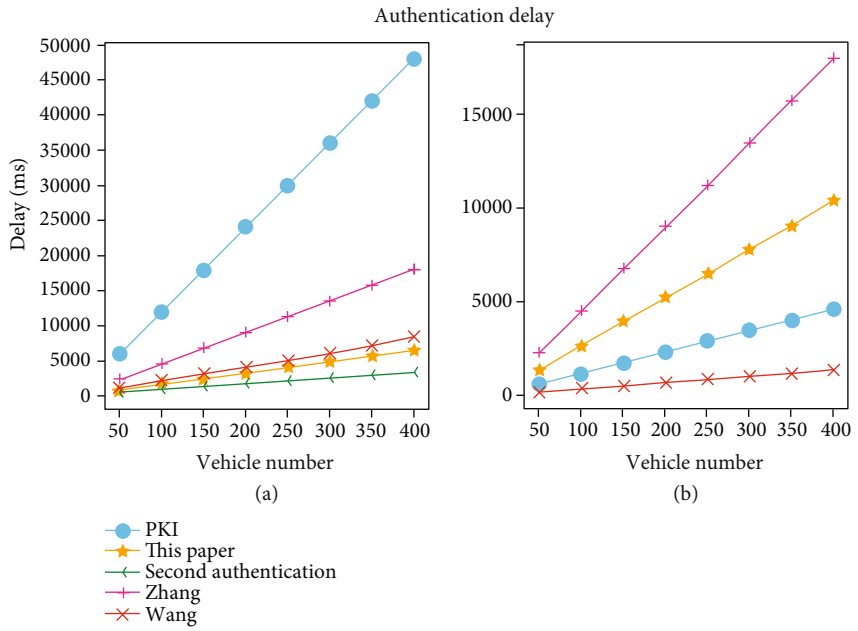


Figure 3: System delay. (a) Two-way authentication delay in a different number of vehicles. (b) Key agreement delay in a different number of vehicles.

Table 1: Storage cost.

Schema	Storage cost (byte)
PKI	126∗(number of certificates)
SA-KMP	37∗XR (the number of RSUs certified)
He	96 + 64 ∗ XV
Wang	128 + 125 ∗ XR
This paper	96 + 11 ∗ XR + 38 ∗ XV (number of key agreement nodes)

between the synchronous clock information and the timestamp is less than the threshold THR. Only when the time difference is within the allowable range is the message considered valid.

4.6. Prevent Sybil Attack. Sybil attack refers to a malicious vehicle disguised as the identity of a legal vehicle and released distorted information. The system needs to complete the two-way authentication before communication, and both of them need to sign with their own private keys; the attacker cannot obtain the private key information, so it cannot impersonate the legitimate vehicle. At the same time, the malicious behavior of vehicles will be recorded by the voting mechanism of the blockchain system.

4.7. Location Information Protection. LBS service providers manage users by using public keys as pseudonyms and verify the public key information through blockchain. The pseudonym information in this authentication mechanism is difficult to forge.

5. Results and Discussion

The experimental equipment is a computer equipped with Intel Core i5-9400f cpu@2.90 Ghz and 16 GB RAM win10 system processor.

5.1. Communication Cost

5.1.1. Two-Way Authentication. According to the experimental situation, the entire key agreement process requires an interaction between the two parties (excluding broadcast messages), and the total required communication load is 237 bytes. If there is reauthentication within the validity period of the authentication code, the communication cost is 79 bytes. According to the IEEE1609.2 trial standard [4], two interactions are required to achieve two-way authentication in the traditional PKI system, totaling 554 bytes. Zhang et al. proposed an authentication system that requires CA participation [15]. The communication process required a round of interaction between the RSU and the vehicle and a round of interaction between the RSU and the CA, and the cost was 584 bytes. The scheme proposed by He requires

RSU to interact with the vehicle once to complete two-way authentication [13], and the cost is 272 bytes.

5.1.2. Key Agreement. Define the ciphertext length standard as 213 bytes, the signature length is 64 bytes, and the message payload as 67 bytes. Traditional PKI technology requires encrypted communication at the key agreement stage, and the cost is 280 bytes. He uses random polynomials to maintain key security [13], and the communication cost is 280 bytes. The system proposed in this paper conducts asymmetric encryption twice, and the cost is 560 bytes, and only 128 bytes are needed for the second agreement process within the Basic Key validity period.

5.1.3. Message Authentication. During the message authentication process, the traditional PKI system needs to transmit certificates and signature information. The total message size is 257 bytes. The message size in the system proposed by Zhang is 257 bytes. The message cost in the scheme proposed by Wang is 190 bytes. This system manages key information based on blockchain, and the message size is 131 bytes (see Figure 2).

5.2. Delay

5.2.1. Authentication Delay. The traditional PKI technology needs to check the certificate revocation table in real time when completing the authentication. In this system, the key information is managed through the blockchain, which reduces the query time of the certificate. The schema proposed by Wang et al. [8] completes two-way authentication through certificates. The schema proposed by Zhang et al. [6] requires the participation of the CA. Before authentication, the vehicle needs to be authenticated with CA. In this paper, the security is enhanced within the allowable range of time delay and has better delay efficiency when the same vehicle is certified for the second time.

5.2.2. Key Agreement Delay. This system uses a random number-based key agreement strategy. There are two asymmetric encryption operations and three key generation operations during the first key agreement process, but only two symmetric encryptions are required during the second process. The scheme proposed by Wang encrypts the message through an authentication certificate, and there is no key agreement process (see Figure 3).

5.3. Storage Cost. The storage cost mainly calculates the cost of OBU. The storage cost of each schema is shown in Table 1.

6. Conclusions

The Internet of Vehicles is an important application in the field of intelligent transportation and has received widespread attention in the academic community. The issue of privacy protection is the top priority of the Internet of Vehicles system. This paper proposes a privacy protection system for the Internet of Vehicles based on blockchain. It introduces secure user registration, efficient key management schema, two-way authentication based on blockchain, a key agreement algorithm based on random numbers, and anonymous communication technology between vehicle and service provider to ensure the confidentiality of security information. The system uses the decentralization feature of the blockchain to prevent the central failure problem. After comparing with the existing scheme, the system proposed in this paper has better performance in communication efficiency and communication security. However, the consensus algorithm based on proof of work currently used by the system will consume computing resources, and the system's key exchange strategy will bring certain storage pressure.

In the future, we plan to design a more efficient consensus algorithm to replace the proof of work algorithm and design a more reasonable communication data structure to reduce system costs and improve system efficiency and user experience. We will continue to study how to reduce the delay required to query public keys to better support real-time communication.

Data Availability

The data used to support the findings of this study are included in this article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work is supported by the National Key R&D Program of China (2018YFB1402704).

References

[1] M. Priyan and G. U. Devi, "A survey on Internet of vehicles: applications, technologies, challenges and opportunities," *International Journal of Advanced Intelligence Paradigms*, vol. 12, no. 1/2, pp. 98–119, 2019.

[2] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A survey on the security of blockchain systems," *Future Generation Computer Systems*, vol. 107, pp. 841–853, 2020.

[3] M. Bartoletti and L. Pompianu, "An empirical analysis of smart contracts: platforms, applications, and design patterns," in *Financial Cryptography and Data Security*, M. Brenner, Ed., vol. 10323 of Lecture Notes in Computer Science, Springer, Cham, Switzerland, 2017.

[4] Intelligent Transportation Systems Committee, *IEEE trial-use standard for wireless access in vehicular environments(-WAVE)-security services for applications and management messages*, IEEE Std, 2006.

[5] W. Jiang, H. Li, G. Xu, M. Wen, G. Dong, and X. Lin, "PTAS: privacy-preserving thin-client authentication scheme in blockchain-based PKI," *Future Generation Computer Systems*, vol. 96, pp. 185–195, 2019.

[6] A. H. Salem, A. Abdel-Hamid, and M. A. el-Nasr, "The case for dynamic key distribution for PKI-based VANETs," *International Journal of Computer Networks & Communications*, vol. 6, no. 1, pp. 61–78, 2014.

[7] X. Y. Guo, *Research on the Privacy Preservation in Security Communication for VANET*, Shenyang Aerospace University, 2017.

[8] W. Qinglong, Q. Rui, F. Na, and D. Zongtao, "An efficient conditional anonymity authentication scheme for VANETs," *Journal of Beijing Jiaotong University*, vol. 43, no. 5, pp. 80–86, 2019.

[9] J. Wei, X. Wang, N. Li, G. Yang, and Y. Mu, "A privacy-preserving fog computing framework for vehicular crowdsensing networks," *IEEE Access*, vol. 6, pp. 43776–43784, 2018.

[10] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," 2008.

[11] N. Lasla, M. Younis, W. Znaidi, and D. B. Arbia, "Efficient distributed admission and revocation using blockchain for cooperative ITS," in *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, Paris, France, February 2018.

[12] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain-based decentralized trust management in vehicular networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1495–1505, 2018.

[13] W. He, *Research on Key Management Mechanism of Internet of Vehicles Based on Blockchain Technology*, Xidian University, 2019.

[14] H. J. D. Lopez, M. Siller, and I. Huerta, "Internet of vehicles: cloud and fog computing approaches," in *2017 IEEE International Conference on Service Operations and Logistics, and Informatics, SOLI*, Bari, Italy, 2017.

[15] J. Zhang, F. Li, R. Li, Y. Li, J. Song, and Q. Zhou, "Research on identity authentication based on elliptic curve encryption algorithm in V2X communication," *Automotive Engineering*, vol. 42, no. 1, pp. 27–32, 2020.