Wireless Communications and Mobile Computing

Learning Methods for Next Generation Cyber-Physical Systems 2021

Lead Guest Editor: Wei Wang
Guest Editors: Keping Yu, Peng Li, and Han Liu

Learning Methods for Next Generation Cyber-Physical Systems 2021

Wireless Communications and Mobile Computing

Learning Methods for Next Generation Cyber-Physical Systems 2021

Lead Guest Editor: Wei Wang
Guest Editors: Keping Yu, Peng Li, and Han Liu

Chief Editor

Associate Editors

Academic Editors

Contents

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9804643, 1 page
https://doi.org/10.1155/2023/9804643

Retraction

Retracted: Marketing Method and System Optimization Based on the Financial Blockchain of the Internet of Things

Wireless Communications and Mobile Computing

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

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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

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

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

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

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

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

References

[1] C. Yan, J. Zhu, Y. Ouyang, and X. Zeng, "Marketing Method and System Optimization Based on the Financial Blockchain of the Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 9354569, 11 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9846860, 1 page
https://doi.org/10.1155/2023/9846860

Retraction

Retracted: Traffic State Detection Based on Multidimensional Data Fusion System of Internet of Things

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] J. Zhang, W. Zhu, X. Wu, and T. Ma, "Traffic State Detection Based on Multidimensional Data Fusion System of Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1374186, 12 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9786547, 1 page
https://doi.org/10.1155/2023/9786547

Retraction

Retracted: Optimization of Intelligent Management and Monitoring System of Sports Training Hall Based on Internet of Things

Wireless Communications and Mobile Computing

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

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. Qian, "Optimization of Intelligent Management and Monitoring System of Sports Training Hall Based on Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1465748, 11 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9896460, 1 page
https://doi.org/10.1155/2023/9896460

Retraction

Retracted: Mathematical Modeling and Simulation of Wireless Sensor Network Coverage Problem

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] J. Guo, "Mathematical Modeling and Simulation of Wireless Sensor Network Coverage Problem," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5537004, 12 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9849025, 1 page
https://doi.org/10.1155/2023/9849025

Retraction

Retracted: Security Control Technology and Simulation of Network News Communication under the Environment of Internet of Things

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] X. Zhou, J. Wang, and X. Zhou, "Security Control Technology and Simulation of Network News Communication under the Environment of Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 2730916, 10 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9861829, 1 page
https://doi.org/10.1155/2023/9861829

Retraction

Retracted: Combined GNSS/SLAM-Based High Accuracy Indoor Dynamic Target Localization Method

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] X. Zhang, S. Zhang, Q. Li, X. Wang, and S. Sun, "Combined GNSS/SLAM-Based High Accuracy Indoor Dynamic Target Localization Method," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 8380869, 12 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 7964409, 9 pages
https://doi.org/10.1155/2021/7964409

Research Article

Structural Balance of Social Internet of Things Networks with Ambiguous Relationships

Hongfei Zhang[1], **Li Zhu**[1], **Haifeng Du**[2], **Li Zhang**[3], **Kaiqi Zhang**[4], **Yutian Yan**[2] **and Chaobo Wang**[5]

[1]*School of Software Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi Province 710049, China*
[2]*Center for Administration and Complexity Science, Xi'an Jiaotong University, Xi'an, Shaanxi Province 710049, China*
[3]*Xi'an Physical Education University, Xi'an, Shaanxi 710068, China*
[4]*School of Economics and Management, Chang'An University, Xi'an, Shaanxi 710054, China*
[5]*Shaanxi Labor College, Xi'an, Shaanxi Province 716000, China*

Correspondence should be addressed to Li Zhu; zhuli@xjtu.edu.cn

Received 29 April 2021; Accepted 25 June 2021; Published 1 August 2021

Academic Editor: Wei Wang

In the social Internet of Things, social networks can be built among smart objects or between smart objects and people just like human beings. One of the factors that determines the effect and efficiency of service matching in SIoT is the structure of social networks. In this paper, we exploit the theory of structural balance in signed networks to optimize SIoT network structures to provide a friendly and stable environment as well as a solid foundation for service matching of social Internet of Things. Next, besides being friends or enemies, which are traditional relationships of structural balance or structural changes in signed networks, in our research, we introduce the ambiguous relationship which is the certain state between the hostile and the friendly status and we discuss the meaning and significance of ambiguous relationship in dynamical changes of structural balance in SIoT networks. Based on previous studies, we apply an enhanced objective function and a modified approach concerning the ambiguous relationship towards the dynamical change process. Experiments show that our approach is more effective and efficient than former studies in optimizing dynamical evolution of structural balance in signed networks of SIoT.

1. Introduction

Social Internet of Things (SIoT) originates from the intention of integrating and applying the concept of social network to the solution of Internet of Things. Since then, Internet of Things has evolved from primitive interaction between smart entities to integrated intelligent cross-correlation with human elements. With the participation of people, the essence of Internet of Things is combined with the substance of social networks.

In IoT or SIoT networks, one of the most important issues is how to match people with smart devices or the services provided by them perfectly. However, the collaboration among people and things in social Internet of Things contributes to greater complexity than either social networks or Internet of Things alone. In addition, the heterogeneous nature of Internet of Things and large-scale context data also make computing and matching more sophisticated [1]. Due to the attributes of free alternation and open network circumstances, when we regard people and smart things in SIoT networks as nodes, users will face a lot of risks in choosing the nodes they want to communicate with.

Thus, a trust mechanism is proposed to solve the dishonesty such as fraud, and trust itself has become a kind of judgment basis to evaluate whether the interactive node is credible. Nevertheless, with the development of social Internet of Things, the number of nodes connected to the network increases exponentially, which leads to a huge search space [2]. As a result, traditional service matching methods based on computing the trust value and matching degree between

each pair of nodes will lead to a more complicated calculation and high cost.

Because of these difficult challenges, we consider new solutions from a higher dimension. That is, we can improve the efficiency and environmental friendliness of people-things or people-services matching by adjusting the overall state of the network.

Obviously, the hostile and unbalanced network environments will damage the interests of nodes within the network. Apart from this, network balance is a key issue in traditional social networks [3] and Internet of Things in which trust can be sustained and the environment can be stable. Then, it is necessary to develop a balanced environment evaluation mechanism to suppress these unfriendly or uncoordinated behaviors and ensure a more stable and secure network environment. Thus, the balance problem in the social Internet of Things is worthy of our serious study.

Because things are essentially operated by human beings, the concept of relationship in social networks can be integrated into the conceptual structure of Internet of Things. Trust is the foundation of interaction between nodes, and it exists in any pair of nodes reflecting the reputation and behavior of nodes in the network. People's interaction is based on trust. Trust in the network can be equivalent to whether it is a relationship of friends or not. As we all know, friendship represents trust and vice versa. There are large numbers of locally interacting agents connected by social and economic relationships which include like or dislike and respect or disrespect for people in human societies [4]. Some social relations are assigned as positive or negative [5]. The classic structural balance theory has always been a focus for social psychologists and sociologists [3]. It is a fundamental social process and the structure of affective relations of social actors towards one another [6].

We use the theorems of structural balance as the foundation for all kinds of signed social networks to firstly construct the partitions through some optimized criterion functions towards the balanced state; secondly, we measure the extent of how the structure is balanced; and thirdly, we find the factors that can shift the network to balanced states [3]. The points or vertices in social networks denote people, and the lines or edges represent acquaintance [7]. The interplay which consists of positive and negative interactions defines the smooth functioning of society [8].

Due to the fact that most of the signed networks in the real world are unbalanced, amounts of approaches are proposed to ascertain the level or degree of unbalance in signed networks; Marvel et al. introduced a numerical model to indicate the potential function [9] in which the negative products of triangles are summed up and divided by the quantity of trigons. In [8], the subsection of unbalanced triangles is related to the feature values of the adjacency matrix of the networks, and a spectral algorithm is provided to compute social balance. With that efficient algorithm above, the top-n eigenvalues instead of all of them are used to calculate the subsection of the disequilibrium trilateral. In [10], the lopsidedness of the meshwork is surveyed by the magnitude of the rims in the disequilibrium trigons, cycles, or networks. An energy function is presented in [11] to reckon overall equilibration in the signed meshwork. The lowest value of the energy function equals to the minimal quantity of negative signs making the network unbalanced, that is to say, the meshwork is trim when the value of energy function is zeroth. Finally, amid [12], an effectual heuristic approach appeared to improve the energy function, in which the obvious chaos in the network is deleted through the metamorphosis; meanwhile, the scalar of nonpositive edges is reduced correspondingly.

We also focus on the approach a network evolves to the balanced state. Among dynamical evolvement of signed networks, the positive edges transform to the negative ones or vice versa to make unbalanced ones balanced. In [13], Antal et al. produced some discontinuous models for the dynamic state in which the sign of the edge remains as it is unless it increases the number of balanced triangles. But there are the processes called "jammed states," in which the model may be trapped in unbalanced local minimums [13]. Recently, Marvel et al. [14], Traag et al. [15], and Summers and Shames [16] have studied continuous-time models. For instance, the model in [14] converges to the balanced state in finite time. However, the conclusion does not cover cases for nonsymmetric initial conditions. Therefore, in [15], a general model is proposed, which leads to the state of social balance. Besides that, the active external influence in dynamical models is studied in [16].

Social networks can evolve in convoluted ways. Friends can become enemies and vice versa [17]. Thus, all the methods above originated from this hypothesis.

However, we have noted that there are three flaws with them in SIoT networks.

Firstly, in the real world social network, the ties between individuals are complicated, and it is quite probable to imagine that two strangers are neither friends nor enemies, so we can get the conclusion definitely that not being friends does not lead to being an enemy inevitably and vice versa. Especially in the social Internet of Things, whether the two sides are really friends does not determine whether they can share the understanding, concept, and use of intelligent devices or services.

Secondly, in traditional social networks, people may encounter completely unfamiliar users who have no previous interaction experience with them before [18]. The existence of these users threatens the feeling of safety to some extent, but from the perspective of social Internet of Things, it also expands the scope of interaction of people, so that people in the same field or with the same interests establish more and stronger trust relations. For example, the novel corona virus pneumonia spreads throughout the world during this period and people in most countries of the world have been affected by city bans and home quarantine orders. Economic growth may decline and cultural exchanges may be suspended, but the use of the Internet, social media, and intelligent objects has not subsided yet has grown strongly. Massive data obtained from the selection, evaluation, and application of various intelligent objects by people who are related to various attributes contribute to the stability and connectivity of the social network. However, in the real world, these people do not have to be friends. It is even possible that they will

not meet for the rest of their lives. Of course, they are not hostile at all. Therefore, this absolute binary-like, strict distinction between friends and enemies does not make much sense in SIoT.

Finally, even if traditional methods are capable of shifting unbalanced signed networks into balanced states theoretically, but in practice, it is not convenient in realities. Sometimes, it is even impossible to convert the signs of edges in signed networks to their opposite directions.

The innovation of this paper is as follows:

(1) Originating from the flaws above, we introduce the third kind of connection in SIoT networks which is called the ambiguous relationship

(2) We suggest that the forms of structural balance of SIoT networks stem from how the ambiguous relationship shifts the state of balance in the network

(3) By applying the concept of structural balance in signed networks into SIoT networks, through establishing the dynamical evolving mechanism of structural balance in SIoT networks which is more complete and more efficient than that in the mode of traditional social networks, we provide a more amicable and stable environment in order to build a solid foundation for service matching of social Internet of Things for further studies

2. Background

2.1. Trust in SIoT. Each node has its own credibility just as everyone has his reputation, and here comes the concept of trust. The existence of trust is the basis for the delegator to judge whether the entity which provides the information or service is trustworthy. In social networks, the trust degree of an object to another entity providing a specific service is calculated according to the behavior of the service provider in a specific stage and in a designated environment related to the service [19].

In the network of social Internet of Things, each node plays the role of trustor or trustee and establishes social relations independently, so as to carry out service-based interaction. The trustor generates and delegates the task, and the trustee receives the task and provides services for the successful completion of it. In detail, the trustor is a conscious subject with its own goals, needs, and attitudes towards other objects and their behaviors. Based on the trust of other objects, the cognition of the specific environment, and the prediction of the results, the trustor can generate and delegate the task as well as evaluate the related results. The trustee is a cognitive object or device equipped with some information or abilities that can produce a certain effect on the results of its behavior. In a sense, the trustee is not under the direct control of the trustor. The actions of both the trustor and the trustee must be consistent with their trust relationship. Both sides establish the cooperation relationship based on trust and complete the whole process from task initialization to service acceptance throughout with the element of trust.

However, how to find the most suitable trustee to perform the corresponding task is the key problem when the trustor assigns the task. Traditionally, trust degree or trust value is generally used as a measure to select a suitable trustee. Trust values are usually normalized to [0,1], and it concludes two parts called capability and honesty which determine two fundamental aspects of trustworthiness of both the trustor and the trustee. The trustor can use these two parts to evaluate the qualification of trustee and judge whether the cooperation can be obtained. For the trustee, capability refers to the hardware conditions such as the number of related devices equipped, the performance of each device, and the involved information or intelligence, while for the trustor, it refers to the ability of the subject to use the object or device and to provide the resources available for the trustee. A competent trustor (or trustee) may also deceive the other side and behave badly or even maliciously, and an incompetent candidate may also claim to be qualified for the job. Therefore, it is necessary to have an honest index to evaluate the authenticity of the trustor and the trustee, as well as the importance they attach to the cooperative relationship. To summarize, we need to combine the capability and honesty element to constitute the trust value.

There is another important issue that the trust value is task related. When the same pair of trustor and trustee handles different tasks, there will be different trust values. The most important reason for the above conclusion is that each task has its unique attributes and specific targets only belonging to this service. To achieve the task goal, different equipment and different activities will be used.

In SIoT, trust is a process in which a trustor entrusts a task with a certain degree of risk to a trustee equipped with appropriate devices or abilities in a specified context based on the positive expectation for credibility. The trustee acts and evaluates the trustor, while the trustor also assesses the other side during the same period. Finally, results of the service and the trust evaluation of the whole process will in turn affect the attitudes of both sides.

SIoT ensures the scalability of the network and can effectively help the discovery of services [20]. The trust management of nodes in the SIoT network can be achieved by the interaction experience of smart objects. Although trust seems to have nothing to do with intelligence, the realization of intelligence must be realized through a strict trust management mechanism.

As we discussed earlier in the introduction section, the level of trust in the SIoT network is based on the propagation of trust. In addition to the length of trust propagation route, it is also affected by different aggregation strategies of trust values of multiple trust propagation paths. Although evaluating trust values based on all trust propagation paths provides more accurate trust inference values, it is excessively time-consuming in large-scale SIoT or social networks.

Besides, trust can be divided into direct trust and indirect trust [21]. To evaluate the entity trust in the SIoT network, both direct trust and indirect trust should be considered comprehensively. The direct trust is stored locally by the trustor, and the indirect trust value can be obtained from the third-party node. Therefore, when we need to obtain an

indirect trust value, it is necessary to query the relative information of the IoT service provider, and this process will cause network congestion.

In the process of interaction, nodes often encounter unfamiliar ones. In this case, the cooperation needs to be introduced by the middleman, i.e., the intermediate node who plays the role of trust messenger. Therefore, the transitivity of trust makes more cooperation possible in SIoT networks. As it happens, in the balanced network state, friend relationships are also transitive, because if one person's friend is the mate of another, they will all become compadres.

To sum up, there are many problems and limitations to optimize the network structure performance of social Internet of Things purely through a trust management system. Therefore, we introduce the idea of a network structural balance with the relations of friends and enemies to utilize the essence and significance of the trust management mechanism to provide a more favorable network environment for matching people with smart devices or the services through improving optimization performance and efficiency of dynamical evolving of structural balance in SIoT networks.

2.2. Structural Balance. From the structural balance theory of Heider, there are four situations in a signed network which conclude three nodes and three edges between them, as shown in Figure 1. Each network in the triangle form must have 1 or 3 positive edges to be balanced. The state of Figures 1(a) and 1(c) is balanced, while the situation of Figures 1(b) and 1(d) is unbalanced. Generously speaking, in a balanced basic network composed of three nodes, there must be even negative edges, while in an unbalanced basic network, there are odd negative edges.

Based on the classic balance theory of Heider, the state of a signed network is balanced when all the triangular basic networks of nodes within it are balanced. Another definition is that the network is structural balanced only if it does not contain cycles of negative product.

Furthermore, Iacono et al. demonstrated that the definition of the triangle- and cycle-based theory is equivalent [12]. That is, the structure of a signed network is in equilibrium, when any edges in it is positive, or all the nodes can be partitioned into two cliques where nodes in the identical clique can be seen as friends. Meanwhile, nodes come from different cliques can be tagged as enemies. We can briefly describe this definition in Figure 2.

Based on our theory, besides hostile or friendly relationships in SIoT networks, there is a mild or neutral mode, and we call it the ambiguous relationship. It is between the hostile and the kindness situation, and it has the opportunity to be altered towards any of the two standard bipolar states when external forces or elements are exerted. So far, the so-called ambiguous relationship has not caught too much attention in this field of research. However, an ambiguous relationship counts much in SIoT, making it possible to predict temporary trust between unfamiliar nodes without interaction experience. If an ambiguous relationship cannot be defined, strangers will never be able to evaluate whether the other side is trustworthy at least in information, experience or advice, and strangers will always remain unfamiliar, and they cannot form new interactions. This will lead to low connectivity of the SIoT network, and tasks and services among the network cannot be handled properly and effectively. Therefore, ambiguous relations are fundamental for social Internet of Things.

Furthermore, we can classify the conversion process into two kinds: friendly relationship to none friendly (hostile) and vice versa. Each kind mentioned above may spend too much cost such as resources, time, energy, or power in reality. Although nothing is impossible in this world, in the real world, anything could be done only with the right prize.

Moreover, according to [22], when dealing with the relationship between the owners of IoT nodes sharing similar social interests, there are three kinds of social relations: friendship, social association, and interest group communication, which are used to measure social similarity and to filter trust feedback based on them. This theorization proves that in social Internet of Things networks, in addition to friend relationships, there are two other kinds of relationships between nodes that can be exploited. Besides, if we suspect whether the new relationship dimension exists or not, in some deeper level, the comprehensive meaning of the combination of the two other relations can correspond to the ambiguous relationship we proposed in this paper.

Therefore, it is improper to sign the relation between two individuals just as a friendly or a hostile relationship and deny the existence of the third connection which is neither friendly nor hostile in reality. Consistent with the traditional theory that "1" indicates friendly relationship and "-1" denotes hostile connection, we introduce "0" to represent ambiguous relationship, which makes our model match the real world system more appropriately.

However, according to the hypothesis of Heider, structural balance makes sense only with fully connected networks. We suggest that the "0" edge represents the edge which cannot be defined as "1" or "-1" state, rather than having no edge at all. Once we appreciate the existence of the "0" relationship, we see how this mechanism can establish structural balance other than simply shifting "1" to "-1" or vice versa.

In this essay, with the introduction of the "0" edge, we transform the unbalanced signed networks of SIoT to balanced states only by one step, which makes our method more economical, efficient, and realistic. The "0" edge signifies the ambiguous relationship of nodes in SIoT networks that is neither friendly nor hostile but falls in between the two extreme relations.

3. Structural Balance Evolving with Ambiguous Relationships

3.1. Energy Function. The generally accepted relationship types in social Internet of Things can be divided into three categories: human-object relationship, object-object relationship, and human-human relationship. In the human-object relationship, users search for services provided by smart objects. In the next step, a new search mode will naturally arise, that is, in order to meet the needs of human users, intelligent objects will actively search for other objects that can

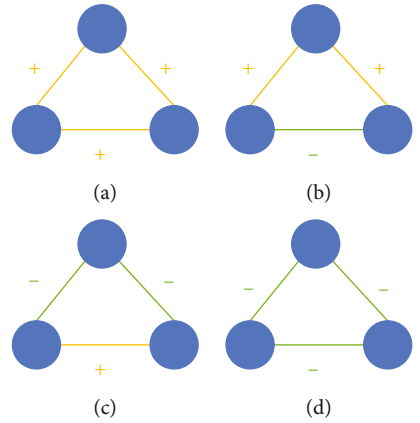


Figure 1: Four modes of basic signed networks.

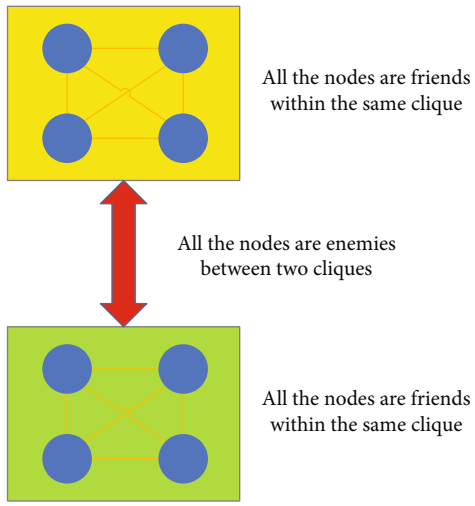


Figure 2: Same clique: friends. Different cliques: enemies.

provide relevant composite services. From this perspective, all intelligent objects in social Internet of Things can cooperate with other people or things and establish social relationships, just as human groups do. Therefore, the services and information provided by each intelligent object can be automatically processed by any node in the whole social Internet of Things system.

In the process of node interaction, the two sides will judge whether the other party is trustworthy or the specific value of its trustworthiness, respectively, according to their own experience and the amount of resource reserves. The availability of empirical data may affect the trustworthiness of a node persistently. Moreover, even if they have experienced the same events, there will be differences in all kinds of benefits obtained by both nodes, which leads to a dissimilar evaluation value of the two sides. In other words, trust has the natural attribute of asymmetry.

Next, trust can be divided into two types: identity trust and behavior trust. Identity trust, also known as static trust, at the technical level, refers to the application of encryption, authentication, and other traditional security mechanisms to authorize and verify the validity of entities or nodes in the network. From the perspective of social networks, identity trust is the historical performance of the subject, which makes it have the corresponding word-of-mouth figure in a certain field. High credibility is the performance of their own identity trust. Behavioral trust, also known as dynamic trust, refers to persistent adjustment of trust relationship based on past interaction experiences. This is the cumulative result of the historical evaluation given by the service user according to the behavior of the service provider. In the trust relationship, it includes the trustor who enjoys the service and the trustee who performs the service relative operation. After the end of the service, both sides will evaluate the other party based on the implementation process and the final results of this cooperation. The evaluation situation will in turn affect the trust value between the two sides and the external credibility of both entities. Therefore, the degree of trust is not static and fixed.

The result of the trustee's action can be described in further detail as it includes not only the expected harvest but also a series of negative effects and the price that must be paid regardless of success or failure. Only by comprehensively calculating the cost of task execution, benefits brought by the action, and the completion degree of expected results can we determine whether the service results are positive or negative.

In addition, the mutual trust between two nodes of the interaction and the trustworthiness of a node in the whole group changes dynamically. If a pair of nodes cooperates with each other frequently, the trust evaluation value of each node will increase gradually. But if the two sides lack interaction or contact and have stopped cooperation for a period of time for some reasons, the trust value will decrease with the passage of time. One of the main reasons is that the judgment data of each node is still in the memory of the past, and there are no real-time source data to help judge the credibility of both nodes, resulting in the decrease of the trust value.

Thus, we can find out that the energy function to calculate the trust value between nodes must be an accurate and complex model with typical and prominent dynamic characteristics.

However, in social Internet of Things, because we do not need to strictly calculate the degree of trust and distinguish

the level of trust, so the excessive accurate and complex trust calculation and management model has become a great burden for the service matching process of social Internet of Things.

Therefore, in SIoT, we use the network of friends and hostility relation of the network structural balance theory to replace the traditional trust relationship network. The relational network between friends and enemies can clearly express the meaning of trust and distrust, and at the same time, it does not focus on detailed trust values. Therefore, the application of the network balance theory in this field is equivalent to the directional abstraction and fuzzification of the trust relationship network, which retains the elements we need and avoids the part of the trust network that is not suitable for social IoT service matching.

In [11], an energy function is presented to calculate the degree of overall balance towards signed networks. By assigning "+1" or "-1" to all the nodes in social IoT, we could compute global balance through the energy function as follows:

$$H(s) = \sum_{i,j} \frac{1 - E_{ij} s_i s_j}{2}, \tag{1}$$

where $H(s)$ equals to the summation of the number of edges leading to the unbalanced state or the unbalanced energy value in the whole SIoT network and it traverses all the couples of nodes in adjacency. $E_{ij} \in \{+1 \| -1\}$ presents the edge between nodes s_i and s_j. $S = [s_1, \cdots \cdots s_n]^{\mathrm{T}}$ $(s_i \in \{+1 \| -1\})$ signifies the nodes in the network. When E_{ij} presents friendship, if nodes s_i and s_j share the same sign (positive or negative), every corresponding determined value of (1) equals zero, and it means that there is no unbalanced or discordant energy between nodes of SIoT in this area. However, when E_{ij} presents hostility, every term in (1) is "+1," and the degree of unbalance increases. Likewise, when the signs of s_i and s_j seem opposite to each other, every corresponding determined value is "0" under the condition that E_{ij} presents rivalry and the value is "+1" otherwise. In the end, if the summation value of (1) is still "0," the whole network of SIoT correlated can be seen as balanced.

Nonetheless, with the introduction of our method with "0" value edge, problems arise. The fatal defect of (1) lies in the fact that when one edge which causes the unbalance of the network is fuzzily processed, the value of E_{ij} equals to 0, but at the same time, the term in $H(s)$ comes to 0.5, and this number conflicts with the theory that $H(s)$ suggests that the number of edges which arouse the whole system becomes unbalanced in that there are no half edges amid the SIoT network background.

Therefore, in order to present the degree of overall balance of SIoT networks, we can take advantage of the formula in [23] to upgrade the optimization approach to

$$H(s) = \sum_{i,j} \frac{|E_{ij}| - E_{ij} s_i s_j}{2}. \tag{2}$$

Table 1: Detailed data related to the experiment of each signed network.

Network for experiments	Number of nodes	Number of edges	Number of positive edges	Number of negative edges
SPPN	10	45	18	27
EGN	330	779	515	264
MN	697	1425	947	478

In comparison, when the fuzzy processing is proposed, $E_{ij} = 0$, and meanwhile, $|E_{ij}| = 0$, the term in the equation is "0" too. Under this circumstance, value "0" represents that since the edge stands for ambiguous or neutral relationship now, there is nothing it can do to make the whole SIoT network unbalanced.

3.2. Quantifying in Dynamical Evolution. We want to locate the set of edges containing the edges that should be made ambiguous to make the SIoT network balanced in dynamical evolution, and we want to keep the number of these edges to a minimum value. So a standard is necessary to ensure that the signed SIoT network can evolve to balanced states.

Equation (2) is adopted as the criterion since we can check whether the SIoT network is balanced by computing the energy function to find out if the final value equals zero. In addition, a parameter $\lambda \in [0, 1]$ is introduced into (2) to show that different edges with related signs are processed with different probabilities [24]. Via shifting the numeric value of λ, varying amount of edges involved in neutralization can be obtained and processed.

The formula of the optimization problem of dynamical evolving of structural balance in SIoT networks is shown below:

$$\begin{aligned} \min \ & F(x) = \lambda x_+ + (1 - \lambda) x_-, \\ \text{s.t.} & H(s) = 0. \end{aligned} \tag{3}$$

According to the previous definition, λ is the parameter to adjust the probability of neutralization towards edge signs, while x_+ is the number of positive signs involved in the ambiguous state of the sign direction and x_- indicates the amount of negative ones. $H(s)$ in (3) and (2) are identical.

3.3. Objective Function and Representation. To optimize (3), an algorithm which comes from [24] is utilized in this section. The algorithm includes the following steps: firstly, the initial population is generated by initialization function, and then, the corresponding population is selected from the parent and child population. The next step is mutation, cross-selection, and population regeneration. Finally, the termination condition function is used to determine when the algorithm can be terminated. In this thesis, the evolutionary algorithm based on MAs (memetic algorithms) is modified and reused to change unbalanced SIoT networks into balanced states dynamically with an enhanced mode.

Table 2: Results of 10 runs of variant values of λ and μ. "NPP" denotes the number of the positive edges being neutralized while "NNP" is the number of negative ones. "VA" denotes the value of the object function and "EN" is the result of the energy function.

Networks	Index	$\mu = 0.1$		$\mu = 1$		$\mu = 5$		$\mu = 10$		$\mu = 100$	
		$\lambda = 0.3$	$\lambda = 0.6$	$\lambda = 0.3$	$\lambda = 0.6$	$\lambda = 0.3$	$\lambda = 0.6$	$\lambda = 0.3$	$\lambda = 0.6$	$\lambda = 0.3$	$\lambda = 0.6$
SPP	NPP	0	0	0	0	0	0	0	0	0	0
	NNP	0	0	2	2	2	2	2	2	2	2
	VA	0.2	0.2	1.4	0.8	1.4	0.8	1.4	0.8	1.4	0.8
	EN	2	2	0	0	0	0	0	0	0	0
EGN	NPP	140.8	134.2	158.5	163.3	169.9	171.7	151.7	169.1	170.2	170.7
	NNP	59	75.1	136.4	136.8	133.9	135.6	144.3	135.4	135.4	133.1
	VA	105.8	132	143	152.7	144.7	156.5	146.5	155.6	145.8	155.7
	EN	223	214.9	0	0	0	0	0	0	0	0
MN	NPP	319.7	305.9	252	267	281.8	281.1	284.7	277.8	277.8	286.5
	NNP	137.7	161.8	270.7	255.9	260.8	257.5	260.1	260.7	261.7	249.6
	VA	232.4	287.3	265	264.2	267.1	271.7	267.5	271	266.5	271.7
	EN	401.2	390.6	0	0	0	0	0	0	0	0

The object is to minimize the value of the objective function in (3). We use the function

$$\min f(x) = \lambda x_+ + (1 - \lambda)x_- + \mu H(s). \quad (4)$$

λ is the weighting parameter, and μ works as the penalizing factor for forming the constraint.

In this algorithm, every individual is encoded as a string: $X = \{X1, X2, X3\}$, where X1 denotes positive edges, X2 represents negative edges, and X3 indicates notes. Their assignment all lies in $\{-1,+1\}^N$ in which N stands for the number of edges and notes. When $Xi \in \{X3\}$, Xi represents the sign of the ith node, and when $Xi \in \{X1\}$ or $\{X2\}$, $Xi = -1$ indicates that the direction of the sign for the edge is neutralized while $Xi = 1$ means that the direction of the sign keeps the same state as before. Node n1 is one of the neighbors of n2 when they are linked by a positive edge, and the sign of every gene in the individual is designated "+1" or "-1" with a random chance.

The initialization process and the operating measure which consist of the crossover and mutation operation can be found in [24].

3.4. Modified Local Search Procedure. Local search procedures including hill climbing [25] and simulated annealing [26] are usually combined with evolutionary algorithms to prevent the algorithm from falling into the local-best solution. In this procedure, every gene is searched by selecting and neutralizing the sign randomly. When the changing leads to the decrease of the value in (3), the sign of the gene is neutralized permanently. The algorithm is described in [24] in which we only process the best samples in the group based on the modified process after the crossover conduct and mutation operation.

4. Experiments

We test our approach on three real signed network datasets, and the experimental outcome demonstrates that our method has a better effect in optimization towards the dynamical evolving of edge signs than the algorithm from [24].

4.1. Experimental Setup. Our model and approaches are realized through MATLAB. We conduct the experiments on a 3.8 GHz computer with 8 GB RAM. The OS on the computer is Windows 10 of Microsoft. The parameters used in the experiment are identical as [24] so that we can make sure that the contrast is accurate and persuasive. The parameters include the number of iterations, sizes of population, mating pool, the tournament, and probabilities of the crossover and mutation operations. They all are fundamental parameters in memetic algorithms.

4.2. Experimental Result. The effect of the model in our paper is examined on three real signed network datasets: Slovene Parliamentary Party Network (SPPN), EGFR (EGN), and Macrophage (MN). These three datasets are all used as experimental networks in [23–25] which have strong correlation with the model and experiment in this paper, so they are more suitable for the experimental scenario. The detailed data related to the experiment of each signed network are shown in Table 1.

Slovene Parliamentary Party Network (SPPN): a group of experts on parliamentary activities created the SPPN dataset in 1994 [27]. The network presents the relationship between different network subsets. Nodes represent the subnetworks, while edges represent the hostile or friendly relationship between subnetworks.

EGFR (EGN): created by [28] with the epidermal growth factor receptor pathway. Edges can be considered as transcription factors. The positive and negative of edges represent the positive and negative of control in the physical sense, respectively.

Macrophage (MN): obtained by [26], and the data came from the molecular reciprocity map focusing on a macrophage. Edges are used to describe the interaction of macrophage molecules in the field of biology. The positive edge

shows activation action, while the negative edge indicates translocation or segregation operation.

There are two variable parameters of λ and μ in order to determine the variation. First, we explore the impact of μ in the goal function. In Table 2, we calculate the average values of the algorithm for $\mu = 0.1$, 1, 5, 10, and 100.

In Table 2, we find that when $\mu = 0.1$, the energy function $H(s) = 0$ can never be achieved on all the networks. The unbalanced network cannot evolve to the balanced one, i.e., the results of the objective function (3) do not satisfy what we need in dynamical changes of structural balance. And for $\lambda = 0.3$ and $\lambda = 0.6$, our algorithm can always get the minimum value of the function on three signed networks when $\mu = 1$. We run the algorithm 10 times for different parameter values of λ for each network.

In our approach, every solution represents a kind of balanced structure of the signed network evolved from the unbalanced state. According to the approach and results in [24], we find that although the processes in the two systems are all alike except the fact that we exclusively change the target sign into "0" instead of shifting it to "+1" or "-1," the effect and result of the dynamical evolution looks better in our data, which means that there are less edges to be changed under the same circumstance compared with that in [24].

5. Conclusion

In SIoT, social networks can be established among intelligent objects or between intelligent objects and human beings. How to get the most appropriate service efficiently and accurately is a fundamental problem for intelligent objects, applications, and users in SIoT networks. Among many factors that can determine the effect and efficiency of service matching in SIoT, the structure attribute of SIoT networks is seldom considered yet. While structural balance is a significant issue in traditional studies of signed networks, in this paper, we exploit the theory of structural balance to optimize SIoT network structures to provide a friendly and stable environment for the service matching of social Internet of Things. Then, based on the nature and features of SIoT, we introduce the ambiguous relationship which is the certain state between the hostile and the friendly status, and we discuss the meaning and significance of an ambiguous relationship in the dynamical shifting of structural balance in SIoT networks. Furthermore, we take advantage of the optimized dynamical evolution of structural balance to shift the SIoT network structure by neutralizing the sign of edges. Based on previous studies, we apply an enhanced objective function and a modified approach concerning the ambiguous relationship towards the dynamical change process. Experiments show that our approach is more effective and efficient than former studies in optimizing the dynamical evolution of the structural balance in signed networks of SIoT. In the future, we will continue to explore and study other elements and methods in the network structure that can enhance the efficiency and effectiveness of SIoT service matching.

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.

Acknowledgments

The work is supported by the National Key Research and Development Program, China (No. 2019YFB2102500), and the Natural Science Basic Research Program, Shaanxi (No. 2020JQ-855, No. 2019-JQ.531).

References

[1] D. Hussein, S. Park, S. N. Han, and N. Crespi, "Dynamic social structure of things: a contextual approach in CPSS," *IEEE Internet Computing*, vol. 19, no. 3, pp. 12–20, 2015.

[2] M. Nitti, L. Atzori, and I. P. Cvijikj, "Friendship selection in the social internet of things: challenges and possible strategies," *IEEE Internet of Things Journal*, vol. 2, no. 3, pp. 240–247, 2015.

[3] P. Doreian and A. Mrvar, "A partitioning approach to structural balance," *Social Networks*, vol. 18, no. 2, pp. 149–168, 1996.

[4] M. Szell, R. Lambiotte, and S. Thurner, "Multi-relational organization of large-scale social networks in an online world," *Proceedings of the National Academy of Sciences*, vol. 107, no. 31, pp. 13636–13641, 2010.

[5] P. Doreian and A. Mrvar, "Partitioning signed social networks," *Social Networks*, vol. 31, no. 1, pp. 1–11, 2009.

[6] N. P. Hummon and P. Doreian, "Some dynamics of social balance processes: bringing Heider back into balance theory," *Social Networks*, vol. 25, no. 1, pp. 17–49, 2003.

[7] M. E. Newman, "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 404–409, 2001.

[8] E. Terzi and M. Winkler, "A spectral algorithm for computing social balance," in *Algorithms and Models for the Web Graph*, pp. 1–13, Springer, 2011.

[9] S. A. Marvel, S. H. Strogatz, and J. M. Kleinberg, "Energy landscape of social balance," *Physical Review Letters*, vol. 103, no. 19, article 198701, 2009.

[10] P. Anchuri and M. Magdon-Ismail, "Communities and balance in signed networks: a spectral approach," in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 235–242, Istanbul, Turkey, 2012.

[11] G. Facchetti, G. Iacono, and C. Altafini, "Computing global structural balance in large-scale signed social networks," *Proceedings of the National Academy of Sciences*, vol. 108, no. 52, pp. 20953–20958, 2011.

[12] G. Iacono, F. Ramezani, N. Soranzo, and C. Altafini, "Determining the distance to monotonicity of a bio-logical network: a graph-theoretical approach," *IET Systems Biology*, vol. 4, no. 3, pp. 223–235, 2010.

[13] T. Antal, P. Krapivsky, and S. Redner, "Dynamics of social balance on networks," *Physical Review E*, vol. 72, no. 3, article 036121, 2005.

[14] S. A. Marvel, J. Kleinberg, R. D. Kleinberg, and S. H. Strogatz, "Continuous-time model of structural balance," *Proceedings of the National Academy of Sciences*, vol. 108, no. 5, pp. 1771–1776, 2011.

[15] V. A. Traag, P. Van Dooren, and P. De Leenheer, "Dynamical models explaining social balance and evolution of cooperation," *PloS One*, vol. 8, no. 4, article e60063, 2013.

[16] T. H. Summers and I. Shames, "Active influence in dynamical models of structural balance in social networks," *EPL (Europhysics Letters)*, vol. 103, no. 1, article 18001, 2013.

[17] T. Antal, P. L. Krapivsky, and S. Redner, "Social balance on networks: the dynamics of friendship and enmity," *Physica D: Nonlinear Phenomena*, vol. 224, no. 1-2, pp. 130–136, 2006.

[18] C. Jiang, S. Liu, Z. Lin, G. Zhao, R. Duan, and K. Liang, "Domain-aware trust network extraction for trust propagation in large-scale heterogeneous trust networks," *Knowledge-Based Systems*, vol. 111, pp. 237–247, 2016.

[19] D. Olmedilla, O. F. Rana, B. Matthews, and W. Nejdl, "Security and trust issues in semantic grids," in *Semantic Grid: The Convergence of Technologies*, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2006.

[20] L. Atzori, A. Iera, and G. Morabito, "SIoT: giving a social structure to the internet of things," *IEEE Communications Letters*, vol. 15, no. 11, pp. 1193–1195, 2011.

[21] F. Azzedin and M. Maheswaran, "Evolving and managing trust in grid computing systems," in *IEEE CCECE2002. Canadian Conference on Electrical and Computer Engineering. Conference Proceedings (Cat. No.02CH37373)*, Winnipeg, MB, Canada, 2002.

[22] I.-R. Chen, J. Guo, and F. Bao, "Trust management for SOA-based IoT and its application to service composition," *IEEE Transactions on Services Computing*, vol. 9, no. 3, pp. 482–495, 2016.

[23] D. Haife, H. Xiaochen, W. Shanfeng, G. Maoguo, and W. Feldman Marcus, "Optimizing transformations of structural balance in signed networks with potential relationships," *Physica A: Statistical Mechanics and its Applications*, vol. 465, pp. 414–424, 2017.

[24] S. Wang, M. Gong, D. Haifeng, L. Ma, Q. Miao, and W. Du, "Optimizing dynamical changes of structural balance in signed network based on memetic algorithm," *Social Networks*, vol. 44, pp. 64–73, 2016.

[25] M. Gong, B. Fu, L. Jiao, and H. Du, "Memetic algorithm for community detection in networks," *Physical Review E*, vol. 84, no. 5, article 056101, 2011.

[26] K. Oda, T. Kimura, Y. Matsuoka, A. Funahashi, M. Muramatsu, and H. Kitano, "Molecular interaction map of a macrophage," *AfCS Research Reports*, vol. 2, no. 14, pp. 1–12, 2004.

[27] A. Ferligoj and A. Kramberger, "An analysis of the Slovene parliamentary parties network," *Developments in Statistics and Methodology*, vol. 12, pp. 209–216, 1996.

[28] K. Oda, Y. Matsuoka, A. Funahashi, and H. Kitano, "A comprehensive pathway map of epidermal growth factor receptor signaling," *Molecular Systems Biology*, vol. 1, no. 1, 2005.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9849025, 1 page
https://doi.org/10.1155/2023/9849025

Retraction

Retracted: Security Control Technology and Simulation of Network News Communication under the Environment of Internet of Things

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] X. Zhou, J. Wang, and X. Zhou, "Security Control Technology and Simulation of Network News Communication under the Environment of Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 2730916, 10 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 2730916, 10 pages
https://doi.org/10.1155/2021/2730916

Research Article

Security Control Technology and Simulation of Network News Communication under the Environment of Internet of Things

Xudong Zhou,[1] Jing Wang ,[1] and Xiao Zhou[2]

[1]*School of Journalism and Communication, Wuhan University, Wuhan, Hubei 430000, China*
[2]*China Unicom Henan Branch, Zhengzhou, Henan 450000, China*

Correspondence should be addressed to Jing Wang; wangjing93@whu.edu.cn

Received 26 April 2021; Revised 29 June 2021; Accepted 5 July 2021; Published 19 July 2021

Academic Editor: Keping Yu

Internet of Things is an application of network news communication technology. Based on the Internet, it uses physical access technologies such as radio frequency tag and wireless sensor network news communication and network news communication information transmission technology to build a network news communication information system that can cover people and things. In the physical layer, the relative position of the object is calculated by using multipoint cooperative localization, so as to determine the minimum anonymous region. Generate and maintain the anonymous tree topology on the network news dissemination layer, and provide storage management support for multiple anonymous groups. In the application layer, the object determines the corresponding anonymity degree according to the identity and uses the frame structure to construct and return the new anonymous group consistent with the existing anonymous group, which can prevent the persistent multiprecision query attack. A real-time control method for intrusion response of a security control system is designed, which includes two stages: response task generation and integrated scheduling. An intrusion response task set generation method based on an improved nondominated sorting genetic algorithm is presented, and a distributed integrated task scheduling and optimization algorithm based on a genetic algorithm and a directed acyclic graph is designed. The numerical simulation results show that this method can quickly and smoothly implement the response strategy of information security intrusion without affecting the normal execution of system tasks.

1. Introduction

With the development and progress of social economy, the Internet of Things has been applied in various important fields. Nowadays, driven by the development of various emerging technologies, such as the rapid development of Internet technology, network communication technology, and radio frequency technology, China's Internet of Things technology is gradually improving and progressing. IoT technology is no longer limited to the previous level and function but is now applied in a more extensive and detailed range [1]. The Internet of Things technology is involved in various aspects such as transportation, military, and financial fields. These fields are closely related to the development of a country. Therefore, only by doing a good job in network information security can the development of a country's economy be better guaranteed [2]. Analyzed at the technical level, the Internet of Things is similar to the Internet, so the information security problems existing in the Internet will also be reflected in the Internet of Things. The Internet of Things will be an important part of information transmission in the future, so it is very important to study the security control technology of network information transmission. The information security of the Internet of Things is a permanent topic in the research of the Internet of Things technology.

The IoT control system is widely used in rail transit, medical and health care, aerospace, industrial manufacturing, disaster, military, and other fields. However, the Internet of Things is a double sword; it brings convenience to people in the control system but also brings some security risks. For a long time, the control system in production and manufacturing industry is a closed and dedicated architecture. However, when the Internet of Things is combined with the control system, the control system will change from the

original closed system to the open system, which is easy to be attacked by hackers, causing security problems such as secret leakage. Therefore, the control security in the Internet of Things environment could not be ignored.

At present, there are many methods to control network information security. However, through investigation, it is found that the existing control systems or methods all have certain limitations when facing the new characteristics of big data. For big data era of network information safety control work faces new situation, only the more practical and reasonable network information security control mechanism can ensure the safety of the huge amounts of data transmission and large database storage security, and only to establish perfect network information safety evaluation system can provide effective standard for network information security work.

2. Related Work

The perceptual layer is an important part of the Internet of Things, which is different from the Internet. The security research on the perceptual layer of the Internet of Things in academia is mainly to expand the existing work, such as data security in sensor network WSN [3], key management in the dynamic ad hoc network [4, 5], and authentication of RFID [6] and privacy [7]. But there is little work that involves various kinds of terminals and network after the integration of new scenarios (such as heterogeneous network integration of data transmission, without authorization, short interactive data interaction, and terminal capacity limited interaction), the security issues, or new problems arising from the Internet of Things (such as attacks of the attacker quantity dominant in the access network). Visual perception layer security issues involved in the information security are more complex than traditional, but due to the restriction of existing production cost and project income, sensing network has not been enough attention, such as Chen listed in the domestic Internet of Things mentioned in the standard [8] perception layer standard which is far less than the application layer and transport layer, also far less than the other two layers of related research work. It is necessary and urgent to study the security of a visible perceptual layer.

The current research on trust in the Internet of Things is not satisfactory. On the one hand, in the distributed environment, the existing trust model only focuses on a certain domain and lacks generality [9]. On the other hand, there is still a lack of research on the trust model specifically for the Internet of Things, so it is urgent to reunderstand and construct the trust model for the Internet of Things. For example, in the perception layer, Garau et al. believed that the trust mechanism had the following challenges [10]: (1) lack of facility support, (2) limited node resources, (3) fragile wireless channel, and (4) multisource.

There are the following challenges in establishing trust mechanisms for the Internet of Things: large-scale scenarios lead to unknown patterns of node interaction. Distributed environment makes the previous centralized security strategy for objects and terminals lose its effect. Heterogeneous environments make it impossible to use a uniform trust mechanism. Research on trust mechanism under the environment of Internet of Things is still quite deficient, and relevant research is not in-depth enough. There is a reputation evaluation system prototype [11] in the proposed security model of the Internet of Things, which includes the reputation management of perceived nodes, terminals, and users but does not include the reputation management of institutions. Obviously, the heterogeneous multisource nature of the Internet of Things will lead to malicious behaviors of institutions. In contrast, Rahim proposed the solution of the trusted Internet of Things [12] to study the trust relationship between institutions (EPCIs) with a sociological approach. However, they ignored the node trust in the perceptual network and did not consider the dynamic, heterogeneous, and large-scale characteristics of the underlying environment.

The Internet of Things is composed of various supporting technologies, such as object-oriented RFID communication, mobile computing for intelligent terminals, ad hoc network for sensing nodes, sensor network, and Internet-oriented service applications, social networking, and cloud computing technologies. The trust requirements of different subjects in these heterogeneous environments are also different. However, trust itself [13, 14] is a subjective and fuzzy concept, which depends on the interaction factors of subject, object, and environment [15]. There is no precise formal definition, and there is no single model that can accurately describe the trust value of subject, although there are some attempts to theoretically unify the trust between subjects in different networks [16]. But the concrete calculation model is not given. In addition, trust value is obtained through direct or indirect analysis of the attributes or behaviors of the subject, but if the interaction process of different subjects is different, it will bring great difficulties to the evaluation of trust value.

At present, the model-based fault recovery control method has been widely applied in distributed systems and autonomous computing fields [17, 18]. Vijayalakshmi et al. [19] studied the model-based adaptive self-healing system in autonomous computing, introduced the self-detection, self-reconstruction, and self-healing processes, and provided the types and recovery measures of software failures, the failure specification, and self-healing framework based on a model, and the aspect-based self-healing process and steps based on model. This paper studies the self-healing system based on the structural model, summarizes the basic elements that the dynamic fault repair system needs, describes the formal language of the structural model, gives the application tools and related technical specifications, and gives the self-recovery flow diagram of the system. In view of the software development of an embedded system and the application of the structural model, the structure description of platform-dependent and platform-independent parts is given, and the realization of instantaneous fault recovery of the structural model in an embedded system is introduced with application cases [20, 21]. Zavala et al. [22] introduce the requirement of structural model style for a self-healing system and the method of a self-healing system based on style. From these studies, it can be seen that the fault model construction in the general computing field is mainly based on the structural model, while the safety control field has

more requirements on the management of the fault model and pays more attention to the safety and reliability at the system level.

In terms of the modeling and analysis of the safety system in the field of safety control, an in-depth study was conducted, and a Markov model was proposed to quantitatively calculate the reliability index of the safety system [23]. Aiming at the problem of fault detection in discrete time networked systems, a binary random sequence method is proposed to describe the random time delay of observation signals caused by networks [24]. A real-time reliability analysis and prediction technology was proposed to evaluate the reliability of equipment in use, which analyzed the reliability of equipment according to the input and output of equipment or state detection, and made real-time life prediction to provide technical support for the predictive maintenance of equipment [25]. At the same time, many researchers have studied the instantaneous fault recovery control problem based on a model. In the literature [26], an analytical model is proposed to deal with the sensor fault problem in the wire control system. Wen et al. [27] established a dynamic optimization method for cost assessment, combine multiple control objectives of the system, and comprehensively deal with actuator failures to ensure the stability of the wire control system. Other studies, from the perspective of controller design, enhance the robustness of the system and improve the ability to resist instantaneous failures, and put forward many control algorithms with fault-tolerant function [28–30]. The above studies have played a positive role in the development of fault-tolerant control, but most of them are based on the internal knowledge of some aspects of the system, so it is difficult to deal with instantaneous faults. Therefore, it is necessary to study the knowledge structure of the safety control system from the perspective of the system, build a model based on the characteristics of the system from the perspective of multiple fields, and then solve the instantaneous fault processing problem based on this "external" model. On the other hand, the fault-tolerant control of the security control system must also consider the problem of real time, and reasonable task scheduling is an important means to ensure the real time in the design of a networked control system.

3. Function Design of the Network News Communication Security Control System under the Environment of Internet of Things

3.1. Functional Safety Guarantee Structure Based on Instantaneous Fault-Tolerant Control. The security control system is a typical distributed structure, and its abstraction level can be divided into node level and system level [31–33]. The propagation directions of instantaneous faults in the safety control system are as follows: the propagation between components in the same abstraction level and the propagation from lower abstraction level to higher abstraction level. As mentioned earlier, transient failures occur at the node level and, if uncontrolled, will eventually propagate to the system level. Therefore, the best way to deal with it is to deal with all instantaneous failures in a timely manner at the node level. However, on the one hand, because the instantaneous fault type is changeable and the cause is complex. On the other hand, the node level needs a fast fault processing mechanism, which mostly adopts the fault knowledge-based processing mode. Therefore, it is almost impossible to handle all instantaneous failures at the node level. The hierarchical instantaneous fault-tolerant control structure of the safety control system proposed in this chapter is shown in Figure 1, including node-level components and system-level components. At the node level, the instantaneous fault is detected by the fault feature-based detection method, and the fault recovery is carried out by the preset recovery strategy. Faults that fail to be detected and handled at the node level (including unknown types of faults that fail to be handled and known types of faults that fail to be detected) will eventually affect the control quality or functional task structure of the system after the evolution of fault propagation. For this type of failure at the system level through the performance evaluation model based on the system and the function structure model of anomaly detection methods for testing, for detecting the abnormal, respectively, the fault-tolerant control based on sliding mode and system task allocation system with the method of reconstruction, realize the fault recovery, and ultimately guarantee the safe and stable operation of the system function.

At the node level, transient faults are divided into execution platform-dependent faults and application-dependent faults. The fault-tolerant control for the execution of platform-related faults is usually realized by the fault-tolerant control system embedded in the platform. Now, the microprocessor developers have relatively mature solutions for these kinds of faults. However, the application of related instantaneous fault needs to be considered in the design of a safety control system. In order to detect application-related instantaneous faults, Figure 2 presents the general representation methods of the state transfer model and data flow model of the safety control system. These states and data are selected according to the application design requirement specification of the intelligent node, which can directly reflect the characteristics of application-related instantaneous faults within the intelligent node.

3.2. Security Control Model of Internet of Things News Transmission against Bandwidth Consumption Attack. Figure 3 shows the WSN application scenario of security control. As can be seen from the figure, wireless sensor nodes collect equipment and surrounding environment information, connect with industrial core network through industrial access private network, and send the collected data back in real time and reliably. The application layer builds the application platform of various industrial businesses according to the specific application background of industrial control. Each application platform system provides fine-grained management and control on the basis of large amounts of data acquired by sensing means.

The main goal of the multilevel detection and early warning model of bandwidth consumption attack is to detect the

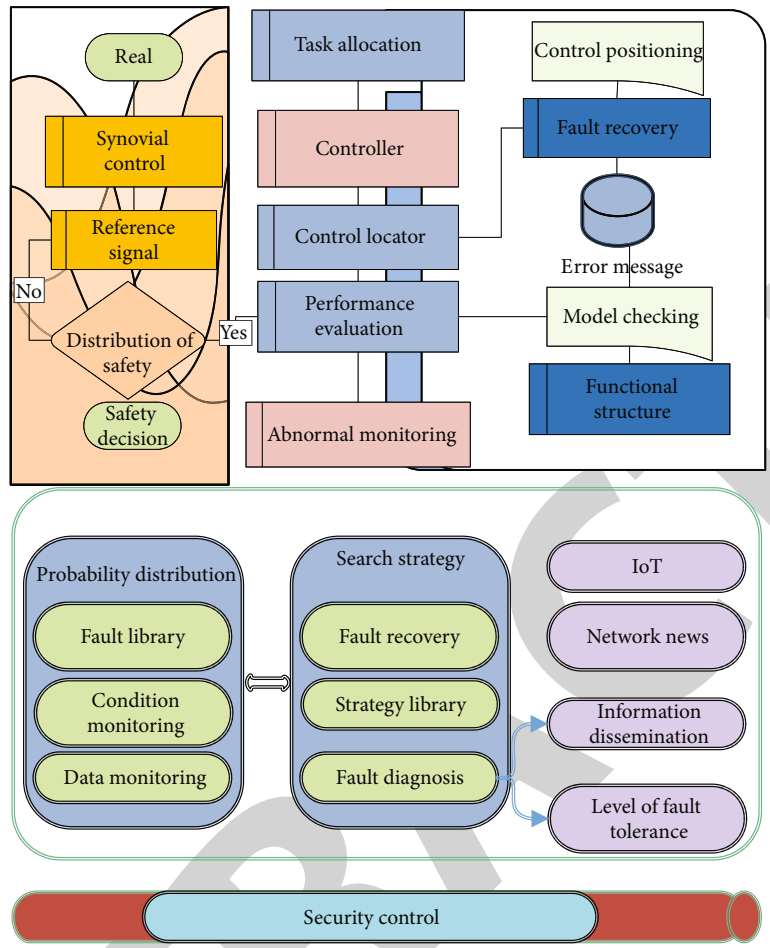


Figure 1: Hierarchy fault-tolerant control mechanism of network information communication security control system under the Internet of Things environment.

abnormal nodes attacked in WSN after the occurrence of bandwidth consumption attack, to give early warning according to the overall attack degree of WSN, and to take different levels of defense measures according to different levels of early warning information. Under this objective, the multilevel detection and warning model of bandwidth consumption attack must be able to complete the following functions:

The multilevel detection and early warning model system against bandwidth consumption attack must be able to detect the response time of nodes in WSN in real time and judge whether the nodes are subjected to bandwidth consumption attack according to the response time.

The multilevel detection and early warning model system for bandwidth consumption attacks should have multilevel early warning function. The system can count the number of nodes attacked in WSN and then infer the overall risk coefficient of WSN, judge the risk level, and finally give the corresponding warning tips.

The multilevel detection and early warning model system for bandwidth consumption attack should be able to take corresponding response measures according to the warning tips of different levels of WSN, so as to further expand the impact of antiattack, so as to ensure that the nodes can provide normal services to the outside to a certain extent.

In this section, a multilevel detection and early warning algorithm based on response time is proposed for bandwidth consumption attack in the WSN environment. In a multiple-level detection algorithm based on response time, monitoring device M plays an important role, the process of detecting early warning and the determination is done by the monitoring device M, and sensor nodes mainly auxiliary monitoring device M have done some related functions, such a design as far as possible to reduce the energy consumption of nodes and prolong the life span of the network.

The core idea of the algorithm is that the monitoring device M sends a request packet to the node in WSN at the time interval of T. The function of the packet is not only to return the total response time from the monitoring device M to the target node but also to calculate the node forwarding hierarchy and form the node forwarding hierarchy matrix. Obviously, the node forwarding hierarchy matrix changes with time. At the same time, the monitoring device detects the response time in the node forwarding hierarchy matrix, determines whether the sensor node is attacked by bandwidth consumption, and gives different levels of warning

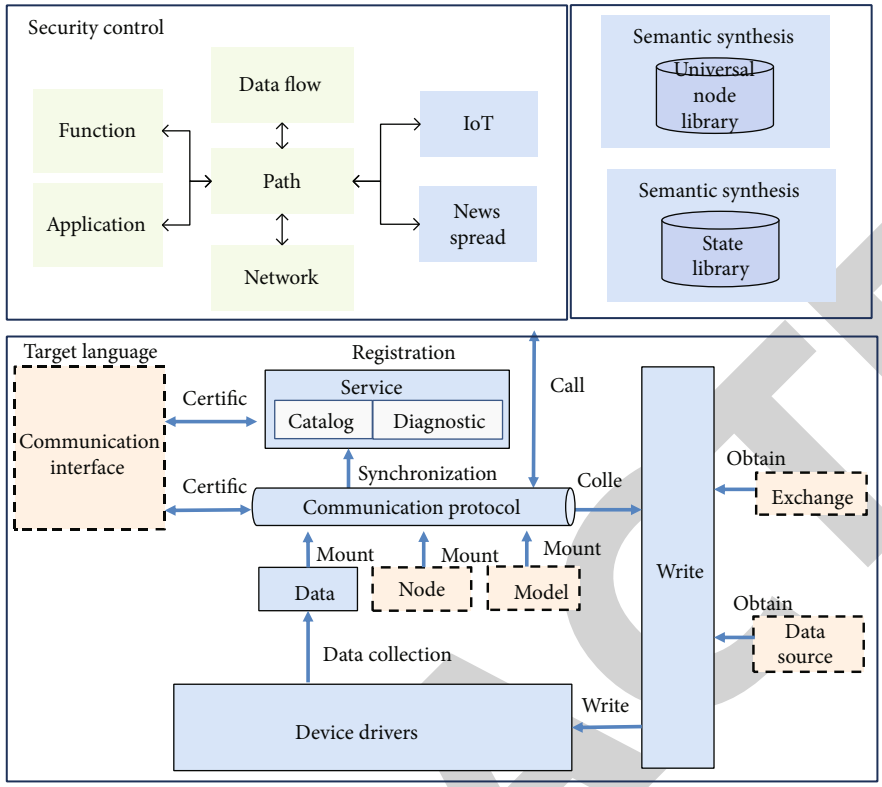


Figure 2: General node model of the security control system.

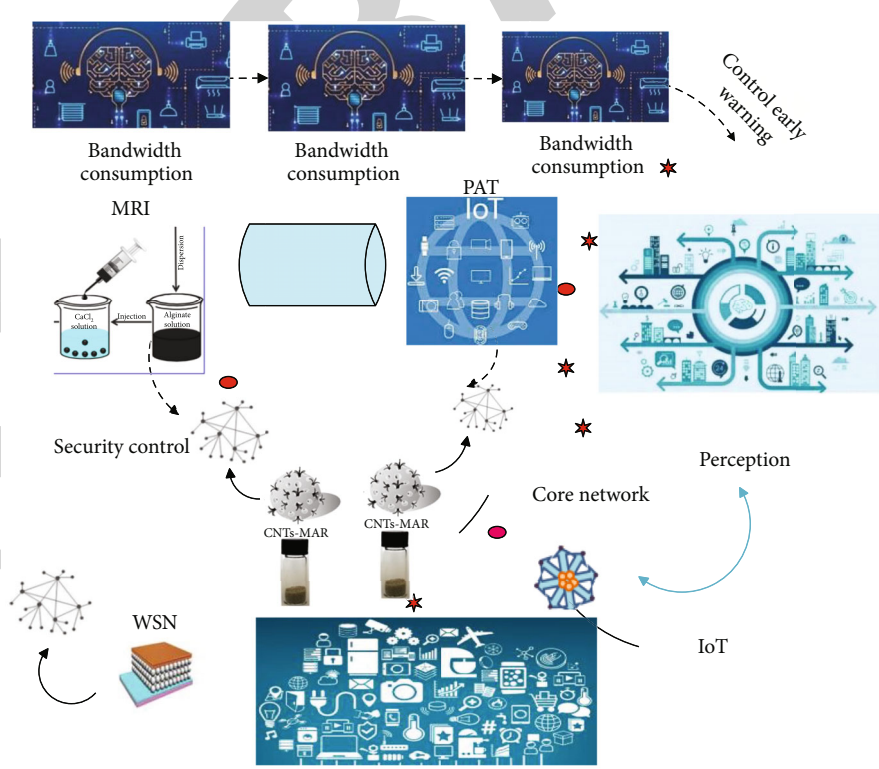


Figure 3: Security control WSN application scenarios.

```
For()
{if(the node is active) {
   If (node response time RT E normal response time range){
   Forward and process packets:
   (Node response time RT exceeds closed value)
   Detect whether the node is under malicious attack;
   If (node is maliciously attacked){
      Reassessing the alert situation;
      Multi-level early warning;
      Take defensive measures; }
   else{
   Forwarding and processing packets; }
}wait(t);
```

Algorithm 1: Security control model of network news dissemination based on bandwidth consumption attack on the Internet of Things

prompts according to the different degrees of attack. The influence coefficient is

$$P = \frac{P + P^0}{P - P^0},$$
$$\chi = \frac{1 + \chi^0}{1 - \chi},$$
$$\beta = \frac{1 + C + H}{Q + C + H}, \quad (1)$$
$$\begin{cases} \forall -\gamma P - \beta P - \chi = 0, \\ \gamma E - \beta E - \chi = 0. \end{cases}$$

According to the above thoughts, warning prompt illustrates the node RT (response time) but may not be getting due to the causes of the node itself, but in the process of forwarding the path of the packet, when the target node appears unusual. It is necessary to check the nodes in the forwarding process from small to large in accordance with the forwarding level of nodes, so as to avoid the error of normal nodes and improve the accuracy of detection.

In addition, based on the response time of the attack more bandwidth consumption level detection of the early warning algorithm, in addition to the attack of detected early warning, severity will against network attacks take different defensive measures, and here are based on the response time of the attack more bandwidth consumption level detection warning algorithm the main logic description as shown in Algorithm 1:

The probability distribution of many random variables in production and scientific experiments can be approximately described by normal distribution, whom the central limit theorem from theory of normal distribution condition: if decided to the results of a random variable is the sum of a large number of small, independent random factors, single function and each factor that are relatively uniform are small, there is not a factor that can serve as the leading role, and overriding the random variables is generally similar to normal distribution. From a statistical point of view, the errors of various measurements are generally normally distributed. Therefore, it is considered that the results of multiple measurements of RTH should also follow normal distribution, and the normal distribution curve of RTH is shown in Figure 4.

3.3. Network Personal Information Security Control Technology and Its Characteristics

3.3.1. Biometric Identification Technology. Biological identification technology is through the computer and biological optics, acoustics, biological sensors, and biological principles of statistics and other high-tech means of closely combined, use human inherent physiological features, such as fingerprint, face, and iris, and the psychological characteristics of behavior, such as handwriting, voice, and gait, characteristics to effective identification and appraisal of personal identity. Because the biometric fingerprint identification of human body has a variety of nonreplicable biological characteristics, the security factor of this fingerprint identification technology has a great improvement and enhancement compared with the traditional biological identity authentication fingerprint identification mechanism. The identification features of retinal biological fingerprint applied to human body mainly include fingerprint, voice, aperture, retina, palmprint, and skeleton. Among them, retinal fingerprint has attracted much attention in the academic world due to its incomparable characteristics such as uniqueness, stability, and reproduction. In addition to the technology of retinal fingerprint recognition, the application and research of retinal fingerprint recognition technology and advanced signature fingerprint recognition technology have also made remarkable progress and achievements in recent years.

3.3.2. Network Security Vulnerability Scanning Technology. Network security vulnerability detection and safety risk assessment technology, because of its characteristic, can accurately predict the network main body possibility of various user network attacks, and specifically against emerging or over the safety of network attack behavior and possible risk consequences, and in recent years, by the network security technology attaches great importance to the industry. The security technology can help administrators identify and test the monitoring system, analyze the possible factors and indicators of the system resources being attacked by hackers, understand the security vulnerabilities of the network monitoring system, and evaluate all possible security risks. Network security vulnerability scanning technology, network firewall, intrusion system, and detection network monitoring system cooperate with each other, which can effectively protect and improve the quality and security of the whole network. By scanning the security vulnerabilities of the whole network, network administrators can timely understand the security settings of the whole network, the operation of the system, and the application services, timely discover the network security vulnerabilities, and objectively evaluate the risk and level of the whole network. Therefore, the network administrator can usually timely correct and find the network security loopholes and various wrong security settings in the process of system operation according to the network

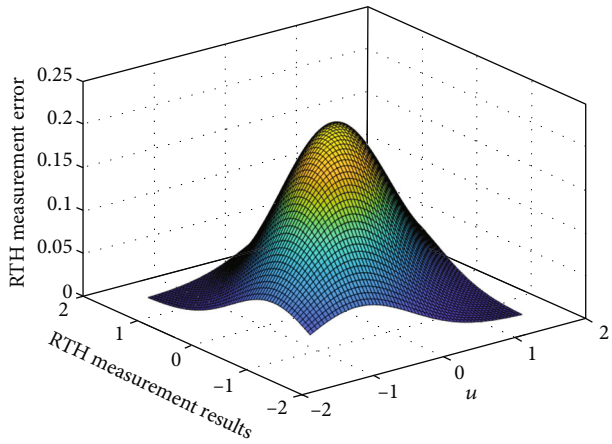


Figure 4: Normal distribution curve of RTH.

scanning status and results and guard against the system before the network is attacked by hackers. When the cluster gateway is determined, the remaining M-N nodes converge to the nearest gateway:

$$P_{ip} = L_{ip} + (1 - \chi)P_J,$$
$$P_{ip} = L_{ip} + L(1 - \chi)p. \quad (2)$$

The certification interval is

$$T_P = \min\left(p_0, \frac{y}{p_o}\right),$$
$$E(t) = \int_0^t \frac{1}{1-p}\sqrt{\frac{\chi^2 - p^2}{p_0}}dt. \quad (3)$$

3.3.3. Detection Process of Abnormal Permission Configuration

Step 1. Clustering: clustering UPA to obtain similar user sets.

Step 2. Preprocessing of clustering results: before detecting abnormal permission configuration, the clustering results should be preprocessed first. For each class cluster, its characteristic pattern vector is constructed. For the elements in the feature pattern vector, if the column vector elements of the corresponding position in the class are all 1, then the position in the feature vector is set to 1, and all other positions are set to 0.

Step 3. Exception privilege configuration rule matching: in this stage, the exception privilege configuration candidate set is screened according to the preset exception privilege configuration mining rules, and the specific rules are as follows.

Rule 1. If a permission is granted only to users in a class, and the percentage of users granted the permission in the class is less than the threshold, mark this configuration as a correct configuration.

Table 1: Introduction of experimental data sets.

Data sets	Number of users	Access number
Healthcare	45	47
University	495	54
Emea	34	3146
Firewall 1	362	1427
Firewa112	324	1169

Rule 2. If a privilege is granted to users in more than one class, and if the proportion of users granted the privilege in the current class is less than the threshold value, then the intersection of the characteristic pattern vectors of the other classes contains the privilege, except this class.

Rule 3. If the proportion of users in a class that has not been granted a certain privilege to all users in the class is less than the threshold value, such configuration is defined as negative exception privilege configuration, these privilege configurations are added to the exception privilege configuration candidate set, and the corresponding position is set as t in the privilege configuration matrix.

Step 4. Cross clustering: after clustering according to the user's one permission matrix and mining the abnormal permission configuration, the row and column of the matrix are exchanged to obtain the transpose matrix of the matrix. Then, cross clustering is carried out according to the clustering algorithm proposed in this chapter. After the clustering results are obtained, the exception permission configuration candidate set is constructed according to the rules mentioned in Step 3.

Step 5. Exception privilege configuration decision: according to the above steps, two candidate sets of exception privilege configuration can be obtained, and the intersection operation of the two sets can be performed to obtain the common set. Define the elements in this common collection as the final exception permission configuration, and update the corresponding elements according to the appropriate modification principles.

Step 6. After dealing with all the clustering results, be able to get a new user access matrix, then the matrix as abnormal access configuration mining framework proposed in this chapter the input and iteration, until no abnormal access configuration is detected or cross the abnormal access configuration of the clustering algorithm to stop the candidate set intersection is empty.

4. Experimental Simulation

Firstly, the data set and evaluation method used in the experiment are introduced. Then, several groups of experiments are designed to compare and analyze the performance of the algorithm in this chapter from different perspectives. The data sets are shown in Table 1.

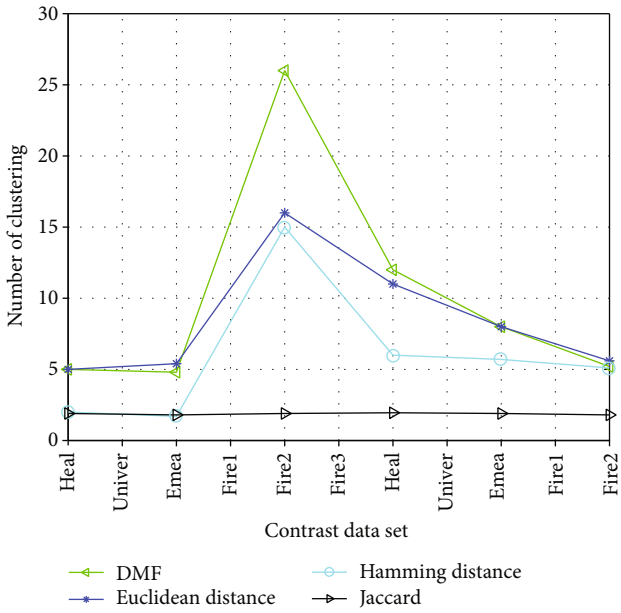


Figure 5: Comparison of the number of clusters.

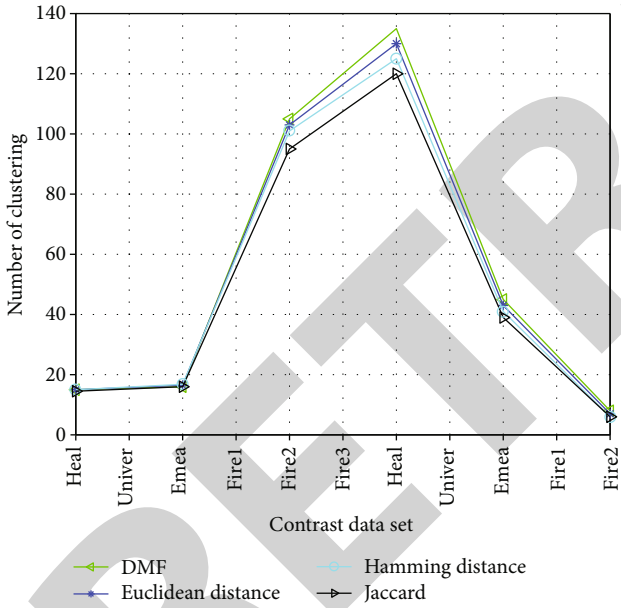


Figure 6: Comparison diagram of the number of generated roles.

In order to verify the effectiveness of DMF mentioned in this chapter, it was compared with Euclidean distance, Hamming distance, and Jaccard distance. As shown in Figure 5, compared with other distance functions, the DMF proposed in this chapter generates more clusters and can be used to distinguish users more effectively. At the same time, as shown in Figure 6, although the clustering algorithm proposed in this chapter increases the number of clustering, the number of roles finally generated does not increase significantly. This indicates that the clustering results under the traditional method are not accurate enough, and multiple roles may be discovered in a single class.

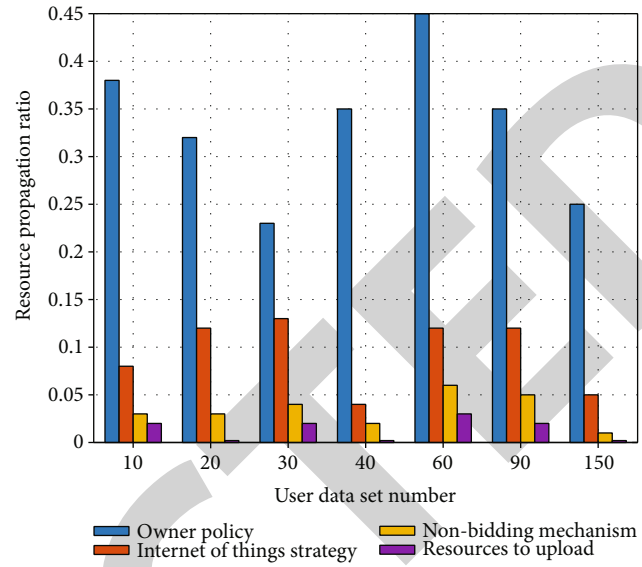


Figure 7: Comparison diagram of the proportion of network news communication resources under the environment of the Internet of Things.

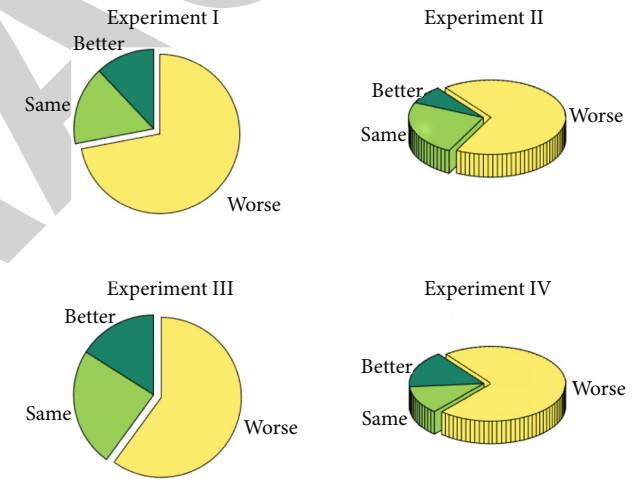


Figure 8: User satisfaction analysis diagram of security control.

As shown in Figure 7, this chapter takes some compensation measures. While increasing the proportion of resource transmission, it does not increase the loss caused by privacy leakage. However, the traditional nonbidding scheme requires users to voluntarily give up part of their privacy, which will still cause privacy loss while increasing the proportion of resource transmission.

In order to prove the universality of the scheme proposed in this chapter, we analyzed the data of all participants, presented the calculated results to participants, and collected their feedback. As shown in Figure 8, about 88% of participants believe that the strength of user relationships calculated by the scheme in this chapter can more accurately reflect their real relationships. According to the above conclusions, compared with the traditional scheme based on simple statistics, the calculation scheme of user relationship strength

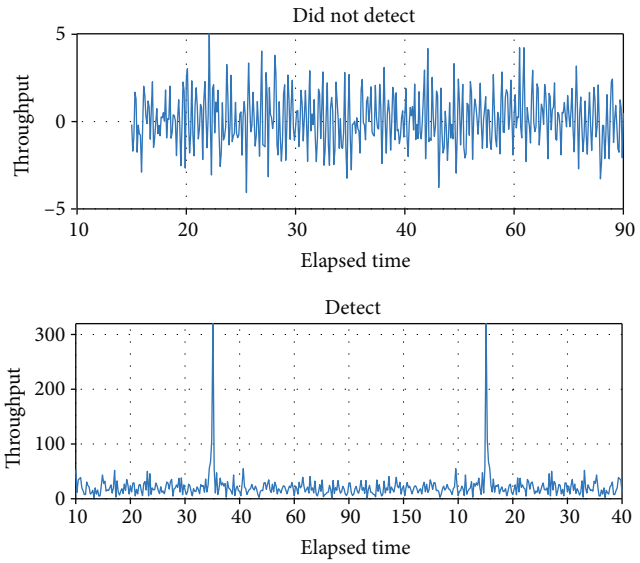


Figure 9: Throughput of the guard module without deep packet detection.

based on communication diversity proposed in this chapter can more accurately reflect the real relationship strength among different users.

In addition to the functional testing of the deep packet detection protection module, a stability test is also needed to compare the throughput and response time of the MATT direct transmission with the deep packet detection. In this experiment, the IxChariot throughput testing tool was used to test the performance of the Matt deep packet detection module. This testing tool supports the throughput and response time testing of various protocols and also supports the testing of customized application layer data. When the deep packet detection function is not added, the throughput of the protection module is shown in Figure 9.

5. Conclusion

By analyzing the characteristics of the Internet of Things, this paper analyzes the standard architecture of the industrial control of the Internet of Things and analyzes the interference factors in the Internet of Things environment from the perspective of the general control system model of the architecture and network environment and the security issues of the controlled system. The general control system model and the security model of the general control system under the environment of the Internet of Things are proposed. In view of transmission security, the security control model of Internet of Things news communication against bandwidth consumption attack is studied, and a detection algorithm of bandwidth consumption attack based on node response time is proposed. Under the guidance of the detection algorithm, a monitoring, analysis, and early warning model for bandwidth consumption attacks is constructed. Finally, through the analysis of simulation experiment and real user experiment results, it is proved that the proposed scheme can achieve dynamic and accurate resolution of coexisting policy conflicts. Security controls on the Internet of Things infrastructure will likely extend to the entire mobile network 10 times or even 100 times more than the Internet, which means the entire Internet of Things market of information security may also be expanded tens of times. The next step will be based on the coexistence strategy conflict resolution of communication diversity: the element types used in communication diversity calculation can be expanded to increase the universality of the scheme and improve the calculation accuracy of user diversity.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Laghari and M. A. Niazi, "Modeling the internet of things, self-organizing and other complex adaptive communication networks: a cognitive agent-based computing approach," *PloS One*, vol. 11, no. 1, pp. e0146760–e0146772, 2016.

[2] V. Rohokale and R. Prasad, "Cyber security for Intelligent-World with internet of things and machine to machine communication," *Journal of Cyber Security and Mobility*, vol. 4, no. 1, pp. 23–40, 2015.

[3] M. Durresi, A. Subashi, A. Durresi, L. Barolli, and K. Uchida, "Secure communication architecture for internet of things using smartphones and multi-access edge computing in environment monitoring," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 4, pp. 1631–1640, 2019.

[4] S. Jaloudi, "Communication protocols of an industrial internet of things environment: a comparative study," *Future Internet*, vol. 11, no. 3, pp. 66–86, 2019.

[5] Y. Jeong, "An emergence of network agenda-setting theory: a comparative analysis of networks of issue attributes between news media and online discussion forums," *Korean Journal of Journalism & Communication Studies*, vol. 59, no. 3, pp. 365–394, 2015.

[6] N. M. Martínez, J. J. Olivencia, I. Mac Fadden, and E. O. Olmedo, "Skills in the use of Technologies of Information and Communication of the teachers (2.0) under the scope of university studies," in *European innovations in education: research models and teaching applications*, E. L. Meneses, Ed., vol. 82, pp. 200–213, AFOE. Asociación para la Formación, el Ocio y el Empleo, 2017.

[7] I. Bladek and K. Krawiec, "Counterexample-driven genetic programming: heuristic program synthesis from formal specifications," *Evolutionary Computation*, vol. 26, no. 3, pp. 441–469, 2018.

[8] G. Chen, "Model innovation of network news communication based on interactive analysis," *Revista de la Facultad de Ingenieria*, vol. 32, no. 9, pp. 271–277, 2017.

[9] S. T. Campbell, "The dynamics of handcart as a means of informal transportation in support of logistics and tourism,"

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 7057513, 13 pages
https://doi.org/10.1155/2021/7057513

Research Article

Improved Height Estimation Using Extended Kalman Filter on UWB-Barometer 3D Indoor Positioning System

Ji Li,[1] Yepeng Wang,[1] Zhuo Chen,[2] Linlin Ma,[3] and Suqing Yan[1,4]

[1]*Guangxi Key Laboratory of Image and Graphic Intelligent Processing, Guilin University of Electronic Technology, Guilin 541004, China*
[2]*University of British Columbia, Vancouver, Canada*
[3]*Zhengzhou Locaris Technology Co., Ltd., Zhengzhou 450000, China*
[4]*Guangxi Key Laboratory of Precision Navigation Technology and Application, Guilin University of Electronic Technology, Guilin 541004, China*

Correspondence should be addressed to Yepeng Wang; yepengwang1@163.com and Suqing Yan; yansuqing@guet.edu.cn

Received 30 April 2021; Revised 19 June 2021; Accepted 1 July 2021; Published 19 July 2021

Academic Editor: Wei Wang

Indoor 3D positioning system requires precise information from all three dimensions in space, but measurements in the vertical direction are usually interfered by sensors properties, unexpected obstructions, and other factors. Thus, accuracy and robustness are not guaranteed. Aiming at this problem, we propose a novel sensor fusion algorithm to improve the height estimation for a UWB-barometer integrated positioning system by introducing a pseudo reference update mechanism and the extended Kalman filter (EKF). The proposed fusion approach effectively helps with sensing noise reduction and outlier restraint. The results from numerical experiment investigations demonstrate that the accuracy and robustness of the proposed method achieved better improvement in height determination.

1. Introduction

With mobile networks' development, people's demand for positioning and navigation has increased rapidly, especially in industrial applications with complex indoor facilities, such as intelligent power supply stations, urban underground trenches, and petrochemical plants [1]. Therefore, it has definite practical meaning to perform real-time location monitoring and tracking for moving targets in these scenarios. At present, ultra-wideband (UWB) network, radio frequency identification technology, and laser scanner are popular methods that provide indoor position information [2–5]. The UWB has high accuracy, strong stability, good antimultipath effect, low transmitting power, and low radiation [6]. Theoretically, such advantages make it capable of acquiring high precision three-dimensional position information in a spacious place. However, there are many occlusions in natural industrial environments, and the inside elevation is usually limited. These existing factors block the UWB signal in the vertical direction and significantly reduce height estimation accuracy [7].

Different from the UWB network, a barometer does not render three-dimensional location information. It only determines elevation through differential pressure calculation and is widely used for outdoor field applications [8]. If this barometer feature can be applied to UWB localization, it will be an excellent complement. Barometers' working manner has no defective impact on the UWB network. Moreover, its height estimates can make up when UWB failed to fetch height information and improve system sensing reliability [9, 10]. Usually, people first calibrate barometers to the value of mean sea level [11]. However, unlike outdoor sensing, the atmosphere inside a room changes little regarding the absolute sea-level reference raising another challenge. One of

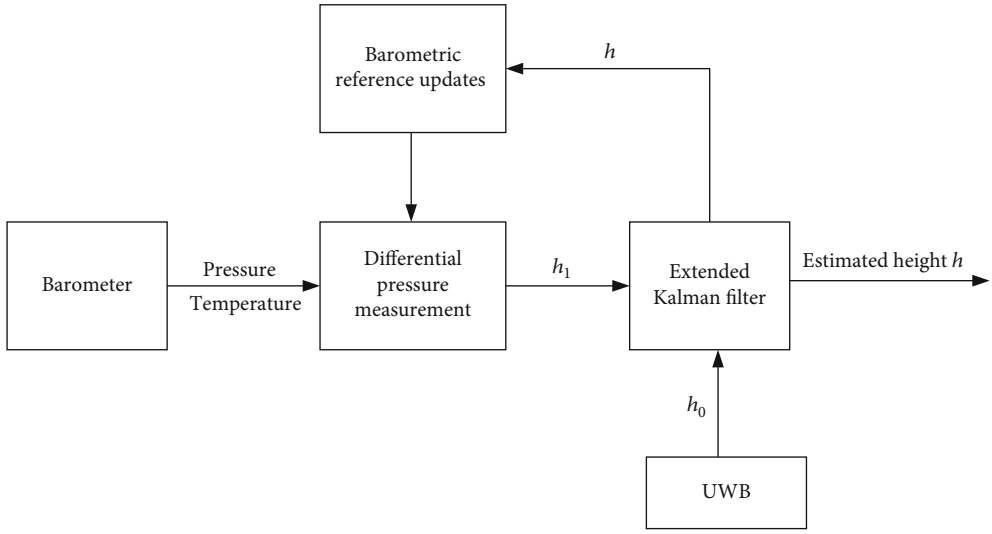


Figure 1: Sensing system overview.

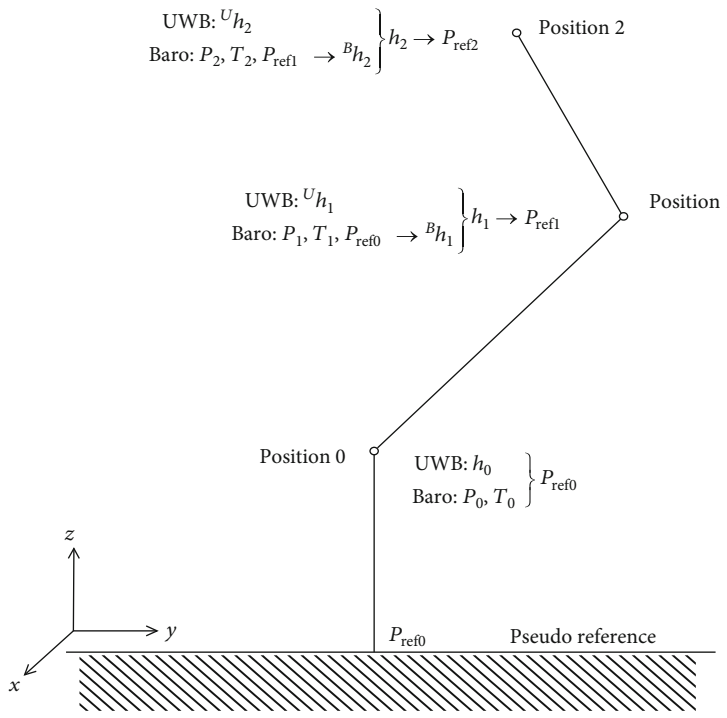


Figure 2: Pseudoreference updates.

the solutions is to set up a new pressure reference, then conduct differential pressure measurements based on this reference [12–14].

Many works have been done on this topic. Using differential calculation by two barometric sensors can make forecasted data more accurate [15, 16]. However, using two barometric sensors for differential calculation also has some disadvantages. For instance, barometric sensors require effective and frequent calibration to avoid long-term drifts, which affects the final height estimation accuracy [17]. Thus, height estimation's long-term accuracy and stability relying on individual pressure sensors are poor [18]. Considering these shortcomings of a single technology, supplementing various measurement advantages can achieve better position accuracy. The vertical height is underestimated due to the nonlinearity caused by the downward integration of the rotating accelerometer [19]. However, extended Kalman filter (EKF) can be used to apply the nonlinear Gaussian distribution [20], then joint EKF can deal with the nonlinearity problem to stabilize the height estimate.

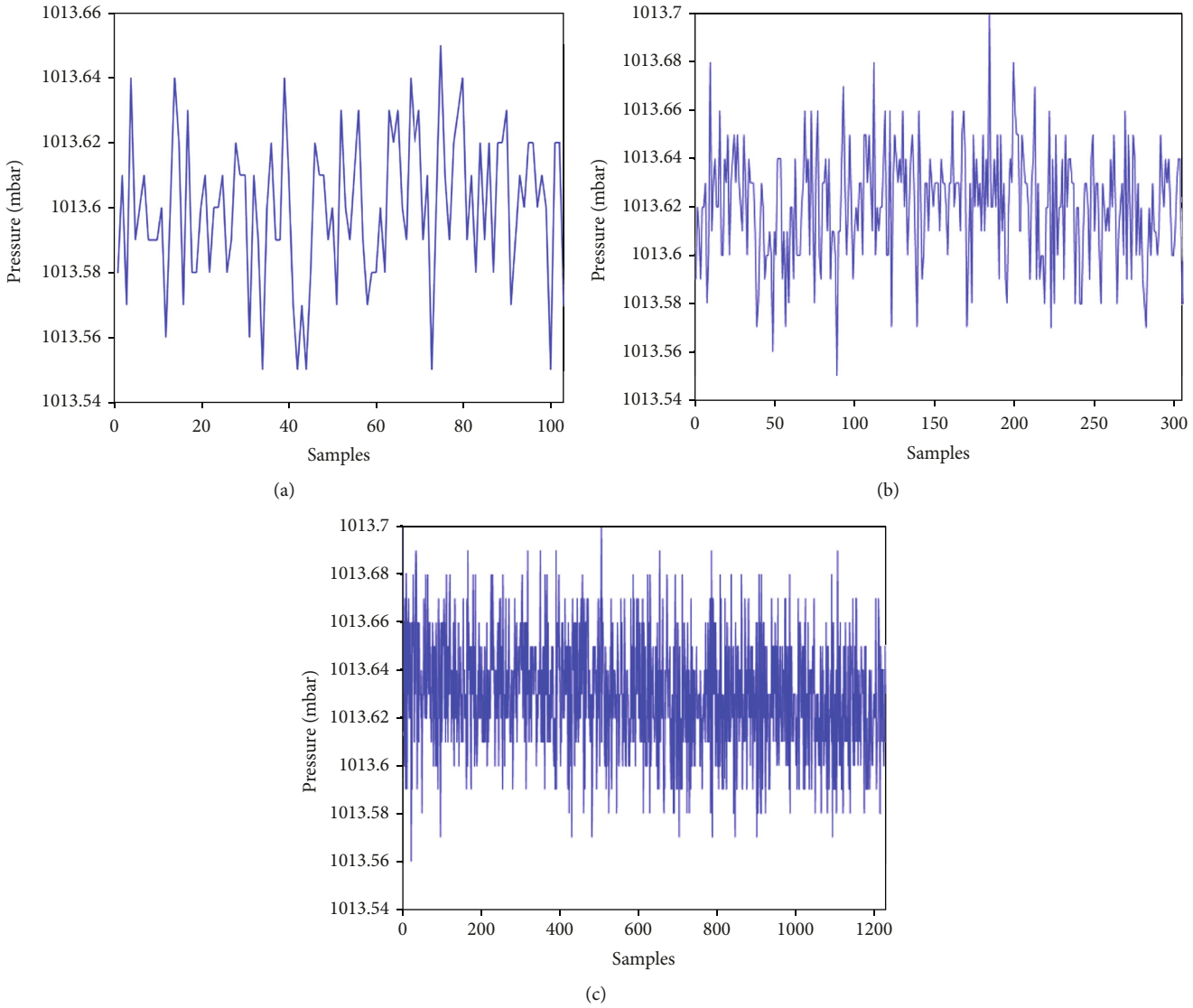


Figure 3: Static measured pressure with barometer for different durations. (a) Time duration is 10 seconds, (b) time duration is 30 seconds, and (c) time duration is 120 seconds.

The main contributions in this paper are as follows.

To improve accuracy and robustness for indoor height estimation, we propose a barometer-integrated UWB height estimation method. The proposed method has greatly improved the accuracy and stability of the estimated height.

During the height estimation, a pseudopressure reference update mechanism is proposed, and the scheme enables the barometer to achieve better performance in indoor environments.

The remainder of this article is organized as follows. The proposed barometer-integrated UWB height estimation method is depicted in Section 2. Section 3 illustrates the experiment's analyses and results. Section 4 shows conclusions.

2. Proposed Method for Height Estimation

Beacons in the 3D positioning system are installed on the floor and ceil two layers. Positioning can be achieved by sampling the selected beacons. However, when the elevation difference is limited and the floor beacons are blocked, the fluctuation of elevation estimation is relatively large. It is not easy to distinguish the upper and lower base stations on some occasions, so that the height estimation cannot be determined. Meanwhile, higher accuracy can be achieved for a reasonable open topological structure. Once occurring, outliers bring a tremendous challenge for positioning.

Table 1: Static barometer reading average and standard deviation in 10 s, 30 s, and 120 s.

Period	10 s	30 s	120 s
Ave (mbar)	1013.600	1013.621	1013.630
Std (mbar)	0.0231	0.0238	0.0243

Decimeter level altitude estimation can be gained through the double barometer differential method when high precision air pressure sensors are adopted. However, the

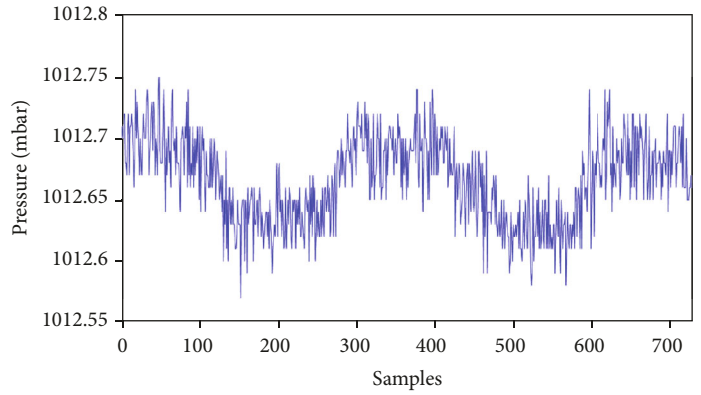


Figure 4: Measured pressure for barometer moving up and down within 50 cm.

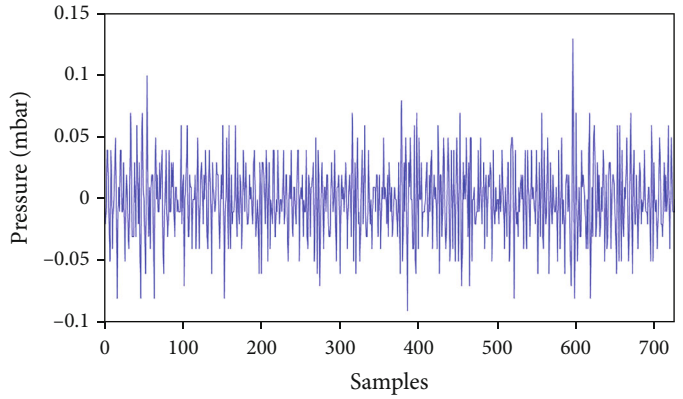


Figure 5: Pressure difference between two adjacent pressures.

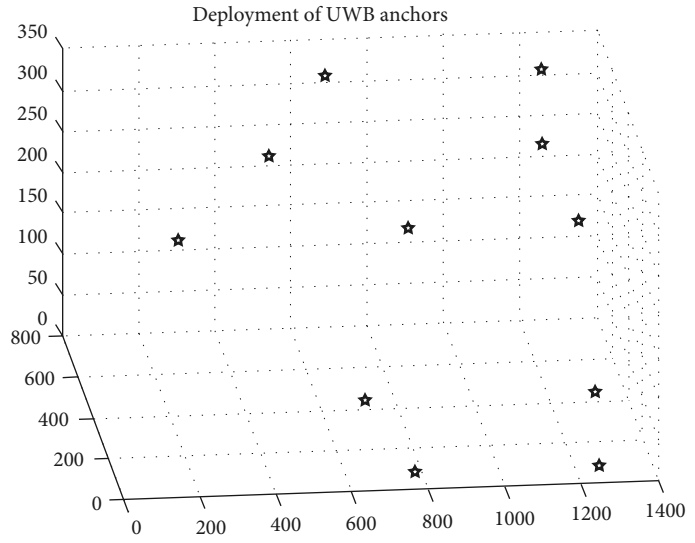


Figure 6: The deployment of 11 UWB anchors in the room with $14 * 8 * 3.5\,m^3$.

pressure deviation between barometers is not stable for a long time, so the deviations between barometers need to be corrected at intervals.

UWB signal has strong penetration ability and good anti-multipath performance, especially in indoor or building intensive environments. It can avoid the shielding effect, and the short-term stability of the barometric pressure sensor can benefit to assist the height estimation.

Aiming to improve the accuracy and reliability of measurement in a vertical direction, a novel high-precision

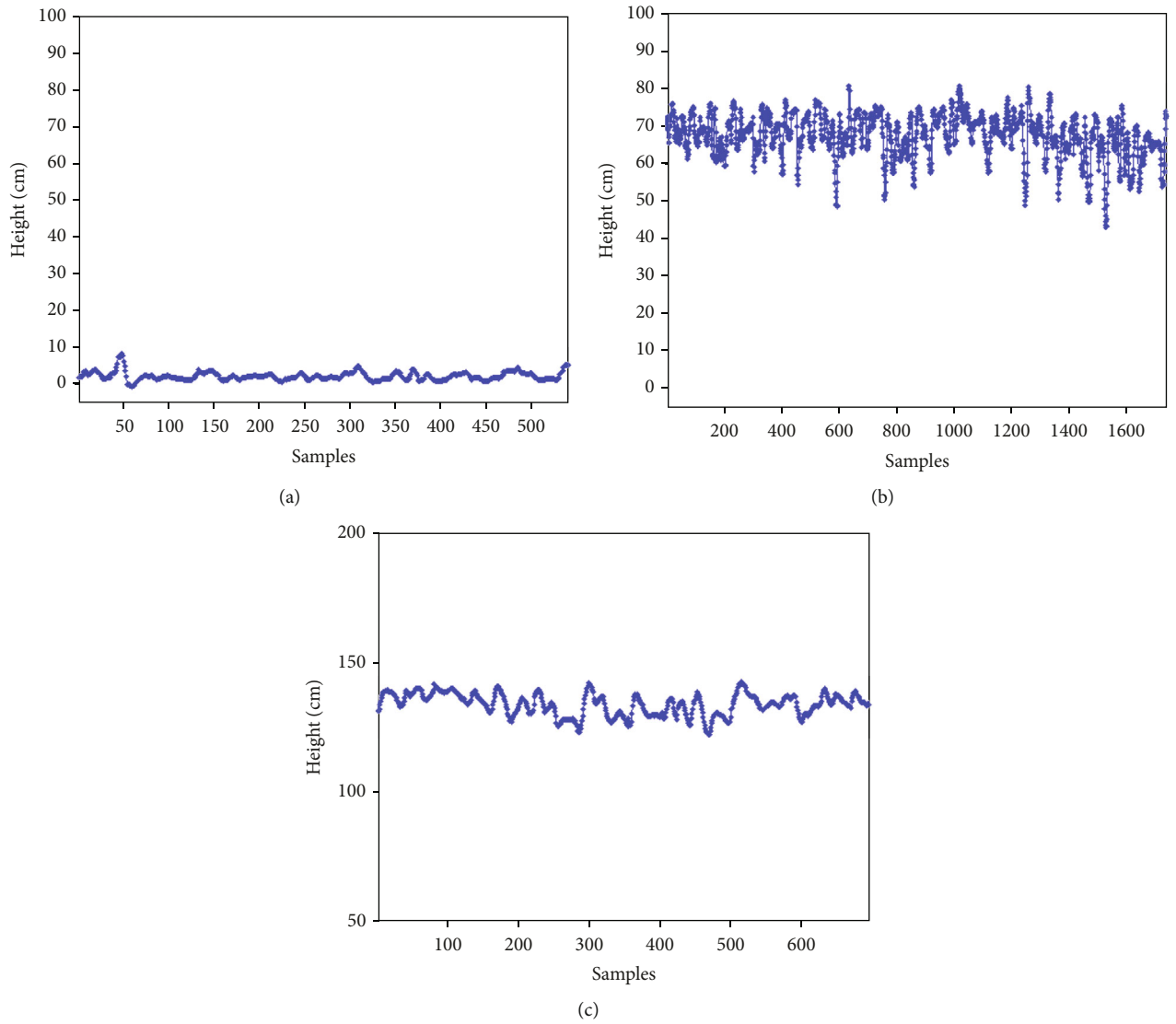


Figure 7: The estimated height with UWB anchors at different heights: (a) height = 0 cm, (b) height = 76 cm, and (c) height = 150 cm.

three-dimensional positioning method based on the fusion of ultrawideband and barometric pressure sensors is proposed. In the proposed method, the single barometer is adopted to eliminate device differences. In this section, we first introduce the general mechanism of EKF, then we describe how the barometer reference update and data fuse in the proposed method, as illustrated in Figure 1.

2.1. Extended Kalman Filter. EKF is used for nonlinear systems and noise models, which is a kind of the general Kalman filter (KF) [21]. In the EKF, the state transformation model and observation model are formulated by nonlinear functions. These two nonlinear models are described as state transition model:

$$x_k = f(x_{k-1}, u_{k-1}, s_{k-1}) \quad s_k \sim N(0, Q), \tag{1}$$

Table 2: Static barometer reading average and standard deviation at different heights.

Ground truth	0 cm	76 cm	150 cm
Ave (cm)	1.99	67.08	133.61
Std (cm)	1.15	5.75	4.25

and observation model:

$$z_k = w(x_k, v_k) \quad \sim v_k N(0, R), \tag{2}$$

where $f(\cdot)$ and $w(\cdot)$ are the nonlinear functions with x_{k-1}, u_{k-1}, s_{k-1}, and x_k, v_k, respectively. s_k is system process noise and zero-mean Gaussian process with covariance Q, and v_k is observation noise and zero-mean Gaussian process with covariance R. u_k denotes the input at time $k-1$. x_k and x_{k-1} represent the state variable at time k and $k-1$, respectively. z_k is the measurement at time k.

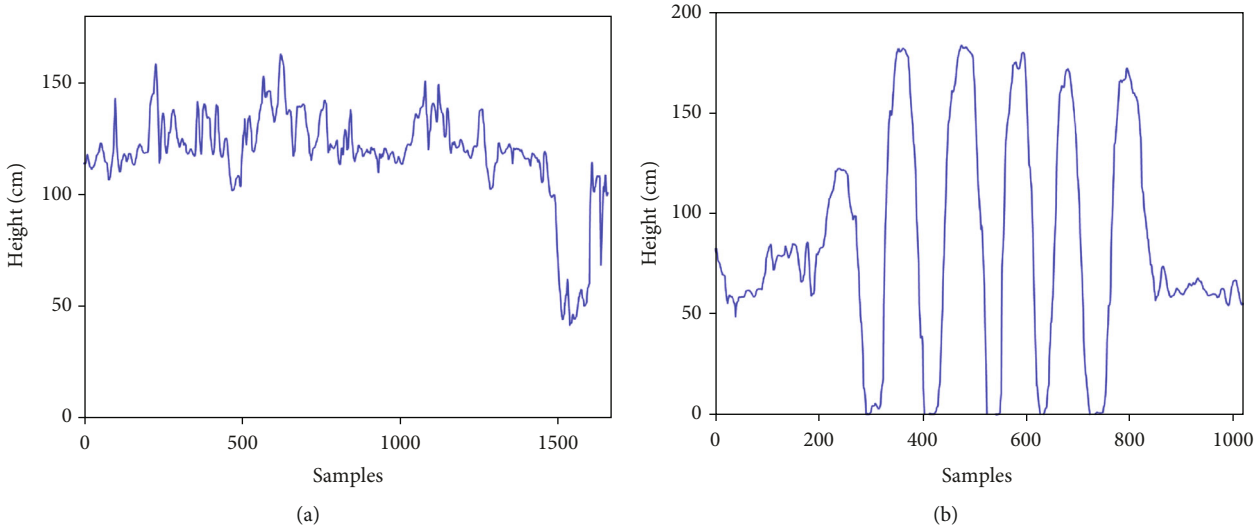


Figure 8: The estimated height of moving target: (a) target move at 130 cm height and (b) target moves up and down.

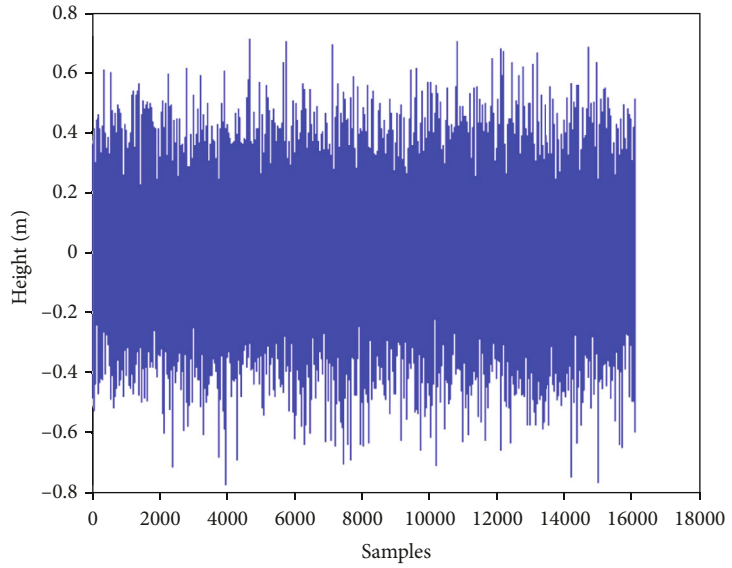


Figure 9: Simulated UWB data at 0 cm height.

Linearizing the two nonlinear Equations (1) and (2) by a first-order Taylor series approximation at $\hat{x}_k$, then we get the linear equations as follows:

$$x_k = f(\hat{x}_{k-1}, u_{k-1}, s_{k-1}) + F_{k-1}(x_{k-1} - \hat{x}_{k-1}) + B_{k-1}s_{k-1}, \quad (3)$$

$$z_k = w(\hat{x}_k, v_k) + W_k(x_k - \hat{x}_k) + V_k v_k, \quad (4)$$

where F_{k-1} and B_{k-1} are the Jacobian matrices of $f(\cdot)$ with respect to $\hat{x}_{k-1}$ and s_{k-1}, and W_k and V_k are the Jacobian matrices of $w(\cdot)$ with respect to $\hat{x}_k$ and v_k.

In the height estimation, we suppose the process noise is zero, therefore, $f(\hat{x}_{k-1}, u_{k-1}, s_{k-1})$ in Equation (3) is the estimated value $\hat{x}_k$ at time k. $w(\hat{x}_k, v_k)$ in Equation (4) is the estimated value $\hat{z}_k$ at time k. Substitute $\hat{x}_k$ and $\hat{z}_k$in the Equations (3) and (4), state transition model and observation model in the EKF can be expressed the linear function as follows:

$$x_k = \hat{x}_k + F_{k-1}(x_{k-1} - \hat{x}_{k-1}) + B_{k-1}s_{k-1}, \quad (5)$$

$$z_k = \hat{z}_k + W_k(x_k - \hat{x}_k) + V_k v_k. \quad (6)$$

EKF iterates in two major steps, the state prediction can be gained as Equations (7) and (8) and the state update as Equations (9) to (11).

$$\hat{x}_{k|k-1} = f(\hat{x}_{k-1}, u_{k-1}, 0), \quad (7)$$

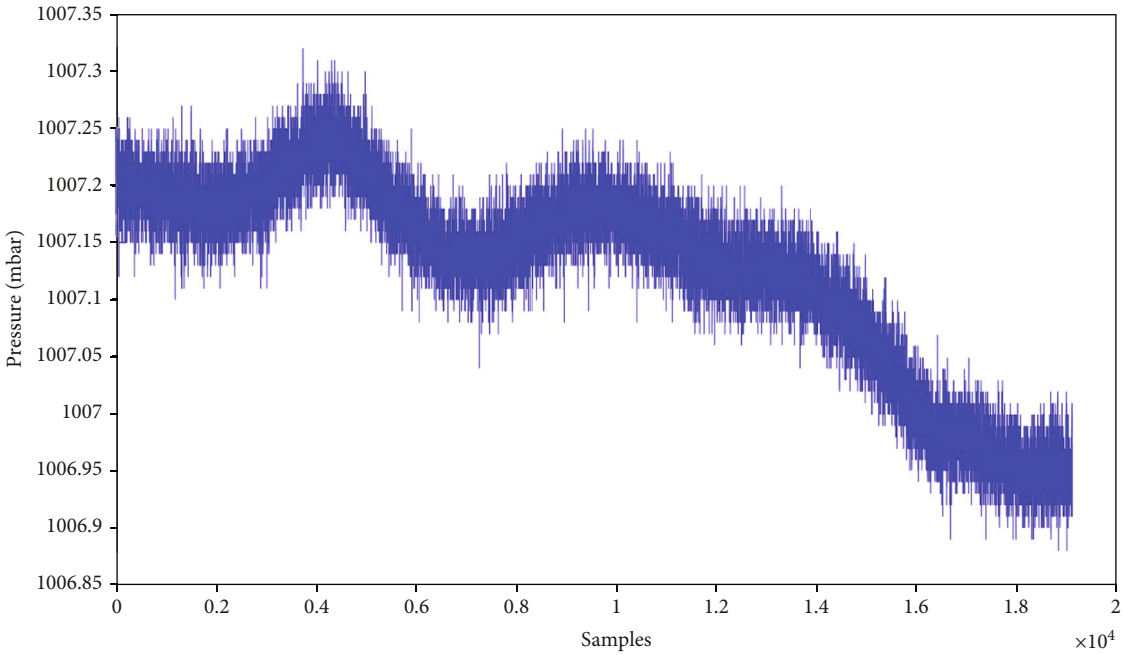


Figure 10: Measured pressure at 0 cm with barometer.

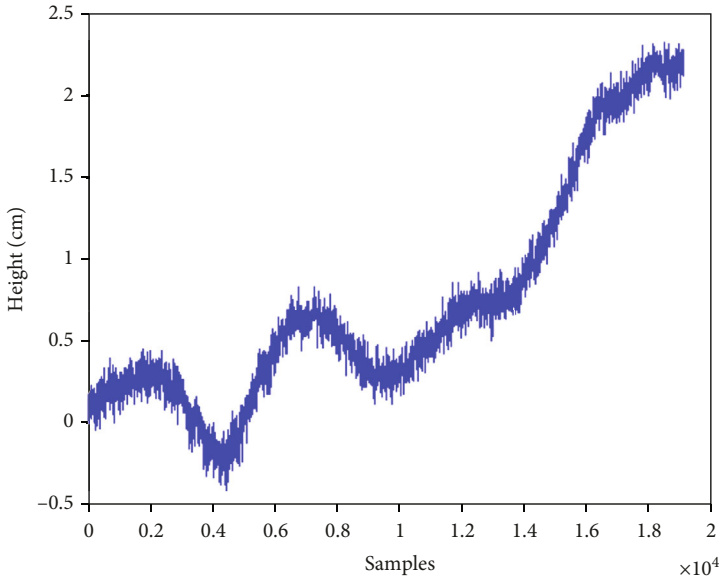


Figure 11: Estimated height with only barometer pressure sensor at 0 cm.

$$P_{k|k-1} = F_{k-1} P_{k-1} {F_{k-1}}^T + B_{k-1} Q {B_{k-1}}^T. \quad (8)$$

In the prediction step, process noise is unknown, state prediction $\widehat{x}_{k|k-1}$ is calculated by supposing s_k is zero. Its covariance $P_{k|k-1}$ at time k is calculated through Equations (8). Q denotes the estimated process noise covariance, which is fixed in the paper. In this paper, Q is the unit diagonal matrix.

$$K_k = P_{k|k-1} W^T \left(W P_{k|k-1} W^T + V_k R V_k \right)^{-1}, \quad (9)$$

$$\widehat{x}_k = \widehat{x}_{k|k-1} + K_k \left(z_k - W\left(\widehat{x}_{k|k-1}, 0 \right) \right), \quad (10)$$

$$P_k = (I - K_k W) P_{k|k-1}. \quad (11)$$

In the update step, the measurement at time k is taken into account. K_k is the Kalman gain, and it will converge as the filter iterates. Outputs of EKF are the optimal state $\widehat{x}_k$ and its covariance P_k which will be considered as inputs of the next iteration. The estimated state $\widehat{x}_k$ is the height h_k.

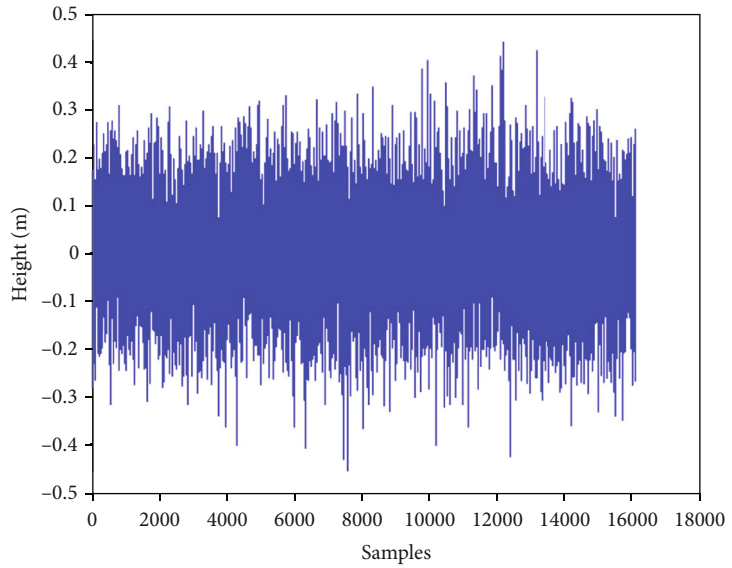


Figure 12: Estimated height at 0 cm with fusion UWB and barometric.

We deploy a barometer to estimate the system state relying on the nonlinear model as Equation (10) shows, and a UWB network to observe the system. Therefore, EKF is a feasible approach to handle this nonlinear-Gaussian fusion problem.

2.2. Pseudobarometeric Reference Establishment and Update. As we all know, the atmospheric pressure will decrease with the increase in altitude, and barometers can be used to obtain altitude information. However, barometers are plagued by an unpredicted environment whose accuracy could be at 0.5 m, but it can provide short-term stable measurement. By considering UWB and barometers' complementary features, a 4-step mechanism to establish and update barometric reference is realized in the paper as shown in Figure 2.

Step 1. Initialization: determine the first height h_0 by trilateration algorithm according to UWB anchors.

Step 2. Reference pressure establishment.

The pressure P_0 is determined from h_0. Reference pressure is updated through Equation (12).

$$P_{\text{ref0}} = 10\wedge\left(\frac{h_0 - h_{\text{ref}}}{18410} * \frac{273.15}{273.15 + T_0}\right) * P_0, \tag{12}$$

where P_{ref0} is the calculated pressure at the pseudoreference level, and T_0 denotes the temperature at the initial measuring point, h_{ref} equals to zero.

Step 3. Height determination.

With the reference pressure P_{ref0}, the pressure P_1, and temperature T_1 measured by barometer at position 1, the estimated height ${}^{B}h_1$ can be determined through Equation (13). Here, the superscript letter B indicates estimation inferred by barometer data.

$${}^{B}h_1 = 18410 * \left(1 + \frac{T}{273.15}\right) \lg \frac{P_{\text{ref0}}}{P_1} + h_{\text{ref}}, \tag{13}$$

where T is an average of T_1 and T_0.

Table 3: Estimated height results at 0 cm.

Methods	UWB	Barometer	UWB-Baro-EKF(proposed)
Ave (cm)	0 (GT)	0.82	0.16
Std (cm)	20.06	14.42	10.69

Step 4. Reference update.

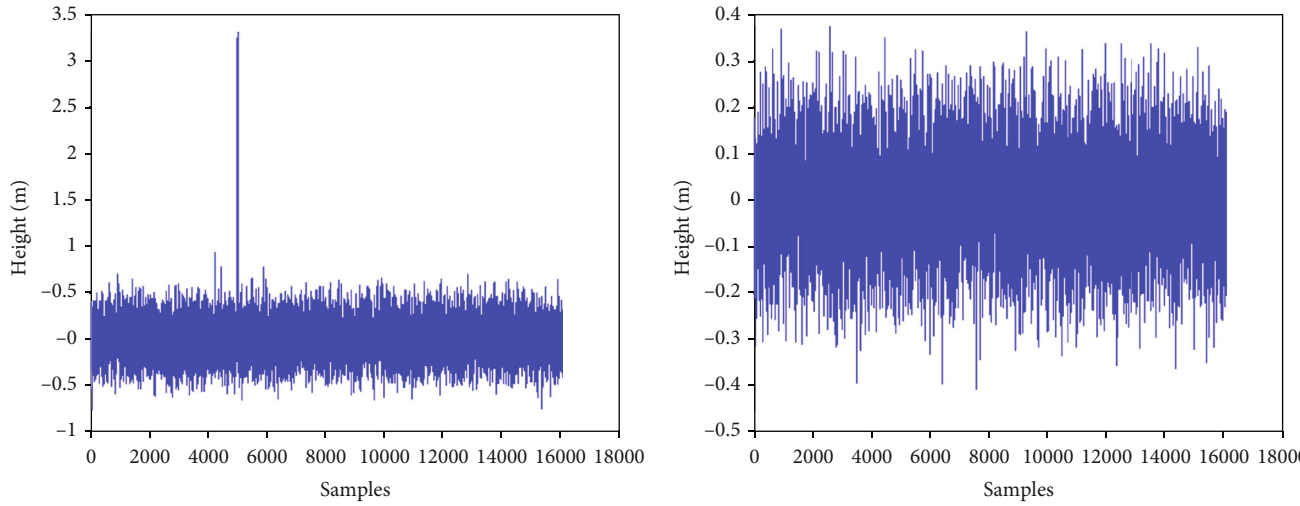
At position 1, UWB determines a ${}^{U}h_1$ through trilateral calculation. By combining ${}^{U}h_1$ and ${}^{B}h_1$, the EKF outputs an optimal estimation height $\widehat{h}_1$ with covariance H_1. Similar to the process in step 2, we calibrate the reference pressure based on h_1 through Equation (9).

According to EKF theory, the accuracy of h_1 is higher than h_0, therefore, the updated reference pressure P_{ref1} is closer to its true value.

3. Experimental Results

The experiments are mainly conducted in three parts. In the first two parts, the stability and accuracy of the applied barometer and UWB are tested, then the proposed fusion method is tested in the last part.

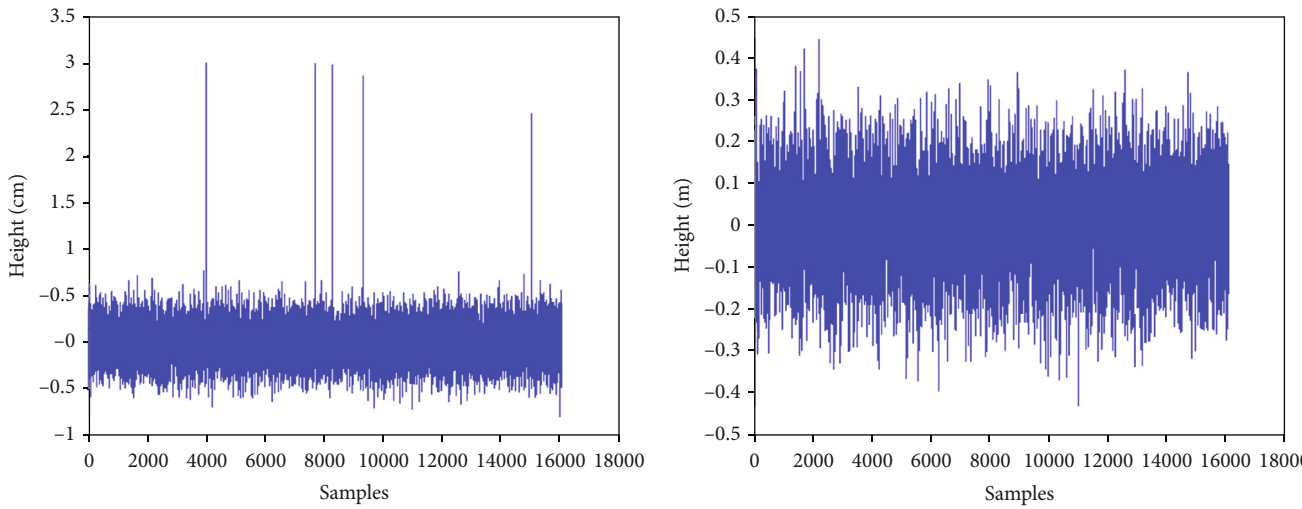
3.1. Barometer. In the experiments, the barometric pressure sensor (Swiss) MS5611 is used to measure the air pressure and has a high-resolution barometric pressure sensor with SPI and I^2C bus interface introduced by MEAS (Switzerland).



(a) UWB simulation dada generated with single outlier

(b) Estimated height with proposed method

Figure 13: Original UWB data with single outlier and estimated height with the proposed method.



(a) UWB simulation dada generated with outliers

(b) Estimated height with the proposed method

Figure 14: Original UWB with outliers and estimated height with the proposed method.

First, we test the barometer's static measurement by setting it at a certain level and collect atmosphere pressure data in 10 s, 30 s, and 120 s shown in Figure 3. With measuring time increases from 10 seconds to 120 seconds, the range where data fluctuates becomes larger. Table 1 reflects the average measurements and standard deviations. There is only little figural change on moderate pressure, suggesting that the barometer can maintain short-term stability. For long-term measurement, a periodical calibration is required because readings begin to disperse after 40 seconds.

We also test the barometer's dynamic measurement by the 50 cm range vertical reciprocating movement. Figure 4 shows the measured pressure as the barometer moves up and down within 50 cm. Meanwhile, we do several times experiments continuously to sample pressure. Figure 5 depicted the pressure difference between two adjacent pressures. It can be seen that the barometer can effectively identify the up and down motion by measuring atmosphere pressure and also have better stability. Moreover, the feature of accordant pressure difference can help with outlier removal and data refinery.

3.2. UWB Height Estimation. To validate the location performance adopted by UWB, we conducted experiments on static and moving targets, respectively. An indoor UWB network with 11 anchors is installed in the room with $14 * 8 * 3.5\ m^3$. The anchor deployment is distributed as Figure 6.

Figure 7 shows the location results of static targets at 0 cm, 76 cm, and 150 cm heights, respectively. It is shown that the positioning results fluctuate around the real value, especially at the height of 76 cm, and the accuracy can reach decimeter level. When the static target is at 0 cm,

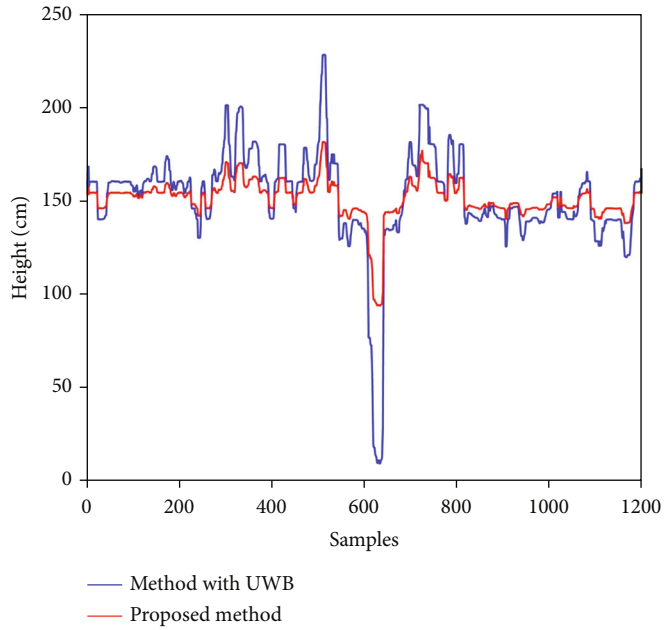


Figure 15: Height estimation between the proposed method and method with only UWB at 150 cm.

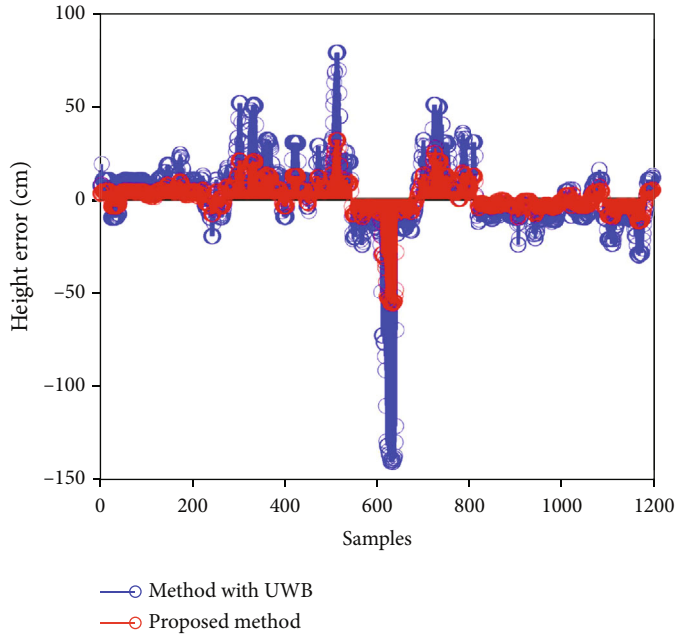


Figure 16: Height error between the proposed method and method with only UWB.

the estimated height error is 1.99 cm and variance is 1.15 cm. The estimated height error at 76 cm is 8.92 cm and its variance is 5.75 cm. When the target lies at 150 cm, the estimation height error is 16.39 cm and its variance is 4.25 cm. Table 2 depicted the estimated height and variance.

Meanwhile, two experiments for a moving target are conducted. One volunteer holds a tag at 130 cm height and walks around. Other volunteer holds the tag up and down. Figure 8 described the positioning results of the moving target. It is shown that the positioning accuracy reaches 30 cm with only receiving UWB signals when the anchors are installed in a reasonable and open topological environment.

Table 4: Estimated height results at 0 cm.

Methods	UWB	UWB-Baro-EKF (proposed)
Ave (cm)	151.77	150.677
Std (cm)	0.799	0.319

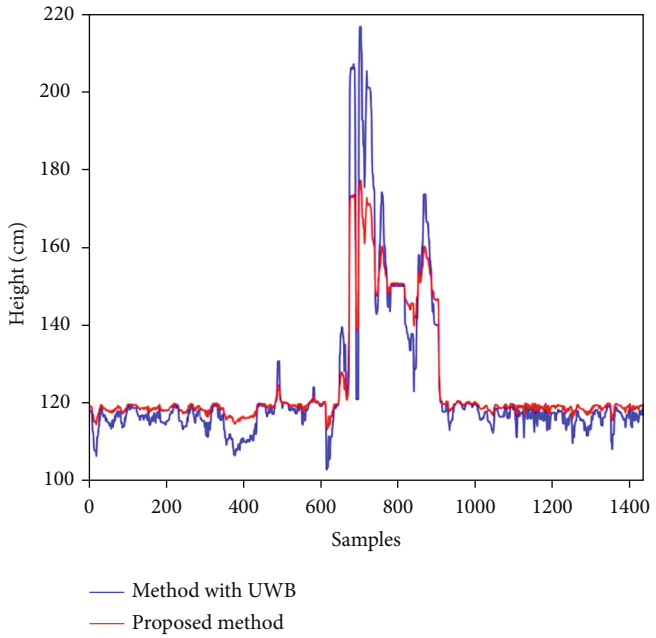


Figure 17: Height estimation between the proposed method and method with only UWB at different height locations.

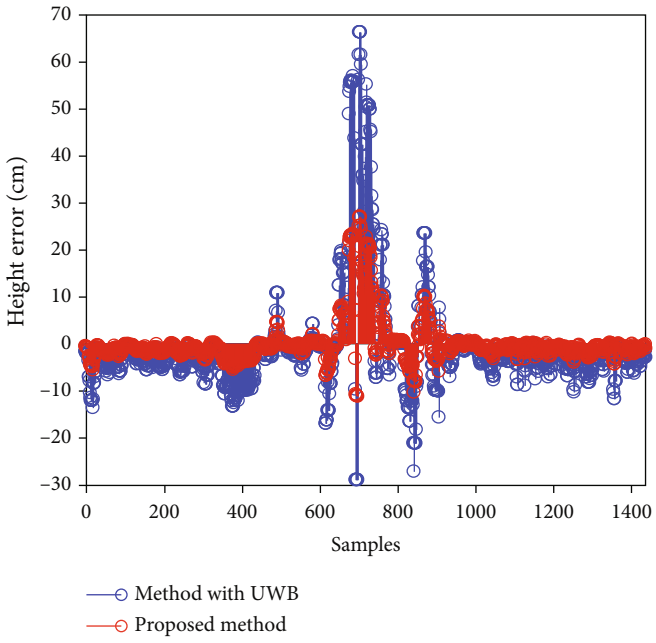


Figure 18: Height error between the proposed method and method with only UWB at different height locations.

However, when the target is located in an unreasonable topology, the estimated height will occur on outliers, especially for 1-2 meters shown in Figure 8. The estimated height only used UWB technology will be deteriorated.

3.3. Barometer-Assisted UWB Height Estimation. Assuming the simulated height estimates from UWB is a zero-mean Gaussian with standard deviation, we randomly generate the estimated height with zero-mean and 20.06 cm variance as shown in Figure 9. The barometric pressure sensor (Swiss) MS5611 is located at 0 cm height, and its measured data are shown in Figure 10. Figure 11 depicted the estimated height with only barometer pressure sensor.

Following the proposed pseudo reference update and EKF fusion, the standard deviation of estimated height reduces from 20.06 cm to 10.69 cm as Figure 12. Table 3 includes the results of all three sets, indicating the significant accuracy improvement.

In the robustness test, we consider errors over 1 m from single sensor as outliers and add such outliers randomly into

UWB simulation data. Figure 13(a) shows the UWB data with single outlier. Figure 13(b) shows the estimated height result with our proposed method. Figure 14(a) shows the UWB data with synthetic outliers. Figure 14(b) shows the estimated height result with our proposed method. It is evident that output data are more concentrated, and outliers are restrained.

Experiments show that the proposed method improved the height estimation performance and solved the problem that height positioning cannot be realized on some occasions. The short-term stability of barometer is fully utilized to assist UWB height estimation. Our method can suppress the abnormal value of height estimation in pure ultrawideband three-dimensional positioning, whether it is an isolated outlier or a continuous outlier. The requirement for frequent calibration of two differential pressure stations in traditional differential altimetry is avoided.

In addition, to verify the effectiveness of the fusion algorithm in the actual environment, the target located in the same height and different heights, respectively, are tested. First, the target is located at 150 cm in height. The height estimation and error are shown in Figures 15 and 16.

Figure 15 demonstrates that the positioning accuracy adopted by the proposed method is more improved than the single UWB method. When the samples are about 650, continuous outliers appear in UWB, and the proposed method can effectively restrain the outliers. Figure 16 shows the height errors between the proposed method and the single UWB method.

Table 4 depicts the average estimated height and variance with the proposed method and only UWB.

We also test the different heights of the target experiments. First, the target is installed at 120 cm, then walking by hand at 150 cm height, and then placed at 120 cm. The height estimation and error in a state of moving are shown in Figures 17 and 18.

The results show that the height accuracy of the proposed methods is better when the target is located in the same high position. However, when the target transfers from one altitude to another, the positioning accuracy will deteriorate, and outliers will appear with only UWB technology. Through the proposed method, the height estimation can keep up with the target position. The proposed method has better performance than the single UWB method.

4. Conclusions

This paper proposes an altitude estimation method, which is composed of UWB system and barometers. It can make a reliable and accurate estimation. In this paper, some experiments' data have proved that this method can enhance height direction measurement accuracy. In addition, a dynamic indoor pressure reference is taken into account and has been proven capable of providing more accurate sensing in complex indoor situations. The improved robustness of the EKF also has some limitations. However, it is still crucial to reliably detect and remove outliers from UWB raw data.

Data Availability

After discussion between all authors, the experimental data are temporarily unavailable. If readers want to study this field further, readers can ask the author for data.

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. 61772149, 61936002, 2018AAA0100300, 6202780103, 62033001, and 62061010) and Guangxi Science and Technology Project (Nos. AD18216004, AA18118039, 2019GXNSFFA245014, 2019GXNSFBA245072, and AD18281079).

References

[1] A. Zhen-Peng, S. Hu-Lin, and W. Jun, "Classify and prospect of indoor positioning and indoor navigation," in *2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*, pp. 1893–1897, Qinhuangdao, China, 2015.

[2] Z. Sahinoglu, S. Gezici, and I. Guvenc, *Ultra-Wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols*, Cambridge University Press, 2008.

[3] Y. Wu, S. Ding, Y. Ding, and M. Li, "UWB base station cluster localization for unmanned ground vehicle guidance," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6639574, 23 pages, 2021.

[4] A. Ren, F. Zhou, A. Rahman, X. Wang, N. Zhao, and X. Yang, "A study of indoor positioning based on UWB base-station configurations," in *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 1939–1943, Chongqing, China, 2017.

[5] V. Renaudin, M. Ortiz, J. Perul et al., "Evaluating indoor positioning systems in a shopping mall: the lessons learned from the IPIN 2018 competition," *IEEE Access*, vol. 7, pp. 148594–148628, 2019.

[6] J. D. Hoi, F. Dijkstra, H. Luinge, and T. B. Schon, "Tightly coupled UWB/IMU pose estimation," in *2009 IEEE International Conference on Ultra-Wideband*, pp. 688–692, Vancouver, BC, Canada, 2009.

[7] Q. Fan, Y. Wu, J. Hui, L. Wu, Z. Yu, and L. Zhou, "Integrated navigation fusion strategy of INS/UWB for indoor carrier attitude angle and position synchronous tracking," *The Scientific World Journal*, vol. 2014, Article ID 215303, 13 pages, 2014.

[8] Y. Son and S. Oh, "A barometer-IMU fusion method for vertical velocity and height estimation," in *2015 IEEE SENSORS*, pp. 1–4, Busan, Korea (South), 2015.

[9] R. Elbakly, M. Elhamshary, and M. Youssef, "HyRise," *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, vol. 2, no. 3, pp. 1–23, 2018.

[10] K. Muralidharan, A. J. Khan, A. Misra, R. K. Balan, and S. Agarwal, "Barometric phone sensors: more hype than hope!," in *Proceedings of the 15th workshop on mobile computing systems and applications*, pp. 1–6, Santa Barbara, California, 2014.

[11] B. Li, B. Harvey, and T. Gallagher, "Using barometers to determine the height for indoor positioning," in *International Conference on Indoor Positioning and Indoor Navigation*, pp. 1–7, Montbeliard, France, 2013.

[12] Z. Xu, J. Wei, J. Zhu, and W. Yang, "A robust floor localization method using inertial and barometer measurements," in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–8, Sapporo, Japan, 2017.

[13] V. Chueh, T. Li, and R. Grethel, "INS/Baro vertical channel performance using improved pressure altitude as a reference," in *2008 IEEE/ION Position, Location and Navigation Symposium*, pp. 1199–1202, Monterey, CA, USA, 2008.

[14] H. Ye, T. Gu, X. Tao, and J. Lu, "SBC: scalable smartphone barometer calibration through crowdsourcing," in *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pp. 60–69, London, Great Britain, 2014.

[15] M. Tanigawa, H. Luinge, L. Schipper, and P. Slycke, "Drift-free dynamic height sensor using MEMS IMU aided by MEMS pressure sensor," in *2008 5th Workshop on Positioning, Navigation and Communication*, pp. 191–196, Hannover, Germany, 2008.

[16] D. E. Bolanakis, "Evaluating performance of MEMS barometric sensors in differential altimetry systems," *IEEE Aerospace and Electronic Systems Magazine*, vol. 32, no. 9, pp. 34–39, 2017.

[17] S. Hyuga, M. Ito, M. Iwai, and K. Sezaki, "Estimate a user's location using smartphone's barometer on a subway," in *Proceedings of the 5th International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments*, pp. 1–4, Seattle, Washington, 2015.

[18] D. E. Bolanakis, K. T. Kotsis, and T. Laopoulos, "A prototype wireless sensor network system for a comparative evaluation of differential and absolute barometric altimetry," *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 11, pp. 20–28, 2015.

[19] D. Banerjee, S. K. Agarwal, and P. Sharma, "Improving floor localization accuracy in 3D spaces using barometer," in *Proceedings of the 2015 ACM International Symposium on Wearable Computers - ISWC '15*, pp. 171–178, Osaka, Japan, 2015.

[20] Y. Xu, G. Tian, and X. Chen, "Performance enhancement for INS/UWB integrated indoor tracking using distributed iterated extended Kalman filter," in *2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS)*, pp. 1–5, Wuhan, China, 2018.

[21] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9786547, 1 page
https://doi.org/10.1155/2023/9786547

Retraction

Retracted: Optimization of Intelligent Management and Monitoring System of Sports Training Hall Based on Internet of Things

Wireless Communications and Mobile Computing

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

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. Qian, "Optimization of Intelligent Management and Monitoring System of Sports Training Hall Based on Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1465748, 11 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 1465748, 11 pages
https://doi.org/10.1155/2021/1465748

Research Article

Optimization of Intelligent Management and Monitoring System of Sports Training Hall Based on Internet of Things

Hui Qian

Ministry of Sports, Sanquan College of Xinxiang Medical University, Xinxiang, Henan 453003, China

Correspondence should be addressed to Hui Qian; 11571092@sqmc.edu.cn

Received 13 April 2021; Revised 21 June 2021; Accepted 26 June 2021; Published 7 July 2021

Academic Editor: Wei Wang

This paper provides an indepth analysis and research on the intelligent management and monitoring system of sports training hall through the Internet of Things. Firstly, it introduces the project background of the sports training base stadium service system, analyses the significance of the system development, and describes the related technology and the method of development. The paper then analyses the requirements of the management business of the system and describes the main business processes of the system with activity diagrams; the paper then analyses the functional requirements and data requirements of the system and describes the system with use case diagrams and thumbnail class diagrams, respectively. Based on this, the thesis designs the overall aspect of the system, giving the corresponding total package diagram, and designs all functions one by one, giving the corresponding class diagram, the corresponding sequence diagram, and the corresponding processing flow diagram. The thesis gives the design of database tables by designing entity property diagrams and entity-to-entity relationship diagrams according to the data requirements and system functions. Finally, the research of the thesis is summarized accordingly, stating what has been done, what results have been achieved, what problems still exist, and what further work needs to be done to make the outlook of the work.

1. Introduction

With the development of networks and the progress of science and technology, people began to pay attention to the interaction between things, which evolved to form the Internet of Things (IoT). At present, IoT industry is still in the primary stage of development, and there is still a lot of space for upward and transformation in terms of technological innovation, standard development, and business forms. The Internet of Things industry is the key development of new strategic industries, and based on this environment, the application of the Internet of Things has a very large potential and space [1]. In recent years, IoT technology has gradually matured, focusing on sensors, software, and other aspects, while the rapid development of supporting equipment for IoT, especially intelligent circuits, transmission networks, and other basic equipment. In terms of the current situation, IoT has a wide range of applications, commercial, agricultural, and service industries, playing an extremely important role in urban construction, environmental protection, urban security, intelligent transportation, etc. This also brings new development opportunities and challenges in various fields, especially in the sports industry, and the use of advanced technology of IoT can bring great technological innovation that can be used to promote the development of sports events [2]. Nowadays, IoT technology has made considerable achievements in the fields of smart cities and the smart industry. Some researchers and entrepreneurs have noticed the great potential of IoT in the field of sports, such as smart stadiums and smart sports apparel. Rowing, as a competitive sport established by the Olympic Games, is mostly played in natural waters, which can improve the cardiovascular and respiratory functions of the human body and play a role in strengthening the body, and is loved by the public [3]. Therefore, there is still a gap in the market for intelligent rowing training equipment that can be filled.

The development of sports events is in the stage of rapid development, but the development of information technology is still relatively lagging, most of the sports event operation only stays in the traditional operation mode, the

organization capacity is limited, the overall operational efficiency is not high, the resource consumption is huge, and these problems seriously restrict the sustainable development of sports events [4]. In the background of the development of Internet information technology, the development of sports events relying on information technology has become a trend, the importance of resource integration and optimization of the whole process of event management has become inevitable, and the Internet of things plays an important role in the field of sports events [5]. This paper selects the application of IOT technology in sports events as the research object, its purpose is to explore the current field of sports events in China, the practical application of IoT technology, and find its shortcomings, and then put forward the corresponding improvement measures, and its research results for all kinds of enterprises and institutions, government agencies, and other institutions engaged in sports events related industries in the research and development of IoT technology, application research, and management personnel [6]. The research results will provide some reference value for various enterprises, government agencies, and other institutions engaged in research and development, application research, and management of IoT technology in sports event-related industries [7].

As an important part of the sports industry, sports events also show that people have the spirit of continuous innovation and challenge, and modern sports events are one of the important symbols of current social progress, civilization development, and economic strength. In today's social life, the process of economic integration and the rapid development of information technology make the combination of the Internet of Things and sports become an important development direction for the development of the sports industry, and "Internet +" sports continue to ferment big data, VR, artificial intelligence, etc. which will be deeply integrated with sports. In the management of large sports events, the full use of IoT technology, to further enhance the level of sports events, to drive the competition management, information management, venue management, and other event-related industrial progress, as well as to promote the construction of sports culture and other aspects, has important practical significance. The research on the application of IOT technology in sports event management is relatively small, and China's research on the application of IoT in this area is still in a blank stage. This research on the application of IOT technology in sports events has yielded theoretical results, which has certain theoretical significance to fill the gap in this area of research.

2. Related Work

In terms of essential properties, the primary function of large stadiums should be to provide venue space, where the main bearing objects are sports and performing arts activities [8]. Based on the historical background of the development of sports events, whether the state invests in the construction or the society raises funds to build large stadiums, the original purpose of the service includes meeting the demand of watching sports competitions [9]. Javed et al. focus on the public service function of large stadiums from six aspects: fitness facilities, sports organization, physical fitness monitoring, fitness guidance, sports activities, and sports information services [10]. Aithal et al. believe that public stadiums should have six basic functions, such as education, competition, training, viewing, testing, and leisure [11]. Kharel proposed that in nature, the basis of its cognition needs to be based on objective experience, with natural science as the bridge; therefore, drawing on this idea, the research perspective of this paper is to take empirical evidence as the starting point, combining their own practical experience and theory, focusing on the application and development of IoT technology in the industrial field, in the application of the industrial field that needs to recognize the advantages of IoT technology [12]. The convenience and openness of enterprise products as well as services which cannot only promote the further development of industry but also effectively enhance the competitiveness of products [13].

The application of IOT technology to related industries can also reduce the cost of production materials, which is of great significance to the development of industry, and in the future, it is also necessary to continuously increase the depth and breadth of the application of IoT [14]. On the current situation of the application of the Internet of things, the application of Internet of things technology is mostly in the actual production of the application; so, the Internet of things technology based on relevant research needs to strengthen the combination of research and development of technology and industrial development [15]. At present, the application of IOT technology is mostly in the city, and the application in the countryside is very little, which requires government coordination and development. The government needs to increase the investment in the research and application of IoT technology, to develop the application of IoT technology in a balanced way, to regulate the balance of urban and rural development, and to actively develop the application of agricultural IoT technology [16]. In the business requirements, the business aspects are described accordingly, and the main corresponding business processes are drawn. In the functional requirements, a detailed analysis of the use case diagram with the business functions is done through the analysis of roles [17]. In data analysis, the required data is described. In the nonfunctional requirements, the corresponding environmental, performance, and security requirements are described. In the detailed design, the corresponding submodules are designed in functional detail with corresponding package diagrams, corresponding class diagrams, corresponding sequence diagrams, and corresponding activity diagrams, and the database tables are designed in detail with entity attributes and entity relationships.

In summary, although the academic research on IoT technology is relatively active with many published articles, there are only a few articles that apply IoT to sports events, and many important core factors have not been covered. Therefore, in this paper, we want to study all aspects of IoT technology indepth and analyse the importance and significance of IoT technology in sports event management from national policies, residents' ideology, and practical actions. In our country, the application of IOT technology in sports

event management, although the application of IoT in sports events is not yet widespread, in terms of research content, especially in the context of the rapid development of information technology, the proportion of IoT technology penetration into sports events will be increasingly important; therefore, the research on the application of IOT technology in sports events is particularly important. The construction of a sports training base service management information system cannot only enhance the utilization rate of sports venues and improve the working style of staff but also respond to the service status of the venues in real-time, effectively manage the members, save the cost of venue services, and make the venue services run efficiently, thus improving the economic benefits of the base venues.

3. Optimization Analysis of Intelligent Management and Monitoring System of Sports Training Hall with the Internet of Things

3.1. IoT-Based Sports Training Hall Design. Broadly speaking, the Internet of Things (IoT) emphasizes the use of bar codes, radio frequency identification (RFID), sensors, global positioning systems, laser scanners, and other information sensing devices, following the agreed protocols, to achieve human and human, human and object, and object and object in any time and any place connection (anything, anytime, anywhere), and then can exchange information and communication, to achieve intelligent identification, positioning, tracking, monitoring, and management of a huge network system. The Internet of Things (IoT) is a network in which ordinary objects are interconnected through the medium of the Internet and telecommunication network consultation with the support of new media and technology. As for the problem of design theory in social sports guidance, the IoT technology will be used as an effective path for effective guidance with the help of IoT, because IoT can effectively solve the problem of information collection and utilization rate underground, can realize offsite network connection, realize intelligent management service of multiple sports resources, and implement management and automatic detection in a faster and more convenient way, so that people can participate more conveniently [18]. Internet is the basic element of the IoT technical architecture, there is indispensability, under the loop of the Internet to ensure that the information flow collected by the object is safe, accurate, real-time transmission is the basic guarantee, especially in 5G began to emerge today, and information security is heavy.

The other two important applications of IoT are sensors and intelligent information data processing, which are captured by various identification technologies from the sensor side and then analyzed in the data service management side, and the data is processed and processed to classify, store, and feedback the effective data to meet the different needs of different users. Through the application of computer technology, IoT can control massive data in a timely and effective manner to achieve the connection and communication between people and things, and things and things. The perception layer as the name implies is to perceive things, and it includes such as RFID tags and readers, cameras, infrared, two-dimensional code tags, and another perception levels. In the Internet of Things to identify things, to screen things belongs to the perception layer, which is an important core technology to achieve universal perception in the Internet of Things, and is one of the most important external functions in the Internet of Things architecture [19]. The most core and key inside is the coding layer, only through coding, so that each object has a sticky note of its own, to screen objects effectively and quickly, to achieve rapid identification, and make a choice, as shown in Figure 1.

Through the Internet or communication technology, this will be used as a medium to screen objects in the perception layer and achieve interaction through the network layer. For example, Internet, broadband network, network management system, cloud computing platform, and other components, the infrastructure to achieve the Internet of things is wide coverage of the mobile communication network, at present, with the popularity of the 5G era, the future development of the Internet of things will likely become mainstream, and people's lives will become more intelligent and will also drive a large number of enterprises and people from all walks of life to participate in the Internet of things, to promote the Internet of things technology faster. Simply put, after the previous encoding, each object has its business card; in the network layer, the object will be network communication interoperability; then, how to carry out the effective implementation, this is the application layer involved, the realization of its presentation in front of people's eyes, and the most intuitive is also the most familiar to us which is the supermarket shopping, Taobao buy something coordination letter, etc. It can be effectively based on industry needs, targeted to combine ten intelligent applications of the Internet of Things. The development of the Internet of things is to provide a variety of applications, the use of Internet of things technology, industry information technology needs, fully improve the integration of different industries, the development of high-quality low-cost related information resources, and the use of program solutions to provide effective business models and data security.

NB-IoT is a cellular network connectivity technology designed for IoT, which can be deployed directly on existing 2G, 3G, and 4G networks to achieve a smooth transition and upgrade while reducing deployment costs. The performance comparison of the above two solutions is shown in Table 1.

In Table 1, we just compare some parameters of NB-IoT to our IoT. In summary, both modules can meet the functional and performance requirements of this system. The advantage of using the LoRa module as a communication module is that the communication band is free, no need to be restricted by operators, the signal is only good or bad related to the separate deployment of the gateway, its controllability is greater, the data transmission and reception do not need to go through the operator network, and the privacy is better. However, the LoRa module adopts a self-assembled network structure and builds its gateway, which is unique in dealing with the signal interference and network overlap but cannot provide the same quality of service as the NB-

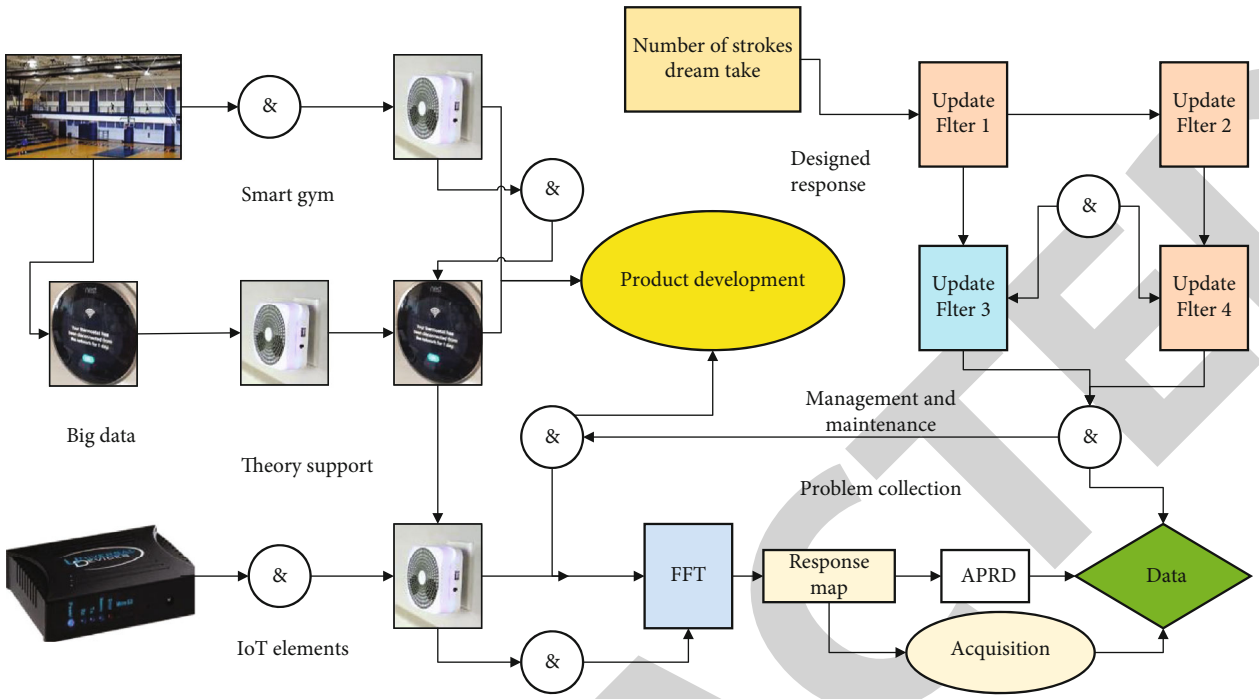


Figure 1: Framework of IoT-based sports training hall.

Table 1: NB-IoT vs. LoRa module performance comparison.

Long-range IoT technology	NB-IoT	LoRa
Transmission distance	<22 km	<14 km (city 1-2 km)
Networking method	Based on existing cellular network	LoRa-based gateway (self-organizing network)
Frequency band	Carrier band	150 MHz to 1 GHz
Speed	<100 kbps	0.3-50 kbps
Number of connections	200 k/cell	200 k-300 k/hub
Continuity	Approx. 10 years/AA battery	Approx. 10 years/AA battery
Network deployment	Node	Gateway + node

IoT module, and the selection of LoRa gateway location, deployment method, and postmaintenance is also more difficult. Finally, after analysis and comparison, the BC26 model of the NB-IoT module was selected as the wireless communication module for this system.

The resistance value of the strain gauge changes with the ambient temperature. Since the temperature change is the same as the strain gauge coefficient, the resistance value changes by the same amount. At this time, the output voltage value of the bridge is unaffected or less affected, i.e., the output of the bridge is related to the measured strain, thus acting as temperature compensation.

$$U_0 = Vcc \cdot \left(\frac{R_1}{R_1 - R_2} + \frac{R_3}{R_3 - R_4} \right). \quad (1)$$

The Kalman filter algorithm for state estimation is implemented using a feedback control method. The Kalman filter is divided into two types of state equations: the time-updated state equation and the measurement-updated state equation. The time-updated state equation is an estimation equation that derives the current state variables and error covariance estimates based on the state estimates of the previous moment to estimate the state a priori for the latter moment. The measurement update state equation is a correction equation that combines the a priori estimates and the new measurement variables to construct improvements.

Decision-makers of relevant government departments and venue managers need to be cautious about the location of venues, combine the overall layout of the city to plan the construction of venues, based on facts such as resource scarcity and residents' needs, and cooperate with the surrounding infrastructure construction to give full play to the value of large sports venues in the city's functional services. It has been proved that the construction of large sports venues can further promote urban development and drive the rise of related industries around [20]. However, the return on the benefits of the stadium is a long-term business, often requiring a long period of business development to achieve its economic and social value. Therefore, managers should not be too hasty in the process of managing the venues but

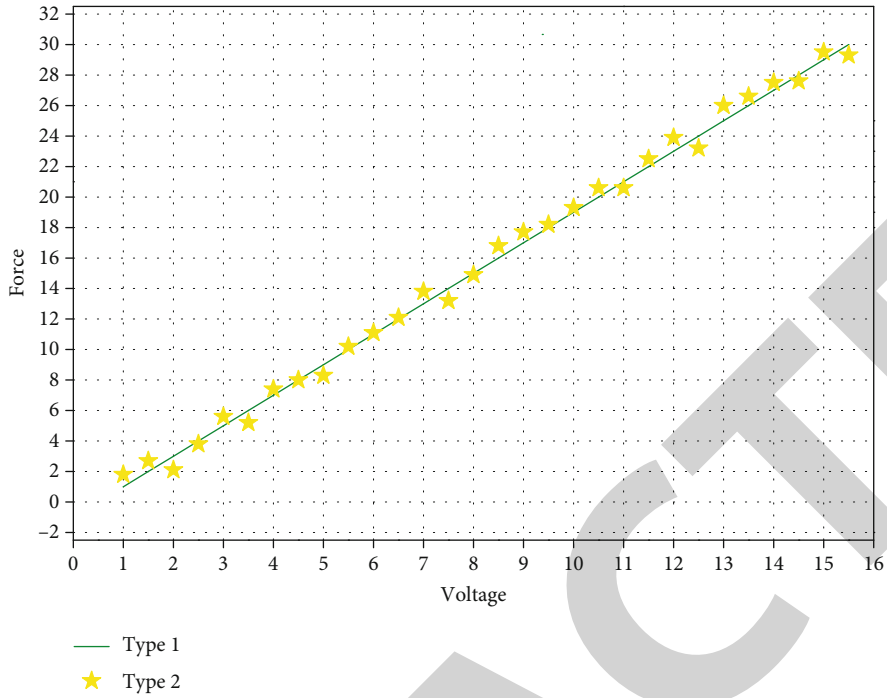


Figure 2: Relationship diagram.

should be steady and proactive in seeking market opportunities, attracting the public with quality and professional services, gradually gathering popularity, and thus driving the development of the surrounding industries of the venues. On the other hand, the convenience of public transportation around the venue as an external resource element indirectly affects the convenience of using the venue, as shown in Figure 2.

In Figure 2, type 1 refers to the relational curve of our IoT structure, and type 2 is the relational curve of the IoT section structure from the literature [8]. For the remote location of the venue, in hosting specific large-scale events and cultural performances, the venue management should actively coordinate with other relevant functional departments, to negotiate the possibility of increasing urban traffic. Also, the venue should take the initiative to sense the urgency of the people and think of the people's thoughts, such as the use of additional buses to pick up the audience and staff; in daily use, you can follow the practice of large shopping malls and entertainment centers, to provide scheduled shuttle buses to pick up residents, to provide as convenient as possible for the users of the venue, should also be an open space, reasonable planning, and price concessions for parking lots. It is also necessary to create a parking lot with wide space, reasonable planning, and favorable price to ensure that people who drive to the venue have no worries and make up for the inconvenience of the venue in terms of location and transportation.

The effective realization of the functions of the venue depends on the full consideration before construction, and the venue managers should pay attention to the functional planning and design of the venue. The planning and design of the venue include two situations before construction and renovation: for the planned new venues, full consideration should be given to the postgame utilization and functional transformation, so that the pregame construction is well thought out, the postgame utilization is in mind, and the transformation can be achieved through demolition to avoid secondary transformation as far as possible, saving human, financial, and material resources costs. For example, Baoneng International Sports and Performing Arts Center fully considers multifunctional use before construction: it has the strongest load-bearing ceiling in Asia, making it the first choice for various cultural and entertainment performances; the main hall is equipped with retractable seats, which is convenient for receiving audiences of different scales; the cement floor is equipped with drainage channels, and the ice room is connected below, which is convenient for the ground to be flexibly transformed into an ice rink to hold related events: the use of mezzanine, sinking, and other. The space design makes the arena have open space for parking, unloading, and loading, and it is convenient for the indoor operation of transportation and signal trucks. For the venues that need to be renovated, we should study excellent cases at home and abroad, listen to the opinions of experienced experts in venue operation and management, venue staff and users, etc., seek the best renovation plan from the needs and practical problems, focus on multiple uses of one venue, and build venues with high practicality and short renovation period.

3.2. Monitoring System Optimization Analysis. Hardware facilities are the necessary material conditions for venues to provide comprehensive functional services, and nowadays, to meet diversified social needs, large stadiums are gradually developing from "body-oriented" single services to "diversified co-existence;" thus, venue managers can guarantee the quantity and quality of sports. Therefore, the venue

managers can guarantee the quantity and quality of hardware facilities, under the premise of enriching the existing venue facilities. First, according to the needs of consumers, the transformation or temporary construction, fully develop the spare space of the venue, set ups a variety of sports facilities, to carry out a wealth of sports competitions, sports training, and other activities. Second, you can quickly change the use of the venue through site layout, lighting erection, etc., to undertake recreational performances and other social activities, and the venue will be used live and diligent, to enhance the efficiency of the use of the venue [21]. It should be noted that regular cleaning and maintenance of facilities and equipment protect the personal safety of consumers and ensure the quality of sports services. The establishment of an effective regulatory mechanism, focusing on the use of venues to monitor the use of venues to avoid abuse and overdevelopment, especially the use of professional competition venues, needs to carefully weigh the pros and cons and do not put the cart before the horse.

In this system, to calculate the optimal estimates of acceleration and angle values, the Kalman filtering algorithm is used to solve the weighting relationship between the estimated and measured values, to produce results that are closer to the true values. The Kalman filter algorithm has five state-step equations. The prediction estimation state equation is shown in Equation (2).

$$X(K|K+1) = A \cdot X(K+1|K+2) - B \cdot U(K+1), \quad (2)$$

$$P(K|K+1) = A \cdot P(K+1|K+2)A^T - Q, \quad (3)$$

where $X(K \mid K-1)$ is the result of the previous state prediction, $X(K-1 \mid K-1)$ is the optimal prediction at the previous moment, $U(K)$ is the control quantity of the present state, A is set equal to 1, and $U(K)$ is equal to 0 in this system, where $P(K \mid K-1)$ is the covariance of the $X(K|K\text{-}1)$ state values, $P(K-1 \mid K-1)$ is the covariance, and Q is the system noise. The equation is updated as shown in Equation (4).

$$X(K|K+1) = P(K-1|K) \cdot Z(K), \quad (4)$$

$$K_g(K) = P(K+1|K) \cdot (P(K+1|K) - R), \quad (5)$$

$$P(K|K-1) = \left(1 + K_g(K)\right) \cdot P(K+1|K). \quad (6)$$

The application of IoT technology is to put new high-tech products and thinking into sports events in the development and construction and operation management stages of sports events. It is effective to control costs, improve efficiency, and enhance service quality. With the advent of science and technology and the Internet era, information perception and information technology such as the Internet of Things, cloud computing, the Internet, and mobile Internet have become the basis of industrial development. The main purpose of sports events is to "provide efficient and secure management for users with the high standard of experience." The development and management of sports event resources use the necessary information technology and means to improve the utilization rate and service quality of sports events through the application of the Internet of Things in the management and development of sports event resources and the analysis of data collected through technology related to sports events to accurately grasp the problems in the process of sports event services. In the management of sports events, the construction of intelligent venues is based on the application of IoT practices in the sports industry important construction. In the intelligent management of sports venues, the information collected can effectively optimize as well as adjust the evaluation index of venues, innovate the content and way of venue service management, promote the standardization of venue services, and promote the construction of national fitness service body and the improvement of policies and regulations [22–26]. The venue center introduces strong Internet technical support, which is combined with the rich front-line operation and management experience of the venue team to promote intelligent services. Integrating mature information technology in all aspects of venue operation and management, the content covers functions such as supervision of national fitness services by government departments, management of large-scale activities at venue sites, management of fitness membership data at venue sites, management of financial assets operation at venues, management of daily operation of venues, sales of goods around venues, and management of sales of goods by partners. In the context of "Internet of things technology Under the background of Internet of Things technology," we build "intelligent venues," improve the intelligent operation management system, make the Internet penetrate all corners of the stadium, and at the same time collect a large amount of data information for analysis, such as booking management data of fitness venues, ticketing and registration information of large activities, cell phone, the big data collection of the client, and website click traces, as shown in Figure 3.

Air membrane stadiums are being used increasingly frequently for their advantages of cost reduction, cleanliness, and energy saving. Air membrane building is a kind of building structure system that uses special building membrane material as the shell, equipped with a set of intelligent mechanical and electrical equipment to provide positive air pressure inside the air membrane building to support the main body of the building, which applies to three ways, such as new stadiums, renovation of outdoor venues, and energy-saving renovation of indoor venues. An air membrane system is built on an existing outdoor stadium site to transform the original outdoor site into an indoor stadium. This allows the renovated venues to be used in different seasons and under various weather conditions, thus improving the sports environment and increasing the opening hours of the venues. And the indoor arena is retrofitted with air membrane energy saving. By reducing the cooling and heating space of the original arena, the cooling and heating costs are significantly reduced, and the total energy costs of the entire arena are significantly reduced. Air membrane stadiums can be used for most sports, such as tennis, basketball, swimming, and hockey. There are many successful cases to learn from, and the construction of an air membrane stadium within 10,000 square meters takes only about 15 days. Modularized intelligent automatic control can be achieved. Research shows that

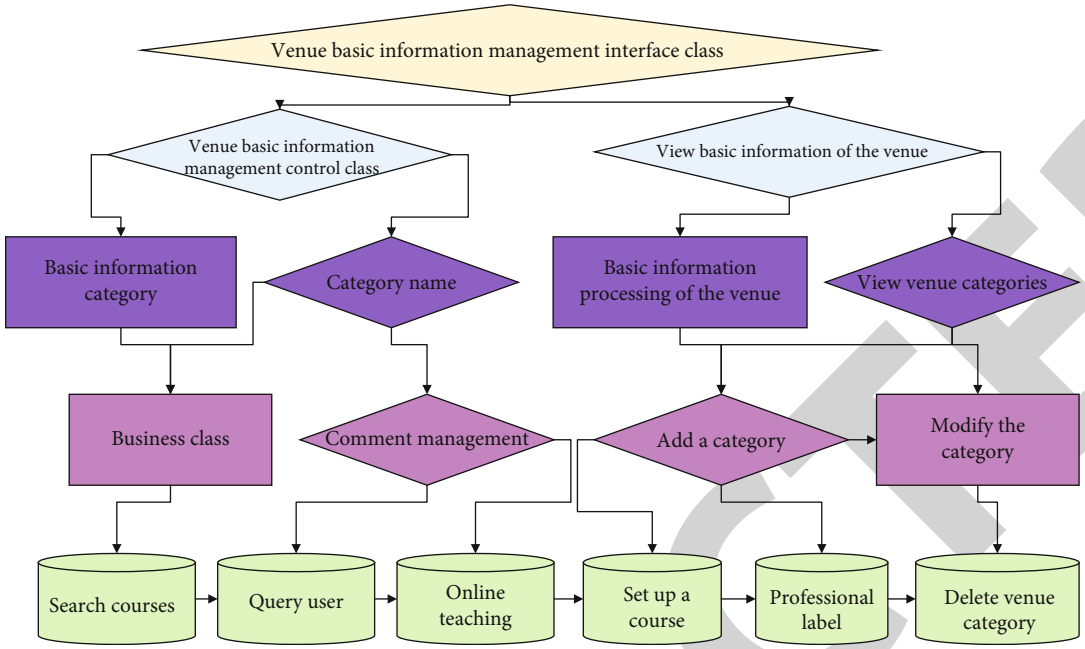


Figure 3: Management monitoring optimization design.

the maximum life span of air film stadiums can reach 30 years, with a maximum span of 130 meters and no beams or columns inside. In the patented anchoring system, easy to disassemble, the whole building body is easy to relocate and reusable. The air membrane arena has unique advantages in purification treatment.

4. Analysis of Results

4.1. System Performance Test Result Analysis. The following simulation plots are obtained by MATLAB, which mainly analyze the trend of the number of sensors as the relay distance, sensor distance, and transmit power vary. The power allocation factor is discussed in the previous section to influence the total system throughput (Sum-throughput) during the study; so, the simulations in this section are performed with the optimized parameters.

The effect of different distances between S and R on the total throughput is represented in Figure 4. For comparison, we also perform the average power allocation (MPA, mean power allocation) protocol (MPA-MRC) based on MRC, the average power allocation protocol (MPA-RT) based on indirect transmission (RT, relay transmission), and the direct transmission (DT, direct transmission) based on mean time allocation (MTA) protocol (MTA-DT). As can be seen from the trend of the curves in the figure, there exists a unique d that allows the system to achieve the maximum throughput. As d increases, the closer to d, the greater the system throughput. This is because the relay position gradually approaches the optimal position, and the energy collected by the relay node decreases, but at this time, the system is still limited by the energy collected by the sensor, the greater the channel gain between the relay and the sensor; so, the system throughput improves; in contrast when d is greater than d, the closer to the sensor node, the smaller the channel gains

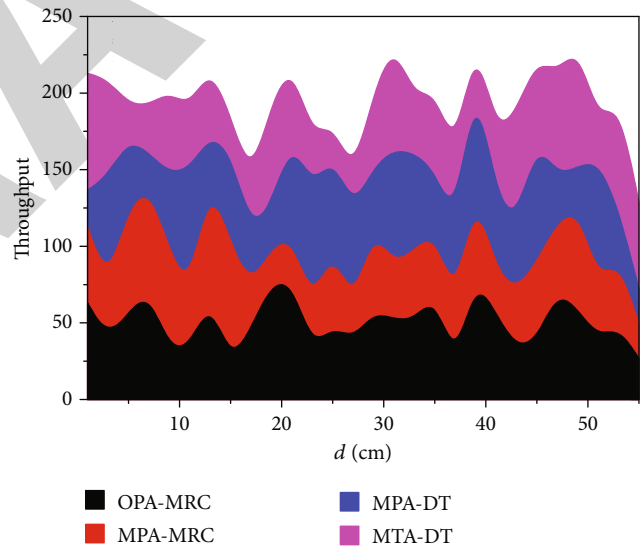


Figure 4: Effect of relay location on total throughput.

between the source node and the relay; at this time, the system is limited by the energy collected by the relay; so, the throughput decreases. The coincidence curves show that for all three schemes except the direct transmission method, the total throughput is the same for the sensor energy-constrained case. Notably, the OPA-MRC scheme consistently has the best total throughput, while the other three schemes are worse, and the performance gap becomes more pronounced as d increases.

In the management of sports competitions, there are loopholes in the rules, and active defeat has become a very common phenomenon, but the pursuit of sports competitions is "fair and just competition," but it is far from enough to rely on sports rules to restrain, which also requires good

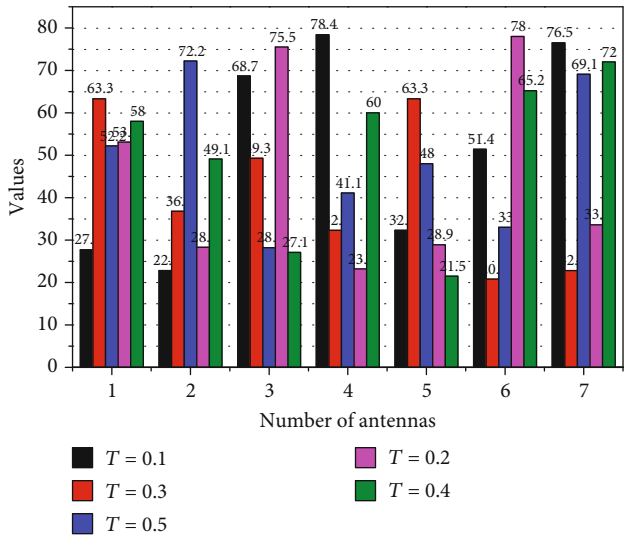


Figure 5: Schematic of the average throughput of a single sensor with T.

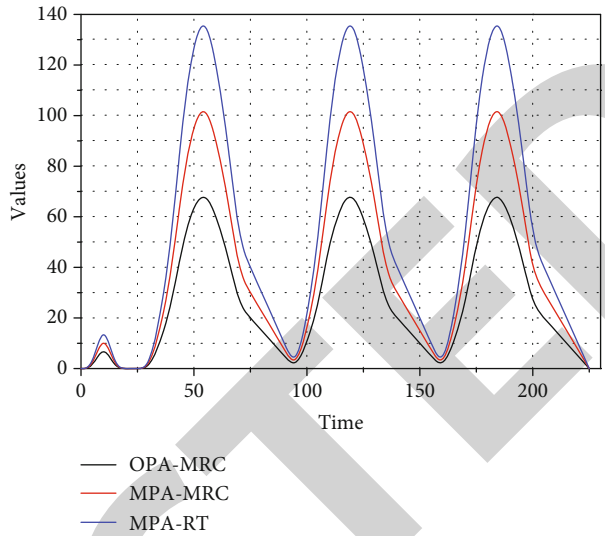


Figure 6: Optimal throughput curve with sensor position id.

professional ethics and sportsmanship of athletes, coaches, and referees. The application of the Internet of Things can improve this situation, the Internet of Things technology can transmit the specific information on the field of play to the computer, whether the athletes have the initiative to seek defeat can be understood and mastered, and the important development goal of the application of the Internet of Things in sports events is to improve the environment of sports competition, comply with the Olympic competition, create a fair competition environment, and promote the healthy development of sports, as shown in Figure 5. Sports events are inseparable from the venues and venues, and the venues of sports events are an important condition for the development of sports events. The application of the Internet of Things makes the information construction of sports events' venues more rapid. At the same time, the state has given a lot of support and help in the information construction of sports venues, and there are more than 10 high-level integrated training venues.

In recent years, many large sports events have been held, and in practice, it has been found that the awareness of comprehensive fitness helps to improve the level of competitive sports. The construction of sports event venues is also being enhanced, and Internet technologies such as cloud computing technology, trinetwork convergence technology, Internet of Things technology, mobile Internet technology, and new flat panel display technology are constantly being integrated into the construction of sports venues, which are dedicated to creating intelligent venues. From the current situation of stadium construction, the construction of intelligent venues is still in the primary stage, and there is still a lot of upside in technology and application conditions, as shown in Figure 6.

When the number of sensors is fixed, the four transmission protocols are also increasing with the increase of the sending power of the source nodes, this is because increasing the sending power of the source nodes, the energy collected by the sensor nodes and relay nodes will increase, and this will increase the power of the sensor nodes and relay nodes in sending the information, leading to the increase of the system throughput. And for the same number of users, the power optimization algorithm in this paper has a greater improvement in the total system throughput compared to the average power algorithm.

4.2. Analysis of Optimization Results. Locust is a distributed user load testing tool, based on python language to write test scripts to load test the server to determine the system that can determine the average response time of the server to user access. Taking the login function as an example, the test URL is http://59.110.142.44:8080/new2/phone/Login, and the user makes 200 cumulative requests to the server with an average response time of 30 ms and the fastest response time of 29 ms with 0 failures. The test results are shown in Figure 7. Policy factors and location traffic contain all the secondary elements which are important elements that play a role in the realization of the integrated functions of the two types of venues; at the same time, the managers of both types venues believe that safety and risk control and the overall quality of managers have an important role in the realization of the integrated functions of the venues. Safety and risk control should be given full attention as a service content of the venue; without the guarantee of safety, the realization of any function is empty talk, the results show that most managers attach great importance to safety and risk control factors, and only a few managers need to strengthen the awareness of safety management. In Figure 7, the graphical nature on the right shows the corresponding comparative trend of the servers.

Managers as the navigator of the venue operation, their advanced management consciousness, outstanding working ability, and efficient working style, will be beneficial to the realization of the comprehensive functions of the venue. Institutions are more concerned about sports event resources, while corporate units are more interested in cultural and performing arts resources, which are also closely

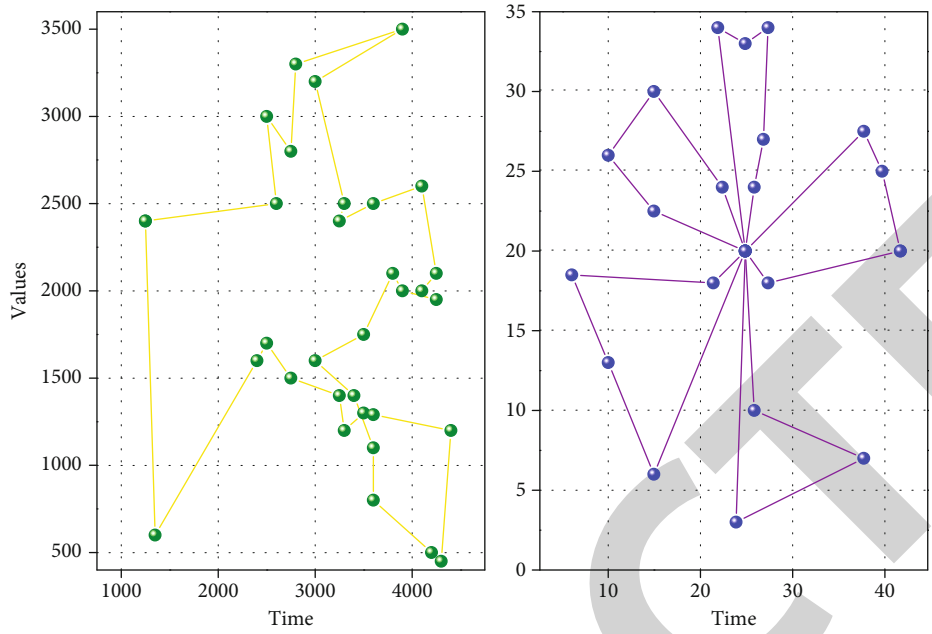


Figure 7: Server average response time graph.

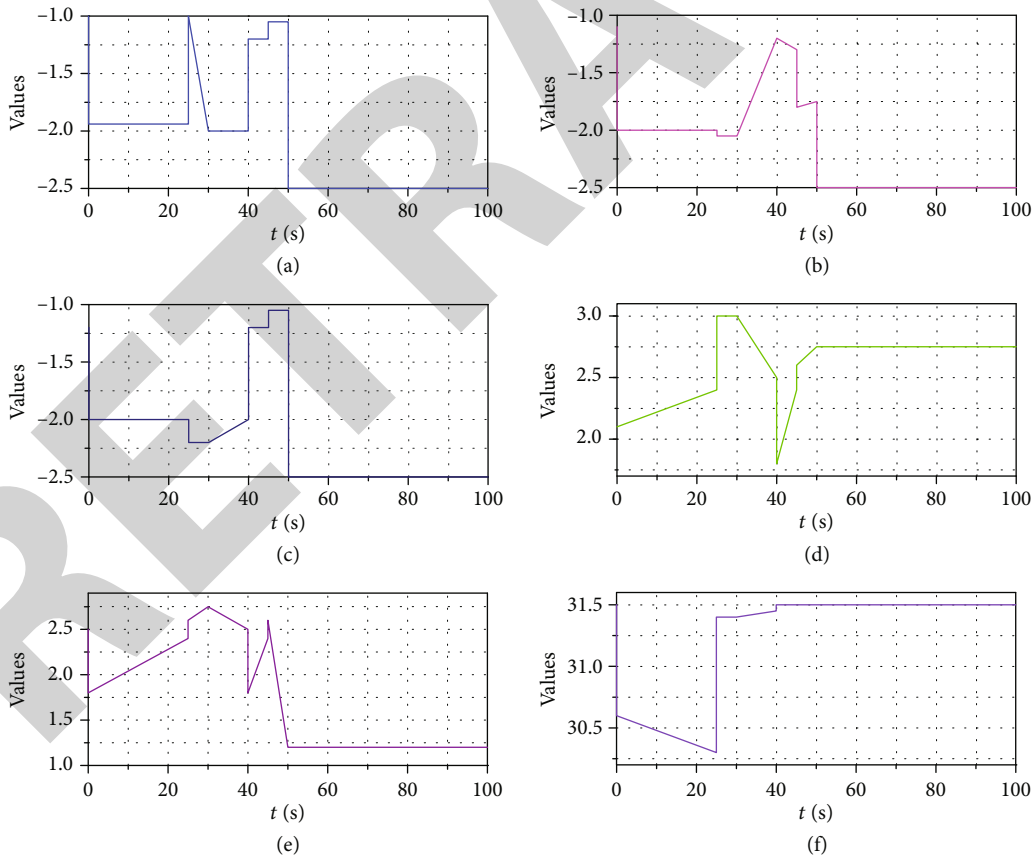


Figure 8: Data analysis chart.

related to the management mode and functional positioning of the venues. At present, the number of cultural and performing arts activities undertaken by corporate-type large venues is significantly higher than that of sports events, and institutions are mainly providing sports services; however, the singularity of functional services will not be conducive to activating the idle resources of the venues and hinder the maximum benefits of the venue, as shown in Figure 8.

The logic of the current system is relatively simple, and the functional complexity of the system is relatively single. If later faced with many collections equipment access, the rowing training system server needs a more efficient framework and logic processing strategy, to ensure the stability and reliability of the connection communication between the oar collection terminal and the rowing training system server, the server, and the handheld terminal, the three. The data in the current system rowing training system server only carry out the simple calculation, processing, and analysis work. In the future, the system can refer to the method of big data analysis, build a professional algorithm model, and analyze the collected sports data comprehensively. Thus, it can provide more professional guidance for rowers.

5. Conclusion

The application of IoT technology has well improved the level and efficiency of sports event management. The application of IOT technology in sports event competition management, event information management, and venue management has built a digital management system and organization, which has comprehensively promoted the level of sports event management, well met the needs of sports event spectators, and provided athletes and coaches with more accurate information and data, so that they can play at a higher level on the field. IoT technology is an innovative change for sports event management, which is very important for the construction of informationization and modernization of sports event management. Sports training base stadium service management system improve the efficiency and make the work simple words, and this paper starts from the actual needs of the ridge sports training base stadium service management from the business process to the use cases, to the data needs, and finally to the environment, performance, security, and other aspects of the demand analysis. From the overall design to the design of each detailed module, the detailed design of each module is carried out through the corresponding package diagram, class diagram of each module, corresponding sequence diagram, and important activity diagram to make the functions more detailed and clearer. Through the attribute diagram of the entities, the relationship diagram is with the entities, followed by the design of the database tables. Our optimization results show a performance improvement of about 10% and an efficiency improvement of about 6% relative to the results of other studies.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The 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

Support of Key Scientific Research Projects in Henan Province Colleges and Universities: study on physiological, biochemical, and immune indexes of intermittent hypoxia training for athletes (subject number: 20A890010).

References

[1] X. Deng, P. Jiang, X. Peng, and C. Mi, "An intelligent outlier detection method with one class support tucker machine and genetic algorithm toward big sensor data in internet of things," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 6, pp. 4672–4683, 2019.

[2] R. Fernandez Molanes, K. Amarasinghe, J. Rodriguez-Andina, and M. Manic, "Deep learning and reconfigurable platforms in the internet of things: challenges and opportunities in algorithms and hardware," *IEEE Industrial Electronics Magazine*, vol. 12, no. 2, pp. 36–49, 2018.

[3] S. Kumar, R. D. Raut, and B. E. Narkhede, "A proposed collaborative framework by using artificial intelligence-internet of things (AI-IoT) in COVID-19 pandemic situation for healthcare workers," *International Journal of Healthcare Management*, vol. 13, no. 4, pp. 337–345, 2020.

[4] O. B. Sezer, E. Dogdu, and A. M. Ozbayoglu, "Context-aware computing, learning, and big data in internet of things: a survey," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 1–27, 2018.

[5] V. Jagadeeswari, V. Subramaniyaswamy, R. Logesh, and V. Vijayakumar, "A study on medical internet of things and big data in personalized healthcare system," *Health Information Science and Systems*, vol. 6, no. 1, pp. 14–20, 2018.

[6] M. Mayer and A. J. Baeumner, "A megatrend challenging analytical chemistry: biosensor and chemosensor concepts ready for the Internet of Things," *Chemical Reviews*, vol. 119, no. 13, pp. 7996–8027, 2019.

[7] M. K. 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.

[8] C. K. Ng, C. H. Wu, K. L. Yung, W. H. Ip, and T. Cheung, "A semantic similarity analysis of Internet of Things," *Enterprise Information Systems*, vol. 12, no. 7, pp. 820–855, 2018.

[9] M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C. H. Youn, "Wearable 2.0: enabling human-cloud integration in next generation healthcare systems," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 54–61, 2017.

[10] A. Javed, H. Larijani, and A. Wixted, "Improving energy consumption of a commercial building with iot and machine learning," *IT Professional*, vol. 20, no. 5, pp. 30–38, 2018.

[11] P. S. Aithal and S. Aithal, "Management of ICCT underlying technologies used for digital service innovation," *International Journal of Management, Technology, and Social Sciences*, vol. 4, no. 2, pp. 110–136, 2019.

[12] J. Kharel, H. T. Reda, and S. Y. Shin, "Fog computing-based smart health monitoring system deploying lora wireless communication," *IETE Technical Review*, vol. 36, no. 1, pp. 69–82, 2019.

[13] K. Pradhan and P. Chawla, "Medical internet of things using machine learning algorithms for lung cancer detection," *Journal of Management Analytics*, vol. 7, no. 4, pp. 591–623, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 7132301, 11 pages
https://doi.org/10.1155/2021/7132301

Research Article

Optimization of Data Processing System for Exercise and Fitness Process Based on Internet of Things

Yiming Wang and Xidan Gong

Physical Culture Institute, Henan Polytechnic University, Jiaozuo, Henan 454000, China

Correspondence should be addressed to Xidan Gong; gxd@hpu.edu.cn

Received 8 April 2021; Revised 19 June 2021; Accepted 26 June 2021; Published 7 July 2021

Academic Editor: Wei Wang

In the digital network era, people have higher requirements for physical fitness. In the future, physical fitness requires not only good fitness equipment and fitness environment but also more convenient and intelligent health management, service guidance, social entertainment, and other refined fitness services. The innovation of sports and fitness equipment for the digital network era will definitely depend on the development of information technology and network technology. Based on the cutting-edge Internet of Things technology, this thesis focuses on the development and application of a new generation of digital fitness equipment adapted to future development, advocating the new concept of seamless integration of fitness exercise and information services through human-oriented systematic design thinking and providing implementable solutions to realize the science, convenience, and life of public fitness. This thesis uses modern science and technology, especially the Internet of Things (IoT) technology, to fully meet the diversified fitness needs of the fitness crowd as the guide; IoT digital fitness equipment design and application research was newly generated, using a variety of research methods to explore the functional design and application of IoT fitness equipment; the goal is to create a more intelligent and three-dimensional IoT fitness service model in the future. Through the application research of intelligent devices in IoT fitness equipment, the realization of the functions of identity identification, environment perception, and data transmission of IoT fitness equipment is made faster. Intelligent devices can become the interaction channel between fitness service personnel, fitness equipment, and fitness users and also reduce the development cost of IoT fitness equipment. The construction of an IoT fitness cloud service platform and data management system integrates the application of IoT, cloud computing, mobile communication, and other technologies to make IoT fitness service supply remote, real-time, and diversified. While providing convenient and value-added fitness services for fitness people, it also brings sustainable development space for the health service industry.

1. Introduction

The sports and fitness equipment industry is an important part of the sports industry; at present, fitness equipment product design and development mainly to imitate similar foreign products in the road of independent innovation are slow; coupled with the worldwide fitness equipment product innovation which has not yet appeared, product homogenization is serious and has been difficult to adapt to today's network; information technology and intelligent product design have also become the urgent need to address the fitness equipment industry. This has become a pressing problem for the fitness equipment industry. This thesis is aimed at realizing the application of IoT technology in fitness equipment design, with health as the guide; fully meeting the diversified fitness needs of fitness people as the core; and exploring the application of IoT technology in digital fitness equipment design, not only to provide ideas to overcome the multiple contradictions in the current fitness equipment industry but also to provide the most optimal solution to promote the innovative development of the sports and fitness equipment industry [1]. In modern social life, portable smart devices such as smartphones, tablets, and watches based on digital networked services are widely used, and the integration of

network-based product application services with smart devices and wearable devices has been more emphasized, with network services more closely connected to real life and the content of services more refined [2].

As people's living standards and health concerns continue to improve, mass fitness is moving in the direction of specialized, scientific, and refined needs. Traditional fitness equipment has a single function, has poor interactivity, lacks guidance, has other drawbacks, and has been difficult to meet the growing diversified fitness needs of people in the era of network information [3]. The innovative development of these fields has provided a more mature environment for the application of IoT technology. The IoT can be extended and expanded to many sports-related fields, especially for the construction of future networked sports and fitness service systems. The digital fitness equipment based on IoT technology proposed in the thesis will play an important role in the future networked fitness service systems, which has high theoretical research significance and practical value [4]. In the digital network era, people also have higher requirements for physical fitness, and the future physical fitness not only needs good fitness equipment and fitness environment but also needs more convenient and intelligent health management, service guidance, social entertainment, and other refined fitness services. The innovation of physical fitness equipment for the digital network era will definitely depend on the development of information technology and network technology [5]. Based on the cutting-edge IoT technology, this dissertation focuses on the development and application of a new generation of digital fitness equipment adapted to future development, advocating the new concept of seamless integration of fitness exercise and information services through human-oriented systematic design thinking and providing implementable solutions to realize the scientific, convenient, and lifelike aspects of public fitness. The research of the thesis actively seizes the historical opportunity of the development of the Internet of Things industry and actively responds to the objective laws and requirements of the development of the Internet of Things industry and the sports and fitness industry, which can both promote the innovative development of the sports and fitness equipment manufacturing industry and help to improve the theoretical research of digital sports and digital fitness Internet of Things [6]. The recommendation system can tap the association between user preferences and information based on the individual differences of each user, such as the basic information of the user and the behavior habits of the user, and help the user find what he or she really needs by actively pushing information to the user. For this reason, it is of great importance and far-reaching significance to successfully introduce recommendation technology into motion monitoring systems to help improve user experience, enhance users' reliance on the system, and build user stickiness, as shown in Figure 1. However, recommendation technology often relies on large-scale data computation to explore the association between users and information, while big data is the product of the current Internet era, and the use of big data processing technology to achieve recommendations has naturally become a new research hotspot, and the sports monitoring system under big data is born under such a demand of the times [7–12].

This thesis focuses on the development and application of IoT fitness equipment, taking the digital fitness treadmill based on IoT as a design case to explore the design concept and ideas of IoT fitness equipment, in order to provide a reference solution for future fitness equipment production and design. By exploring the construction and application of new IoT fitness service modes based on IoT fitness equipment, the modern fitness service mode integrated with IoT, cloud computing, and mobile communication technologies is proposed, which makes the construction of a large fitness service system for the national and even global fitness population possible in the next step. The thesis advocates the fitness equipment design thinking based on the Internet of Things, which can drive fitness equipment, traditional fitness services, and other industries to produce a strong application value, to not only inject more vitality into the development of the emerging fitness equipment industry but also provide reference significance to the adjustment of the entire sports industry structure, which is very important for meeting the masses' scientific fitness needs, improving the market position of the sports and fitness industry, optimizing the industrial structure, and promoting the economy. It is of great social significance to meet the scientific fitness needs of the masses, improve the market position of the sports and fitness industry, optimize the industrial structure, and promote economic transformation. The structure of the whole paper can be described as follows.

In Section 2, we introduce the related work of the IoT technologies and the applications in the related fields. In Section 3, this thesis uses modern science and technology, especially the Internet of Things (IoT) technology, to explore the functional design and application of IoT fitness equipment; the goal is to create a more intelligent and three-dimensional IoT fitness service model in the future. In Section 4, through the application research of intelligent devices in IoT fitness equipment, the realization of the functions of identity identification, environment perception, and data transmission of IoT fitness equipment is made faster.

2. Related Work

IoT technologies are widely mentioned and applied in the fields of healthcare, transportation, health, logistics, environmental protection, etc. [13–17]. As early as 1995, Bill Gates' book *The Way of the Future* first proposed the idea of the interconnection of things and things, which was the prototype of IoT. However, this idea was not taken seriously at that time due to the constraints of related technical fields and the lack of products that could prove the prospect of IoT applications. Subsequently, in 1999, Professor Kevin Ashton, director of MIT Auto-ID Center, put forward the more popular concept of Internet of Things (IoT) when studying RFID technology: IoT is based on radio frequency identification technology and equipment, according to the agreed communication protocol and Internet combination, so that the information of items to achieve the Internet of Things is a network formed by relying on radio frequency identification

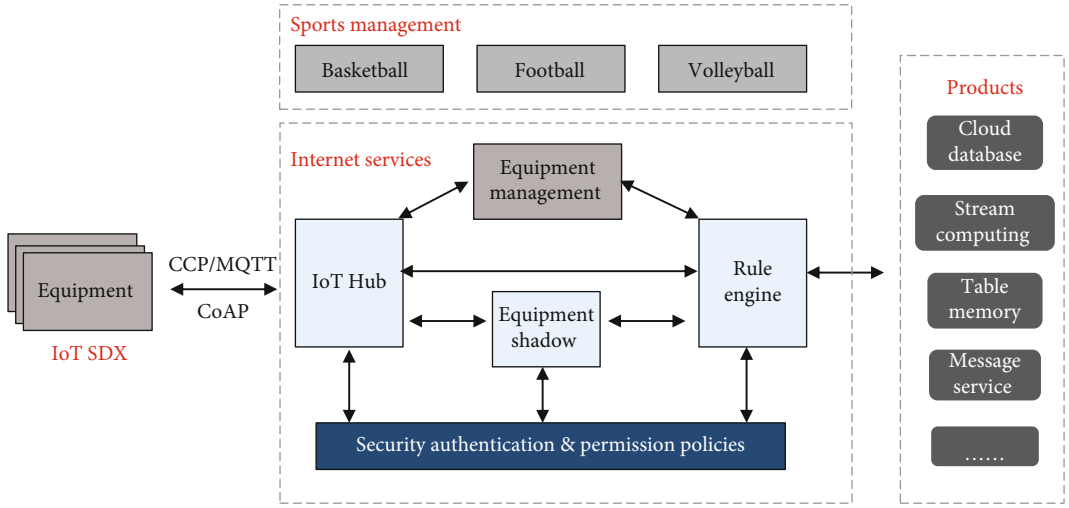


Figure 1: Digital recommendation system based on IoT.

technology and equipment, combining with the Internet according to the agreed communication protocol, so as to realize intelligent identification and management of item information and realize interconnection of item information. Since its focus in the research field is on identification technology, the definition of IoT was not comprehensive enough in the technological context at that time. Subsequently, on November 17, 2005, at the World Summit on the Information Society (WSIS) in Tunis, the International Telecommunication Union (ITU) released the ITU Internet Report 2005: Internet of Things, which formally introduced the more official concept of IoT, which was defined by the reporters as the Internet of Things (IoT) as a network of people to people and people to people by embedding short-range mobile transceivers into various accessories and everyday objects. In September 2009, at the Sino-European Workshop on IoT and Enterprise Environment in Beijing, Dr. Lorent Ferderix, head of the RFID department of the European Commission's Information and Social Media Division, gave the EU the IoT as a dynamic global network infrastructure with self-organizing capabilities based on standards and interoperable communication protocols, where physical and virtual "things" have identity, physical attributes, virtual characteristics, and intelligent interfaces and are seamlessly integrated with information networks. The Internet of Things will constitute the Internet of the future, together with the Internet of Media, the Internet of Services, and the Internet of Enterprises.

In summary, although the concept of IoT is widely proposed, there is no precise and accepted definition of IoT, mainly because IoT is the product of the integration of various technologies including sensor identification, network transmission, data mining, and pervasive computing. It requires the mutual integration of multiple disciplines, and it is closely related to the Internet, mobile communication network, sensor network, etc. Researchers in different fields have different starting points for thinking about IoT, and they cannot reach a consensus in the short term. With the rapid development of mobile Internet and IoT and the introduction of high-performance low-power processing chips, portable measurement devices that can collect data of human daily activities in real time have emerged on a large scale, which has greatly promoted the development of related industries in the field of wearable fitness. The Live Pod is a portable exercise apparatus that can monitor the daily activities of a human body in real time. It can collect fitness indicators including the type of fitness movements, exercise status, fitness time, exercise rhythm, energy consumption value per minute of exercise, the ratio and amount of fat and sugar metabolism during exercise, and the capture of exercise form, realizing the real-time sensing of fitness behavior in human daily life [18]. On July 24, 2010, at the China International Fitness Conference held at the China National Convention Center in Beijing, Xia Gao, the CEO of http://YouBody.com, proposed a new concept of fitness Internet of Things (IoT), and the first generation of http://YouBody.com's product was the IoT electronic scale, which implanted a wireless module in the traditional electronic scale and allowed the data related to the electronic scale to be sent to the website and cell phone background [19]. School of Biological and Medical Engineering, Beijing University of Aeronautics and Astronautics, developed an insole-type detection system for quantitative evaluation and early warning of daily behavioral exercise load and health status of motor function, which realizes the sensing and calculation of human exercise load amount and features low cost, low interference, and low energy consumption, further realizing the monitoring management and intelligent guidance of human health.

The well-known brand Nike also uses sensing and computing technology in basketball and training shoes, which has a built-in Nike+ Pressure Force chip that can record users' movement data and then send it to users' mobile devices through wireless data transmission, which can sense users' jump height, confrontation intensity, and reaction speed, and the collected data can help users improve their skills and further guide them to perform healthy exercises [20]. Thanks to the development of the Internet of Things and mobile Internet technologies, technological changes and upgrades are constantly being made in the field of large fitness equipment development and design. The My Wellness

product launched by Technogym, Italy, combines miniature portable measuring devices with treadmills and strength trainers to realize the monitoring and management of daily life and professional fitness, and fitness data can be shared on the My Wellness platform, while programmed fitness guidance can be obtained [21]. The Iconcept launched by the American Akcome Company takes full advantage of the APP program application of Apple and other smart cell phones to realize the design of treadmill controlled by cell phones, while its cooperation with Google Maps realizes the reproduction of virtual scenes when working out. In addition, there is also the treadmill that can access the Internet launched by Technogym. The treadmill is equipped with a VISIOweb display screen, which can be used for entertainment and leisure while exercising, and at the same time, through the cooperation with social platforms such as Twitter and Facebook, the gym-goers can share and socially entertain themselves while working out [22].

3. Design of Fitness Equipment Based on IoT Technology

3.1. Design Objectives. The design goal of IoT fitness equipment is to use IoT technology to connect fitness equipment with professional fitness service resources in real time, to realize the integration of multiple resources, and to provide interactive interfaces for fitness people, social media, digital game developers, fitness service providers, and other resources [23–26]. The IoT fitness equipment collects and transmits personal fitness data in real time through terminal software, realizes massive fitness data storage and processing technology based on IoT mode, provides corresponding fitness service solutions to different fitness people through server-side data analysis and calculation, realizes automatic control of fitness equipment, and finally realizes personalized and customized service modes of fitness service. The main design content of IoT fitness equipment is the IoT hardware system and the corresponding software service system, as shown in Figure 2. The fitness data is obtained and controlled by the whole IoT system from the vessel data and data analysis and finally result evaluation. At present, the hardware system required sensors, automatic control, and network link technology which has been relatively mature; this thesis is mainly based on the IoT fitness equipment hardware required to achieve the function of the main research, to explore the existing technology to achieve the application of IoT fitness equipment; software service system design is based on IoT fitness equipment to achieve the focus of the networked service model.

In the study of the architecture of IoT, the perception layer, as the lowest layer, is the basis for the realization of IoT functions. In the fitness service system based on IoT fitness equipment, the perception layer is mainly to realize fitness equipment identification and intelligent collection of various fitness information, which is composed of various fitness information collection sensors and sensor gateways. Through the IoT fitness equipment, the fitness of the exercise state, exercise environment, life pattern, and other information monitoring are achieved. Fitness data collection is an important function of the perception layer of the IoT fitness service system. The acquisition of user fitness data is also the basis for personalized services; the application of modern sensor technology can not only better achieve the collection of user fitness data but also provide more data interfaces to the relevant fitness service software developers, to facilitate the development of more interesting and practical fitness service software.

The implementation of automatic control can hand over the development of fitness programs and plans to more professional fitness or health service providers, avoiding the damage to fitness users due to blind fitness or operation errors, as shown in Figure 3. For example, the mainstream fitness treadmill in the market will integrate a number of running fitness models; however, many fitness users are not sure which fitness model they should choose to work out; therefore, remote automatic control will be the future fitness equipment development that needs to be paid attention for the function. At the same time, the core of IoT fitness equipment development is also intelligent control; automatic control function is a good reflection of the intelligent level of IoT fitness equipment, which is an important hardware need of IoT fitness equipment. With the development of IoT, cloud computing, wireless mobile communication, and other technologies, the functions of online fitness, remote interaction, and cloud management can be better applied in the design of IoT fitness equipment. IoT treadmill, as the terminal equipment of IoT fitness service mode, will change the "closed" fitness mode of traditional treadmill exercise in the future and realize the quantification, science, and life of fitness, so that the gym-goers can adhere to it for a long time. Through the construction of a relevant IoT fitness service platform, based on the massive fitness data provided by IoT treadmill and other IoT fitness equipment, the quantitative statistics, mining, and analysis of fitness data can be carried out to establish and promote the construction of scientific fitness behavior universal model and provide personalized scientific fitness service for diversified fitness groups.

3.2. Design of IoT Fitness Equipment Cloud Service Platform. Through the establishment of the cloud service platform, it can give access to a readily scalable resource pool, while using cloud load balancing technology to deal with the high concurrency of the fitness service system and other issues, and take a dynamic approach to deploy various application systems to maximize the efficiency of the utilization of the fitness service business system resources, so as to meet the IoT fitness service workload and business needs. The specific implementation is as follows.

(1) The core framework of the fitness service system is built based on technologies such as KVM virtualization technology and the cloud platform of Open Stack, which enables efficient processing of business such as large fitness data and fitness services

(2) Methods such as complex systems and complex networks are used to achieve scientific fitness adherence and efficient dissemination mechanisms in the system

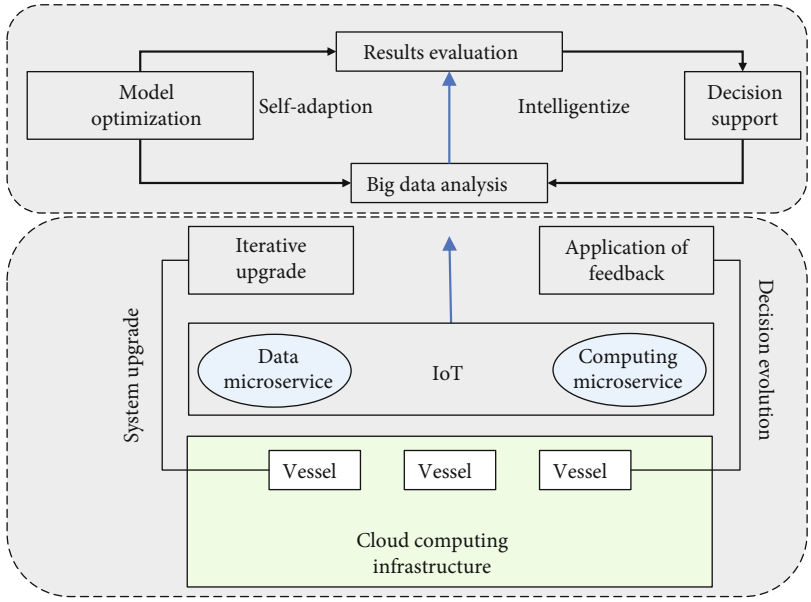


Figure 2: IoT fitness equipment application model.

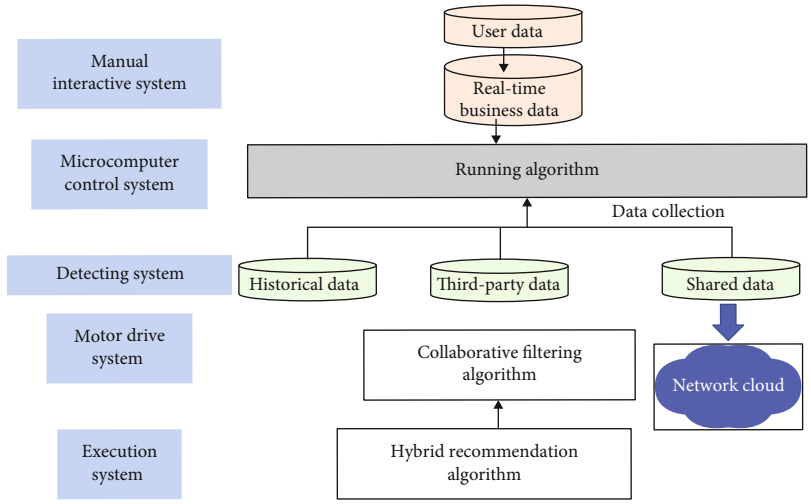


Figure 3: Digital treadmill control system structure diagram.

(3) Using REST and other technologies to achieve real-time data transmission between the Android and Apple IOS-based mobile fitness service client APP and the system server, the data is stored in JSON format, and the REST interface is called to realize the server side to parse the data and realize the transmission and interaction of business data

(4) Spring technology is used to control the business logic of massive fitness data in the system

(5) Mongo DB database technology and cloud load balancing technology are used to build the underlying database

The architecture is shown in Figure 4.

Fitness service is based on the soul of the Internet of Things fitness equipment, network services must have the support of server technology, and the server as the center of data processing is the basis of network fitness service provision. The fitness service system based on IoT fitness equipment provides personalized fitness service for fitness users through the collection, analysis, and calculation of fitness user data; in view of the wide area of IoT fitness service, its supporting server needs to have the ability to meet the huge volume, type, multiple sources of data storage, and calculation in real time, and the fitness service application of computing, storage, and network resource relationship is dynamic change. The relationship between computing, storage, and network resources of fitness service applications is dynamic. For the traditional WEB server cluster solution

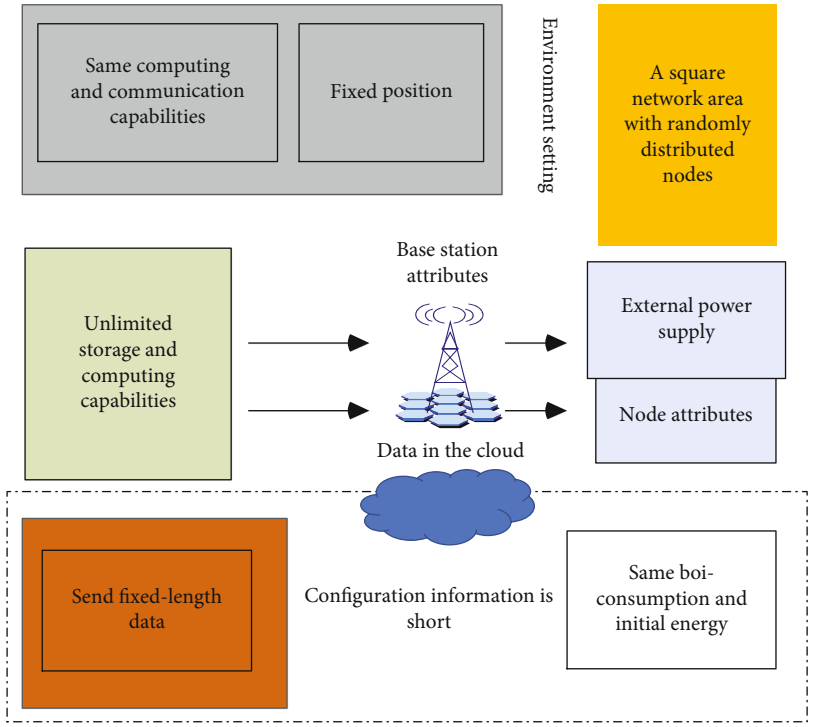


Figure 4: Fitness cloud service platform architecture.

model, it is difficult to maintain and support normal operation around the clock; the ordinary form of WEB server and database system can only be used in the early stage; later, with the growth of the service population, it can only rely on the cloud server model to solve this growth crisis. In the process of IoT fitness equipment application implementation, the amount of data from nothing to have, service data from tiny to huge, and its supporting server will certainly experience the upgrade from local to cloud. Therefore, the establishment of a cloud service platform based on IoT fitness equipment is one of the core tasks to realize IoT fitness services. The construction objectives of the IoT fitness cloud service platform are three things.

(1) Create a standard system of IoT fitness equipment. The use of the Internet of Things, cloud computing technology, and Wifi, GPRS, 3G, and other wireless mobile communication network technology to achieve the Internet of Things fitness equipment user physical condition monitoring, remote fitness guidance, and other service functions is an important measure to ensure the development of the scientific nature of the user fitness activities, through the development of data transmission and processing standards to achieve the universal characteristics of the production of IoT fitness equipment

(2) Construct a sharing platform for IoT fitness services, use IoT technology to realize the direct connection between fitness instructors and fitness participants, digitally integrate the fitness-related exercise instruction service information and so on, and send it to IoT fitness equipment users in real time, targeted and remotely through the network. In the conduct of such fitness services, a number of monitoring instruments can be selected for daily fitness, and the exercise status, energy consumption, and life pattern of fitness participants can be monitored in real time in a lifelike manner to provide basic data support for personalized scientific fitness guidance. The cloud fitness service platform can effectively promote the fitness activities of the fitness crowd. Each fitness organization contains fitness service enterprises, social fitness instructors, fitness personal trainers and fitness participants, and other different roles. The construction of a fitness cloud service platform can effectively expand the communication, coordination, supervision, evaluation, and management among organizations, families, and users and finally form an urbanized "network national fitness club" supported by Internet of Things technology

(3) To build an all-weather fitness service system that can work out anytime and anywhere. The cloud fitness service platform is a public service sharing platform based on cloud computing technology serving the public, which should be compatible with different people and different kinds of IoT fitness equipment, integrate various forms and items of fitness terminals, meet the needs of multilevel and different scale fitness guidance systems, and meet the needs of multiprofessional health information service. Its open sharing mode breaks the information island mode of closed and independent traditional fitness network systems

and adopts cloud computing technology to deploy distributed data centers. It massively reduces the cost of operation, improves the ability to cope with massive data, and expands the scope of system applications. The scattered fitness experts and other resources from different regions and organizations are networked and aggregated using IoT technology to form a large and professional expert fitness service production line, which continuously produces efficient fitness service products and provides them to different regions, thus enabling the public to enjoy the opportunity to receive professional fitness guidance anytime and anywhere and build a 24/7 fitness service system

3.3. Fitness Process Data Processing and Recommendation System. Recommendation system evaluation is an important issue in the field of recommendation. It has a variety of evaluation metrics in different application scenarios, and researchers in the industry have not yet formed a unified scientific standard evaluation system, but generally recognized important metrics have been widely used. Among them, recommendation accuracy is the most used evaluation index, i.e., the degree of agreement between the predicted score (or ranking) of the recommendation system and the actual score (or ranking) of the user, which is mainly divided into three categories: prediction accuracy, classification accuracy, and ranking accuracy. The specific calculation formula of the evaluation index is as follows.

Mean absolute error (M):

$$M = \lim_{n\longrightarrow\infty} \sum_{i=1}^{n} \frac{(p_i - r_i)^2}{N}. \tag{1}$$

Mean squared error (S):

$$S(I) = \frac{\lim_{n\longrightarrow\infty} \sum_{i=1}^{n} P_i(Q) \cap R_Q(I)}{N(x)}. \tag{2}$$

Root mean square error (R):

$$R(I) = \sqrt{\frac{\left|P_i(Q) \cap R_Q(I)\right|}{U}}, \quad D \le 1. \tag{3}$$

Standard mean absolute error (NMAE):

$$N(x, y) = \frac{a_i \cdot b_i(y) \cup b_{i(x)}}{r_{\max} + r_{\min}}. \tag{4}$$

In the formula, N is the number of objects that have been rated by users, $\{p_i\}$ is the set of predicted user rating values, $\{r_i\}$ is the set of actual user ratings of objects, and $r_{\max}$ and $r_{\min}$ are the maximum and minimum values of user ratings, respectively. The most popular recommendations today are based on proximity relationships, including user-based and object-based recommendation algorithms. However, such recommendation algorithms often determine the association between users or objects by calculating the similarity of users or objects first and recommend objects under similar users or similar objects to the target users, which lacks comprehensive consideration of the target user model and also suffers from the naming conflict problem, i.e., objects of the same nature may have multiple different names, while they are similar in nature, but this type of recommendation algorithms cannot discover and exploit this similarity. To address these problems, model-based collaborative filtering recommendation algorithms have become an inevitable alternative to traditional collaborative filtering recommendation algorithms, among which the traditional matrix factorization (MF) technique is widely used. The following paper will elaborate on the SVD decomposition method in the traditional matrix factorization technique. Singular Value Decomposition (SVD) is a generalization of eigenvalue decomposition to arbitrary matrices. Since the user-object rating matrix may not be a square matrix, this matrix decomposition method can be introduced into the recommendation system to achieve the purpose of reducing the dimensionality of the rating matrix. The method assumes that R is an $M \times N$ matrix, which can be decomposed into the product of three matrices as follows:

$$R = \lim_{n\longrightarrow\infty} \sum_{i=1}^{n} a_i \cdot b_i(y), \tag{5}$$

where U is an $M \times M$ square matrix (the vectors of the square matrix are orthogonal and are called left singular vectors), Σ is an $M \times N$ matrix (all elements except those on the diagonal are 0 and arranged in descending order), and VT is an $N \times N$ square matrix (the vectors of the square matrix are orthogonal and are called right singular vectors). First, finding the eigenvalues, one obtains:

$$\left(R^T R\right) p_i(x, y) = b \cdot b_i(x), \tag{6}$$

where v is the right singular vector. In addition, it is also obtained that

$$\delta_i = \sqrt{\alpha_i}, \tag{7}$$

$$u_i = \frac{\sqrt{\alpha_i}}{\delta_i}, \tag{8}$$

where δ is the singular value and u is the left singular vector. It is known from the existing relevant mathematical knowledge that the value of σ decreases rapidly. In other words, the first W singular values arranged from largest to smallest can be taken to describe the matrix approximately, i.e., the matrix Σ of $M \times N$ decreases in dimension to $W \times W$, while the square matrix U of $M \times M$ is transformed into $M \times W$ and the square matrix VT of $N \times N$ is transformed into $W \times N$. In this way, the original matrix can be compressed according to a certain proportion and retain the original matrix information, where $W \ll M$, $W \ll N$, and the degree of reduction of the RW matrix to the original matrix R can be controlled by adjusting the value of W. Here, the closer

W is to N, the closer the product of the three matrices on the right-hand side is to the original matrix R. Decomposing matrices using the SVD method does not allow the existence of blank parts in the matrix to be decomposed. In fact, it only makes sense to apply the SVD decomposition method to the field of recommendation system after we solve the data sparsity problem. In the recommendation scenario, there are a lot of gaps in the scoring matrix and the data is sparse. For this reason, we need to preprocess the scoring matrix, i.e., fill the missing values in the scoring matrix with the help of a global mean, object scoring mean, or user scoring mean. Of course, some normalization of the matrix is also possible. However, even with the fill strategy, in practical application scenarios, the number of users and items is often in the thousands, and the time complexity of the traditional SVD algorithm $O(n^3)$ is obviously unsolvable in the face of such a large order of magnitude.

4. System Optimization Test

A dataset of a certain size is extracted from the file system, and the original file is read using the read function in Python and converted to matrix form by the Data Frame data structure common function under the pandas package. By using the implicit feedback data (user personal attributes, user motion data, and user adoption) and explicit feedback data (user ratings of motion guidance) actually collected by the system for experiments, the root mean square error (RMSE) is used as a performance index to evaluate the recommendation accuracy of the algorithm, so as to investigate the advantages and disadvantages between this algorithm and the SVD decomposition method. Assuming that the difference between users' real ratings and predicted ratings obeys Gaussian distribution, the corresponding loss functions are jointly constructed using the rating matrix R (explicit feedback data) and the user data behavior weighting matrix W (implicit feedback data), and then, the loss functions are optimized by the stochastic gradient descent algorithm (SGD), and the parameter values in the matrix P and matrix Q are calculated and solved using the existing historical rating data, so as to obtain the corresponding matrices P and Q, i.e., matrix P is the user hidden class matrix, which represents the user's preferences, and matrix Q is the motion guidance hidden class matrix, which represents the features of motion guidance. Since the datasets extracted from the HDFS file system in this experiment are .txt files, which need some conversion processing, and the method of machine learning is adopted, a large number of third-party learning libraries need to be used, so the Python programming language is chosen here for implementation. The following is the specific experimental procedure.

In the first step, a dataset of a certain size is extracted from the HDFS file system, and the original file is read using the read function in Python and converted from .txt format to .csv format and then converted to matrix form by the Data Frame data structure common function under the pandas package. The user score table (score_data Frame) contains table id, user id, campaign guide id, and score value, and some of the data are shown in Figure 5.

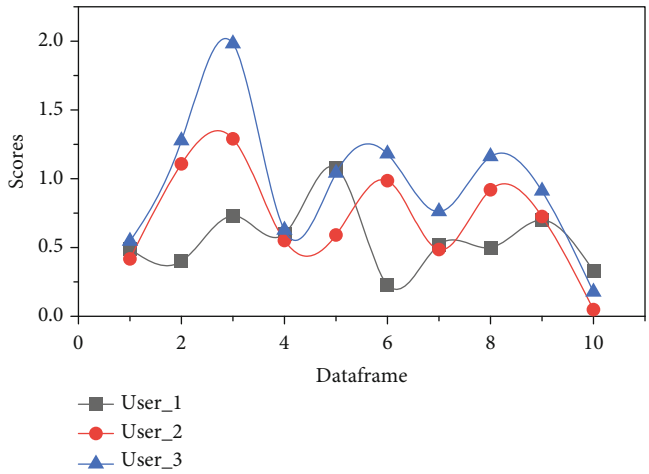


Figure 5: Sports guide rating chart.

In the second step, we calculate the weights of user data behaviors (user's personal information, personal movement data collected by user wearing the bracelet, and behavioral data such as operating APP) to generate the weighting matrix W of user data behaviors and then combine with the scoring matrix R to construct the corresponding loss function. In the third step, the stochastic gradient descent algorithm (SGD) is used to optimize the loss function by finding the partial derivatives of the parameters $p(u, f)$ and $q(f, i)$ to determine the fastest direction of descent, and then, iterative computation keeps optimizing the parameters until the parameters converge. Here, the parameters converge at about 500 iterations, and the iteration ends with the output of two low-dimensional matrices P and Q, as shown in Figure 6.

The experimental results in this paper use the root mean square error (RMSE) as a performance indicator to evaluate the algorithm recommendation accuracy, and the calculation formula is shown above, i.e., the smaller the value, the higher the algorithm recommendation accuracy. The RMSE of the weighted matrix decomposition recommendation algorithm based on the hidden semantic model is calculated to be 0.12, and the RMSE of the SVD decomposition method is calculated to be 0.42. It can also be seen from Figure 7 that the two fine curves match well, and the coarse curve deviates to a greater extent, indicating that the recommendation accuracy of the improved recommendation algorithm is higher, i.e., the recommendation accuracy of the weighted matrix decomposition recommendation algorithm based on the hidden semantic model is better than that of the traditional SVD decomposition method. The reason for this difference is that the singular value matrix decomposition model first needs to mean fill the historical rating data and then perform matrix decomposition to obtain the correlation submatrices, which reduce the model credibility and naturally result in lower recommendation accuracy. On the contrary, the Latent Factor Model (LFM) fully utilizes the user item rating matrix as the explicit feedback data and introduces the idea of weighting to build a weighted matrix of user data behaviors as the implicit feedback data, which deeply explores the user's preferences from both explicit and implicit aspects, making

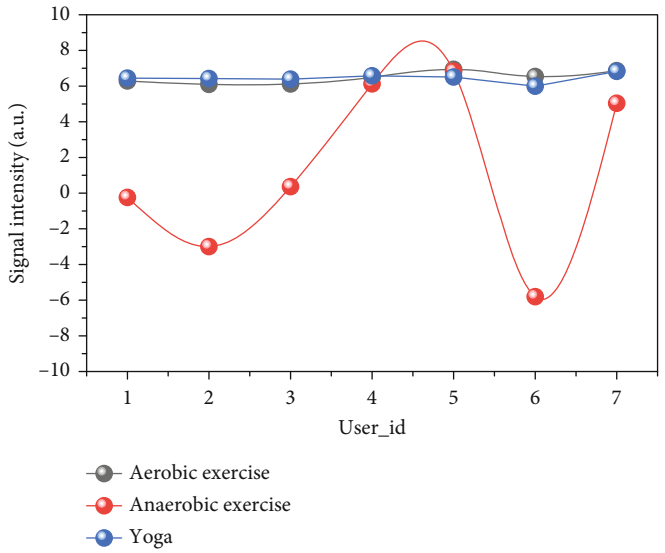


Figure 6: User hidden class matrix.

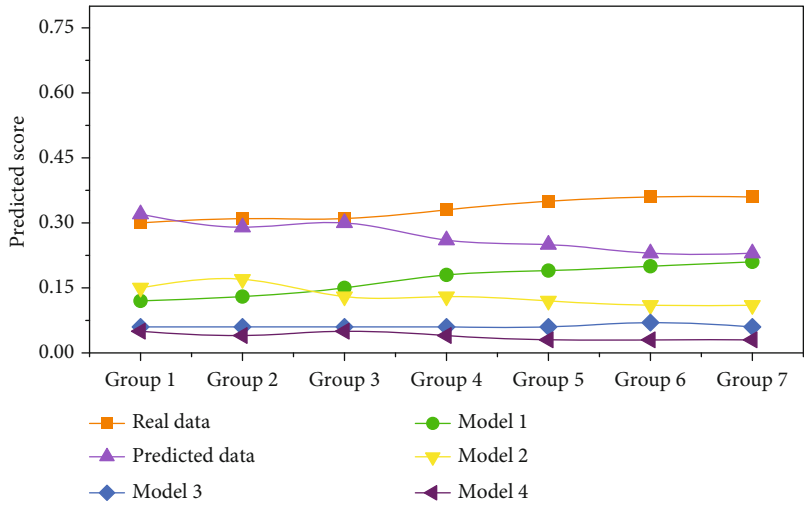


Figure 7: Algorithm predicted score vs. actual score trend.

the improved recommendation algorithm more perfect and the recommendation accuracy higher.

After users use the system for a period of time, the system backend can calculate the corresponding rating value based on the ratio of the user's browsing time for a certain exercise guide to the browsing time for all the exercise guides that have been browsed, which eventually forms the user rating table, i.e., the user-generated display rating data. At the same time, the implicit behavioral data generated by users, such as personal information and sports data mentioned above, are jointly extracted from the server My SQL database to HBase by Sqoop tool. Then, the offline batch computing framework of Map Reduce is used to decompose the user rating matrix into user implicit class matrix and motion guidance implicit class matrix to generate the original recommendation model. Finally, the Spark Streaming streaming computing framework is used to combine the user data behavior weighting matrix generated by the current user behavior data (implicit data), so as to reupdate the user rating matrix (i.e., user prediction rating matrix) and update the recommendation model to get the online recommendation results, and Figure 8 shows the exercise guidance recommendation. The popularity of IoT fitness equipment will strongly contribute to the application and development of related IoT fitness services. From the perspective of practitioners, the growing sports and fitness service industry has given rise to a large demand for sports industry practitioners, and the frontline work of fitness service product providers, that is, fitness service industry practitioners with professional fitness guidance and management knowledge, has made up for a large labor demand gap. More sports instructors, personal trainers, and students from sports colleges and universities are attracted to enter the sports and fitness service industry, and the proportion of the number of people employed in the sports and fitness service industry to the number of people employed in society as a whole will increase significantly. The improvement of the quantity and quality of sports practitioners will play an important role in promoting the

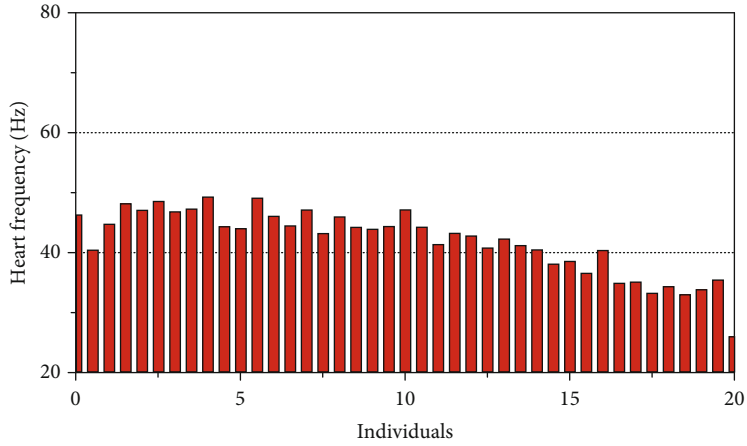


Figure 8: Exercise instruction frequencies.

industrialization of fitness services, improving the quality and level of fitness services, increasing the added value of fitness industry products, promoting the overall improvement of the sports industry, enhancing the core competitiveness of the fitness service industry, and promoting the sustainable development of the fitness industry.

5. Conclusion

The case design of IoT digital treadmill not only takes into account the advantages of traditional digital treadmill but also makes use of sensor technology, embedded technology, and automatic sensing technology to make IoT treadmill a fashionable and modern fitness tool, which provides an important basis for the hardware design of other types of IoT fitness equipment. Through the application research of intelligent devices in IoT fitness equipment, it makes the realization of identity identification, environment perception, data transmission, and other functions of IoT fitness equipment faster. Smart devices can become the interaction channel between fitness service personnel, fitness equipment, and fitness users and also reduce the development cost of IoT fitness equipment. The construction of an IoT fitness cloud service platform and data management system integrates the application of IoT, cloud computing, mobile communication, and other technologies to make IoT fitness service supply remote, real-time, and diversified. While providing convenient and value-added fitness services for fitness people, it also brings sustainable development space for the health service industry. The exercise monitoring system researched and implemented in this paper is set in the application context of outdoor sports parks, where there is a situation of large-scale user use, so improvements are made to the multitask real-time scheduling mechanism of the system on the outdoor smart Bluetooth base station.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The study was funded by the University-level Humanities and Social Science Research Fund of Henan Polytechnic University in 2021: Research on the Influence of China Tennis Tour on the Operation Mode of Chinese Traditional Tennis Matches from the Perspective of Self-created IP (number: SKND2021-22).

References

[1] M. Måløy, F. Måløy, R. Lahoz-Beltrá, J. C. Nuño, and A. Bru, "An extended Moran process that captures the struggle for fitness," *Mathematical Biosciences*, vol. 308, pp. 81–104, 2019.

[2] G. R. Tomkinson, K. D. Carver, F. Atkinson et al., "European normative values for physical fitness in children and adolescents aged 9-17 years: results from 2 779 165 Eurofit performances representing 30 countries," *British Journal of Sports Medicine*, vol. 52, no. 22, pp. 1445–1456, 2018.

[3] G. R. Tomkinson, J. J. Lang, and M. S. Tremblay, "Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014," *British Journal of Sports Medicine*, vol. 53, no. 8, pp. 478–486, 2019.

[4] Z. Guo and X. Yan, "Fitness partition-based multi-objective differential evolutionary algorithm and its application to the sodium gluconate fermentation process," *Chemometrics and Intelligent Laboratory Systems*, vol. 177, pp. 8–16, 2018.

[5] E. Tikkanen, S. Gustafsson, and E. Ingelsson, "Associations of fitness, physical activity, strength, and genetic risk with cardiovascular Disease," *Circulation*, vol. 137, no. 24, pp. 2583–2591, 2018.

[6] J. M. Abdullah and T. Ahmed, "Fitness dependent optimizer: inspired by the bee swarming reproductive process," *IEEE Access*, vol. 7, pp. 43473–43486, 2019.

[7] B. Foroughi, M. Iranmanesh, H. F. Gholipour, and S. S. Hyun, "Examining relationships among process quality, outcome quality, delight, satisfaction and behavioural intentions in fitness centres in Malaysia," *International Journal of Sports Marketing & Sponsorship*, vol. 20, no. 3, pp. 374–389, 2019.

[8] E. Foxall and N. Lanchier, "Generalized stacked contact process with variable host fitness," *Journal of Applied Probability*, vol. 57, no. 1, pp. 97–121, 2020.

[9] N. Watanabe, S. S. Sawada, K. Shimada et al., "Relationship between cardiorespiratory fitness and non-high-density lipoprotein cholesterol: a cohort study," *Journal of Atherosclerosis and Thrombosis*, vol. 25, no. 12, pp. 1196–1205, 2018.

[10] J. M. Scott, E. C. Zabor, E. Schwitzer et al., "Efficacy of exercise therapy on cardiorespiratory fitness in patients with cancer: a systematic review and meta-analysis," *Journal of Clinical Oncology*, vol. 36, no. 22, pp. 2297–2305, 2018.

[11] M.-S. Shin and H. Lee, "Effect of additional firing process after sintering of monolithic zirconia crown on marginal and internal fitness," *The Journal of Korean Academy of Prosthodontics*, vol. 57, no. 4, pp. 321–327, 2019.

[12] T. Tarumi and R. Zhang, "Cerebral blood flow in normal aging adults: cardiovascular determinants, clinical implications, and aerobic fitness," *Journal of Neurochemistry*, vol. 144, no. 5, pp. 595–608, 2018.

[13] X. Liu, Q. Pan, and M. He, "Stochastic dynamics in the fitness-based process which can be on behalf of the standard Moran, local and Wright-Fisher processes," *Journal of Theoretical Biology*, vol. 460, pp. 79–87, 2019.

[14] K. A. G. Kremling, S.-Y. Chen, M.-H. Su et al., "Dysregulation of expression correlates with rare-allele burden and fitness loss in maize," *Nature*, vol. 555, no. 7697, pp. 520–523, 2018.

[15] H. B. D. Silva, L. K. Beura, H. Wang et al., "The purinergic receptor P2rx7 directs metabolic fitness of long-lived memory CD8+ T cells," *Nature*, vol. 559, no. 7713, pp. 264–268, 2018.

[16] E. Dean, "The fitness to practise process demystified: what every nurse and manager need to know," *Nursing Management*, vol. 27, no. 5, pp. 9–11, 2020.

[17] H. Zbinden-Foncea, M. Francaux, L. Deldicque, and J. A. Hawley, "Does high cardiorespiratory fitness confer some protection against proinflammatory responses after infection by SARS-CoV-2?," *Obesity*, vol. 28, no. 8, pp. 1378–1381, 2020.

[18] B. V. Adkar, S. Bhattacharyya, A. I. Gilson, W. Zhang, and E. I. Shakhnovich, "Substrate inhibition imposes fitness penalty at high protein stability," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, no. 23, pp. 11265–11274, 2019.

[19] L. Feng, A. S. Raman, M. C. Hibberd et al., "Identifying determinants of bacterial fitness in a model of human gut microbial succession," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, no. 5, pp. 2622–2633, 2020.

[20] J. C. Jones, P. N. Q. Pascua, T. P. Fabrizio et al., "Influenza A and B viruses with reduced baloxavir susceptibility display attenuated in vitro fitness but retain ferret transmissibility," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, no. 15, pp. 8593–8601, 2020.

[21] G. Taino, E. Oddone, G. Corona, R. Foti, and M. Imbriani, "The fitness to work certificate in a worker exposed to ionizing radiation with an oncological disease: criteria and assessment process," *Radioprotection*, vol. 54, no. 4, pp. 303–307, 2019.

[22] C. J. Watson, A. L. Papula, G. Y. P. Poon et al., "The evolutionary dynamics and fitness landscape of clonal hematopoiesis," *Science*, vol. 367, no. 6485, pp. 1449–1454, 2020.

[23] W. Wei, M. Guizani, S. H. Ahmed, and C. Zhu, "Guest editorial: special section on integration of big data and artificial intelligence for Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2562–2565, 2020.

[24] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the Internet of Things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

[25] J. Han, N. Lin, J. Ruan, X. Wang, W. Wei, and H. Lu, "A model for joint planning of production and distribution of fresh produce in agricultural Internet of Things," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9683–9696, 2021.

[26] S. H. Ahmed, V. H. C. de Albuquerque, and W. Wei, "Guest editorial: special section on advanced deep learning algorithms for industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2764–2766, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5172393, 10 pages
https://doi.org/10.1155/2021/5172393

Research Article

Optimization of Land Resource Information Management System Based on Internet of Things

Taizheng Chen,[1,2] Zhun Chen,[2,3] Mingjie Tian,[1,2] and Xi Wang[1,2]

[1]*College of Geography and Planning, Henan University, Kaifeng 475004, China*
[2]*Henan Key Laboratory of Earth System Observation and Modeling, Henan University, Kaifeng 475004, China*
[3]*School of Philosophy and Public Management, Henan University, Kaifeng 475004, China*

Correspondence should be addressed to Xi Wang; wangxi@henu.edu.cn

Received 13 April 2021; Revised 21 June 2021; Accepted 25 June 2021; Published 3 July 2021

Academic Editor: Wei Wang

The business process model of information management is also optimized, while helping the relevant departments using the system to improve their work efficiency. After research and analysis, this paper shows that the land resource information management system can be divided into four functional modules, which are basic land resource information management, land resource land use approval management, official document information management, and office business management, and these modules can effectively complete the work related to land resource management. The full and reasonable use and protection of land resources can ensure the implementation of sustainable development policies, and the development of land resource management is the basic requirement of national modernization and the inevitable trend in the country, and because land resource management covers a large amount of information involving a wide range of information, including information within the scope of work, it is very necessary to ensure the effectiveness and accuracy of land resources and information. The information system can effectively improve the efficiency and accuracy of land resources and information. Information system can effectively improve the development, management, research, and use of land resources; electronic government system can help land resources to obtain accurate analysis and research and detailed access to the ability to understand the status of resources, in the daily changing data information dynamic understanding of the development of information changes, and grasp the latest trends in land resources, the trend of analysis, and monitoring of market resources, to provide customers with reliable decision support. It is a powerful background data guarantee for customers to provide reliable decision support.

1. Introduction

The level of information development has gradually increased, information technology has led to the rapid development of other industries, the rapid progress of high-tech industries has become step by step an important indicator of a country region in the development as well as the direction, while information technology is also an important assessment indicator of the degree of modernization and the level of economic development [1]. For the development of the entire national economy and life need to effectively use the management of land resources and land natural information, land resource management has become a key task; land resource management has played a very important role in the long-term development of our economy and society; along with the urgent needs of economic development and the use of land resources and mineral resources in the process of increasingly serious problems, these problems cover the illegal use and other key contradictory issues such as the occupation of land. At the same time, the construction level of our cities urgently needs to be more on a higher level, making the shortage of land resources and mineral resources present new challenges in the process of land resource management, so how to improve the speed of land and mineral resource use has become a key issue that urgently needs to be solved nowadays, and land resource management departments need to effectively combine information technology, land information management technology, and resource information management technology. Using the abovementioned technologies can improve the efficiency of land and

mineral resource information management, and the influx of information technology has gradually become the direction of improvement in the development of land security management [2].

With the advent of the new era, the comprehensive management of land resources is currently undergoing a comprehensive reform and comprehensive construction; the promotion is mainly carried out from three aspects of information technology, land information management and scientific management, and the standardized management of land resource information; land resources cover a variety of natural information, through a transparent information management system, to achieve scientific management; transparent system can enhance the land. The transparency system can enhance the efficiency of land resource management, strengthen the degree of information sharing, and face the public development which is more conducive to the power from the private sector. According to each information and demand analysis and considering the current situation of information technology development, this paper proposes a comprehensive management information service system for the construction of urban land resources, by realizing the functional modules within each management business [3]. An information sharing platform is built, while the business process model of information management is optimized and while helping the relevant departments using the system to improve their efficiency; at the same time, in order to avoid relatively isolated data that cannot be effectively used, the use of information systems can effectively use the data, the overall structure of which is a system building model with application requirements as the core [4]. Land resources are the natural treasures of the country's social development and survival, and effective management of the country's land resources can provide a favorable guarantee for the survival and development of the general population, while the effective use of natural resource development can make a remarkable contribution to the economic development of our country [5–10]. The full and reasonable use and protection of land resources can ensure the implementation of sustainable development policies; in the country, the development of national land resource management is the basic requirement of national modernization as well as the inevitable trend, because land resource management covers a large amount of information involved in the vast, including the scope of work of the information, so to ensure the effectiveness and accuracy of land resources and information has a very necessary practical significance.

With the comprehensive arrival of information-based E-government, the management of land resources also puts forward the corresponding challenges and opportunities; information systems can effectively improve the development, management, research, and the use of land resources; E-government system can help land resources to obtain accurate analysis and research, detailed access to understand the status of resources, and the ability to dynamically understand the development of information changes in the increasingly changing daily data and information as well as grasp the latest trends in land resources, the development trend of the analysis, and monitoring of market resources, to provide customers with reliable decision-making support to do a strong backstage data security, strengthen resource management planning, and provide land and resources for the further development of the plan in our country. E-government information management system design and construction need to capture the actual needs after field research, through customizing business processes and system to adapt to business processes, and use a unified modeling language to analyze and model resource needs accordingly, to ensure that the system can meet user needs and design a model structure diagram to meet the software development engineers; the system architecture can adapt to the development of technology, while the system can easily adapt. The system architecture can be adapted to the development of technology, while the system can be easily upgraded and expanded, and the secondary development and postmaintenance of the software product can be treated to ensure that the software product has a long life cycle.

The structure of this paper can be shown as follows.

In Section 2, we mainly introduce the related work about the information technology and application about the management system. In Section 3, we mainly introduce the use of the abovementioned technologies and the influx of information technology. In Section 4, we mainly introduce the system optimization test.

2. Related Work

Information technology and E-government in Europe and the United States took the lead in the development and emergence of the United States and other early developed countries to establish a complete system of business systems; information technology was used to take the lead in the field of government management; in the early 1970s, the United States first used information technology to enter the field of high-speed information management construction, the modern tell highway for scientific management, while in the 1990s, information technology has been widely used in the field of government, and the research experience has been gradually developed comprehensively, as shown in Figure 1. The information technology is used to build a browser-server model of management system, and a high-performance server is used as the back-end processing core at the processing terminal of the system to build a hierarchical and centrally managed information system [11].

E-government has been used in the West for a long time and has been widely developed as well as accumulated profound development experience [12]. In terms of land resource management, American scientists in 1998 had proposed the concept of digital earth, that is, the digital construction and layout of global geographic information; digital earth is the layout of network informationization of land resource information, virtualized storage of realistic land resource information, and, with it, the corresponding birth of geographic information system. With the development of national land informationization, in the late twentieth century, Western countries such as Canada joined one after another in the comprehensive construction of digital earth, and at the moment of the end of the twentieth century,

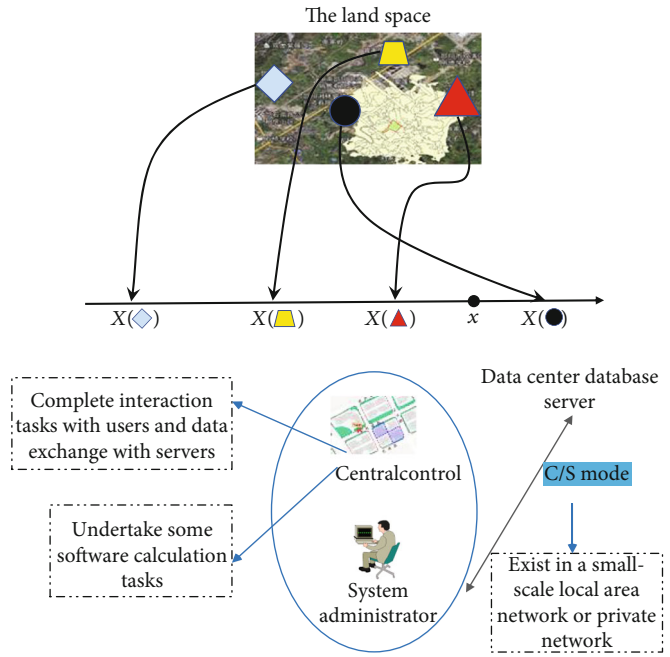


Figure 1: Multiterminal land management system.

various publications, maps, and earth science information from various units have been integrated, so users can query, browse, and evaluate [13].

In the Ministry of Land and Resources Ministry which actively promote the cities and municipalities, the integrated management of land resource center text information system is conducive to the information construction of land resources, construction of land resource information system, and information modeling of land resources which can ensure the effectiveness and efficiency of business; in recent years, land resource management did make sufficient progress and achievements, but compared to the internationalization of land resource information, there is still a certain gap in the development of land information management information technology research which is late, the basic information resources are relatively poor, and national resources and land capital are rich, but the real degree of lack of certain management, the government's limited level of management cognition, and lack of certain investment in reform courage as well as boldness, coupled with the relative lack of talent in this area, for the construction of land resource information technology has not do persistent research and accumulation, the lack of construction experience leads to the lag behind the information technology infrastructure; the overall level of information technology is low [14]. These gaps also need to be improved step by step by subsequent research scholars, and a series of problems currently encountered are mainly shown below. The geographic construction of data-based information is still in the backward stage, the geographic information construction started late, the collection and entry of geographic information have certain problems in the development, and these problems limit the further development of the construction of geographic resource management; at the same time, there is a relative lack of talents in this area, the degree of understanding of the treatment of land resource management is relatively backward, and the theoretical research on the construction of land resources is relatively lacking [15].

At present, research and the development of land resource information platform are a hot topic, a large number of scholars have gradually recognized the importance of the issue, and a large number of research and discussion can help the management of land resources step by step on the right path of development. The effect is obvious; land resource information management is more orderly; the advanced technology to cover all urban construction in the network environment makes full use of [16] dynamic remote sensing monitoring of land use and remote management of cadastral information; in recent years, the land area change data will be used to obtain the most accurate records; the local government has been introduced in a timely manner suitable for the construction of policies; land management information systems have been used in certain areas in the city; advanced technology is currently being promoted and popularized, a strong impetus to the development of land information and technical team and the growth. This paper will focus on the design and implementation of land resource information systems [17]. The information technology research of the time period is late, the basic information resources are relatively poor, the country is rich in resources and land capital, but the real degree of the lack of certain management, the government's limited level of management cognition, and the lack of certain investment in reform courage as well as boldness, coupled with the relative lack of talent in this area, for the construction of land resource information technology did not do persistent research and accumulation; the lack of construction experience leads to lagging behind of information technology infrastructure; the overall level of

information technology is low. Information technology drives the rapid development of other industries and the continuous development of scientific knowledge related to the establishment of scientific development [18].

Due to the implementation of land use in the land management system, the new land management law of the country, arable land protection land use planning, and a series of regulatory measures in all aspects, the rapid progress of high-tech industry has become step by step an important indicator of a country region in the development as well as the direction. Office automation products are being popularized step by step; land information management systems are used to provide good operational tools and interfaces for land resource management [19]. It is necessary to control the scope of document circulation; multifaceted, multidepartmental coordination; and joint consultation to complete the division of tasks to further improve efficiency and ensure the quality of work. In accordance with the requirements of building a unified land resource and land sector, workers share a set of data sources that facilitate the establishment of a scientifically sound data framework, service framework, and runtime environment [20]. System can meet user needs and design a model structure diagram that satisfies software development engineers, and the system architecture can adapt to accommodate the development of technology, while the system can easily upgrade and expand and treat the secondary development of software products as well as postmaintenance to ensure that the software products have a long life cycle [21].

3. System Analysis

3.1. Overall Functional Analysis. The construction level of the city urgently needs to be more on a higher level, which causes the shortage of land resources and mineral resources and puts forward new challenges in the process of land resource management, so how to improve the speed of using land and mineral resources has become a key problem that urgently needs to be solved at present, and the land resource management department needs to effectively combine information technology, land information management technology, and resource information management technology [22–25]. The use of the abovementioned technologies can improve the efficiency of land and mineral resource information management, and the influx of information technology has gradually become the direction of improvement in the development of land security management. The level of information development has gradually increased; information technology has led to the rapid development of other industries; the rapid progress of high-tech industry has become step by step an important indicator of a country region in the development as well as the direction; this technology is not only the result of precipitation but also the pursuit of advanced results; scholars often want to use this new product to take a greater step in their own field, so in the subject of this paper, scholars are seeking effective use of management of land resources and natural information of land; land resource management has become a key task; land resource management plays a very important role in the long-term development of our economy and society; while new phenomena arise, they also cause some new difficulties; new difficulties provide new problems to the field personnel, these problems cover other key contradictory issues such as illegal use and occupation of land.

Therefore, to ensure the effectiveness and accuracy of land resources and information has a very necessary practical significance, in the work process, to ensure the fairness and openness of information, while ensuring the accuracy and transparency of information; the sharing of information should also ensure that all departments can effectively use, while treating the maintenance of the system hoping to strive for simplicity and the establishment of a perfect and reasonable working mechanism and information data flow. The advent of the information age can also be a revolution in the management mode of government departments, which not only means that the business of the Chinese government is gradually transparent and open but also on the other hand can help the relevant government staff to use information technology products to achieve daily affair permitted management and even the use of information technology to build a browser-server mode of management system and the use of high-performance servers in the processing terminal of the system. The information technology is used to build a browser-server mode management system, and a high-performance server is used as the back-end processing core to build a hierarchical and centrally managed information system, as shown in Figure 2.

The level of development of information technology has gradually increased, information technology has led to the rapid development of other industries, and there is an establishment of scientific development of scientific knowledge related to the continuous development; due to the implementation of land use in the land management system, the country's new land management law, arable land protection land use planning, and a series of regulatory measures in all aspects, the rapid progress of high-tech industry has step by step become an important indicator of a country region in the development and direction and mature and advanced management information systems in the application, while information technology is also the degree of modernization, economic development level of important assessment indicators, geographic information systems, and office automation products, which has a step by step process to get popular; land information management system for the management of land resources provides good operational tools and interfaces, using better methods and ways to complete excellent work. Figure 3 shows the business flow chart, after several reviews and judgments to complete the business registration work.

Land resource management has played a very important role in the long-term development of our economy and society; this newly generated social phenomenon often also generates new problems; in several industries, it is the case; the management of land resources will also meet such problems; these problems cover other key contradictory issues such as illegal use and occupation of land. At the same time, the construction level of our cities urgently needs to be more on a higher level, causing the shortage of land resources and

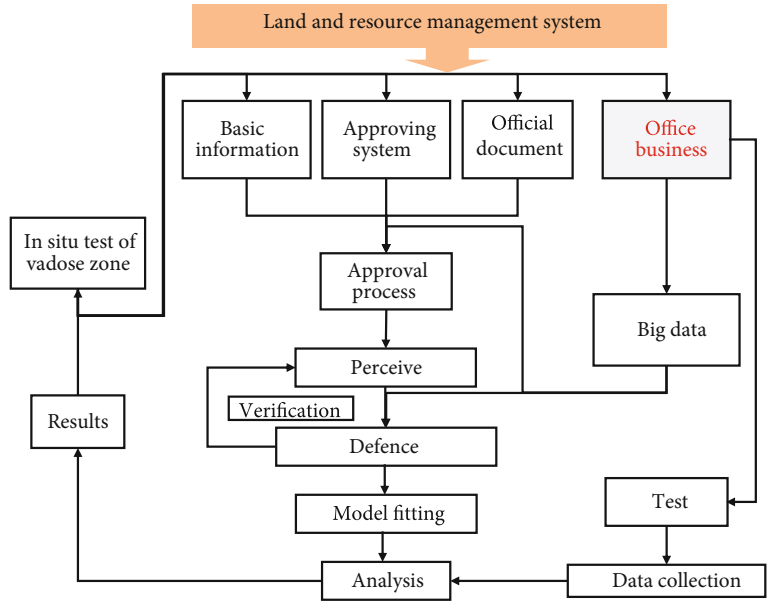


Figure 2: Functional structure of land resource management system.

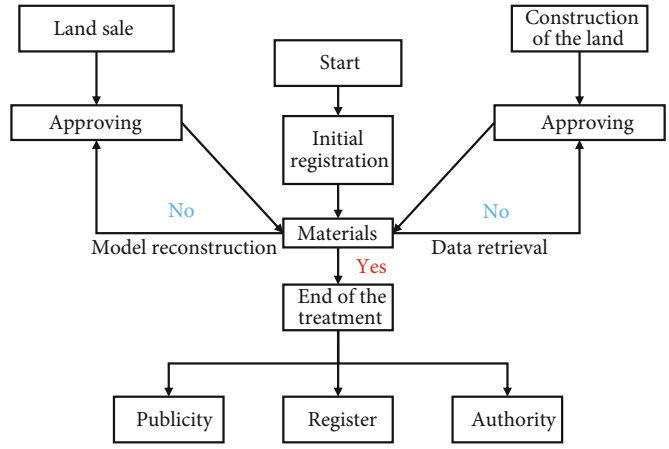


Figure 3: Flow chart of land registration business.

mineral resources; in the process of land resource management, new challenges arise, so how to improve the speed of land and mineral resource use has become a key issue that urgently needs to be solved at present, and land resource management departments need to effectively combine information technology, land information management technology, and resource information-management technology; the use of the above technologies can improve the efficiency of land and mineral resource information management; the influx of information technology is gradually becoming the direction of improvement in the development of land security management.

3.2. Use Case Analysis. Use case analysis is a functional level research study that captures the functionality of the system from a macro perspective. The use cases include its definition, all the necessary behaviors to execute the use case order, the standard behaviors in different deformations, the general behaviors of all the abnormal cases, and the expected responses, without revealing the premise that the internal structure of the system defines a consistent behavior. From a user's point of view, the above cases may be abnormal cases; from the system's point of view, they must be described and attached to the handling of the case. More specifically, the use of the case is not required for functional specification but also shows and demonstrates its requirements in the process of illustration, in UML, using cases represented by an ellipse, in which the execution of each use case is independent from the other use cases, although due to the fact that when the case is shared, objects can generate implicit correlations between use cases for reasons of execution between a use case, each use case described in the vertical functional block, the execution of that functional block can be carried and mixed with other cases together, as shown in Figure 4.

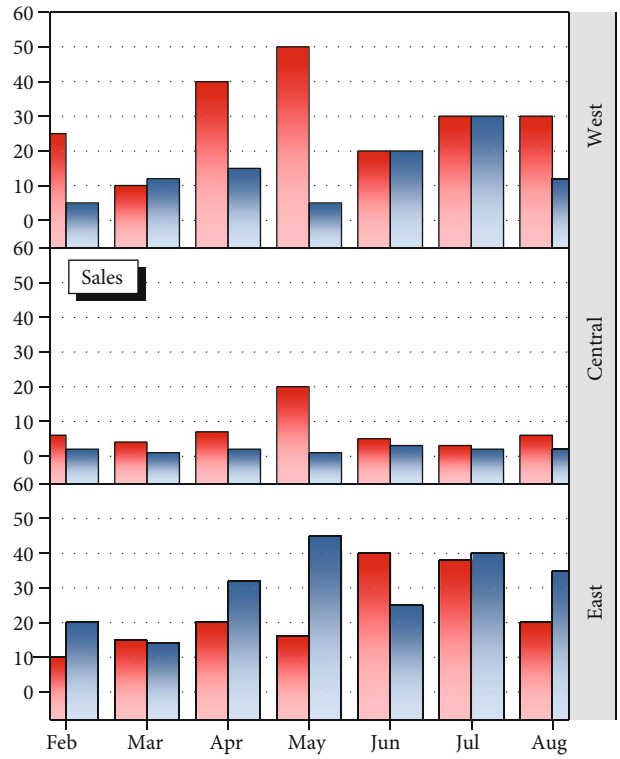


Figure 4: System use case increasement.

The functional functions have a many-to-many relationship with the system participants, where user roles with different privileges can perform the same functional functions, while each class of role users can perform multiple functional functions. There are various complex relationships between functions such as dependency, inclusion, and extension. By subdividing a function into multiple functions, the hierarchical relationship of the functional structure of the use case diagram is realized. In the use case diagram, participants are used for system functions, through the use case diagram that has clear participant system, you can see each participant's clear function at the same time; the use case diagram is very expressive; in the beginning of the do requirement analysis, we can derive the use case diagram and user verification, and thus, further, in each functional module of the subsystem, the aspect of the use case diagram in the system functional modeling is necessary.

Treating the field of land resource information construction has gradually figured out and summarized a set of suitable local land resource management methods, while with the help of today's popular information technology; in the actual work, Dalian City land capital has indeed taken a big step forward, compared to the previous work effectiveness that rose a lot, which has been built and put into use in some pilot cities and county land administration; through the use of information technology system, the effect is obvious; land resource information management is more orderly; and there is advanced technology overlay. Covering all urban construction, in the network environment, makes full use of land use dynamic remote sensing monitoring and remote management of cadastral information; in recent years, the land area change data will be used to obtain the most accurate records; the local government has been introduced in due course suitable for the construction of policies; land management information system has been used in certain areas in the city, where advanced technology is currently being promoted and popularized, a strong impetus to the development of land information and technology team and the growth of information resources of land. This paper will focus on the design and implementation of land resource information system.

Land resource management plays a very important role in the long-term development of both our economy and society. Along with the urgent need for economic development, the process of using land and mineral resources has created increasingly serious problems, which cover other key contradictory issues such as illegal use and occupation of land. At the same time, the construction level of our cities urgently needs to be more on a higher level, making the shortage of land resources and mineral resources present new challenges in the process of land resource management, so how to improve the speed of land and mineral resource use has become a key issue that urgently needs to be solved at present, and land resource management departments need to effectively combine information technology, land information management technology, and resource information management technology; using the abovementioned technologies can improve the efficiency of land and mineral resource information management, and the influx of information technology gradually becomes the direction of upgrading the development of land security management.

At the same time, in order to avoid relatively isolated data that cannot be effectively used, the use of information systems can effectively use data, the overall structure of which is a system construction model with application requirements as the core. Land resources are the natural treasures of the country's social development and survival, and effective management of the country's land resources can provide a favorable guarantee for the survival and development of the majority of people, while the effective use and development of natural resources can make an outstanding contribution to the economic development of our country. The full and rational use and protection of land resources can ensure the implementation of sustainable development policies, and the development of national land resource management is a basic requirement for the modernization of the country as well as an inevitable trend at the national level, as land resource management covers a large amount of information involved in the vast scope, including the scope of work of the information, as shown in Figure 5.

3.3. Database Specification Design. Database detailed design is mainly to conceptualize the attributes of the entities involved in the system, design a suitable type for each attribute, build a database information table for each type of entity, record the changes of data information of the entity during the operation of the system, synchronize with the database in time, and maintain the normal operation of the system. According to the entity relationship model in the conceptual design stage of the database, certain conceptualization work is needed for the database, and some theories

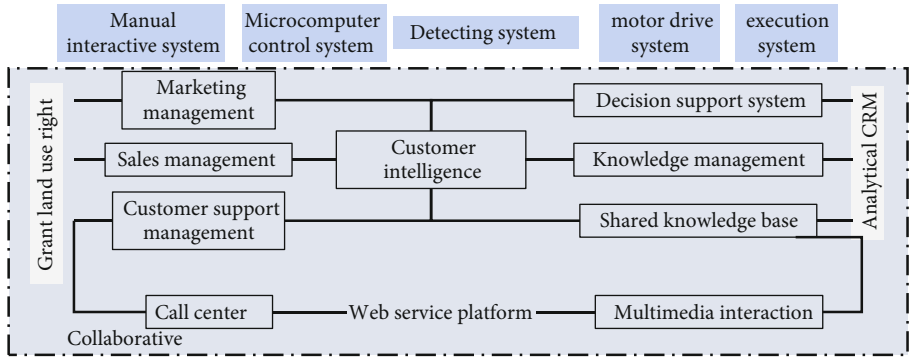


Figure 5: Flow chart of land approval.

exist to study the relationship between database members, which mainly explore and study the relationship between objects, and there are dependency relationships, backup, and organization. The detailed database design is mainly to conceptualize the attributes of the entities involved in the system, design a suitable type for each attribute, build a database information table for each type of entity, record the changes of data information of the entity during the operation of the system, synchronize with the database in time, and maintain the normal operation of the system.

$$\widetilde{x_i} = \lim_{n\longrightarrow\infty} \sum_{i\in q}^{x} \sum_{q=0}^{n} \frac{q(i)+p(i)}{\max\{q,p\}-1}. \tag{1}$$

According to each information and demand analysis and taking into account the current situation of information technology development, this paper proposes a comprehensive management information service system for the construction of urban land resources, which builds an information sharing platform by realizing functional modules within each management business, while optimizing the business process model of information management and while helping the relevant departments using the system to improve their work efficiency, and at the same time, in order to avoid relatively isolated data that cannot be effectively used, the use of information systems can effectively use data, the overall structure of which is a system construction model with application requirements as the core; the core formula of the login verification module of the system is shown below.

$$I(X,Y) = \frac{\lim_{n\longrightarrow\infty} \sum_j^n \sum_i^n p(x,y) \ln_p(i,j)}{h(x+y)}, \tag{2}$$

$$h(x+y) = \ln (x+y)^{1/x+y}, \tag{3}$$

$$Q(x,y) = \frac{m(x+y)}{H(x)+H(y)}, \tag{4}$$

where $I(X, Y)$ is the mutual information between two sets of random variables X and Y. Here, X and Y correspond to two sets of label sequences characterized by clustering results and true labels, and the mutual information is based on the concept of information entropy, the formula of which is shown above; $H(X)$ is the information entropy of random variable X; $p(x, y)$ is the joint probability distribution function between x and y; $p(_{xi})$ is the marginal probability distribution function of xi. The range of NMI is 0~1), and its larger value indicates better clustering and vice versa.

Since land resource management covers a large amount of information and includes all information within the scope of work, it is very necessary to ensure the validity and accuracy of land resources and information, to ensure the fairness and openness of information in the work process, to ensure the accuracy and transparency of information, to ensure that all departments can effectively use it in information sharing, and to treat the system of$Q$and$P$to make the maintenance work simple. And the efficiency (A) of the working surrounding can be improved.

$$Q \times P = \frac{(q+p)}{a+b+c+d}, \tag{5}$$

$$A = \frac{\max(x+y) - \min(x-y)}{\min(x-y)}. \tag{6}$$

4. System Optimization Test

Black box testing is often used to test the overall functionality of software and the functionality of software with graphical interfaces and the external structure of the program, and the black box testing method requires that the testing of the modules that divide the system be explicit and then test the input and output for each module, and then, the output results are actually tested to compare the actual output of the system with the ideal output of the user, and if they are not consistent, then a preliminary decision can be made that a function of the functional module is problematic, with exceptions and usually for testing the exception location method. The next white-box testing method is needed to analyze the detection. In the analysis from a macro-micro perspective, software testing methods can be divided into two main black box testing as well as white box testing; the so-called black box testing functional testing uses the software system that has an opaque black box, from a macro perspective that is only able to operate the input data for the system, to determine the output of the system and whether the user expects the results to be consistent; black box testing mainly

focused on the input and output system, the system from a system. White box testing perspective and testing methodologies corner the code to cover all branches, all statements in the program code for all possible errors, a comprehensive traversal of the test.

Of course, traversing all branches,*c*can simplify the number of test case quality and is a worthy issue for discussion; white box testing has been widely used in software development work, helping developers to strive to find software defects and fix them in time. White box testing is transparent; white box testing mainly comes from testing, testing internal software products from the details, and the correction of problems found during the testing process. In the two phases of the software life cycle, software testing, which usually takes place after writing each module, is also called unit testing. Coding and unit testing belong to the same phase of the software life cycle; the software system does various comprehensive tests after the end of this phase, which is another phase of the software life cycle; the testing phase is very meaningful, as shown in Figure 6.

Test case is the process of software testing scenario test; test cases are used to determine whether the function of the software product is normal; test cases are often selected representative of some typical data, often by data analysts of the system's critical points, failure-prone points for analysis and extraction so as to design test cases, the system according to different functional modules, and test cases to design different programs; there are certain functional modules of the test of software test cases that are often targeted, comprehensive, and representative. Software testing benchmark test case, the structure of the system software product parameter variables, and the case of the execution of the test observe the software operation; the results, if the software preconditions are the same case, can indicate that the functional modules in the normal operation of the use case test need more cases to test and the test of time and withstand the test of time; the system is running stable and is a reliable system; if the test results and the expected results have a discrepancy, showing that there are certain problems with the functional modules, the system needs to be white box tested and tested in more detail to find the problems that exist in the existing system and to solve the program logic problems.

The use of information systems can effectively use data, the overall structure which is a system construction model with application requirements as the core. Information system can effectively improve the development, management, research, and use of land resources; electronic government system can help land resources to obtain accurate analysis and research, detailed access to understand the status of resources, dynamic understanding of the development of information changes in the increasingly changing daily data and information as well as mastering the latest trends in land resources, and the development trend of the analysis and monitoring of market resources, to provide customers with reliable decision support to do a strong backstage data security, strengthen resource management planning, and provide land and resources for the further development of the plan in China, which can provide accurate information services for government managers and at the same time can be in line

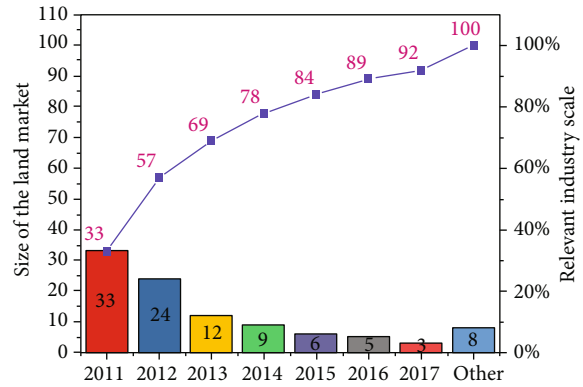


Figure 6: Comprehensive test of systematic induction.

with national development trends. In this paper, we have conducted detailed tests on each module of the system, which treats the user management module of the system as shown in Figure 7.

Select some representative values as feature values to test the functionality of the system. The system feature values need to represent and portray the software features of the system; have a comprehensive understanding of the design pattern, functions, and overall architecture of the system; and test the important modules of the system. The test cases of the system often need to be carefully designed, combined with the information system in different application scenarios for comprehensive testing; test cases often cover the common scenarios of the system as well as some special cases of abnormal scenarios and comprehensive testing of system functionality. Using an effective way to make up for the lack of supervisory power is also an important part of the urban emergency linkage system. In this paper, after in-depth research and analysis, a set of test cases is designed to meet the actual scenarios, and functional tests are conducted on several modules in the system to guarantee the testing quality as well as the credibility of the system.

This software test applies the testing theory of black box and white box, and the system is integrated and comprehensively tested with the software module as the basic unit. Land resources cover a variety of natural information, through a transparent information management system to achieve scientific management; transparent system can enhance the efficiency of land resource management and strengthen the degree of information sharing; facing the public development is more conducive to the power from the private sector, according to each information and demand analysis and taking into account the current situation of information technology development; this paper proposes the construction of integrated management information services for urban land resources In this paper, we propose a comprehensive management information service system for urban land resources and build an information sharing platform by realizing functional modules within each management business. We tested its efficiency improvement, and the results are shown in Figure 8, which shows that the overall efficiency is improved by more than 2 hours compared with the blank control. At the same time, the business process model of

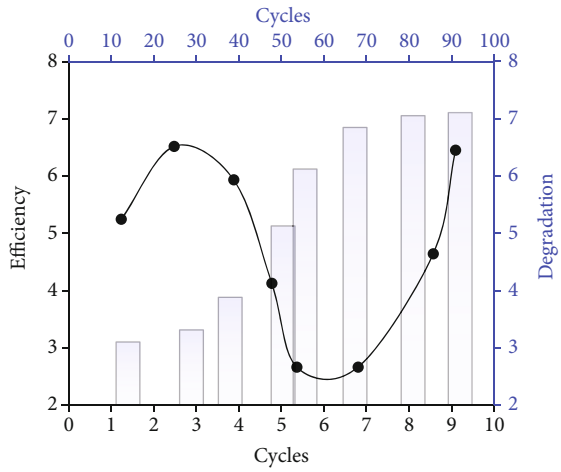


Figure 7: System cycle test performance.

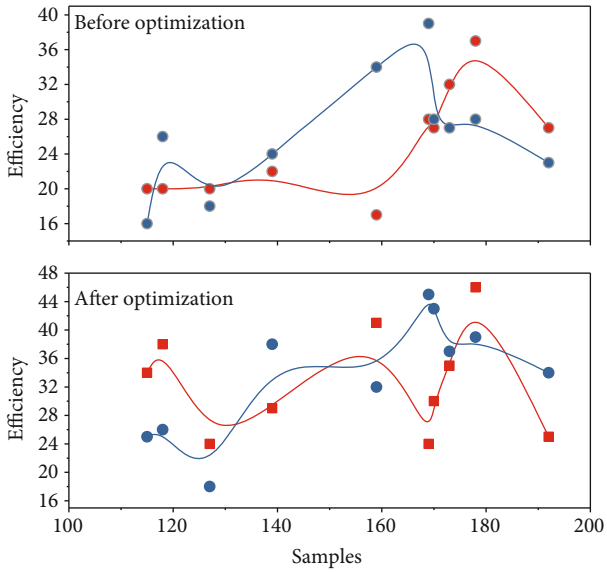


Figure 8: Test efficiency before and after optimization.

information management was optimized, which also helped the relevant departments using the system to improve their work efficiency, and at the same time, in order to avoid relatively isolated data not to be used effectively, the use of information systems can effectively use data; the overall structure of this is a system building model with application requirements as the core.

5. Conclusion

The construction of an integrated management information service system for urban land resources has built a platform for information sharing by realizing functional modules within each management business, while optimizing the business process model for information management and at the same time helping the relevant departments using the system to improve their work efficiency. After field research and analysis, the land resource information management system can be divided into four functional modules, which are basic land resource information management, land resource land use approval management, official document information management, and office business management, and these modules can effectively complete the work related to land resource management. The full and reasonable use and protection of land resources can ensure the implementation of sustainable development policies, and the development of land resource management is the basic requirement of national modernization and the inevitable trend in the country, and because land resource management covers a large amount of information involving a wide range of information, including information within the scope of work, it is very necessary to ensure the effectiveness and accuracy of land resources and information. In the analysis from a macro-micro perspective, software testing methods can be divided into two main black box testing as well as white box testing; the so-called black box testing functional testing is the software system as an opaque black box; from a macro perspective, it is only able to operate the input data for the system, in the daily changing data information dynamic understanding of the development of information changes, and grasp the latest trends in land resources, the trend of analysis, and monitoring of market resources, to provide customers with reliable decision support. It is a powerful background data guarantee for customers to provide reliable decision support.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (41971222); Program for Social Science in Henan Province (2019BJJ019); and First-class Disciplines Training Projects in Henan University (2019YLZDYJ12).

References

[1] C. Ren, Z. Li, and H. Zhang, "Integrated multi-objective stochastic fuzzy programming and AHP method for agricultural water and land optimization allocation under multiple uncertainties," *Journal of Cleaner Production*, vol. 210, pp. 12–24, 2019.

[2] H. Xu, C. Tian, and Y. Li, "Emergency evacuation simulation and optimization for a complex rail transit station: a perspective of promoting transportation safety," *Journal of Advanced Transportation*, vol. 2020, Article ID 8791503, 12 pages, 2020.

[3] R. González-Bravo, J. Mahlknecht, and J. M. Ponce-Ortega, "Water, food and power grid optimization at macroscopic

level involving multi- stakeholder approach," *Energy Procedia*, vol. 153, pp. 347–352, 2018.

[4] S. Hatfield, "Indigenous knowledge of the land and resources for optimization: redefining what management really means," *Ecology*, vol. 99, no. 7, pp. 1701-1702, 2018.

[5] Y. L. Xie, D. X. Xia, L. Ji, and G. H. Huang, "An inexact stochastic-fuzzy optimization model for agricultural water allocation and land resources utilization management under considering effective rainfall," *Ecological Indicators*, vol. 92, pp. 301–311, 2018.

[6] M. Zhao, L. Li, Y. Fang et al., "Optimization of intensive land use in blocks of Xi'an from the perspective of bicycle travel," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 241–249, 2021.

[7] N. Bavrovska and T. Shlikhta, "Land resources of the Zvenigorod district of Cherkasy region: assessment of the state and optimization," *Zemleustrìj, Kadastr ì Monìtorìng Zemel*, vol. 4, no. 4, pp. 53–60, 2018.

[8] M. Anis, A. Idrus, H. Amijaya, and S. Subagyo, "Utilizing coal remaining resources and post-mining land use planning based on GIS-based optimization method : study case at PT Adaro coal mine in South Kalimantan," *Journal of Geoscience, Engineering, Environment, and Technology*, vol. 2, no. 2, pp. 141–148, 2017.

[9] H. Qin, C. B. Andrews, F. Tian et al., "Groundwater-pumping optimization for land-subsidence control in Beijing plain, China," *Hydrogeology Journal*, vol. 26, no. 4, pp. 1061–1081, 2018.

[10] S. Harasimowicz, J. Janus, S. Bacior, and J. Gniadek, "Shape and size of parcels and transport costs as a mixed integer programming problem in optimization of land consolidation," *Computers and Electronics in Agriculture*, vol. 140, pp. 113–122, 2017.

[11] Y. Nie, S. Avraamidou, X. Xiao, E. N. Pistikopoulos, and J. Li, "Two-stage land use optimization for a Food-Energy-Water Nexus system: a case study in Texas Edwards Region," *Computer Aided Chemical Engineering*, vol. 47, pp. 205–210, 2019.

[12] Q. Wang, R. Liu, C. Men, and L. Guo, "Application of genetic algorithm to land use optimization for non-point source pollution control based on CLUE-S and SWAT," *Journal of Hydrology*, vol. 560, pp. 86–96, 2018.

[13] Z. Li, X. Deng, A. Arowolo, Q. Jiang, and H. Yan, "Adapting water scarcity for river basin: optimization of land uses," vol. 1, pp. 19–50, 2019.

[14] N. Xiao and A. T. Murray, "Spatial optimization for land acquisition problems: a review of models, solution methods, and GIS support," *Transactions in GIS*, vol. 23, no. 4, pp. 645–671, 2019.

[15] A. Singh, C. Chhablani, and L. Goel, "Moth flame optimization for land cover feature extraction in remote sensing images," in *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–7, Delhi, India, 2017.

[16] E. V. Panin, I. V. Yaurova, and A. A. Kharitonov, "Improvement of information and technological management of land resources and regulation of land and property relations," *Vestnik Of Voronezh State Agrarian University*, vol. 1, no. 60, pp. 226–233, 2019.

[17] K. S. Harmanny and Z. Malek, "Adaptations in irrigated agriculture in the Mediterranean region: an overview and spatial analysis of implemented strategies," *Regional Environmental Change*, vol. 19, no. 5, pp. 1401–1416, 2019.

[18] Y. A. Maglinets, K. V. Raevich, and G. M. Tsibulskii, "Architecture of the information system of evaluating of the land resources based on processing of spatial data," *Journal of Siberian Federal University: Engineering & Technologies*, vol. 11, no. 1, pp. 52–60, 2018.

[19] N. V. Nechyporuk, "Information support for the land registration: directions for upgrading of statistical reporting," *Sustainability*, vol. 80, no. 1, pp. 24–29, 2018.

[20] M. R. España-Villanueva and L. M. Valenzuela-Montes, "The role of information in plans for progressing in IWLRM," *Land Use Policy*, vol. 67, pp. 327–339, 2017.

[21] P. Danoedoro, "Multidimensional land-use information for local planning and land resources assessment in Indonesia: classification scheme for information extraction from high-spatial resolution imagery," *The Indonesian Journal of Geography*, vol. 51, no. 2, pp. 131–146, 2019.

[22] J. H. Wu, W. Wei, L. Zhang et al., "Risk assessment of hypertension in steel workers based on LVQ and fisher-SVM deep excavation," *Ieee Access*, vol. 7, pp. 23109–23119, 2019.

[23] F. Orujov, R. Maskeliūnas, R. Damaševičius, W. Wei, and Y. Li, "Smartphone based intelligent indoor positioning using fuzzy logic," *Future Generation Computer Systems*, vol. 89, pp. 335–348, 2018.

[24] W. Wei, Q. Ke, J. Nowak, M. Korytkowski, R. Scherer, and M. Woźniak, "Accurate and fast URL phishing detector: a convolutional neural network approach," *Computer Networks*, vol. 178, article 107275, 2020.

[25] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the Internet of Things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9846860, 1 page
https://doi.org/10.1155/2023/9846860

Retraction

Retracted: Traffic State Detection Based on Multidimensional Data Fusion System of Internet of Things

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] J. Zhang, W. Zhu, X. Wu, and T. Ma, "Traffic State Detection Based on Multidimensional Data Fusion System of Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1374186, 12 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 1374186, 12 pages
https://doi.org/10.1155/2021/1374186

Research Article

Traffic State Detection Based on Multidimensional Data Fusion System of Internet of Things

Jinxi Zhang,[1] **Wenying Zhu**,[1] **Xueying Wu**,[2] **and Tianshan Ma**[2]

[1]*College of Transportation Engineering, Chang'an University, Xi'an, Shaanxi 710064, China*
[2]*School of Economics and Management, Chang'an University, Xi'an, Shaanxi, China*

Correspondence should be addressed to Jinxi Zhang; 2018023008@chd.edu.cn

Received 12 April 2021; Revised 29 May 2021; Accepted 8 June 2021; Published 2 July 2021

Academic Editor: Wei Wang

In recent years, the rapid development of cloud computing, mobile Internet, Internet of Things, and other technologies has accelerated the explosive growth of the number and types of data in various industries. As one of the current hot research fields, the multidimensional data fusion system has received widespread attention all over the world. Road traffic and traffic management involve a wide range of participants; a variety of traffic behavior and traffic management means have produced a huge volume of traffic management data and have a very high application value. Therefore, it is of great significance to study the application mode of the traffic management multidimensional data fusion system and realize the asset value transformation of the multidimensional data fusion system. Starting from the safety supervision service of operating vehicles, this paper elaborates the traffic management multidimensional data fusion system platform for the safety supervision service of operating vehicles around the application requirements and architecture of the traffic management multidimensional data fusion system and the results of this research team. Paper with city comprehensive transport hub as the research object uses the Internet of Things technology; establishes a set of BIM model-based intelligent Internet operation management platform, multidimensional collection personnel, and vehicles, such as traffic, equipment, and business data information; realizes the integration of data, correlation, and instructions issued by ability, to support the building intelligence operation management needs of the business; and realizes traffic and transportation status detection.

1. Introduction

With the development of cloud computing technology, Internet of Things technology, and mobile communication technology, the operation and maintenance platform of the intelligent Internet of Things is more and more closely combined with people, and the communication between people and things and between things becomes more and more important. Therefore, the operation and maintenance platform of the intelligent Internet of Things emerges as The Times demand. Intelligent Internet of Things operation and maintenance platform, namely, the intelligent operation and maintenance management based on the Internet of Things, take various application service systems as the carrier and fully integrate management and scientific information technology [1].

In recent years, with the rapid development of urban modernization, TOD (Transit Oriented Development) urban integrated transportation hub projects, which are oriented by public transportation and integrate work, business and residence, etc., have emerged continuously. This paper takes TOD urban comprehensive transportation hub as the research object. Based on the Internet of Things technology, modern information technology means are adopted to build a set of intelligent Internet of Things operation and maintenance management platform based on building BIM model. The establishment of the operation and maintenance management platform of the intelligent Internet of Things realizes the purpose of combining the intelligent operation and maintenance management with customer experience. It collects the data information of personnel, vehicles, traffic, equipment, and business in a multidimensional way and realizes the ability

of data integration, correlation, and command issuing, so as to support the needs of the intelligent operation and maintenance management business of buildings [2–4].

Chen et al. [5] claim that traffic condition detection refers to the data generated during the detection and management of traffic elements such as traffic facilities, transportation equipment, and service objects by public security and road traffic management departments through the Internet of Things and information technology. According to its categories, it can be divided into traffic management government data, operation data, and perception data of the Internet of Things.

Due to the influence of the management system between the traffic and transportation management departments and the public security traffic management departments and between the traffic and transportation management departments and the traffic and transportation enterprises, Soysal et al. [6] claim that the traffic state data has the characteristics of mixed, numerous, and scattered data. At present, the application of big data of traffic condition detection is mainly for the management of their respective functions, the operation of their own enterprises, and the provision of social services to the public, and the data does not form an effective sharing and mining application.

At present, many scholars at home and abroad are looking for more efficient and appropriate research methods on the optimization of logistics transport nodes of supply chain, especially the application of information technology such as intelligent perception in the Internet of Things. Many scholars have applied the technology to the supply chain of manufacturing, logistics, food processing, and other industries. In terms of comparative research in this field abroad, the Internet of Things and intelligent communication technology are applied to the data fusion system; the early domestic research in this field is less; the specific case and application methods are too traditional; so in the Internet of Things multidimensional data fusion system of logistics transport research and research in the Internet of Things technology and logistics transportation control model architecture research, domestic and foreign researches on these two aspects are shown as follows:

(1) *Research on Component Logistics Related Aspects*. Some scholars in this field summarized a lot of relevant research materials from logistics transport nodes. The purpose of the technical research on the application of Internet of Things technology in the component logistics management of data fusion system is to optimize the distribution of resources in the logistics stage and maximize the overall benefit. The current logistics network information technology helps to determine the important factors in the supply chain to optimize and improve the operation performance, improve the operation management process, establish a rapid response mechanism, and gain competitive advantages [7]. The information operation of supply chain management needs the support of modern Internet of Things technology and mobile Internet, especially in the aspect of logistics management, which helps to integrate the information data of upstream and downstream enterprises, establish the logistics information platform, establish effective links between enterprises, and further reduce the management cost [8]. Research on enterprise management and enterprise management mode reform in the supply chain, strengthen the communication between enterprises on the main nodes of the supply chain, realize information sharing, improve the precision of supply chain operation, and give relevant opinions on performance [9]. In the optimization of supply chain management, the most important thing is to reduce the enterprise's own risks by realizing the information technology of the Internet of Things based on the management performance and the enterprise's core competitiveness. In the process of enterprise demand analysis, the functional structure design scheme of the agile supply chain management system is given, and the information technology of Internet of Things is applied to the agile supply chain management [10]. The construction of enterprise intelligent supply chain management framework can realize the optimization of production management, procurement management, logistics management, information management, after-sales service, etc. Based on the understanding of the operation characteristics of supply chain, it is considered that it is a need for the integration of various resources; internal is dynamic and complex, so it is necessary for enterprises or projects to establish a seamless information integration management framework in the process of operation [11]. The material monitoring system based on BIM-GIS can be constructed to track the whole material supply chain in real time and analyze the relationship among logistics supply, resource status, and internal supply [12]. Through the data fusion system on the basis of information sharing platform based on BIM, the establishment of the actual model, the integration management of the two parties can increase the profit. Based on the supply chain management model, this paper analyzes the benefit and cost of housing industrialization and explores the good operation effect of housing industrialization mechanism under the supply chain based on the integration path. Refer to the above literature found that scholars both at home and abroad, the efficient operation of the main logistics nodes in each stage, various information data transmission and access to timely, with Internet information technology as a logistics support, constantly explore the Internet of Things technology based on data fusion system IoT multidimensional data fusion system in logistics stage provides a new solution [13]

(2) *The Internet of Things Plays an Important Role in Practical Project Applications*. Many scholars in the field of the Internet of Things are studying the combination of the Internet of Things and related fields and constructing theoretical models to predict and manage the upstream and downstream of the whole

supply chain and optimize the performance of projects [14–16]. Analyze the supply chain process and establish the quality information management model based on the Internet of Things technology and SOA architecture, so as to realize the dynamic monitoring and management of quality [17]. In order to realize the dynamic monitoring of the data information at each stage of the supply chain, the RFID technology is proposed to provide technical support, and the existing technology is used to improve the inventory management of enterprises [18]. In order to optimize the route of the product logistics stage, the traceability system based on the Internet of Things technology and the Internet of Things network is designed to monitor and manage the products in real time and improve the route problem in the logistics nodes [19]. RFID technology has some defects, for fast moving objects can not be effectively and quickly identified and read, developed an automatic positioning sensor system, using a robot with a reader, combined with Wife network to carry out regular space scanning, to achieve real-time monitoring and positioning of objects [20]. It is difficult to achieve the desired effect in the operation of industrial projects due to the various specifications of accessories. The pipeline traceability system based on the Internet of Things technology is designed, and the technical optimization system is established for effective application in actual projects, so as to provide the basis for the management and decision-making of enterprises. In order to enable retailers and manufacturers to obtain the maximum benefits in the upstream and downstream process of the supply chain, a set of information platform based on two-way design is conceived, based on the Internet of Things technology, and the feasibility of platform establishment is analyzed [21–23]. Enterprises and their assets are faced with the development and application of relevant business software in the supply chain management of this life cycle, and it is suggested to establish an IoT traceability system based on actual projects. In order to prove the feasibility of product real-time monitoring and tracking system, product coding and Internet of Things information technology are managed with the combination of the Internet of Things (electronic coding technology). At the same time, the operation mode and workshop equipment layout method of the system are given, and the demonstration experiment is carried out [24]

2. Architecture of the Multidimensional Data Fusion System of the Internet of Things

2.1. Overview of the Internet of Things and Related Technologies

(1) *RFID and Internet of Things Technology*. Radio Frequency Identification (RFID) technology, also known as Radio Frequency Identification, is a short-range communication technology. Its working principle is as follows: firstly, after the label is put into the magnetic field, it receives the radio frequency signal sent by the reader and sends the product information stored in the chip, or the label initiatively sends the signal of a certain frequency. After the card reader reads the information and decodes it, it is sent to the traceability system server for relevant data processing. Internet of Things is the English abbreviation of the Product Identification Code, which is the global unique and unique identification for all objects in the supply chain. The Internet of Things is usually stored on an RFID electronic tag, which includes an antenna and a silicon chip. When identifying the label, through the communication data connection, we can know the origin information, workmanship process and production date information, etc.

(2) *Sensor Network and Embedded Technology*. Wireless sensor network can be considered to be composed of three parts: data acquisition network, data distribution network, and control management center [25–28]. The main components are sensors, data processing unit, and communication module integrated node. Each node individually forms a distributed network, which is optimized to send the collected data via radio waves to an information processing center, where it is processed by dedicated equipment. Embedded technology is a computer system, typically a firmware, that consists of a single program that performs logic programs

(3) *Positioning System*. Positioning system (GPS) is an interdependent assembly, receiving equipment or positioning components to determine the position of the system, and has three parts: space components a GPS constellation, Ground Part-Ground Monitoring System, and user equipment part a GPS signal receiver. Various positioning systems greatly improve the working efficiency of the Internet of Things and provide interoperable communication for frequent activities of various industries through precise positioning information [29, 30]

(4) *Cloud Computing Technology*. Cloud computing mainly takes the Internet as the center; establishes large capacity storage servers; provides safe, fast, and convenient network services on the basis of open standards and services; and makes the "cloud" of the Internet become the data center and computing center of all walks of life

(5) *IPv6*. IPv6 is the abbreviation of Internet Protocol Version 6, mainly used for network address location, and is a new generation of network protocol, from the data protection process and data storage, with higher confidentiality and fast transmission. Figure 1 is the structure diagram of core technologies related to the Internet of Things

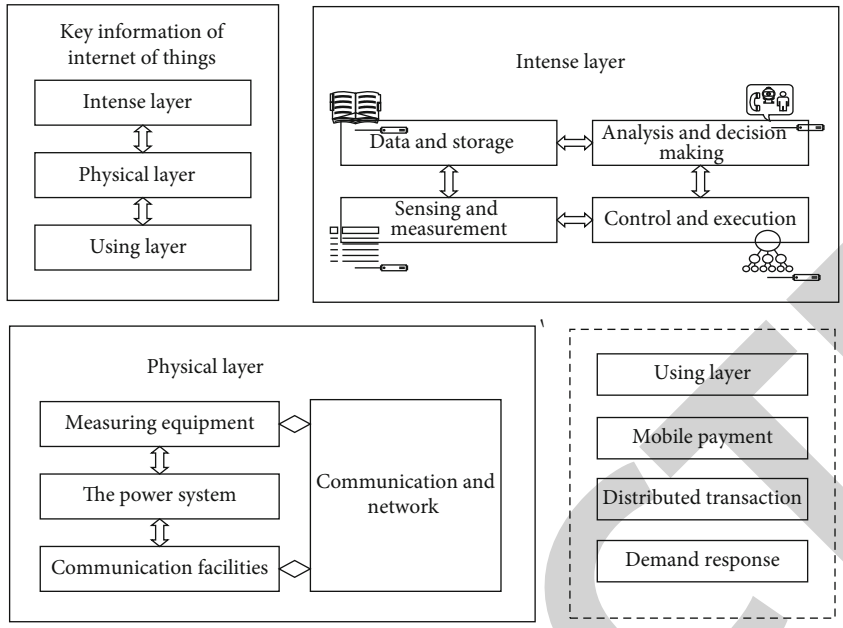


Figure 1: Structure diagram of core technologies related to the Internet of Things.

The application of IoT technology can be divided into several stages: feasibility study stage: (1) + Internet related technologies, plus one for radio frequency technology, E IoT coding techniques such as prefabricated in prefabrication composite beam, column, composite board, panel, prefabricated stair IoT application of multidimensional data fusion system, and design based on component traceability system platform, aiming at improving the efficiency of early preoperation. The preliminary planning of the whole assembly project is systematically and comprehensively analyzed so as to achieve the expected effect in the feasibility study stage. (2) Production-construction stage: the management of the Internet of Things in the design stage is based on the production management of the information system platform, and the transportation, construction, and installation of the multidimensional data fusion system of the Internet of Things are the main stages of management. RFID chip embedded component to produce a construction whole process, the chip is used as the component identification code and the distribution of media, from the material supply, production orders, production, storage, transportation, storage, and construction loading process, such as by collecting and analyzing information as well as the component plus state and properties, and the development of BIM interface in order to realize data integration, to facilitate the overall efficiency analysis, and monitor the entire construction cycle of the data fusion system to ensure overall control of the construction process. (3) Operation and maintenance stage: the subsequent operation and maintenance can be traced during the manufacturing process, and the relevant information of each component can be traced back. Therefore, if a failure occurs during subsequent installation and debugging, it can be found in the management system and recorded. In the later stage, it can also collect user feedback and accept user complaints and maintenance information. Traceability of quality defects can be achieved throughout the full cycle, and a database of IoT components and other products can be established.

The Internet of Things technology is mainly applied in several aspects of logistics node work: (1) order management: the field system automatically calculates the ID number of the multidimensional data fusion system of the Internet of Things labeled with RFID tags and uses MRP (Material Requirement Planning) to query demand information and generate orders intelligently. And according to the materials, suppliers and other information to carry out the actual adjustment of procurement requirements and to meet different enterprises for different business procurement demand management requirements. The production and logistics stage information of the product is stored in the system, which facilitates the visualization of the whole process and reduces the management cost. Table 1 is the structure of the information storage system of the Internet of Things:

(2) Device operation: IPv6 can assign IP addresses to all electromechanical devices in smart buildings. In component traceability system, general users can use terminal devices or web pages to inquire, and terminal managers can make decisions according to the actual situation. The query information includes component name, corresponding company name, component history information, distribution center, supply chain sale distribution information data, and output data statistics as well as third-party monitoring backup data, which helps to pay close attention to whether the transportation work equipment is in normal condition and can ensure that the corresponding type of fault is solved. (3) Monitoring the working state: for the equipment types such as the multidimensional data fusion system of the Internet of Things (IoT), the IoT embedded technology and cloud computing can be used to automatically record the production and transportation time of components. In addition, aggregated

Table 1: Internet of Things information storage system.

System structure	The name of the annotation
Electronic product code	The coding system
The Internet of Things coding standard	Identify target setting code
Radio frequency identification system	The Internet of Things tags
Information network system	
Read/write device	The Internet of Things middleware
Object name resolution service	Entity markup language
	Enter the ID number of the Internet of Things multidimensional data fusion system
Identify the Internet of Things tags	The Internet of Things system software
Locate the corresponding information of the multidimensional data fusion system of the Internet of Things	Describes the information language of the multidimensional data fusion system of the Internet of Things

GPS data and GSM can be used to track the near and far positions of vehicles, helping to accurately locate and measure on the terrain and providing automatic instructions for location modification for logistics transportation through the use of virtual maps. Figure 2 is the schematic diagram of the positioning module.

(4) Save energy consumption: by using cloud computing and embedded technology, the electricity used can be calculated, which can help the factory of multidimensional data fusion system of the Internet of Things to take energy-saving measures. In logistics transportation, according to the technology, the machine can also record and send back the idle time information of energy consumption, so as to adjust the vacancy rate of transport vehicles without having a direct impact on production and transportation, further reducing the consumption cost. (5) Equipment and data operation and maintenance: sensors in machinery and equipment transmit information about vehicle logistics and transportation status and its service or maintenance needs, which helps to repair machinery and equipment, save time and improve the efficiency of the machine and reduce the occurrence of failures.

2.2. Organization Structure Analysis of Multidimensional Data Fusion System. The operation and maintenance platform of the intelligent Internet of Things studied in this paper is based on building BIM model and adopts the management strategy of "platform + ecology" to collect data information of personnel, vehicles, traffic, equipment, and business in multiple dimensions, so as to realize positioning, navigation, and data interactive management of people, vehicles, and things in the building of TOD urban comprehensive transportation hub. And realize the public personalized information push and emergency linkage command and dispatch function, so as to realize the intelligent management of the whole building. The intelligent IoT operation and maintenance platform include the IoT perception layer, the network layer, and the platform application layer. The platform architecture is shown in Figure 3.

(1) Comprehensive perception: building networking is comprehensively arranged. Various sensors, such as temperature and humidity sensors, two-dimensional code tags, RFID tags, readers, cameras, detectors, and other sensing terminals, are used to comprehensively perceive the running situation of buildings in TOD urban integrated transportation hub, so as to establish a full connection between people, vehicles, objects, environment, and services. (2) Real-time communication: the network layer is composed of wide area network, local area network, and 2G/3G/4G/WLAN network. It is the backbone of the whole Internet of Things and is responsible for transmitting and processing the data information acquired by the sensing layer. (3) Deep integration: through connection management platform, multidimensional data fusion system platform, and integrated communication platform, the application layer of the platform fuses various data and communication information in TOD urban comprehensive transportation hub buildings to support multidimensional intelligent analysis and processing of building business of TOD urban comprehensive transportation hub, so as to realize whole-system communication collaboration.

According to the characteristics and actual status of component logistics and transportation, in order to improve the actual efficiency, the Internet of Things application platform on the market is included in the logistics and transportation nodes, as shown in Figure 4. The system platform is beneficial to the operation and maintenance of data information and management information. For logistics nodes, they can take more coordinated measures to solve the problems caused by the cross of component information and have more visualization and low-cost information operation. The system is mainly an auxiliary transportation node.

Based on the production standards of the data fusion system, the new wireless sensor network technology, Baidu Geographic API, 4G and ZigBee network, and other technical supports were adopted to build the model organization framework and get familiar with the specific information of each participant. Ensure supply chain interactivity, reliability, and security throughout the system. The database adopts main information storage and data processing, which is mainly based on raw material information, production and processing log information, sales and distribution information, and

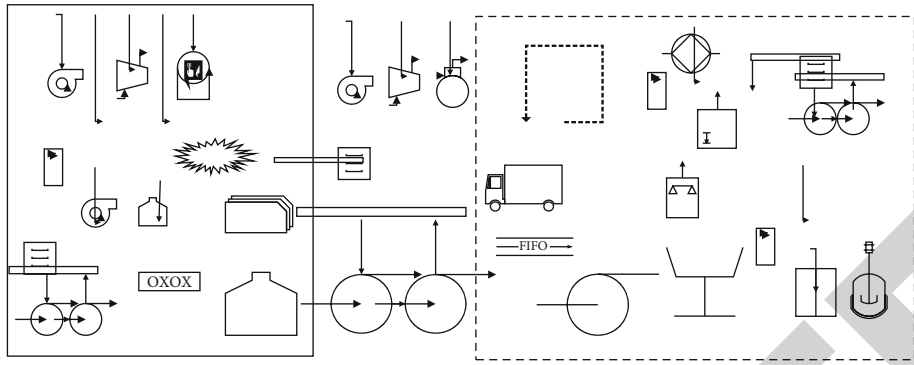


Figure 2: Schematic diagram of positioning module.

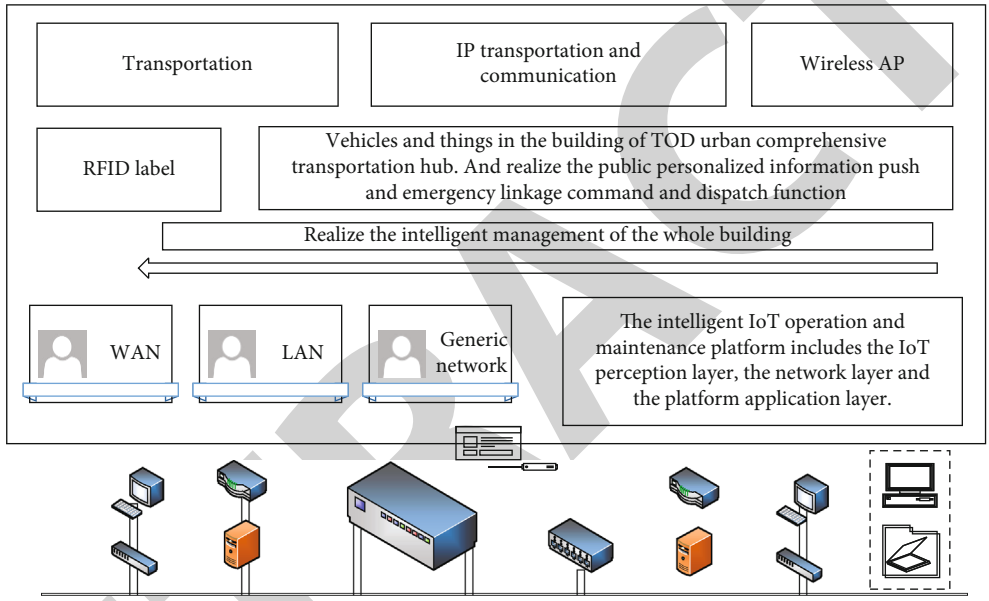


Figure 3: Network platform architecture.

logistics process information of the multidimensional data fusion system of the Internet of Things. Combined with the enterprise management system (ERP), the statistical table of traceability information will be created. The statistical table is mainly the component data information. The statistical table will be connected to the background management and can be modified and updated by the main management organization. RFID technology is optimized and combined with BIM database, and the upgraded RFID technology associated with BIM database can perform receiving, transmission, and feedback functions, so as to establish an intelligent monitoring and traceability system for component supply chain. With the support of RFID&E Internet of Things as the main technical means, the supply chain traceability system of the multidimensional data fusion system of the Internet of Things is built. The traceability of the system is based on the complete processing of RFID information and the standard of the Internet of Things code. The system has three main code formats: E IoT -64/E IoT -96/E IoT -256. According to the size and characteristics of the component system, the project adopts the Internet of Things coding format, as shown in Table 2.

3. Construction of Transportation State Detection Model Based on the Internet of Things

3.1. Research on the Transportation State Detection Model Based on the Internet of Things. According to the internal factors and external factors of the multidimensional data fusion system of the Internet of Things, combined with the feasibility of index data and the actual situation, the index of component logistics transport node model is selected. The main indicators include input index and output index. Data are provided in many aspects, including production data, personnel data, transportation data, and other index data. Based on the above data, a comprehensive and systematic study is carried out on logistics transportation nodes of the multidimensional data fusion system of the Internet of Things. Since the data envelopment method is a quantitative method that is not determined by the weight of each index, there is no need to do the corresponding index definition and other explanations in determining the definition of the index. The index selection is mainly based on the characteristics of components

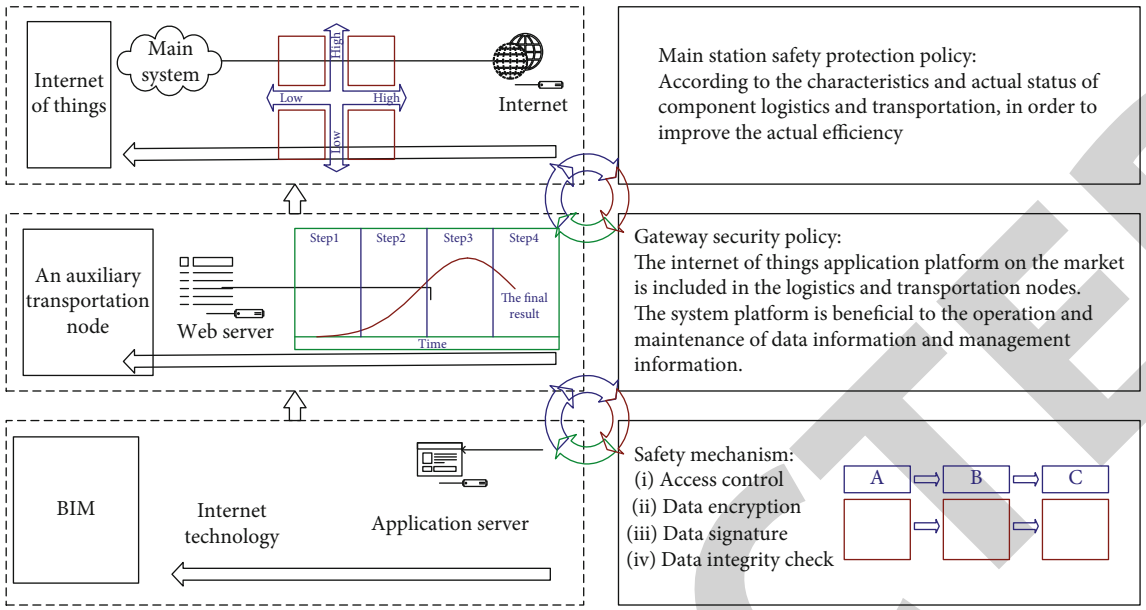


Figure 4: Platform architecture of logistics transport data system.

Table 2: Coding principle.

Project	Management process	Member management	
Basis of logistics system construction model	Production on the basis of	Wuhan City assembly Wuhan People's Government into the Hubei province assembly	
Process information	Regulation implementation measures "notice of building" whole supervision points	The manufacturer has complete components	
Technical basis	Type building construction pipe steps on accelerating the development of building construction quality safety installation	Wireless sensor network, etc.	
Transformation	XXX (header)	XXX (regulation)	XXXXX (text)
	Management	Content	The serial number
Specific information	Baidu geographic API, GPS positioning	Specific information of government test data participants	
Meaning	8 and 24 to 36	The header	
The technical level	Storage level	The management level	
ERP	RFID + BIM	Raw material information	
Plant manager	Production processing log information sale distribution information logistics process information	Logistics enterprises	

themselves, the survey of professionals and the data available for collection, etc.

The node model is mainly used to analyze the transport node control research of the multidimensional data fusion system of the Internet of Things, and the data envelopment analysis method can be used to evaluate the overall efficiency of logistics transportation. This is a relatively effective efficiency evaluation method, which can subdivide and effectively evaluate the efficiency of decision-making units with multiple inputs and outputs (F641). Component supply chain logistics transport node is a typical complex case of multiple inputs and outputs. This method can analyze the effects between different factors through variable returns to scale. Based on this paper, the DEA method is used to evaluate the research on the control of logistics and transportation nodes.

BBC model is mainly used to measure pure technology and scale efficiency in the case of variable remuneration for decision-making units. Generally, it is applied in the analysis cases of variable remuneration for scale with multiple inputs and multiple outputs. However, from the research focus direction of this paper, component production is a product with variable scale and large input and output. Therefore, the BBC model with variable returns of scale should be considered first in the application model.

In this case, a number of DEA models were selected, and BBC model was selected after comprehensive consideration.

Original data	Level1	Level2	Level3	Level4
Series1	1377.85	2457.68	7323.32	2522.46
Series2	1397.42	2332.21	1955.8	2457.68
Series3	1416.99	2206.74	−3411.72	2392.9
Series4	1436.56	2081.27	−8779.24	2328.12
Series5	1456.13	1955.8	−14146.8	2263.34

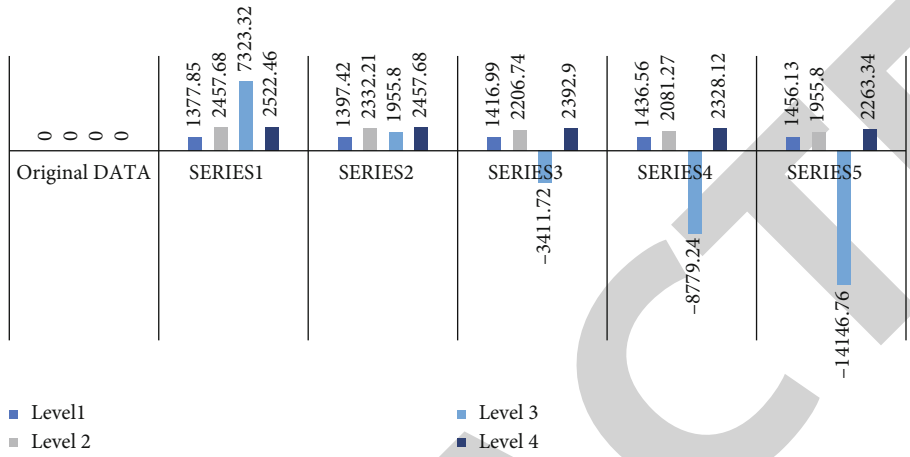


Figure 5: The original data set.

The output-oriented BCC model is the main one with several decision-making units (denoted as N). Each decision-making unit DMU is divided into M inputs, P outputs, and M inputs. Therefore, the output input ratio of DMU should be measured and expressed as:

$$i_l = \frac{\lim_{q\longrightarrow\infty}\sum_{i=1}^{q} v_i \cdot y_{il}}{\int \sum_{i=1}^{q} w_i \cdot z_{il}}. \tag{1}$$

The efficiency value obtained by all DMU using the above weight is limited within the interval [0,1], that is:

$$i_l = \frac{\oint v_q \cdot y_{ql}}{\oint w_q \cdot z_{ql}} \mapsto 1. \tag{2}$$

The efficiency evaluation index represents the empirical efficiency obtained from the multi-index input and multi-index output of the decision-making unit, so the model can be established and optimized according to this principle. Formulas (3) and (4) are the basis of the BBC model:

$$\min \frac{\overrightarrow{\oint v_q \cdot y_{ql}} \Rightarrow 1}{\oint w_q \cdot z_{ql} \longrightarrow {}^{v_q \cdot y_{ql} \oplus w_q \cdot z_{ql}}_{\longrightarrow}\infty}, \tag{3}$$

$$\frac{\overrightarrow{\oint v_q \cdot y_{ql}} \Rightarrow 1}{\oint w_q \cdot z_{ql} \longrightarrow} \overset{y_{ql}}{\underset{z_{ql}}{\rightleftarrows}} 1. \tag{4}$$

Dual model of formula is:

$$\max e \ni E, \tag{5}$$

$$\text{s.t.e} = \oint v_q \cdot y_{ql}, \tag{6}$$

$$\lim_{q\longrightarrow\infty}\sum_{i=1}^{q} v_i \cdot y_{il} = \frac{w_q \cdot z_{ql}}{\sqrt{v_q \cdot y_{ql}}}, \tag{7}$$

$$\lim_{q\longrightarrow\infty}\sum_{i=1}^{q} w_i \cdot z_{il} = \frac{w_q \cdot z_{ql}}{\sqrt{v_q \cdot y_{ql}}}. \tag{8}$$

The judgment of BCC evaluation model can be divided into the following three situations: when $E = 1$ and each optimal solution satisfies the condition that the slack variable is 0 and DMU is effective, then the optimal decision unit of comprehensive efficiency has reached the optimal combination, and the maximum DEA is effective. When $E = 1$ and the slack variable is not equal to 0, the DMU is valid, and then, the DMU is not the best technical efficiency and the best scale at the same time. When $E < 1$, it indicates that the comprehensive efficiency is invalid, and the technical efficiency and scale efficiency have not reached the optimal scale, so it can be adjusted through projection analysis to achieve DEA effectiveness. After the establishment of the model, the index data are selected. In the process of data processing, the index data of the model should be analyzed with the help of the DEAP2.1 software, mainly to obtain the comprehensive efficiency, pure technical efficiency, and scale efficiency as well as their average value.

3.2. Optimization of Traffic State Detection Model and Analysis of Relevant Data. The experimental analysis method of this paper applies the data envelope model for quantitative analysis and obtains the practical and feasible node control method through the quantitative control of the problems existing in the logistics and transportation nodes in the

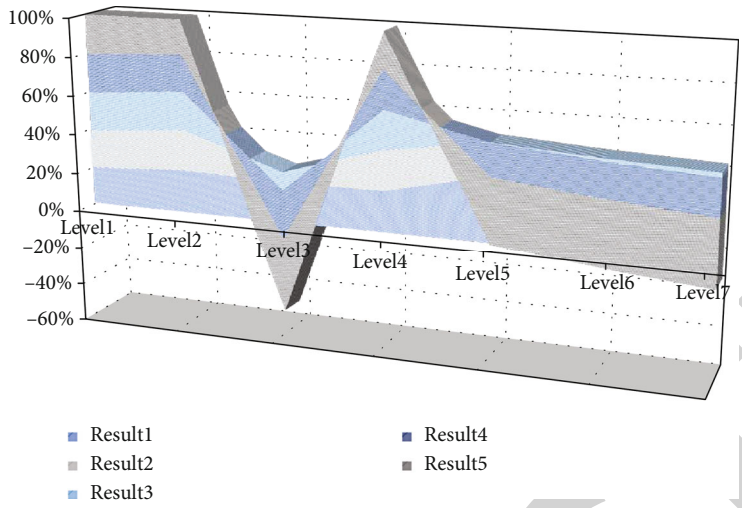


Figure 6: The overall return to scale.

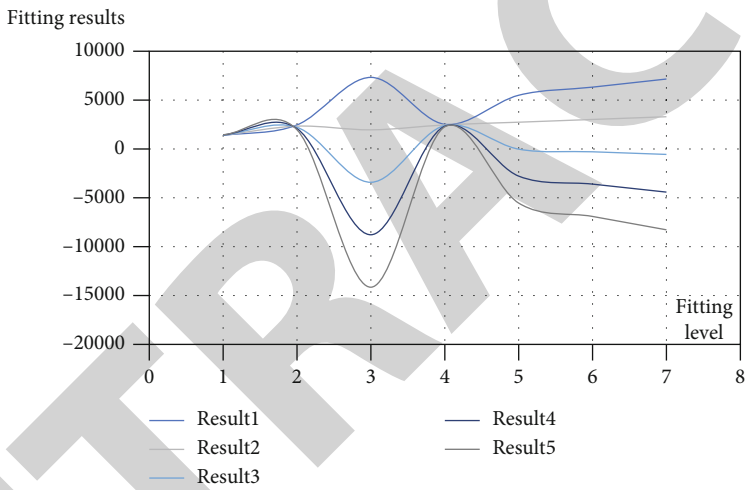


Figure 7: Route planning and vehicle configuration 1.

component supply chain of a component factory. The model data calculation and projection analysis of the three projects draw the following conclusions. This paper analyzes the overall efficiency of the input and output of the multidimensional data fusion system of the Internet of Things in three data fusion system projects of a component factory in three years from 2019 to 2020 from two aspects of efficiency and projection analysis. The original data are shown in Figure 5.

Through the above three projects with a total of nine samples, this paper analyzes their comprehensive efficiency, pure technical efficiency, and scale efficiency and explains their input and output.

(1) *Start with the Comprehensive Efficiency*. As can be seen from the table, the comprehensive efficiency of samples 1, 2, and 3 is 1

In addition, the input and output redundancy rate shows that the 1, 2, and 3 samples are completely effective. The effective scale efficiency of DEA shows a trend of gradual scale rewards. The overall operation control of the project is effective, and no additional supervision is required. But judging from the overall data, the profit of output value is low, and the product plan has been completed more. It can be seen from the table that the comprehensive efficiency of samples 4, 5, and 6 shows an increasing trend. The scale efficiency DEA of samples 4 and 5 is invalid, while the scale efficiency DEA of sample 6 is effective. In sample 4, input factor 1 has redundancy of 750061.00, output factor 2 increases by 6.458, output factor 3 increases by 438.926 + 4243.966 = 4682.89, and output factor 4 increases by 369.851 + 4801.271 = 5171.12. In sample 5, the output of output factor 1 increased by 1127767.44, the output of output factor 2 increased by 8.136, the output of output factor 3 increased by 551.076 + 6331.126 = 6882.20, the output of output factor 4 increased by 557.343 + 4824.242 = 5381.59, and the input of input factor 2 was redundant. Sample 6 has no redundancy in input-output. On the whole, the return to scale of samples 4, 5, and 6 shows an increasing trend. It can be seen from the profit that the actual profit is low, and the overall corresponding return for the multidimensional data fusion system

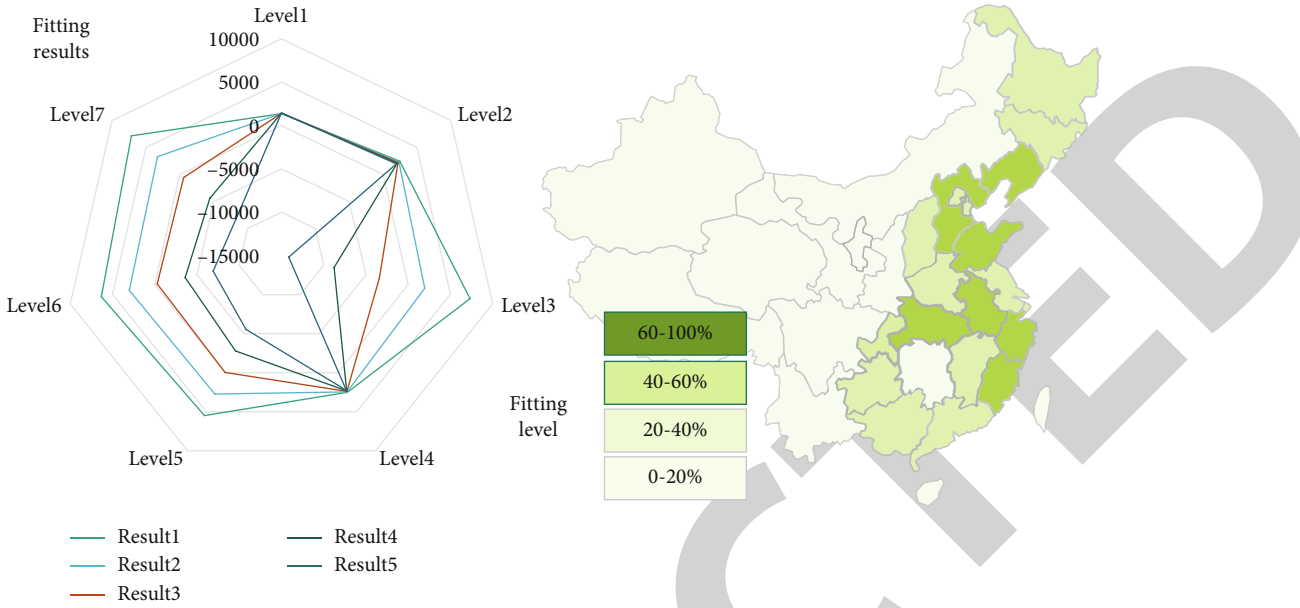


Figure 8: Route planning and vehicle configuration 2.

of the Internet of Things logistics supply chain is low. The project is mainly in loss, mostly with government subsidies. In samples 8 and 9, there is no redundancy in input and output, and the sample value is fully valid. In sample 7, the input redundancy value is 364530, and the overall return to scale shows a gentle trend, as shown in Figure 6.

As is shown in Figure 6, results 1 to 5 correspond to valid samples (1, 2, 3, 5, 7) and their overall return to scale. Among them, result 3 is the most stable with a higher average value, followed by result 2. Results 5 have a high upper limit and a low lower limit, which may obtain the best predicted value but is the most unstable. Results 1 are similar. Result 4 has no analytical advantage.

(2) From the perspective of pure technical efficiency, the pure technical efficiency of samples 1, 2, and 3 remained stable for three years. Therefore, input and output are the effective use of resources, and the Internet multidimensional data fusion system factory has promoted production efficiency in production management and technological innovation, thereby optimizing the project. The pure technical efficiency of samples 4, 5, and 6 shows an increasing trend year by year. In 2018 and 2019, the pure technical efficiency is lower than 1, indicating that there are problems in the management and technology of input and output in the production process of the factory. In 2018, the output profit is in deficit, indicating that the management fails to follow-up with the market, and the component factory fails to keep up with the update of information technology. Most of the staff are workers who have just received simple training. Problems such as poor communication and unclear understanding occur during the management process with managers. Due to the increase of factory output in 2019, input-output resources are not properly used in the process of technical follow-up. Samples 7, 8, and 9 pure technical efficiency except in 2018, the factory is still preparing for the project, so the output is 0, and a small amount of capital is invested, so the pure technical efficiency is 0. Irrational situation occurs in the calculation of input-output resources, which belongs to the sunk cost in the actual process. In both 2019 and 2020, the pure technical efficiency is 1, indicating that resources have been reasonably utilized in the process of input and output of the project, and there has been a great improvement in management and technology. The associated multidimensional data of traffic condition detection will be used for fusion, so as to play a greater role

(3) From the perspective of scale efficiency, the scale efficiency of samples 1, 2, and 3 is 1, indicating that the industrial structure planning of the IoT multidimensional data fusion system factory project is reasonable. By optimizing various configurations, the output unit meets actual production needs. File information update, production route optimization, and internal storage space layout meet actual requirements. The supply chain of component factories is perfect, and the scale effect shows the improvement of scale efficiency, but the return to scale does not actually increase under the condition of the same input. The scale efficiency of samples 4, 5, and 6 showed an increasing trend in three years. In 2018 and 2019, the scale efficiency was not effective by DEA, indicating that in the process of actual production and component configuration, through improved management, personnel training, and technical

optimization, there was a positive trend in data continuously, and the return to scale was also increasing. The scale efficiency of samples 7, 8, and 9 shows an increasing trend year by year. In 2018 and 2019, the scale efficiency is not effective by DEA. In the component supply chain of the multidimensional data fusion system of the Internet of Things, there are some problems in the process of logistics and transportation, such as route planning and vehicle configuration, as shown in Figures 7 and 8

Compared to other literature analysis method, this article through to overall comprehensive analysis of logistics node, PC components from the data, on the basis of quantitative analysis to the actual situation to adjust the direction, through the SWOT matrix analysis of the whole node external quality and internal opportunities and threats, clearly understands the advantages and deficiencies, to the problem put forward the deficiency of the modifications of optimization. As for the integration of the scheme, we adjusted the component logistics and transportation nodes from four aspects of logistics cost, warehouse management, route planning, and Internet of Things technology, so as to achieve the completion of the planned goals. In terms of implementation strategy, from a macropoint of view, the overall node control strategy is constructed, and then, the control process framework is constructed in the control analysis before, during, and after the event, and the steady and continuous follow-up is carried out to ensure that the work is within the reasonable operation range, and the work efficiency of the logistics and transportation nodes is effectively improved.

4. Conclusion

In this paper, the research on logistics transportation nodes of the multidimensional data fusion system of the Internet of Things starts from the micropoint of view, combines the macrooverall planning, innovatively combines qualitative analysis and quantitative data analysis, establishes a system platform and two reasonable models, and makes an in-depth study on the component logistics transportation node control. Component transport node control mainly has many problems in transport node management, logistics cost, warehouse management, route planning, Internet of Things technology application, and so on. In view of this, in the research of component supply chain logistics transportation node control, logistics node management should be familiar with the application field of the Internet of Things. According to the characteristics of the logistics node process, combined with Internet technology, a logistics tracking system is established. A comprehensive analysis of the entire life cycle from production to construction of logistics nodes can be used to establish a traceability system platform. This paper discusses the function, process steps of the system, and how to assist the logistics and transportation node control of the multidimensional data fusion system of the Internet of Things and provides a theoretical basis for the logistics and transportation node control. Establish the component logistics transport node control model. Based on the data envelopment method, the BBC output-oriented model is built. By analyzing the output and input index data, it is concluded that the average comprehensive efficiency of the three projects is 0.932, 0.832, and 0.985. It can be seen that the overall efficiency is not effective by DEA. The overall input and output resources that have not been effectively used include unrealistic planned level profits, less redundant fixed investment, parts distribution batches, low production levels, and storage efficiency. Predictive analysis adjusts the effective indicators of DEA to achieve theoretical rationality and provides corresponding change measures to provide model data support for improvement. The analysis explains the data node control model, plans specific improvements, and optimizes the logistics nodes of the IoT multidimensional data fusion system. Strengths and weaknesses list external and internal opportunities and threats through the SWOT matrix and comprehensively integrate solutions. Decisions are made from four aspects: logistics cost, warehousing management, route planning, and Internet of Things technology. Make adjustment to the component logistics transport node and realize the completion of the planned goal. In the implementation of the strategy, the overall node component control strategy is analyzed from a macroperspective, and the before and after control analysis is performed. Through the control process framework and continuous and steady follow-up, the work is ensured to operate in a reasonable manner, and the work efficiency of the logistics node is improved. From the macro- and microlevel, the node control strategy has been improved in an all-round and multilevel manner.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Jilbab and A. Bourouhou, "Efficient forest fire detection system based on data fusion applied in wireless sensor networks," *International Journal on Electrical Engineering and Informatics*, vol. 12, no. 1, pp. 1–18, 2020.

[2] B. Sliwa, N. Piatkowski, and C. Wietfeld, "The channel as a traffic sensor: vehicle detection and classification based on radio fingerprinting," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7392–7406, 2020.

[3] A. Rega, F. Vitolo, S. Patalano, and S. Gerbino, "A sensor data fusion-based locating method for large-scale metrology," *Acta IMEKO*, vol. 9, no. 4, pp. 136–143, 2020.

[4] W. Ding, X. Jing, Z. Yan, and L. T. Yang, "A survey on data fusion in internet of things: towards secure and privacy- preserving fusion," *Information Fusion*, vol. 51, no. 3, pp. 129–144, 2019.

[5] C. Chen, B. Liu, S. Wan, P. Qiao, and Q. Pei, "An edge traffic flow detection scheme based on deep learning in an intelligent

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9999113, 11 pages
https://doi.org/10.1155/2021/9999113

Research Article

3D Virtual Animation Instant Network Communication System Design

Jing Liu,[1] Qixing Chen,[2] and Xiaoying Tian[1]

[1]*College of Culture and Art, Chengdu University of Information Engineering, Chengdu 610225, China*
[2]*College of Communication, Chengdu University of Information Engineering, Chengdu 610225, China*

Correspondence should be addressed to Xiaoying Tian; txy@cuit.edu.cn

Received 9 March 2021; Revised 15 June 2021; Accepted 18 June 2021; Published 1 July 2021

Academic Editor: Wei Wang

This project uses Openfire to implement a virtual 3D animation instant messaging system, which is easier to use and more expandable. The main work of the client is to implement the Extensible Messaging and Presence Protocol (XMPP) and use XMPP to transmit data to the server side and receive data from the server side, while Openfire is built by the server side to use. To address the problem that the current mainstream face key point localization model is less robust to complex environments, this project adopts a deep learning-based approach to design and implement the face key point localization model, through data preprocessing, model design, and model training, to achieve a robust model that can locate 68 face key points and complete the migration of the model to mobile. The current video communication often suffers from delay and lag, so this project uses face key point data instead of video stream data transmission to reduce the pressure on the network. This topic also uses voice coding and decoding, noise reduction, echo cancellation, and other processing to solve the problems of noise interference and echo interference in voice transmission. This paper also introduces the creation, import, and loading of 3D virtual models, and explains how to use face key point association to drive 3D animation models, how to make the drive smoother and more natural, and using individual face key points as an example.

1. Introduction

Video communication in instant messaging systems usually requires high real time and stability; otherwise, it is prone to data delay, playback lag, and other instability. Due to the influence of an unstable network environment, the data are easily disturbed by various factors during transmission, resulting in the data not being broadcasted properly at the receiving end [1]. The goal of this topic is to design a virtual video chat system combined with virtual reality, which needs to be based on application scenarios, transforming from the original transmission of video data to the transmission of user's face key point data and handling transmission abnormalities [2]. At the same time, because this topic is not based on video streaming instant messaging, but 3D virtual animation video chat, the user sees the expression animation of the virtual animation model during the chat, so the data format transmitted in the network is the dataset of face key points and voice data, and this topic is based on the actual application scenario; there are high requirements for the noise reduction and echo cancellation of voice; the synchronization of voice and animated expressions and speech optimization become the urgent problems in this project [3, 4].

With the rapid development of computer network technology, the advantages of applying virtual reality technology in the field of instant communication have become greater and greater [5]. In a comprehensive and detailed exploration of the virtual reality technology applications that exist in computer communications, virtual reality technology plays an important role in communication networks, including enhancing resource utilization, device redundancy, immersion, interactivity, conceptualization, and holography [6]. Research of combining instant communication with virtual reality is still challenging due to the current problems of poor computing power, unstable data transmission, and large transmission delay on the mobile side. The ability to integrate communication and computing is used in the future 5G mobile communication network with multilevel

computing to solve the problem of the limited computing power of mobile terminals, efficient spectrum sensing, layered coding, and airport adaptive transmission to overcome the problem of unstable mobile channel transmission and delay guarantee mechanism to ensure the delay of mobile ARIVR services. However, research on the application of combining instant messaging with virtual reality technology is relatively routed and still in the exploration stage [7]. The system targets autistic patients and can measure patients' gaze-related index during their interaction with virtual companions, and this index can be mapped to their corresponding anxiety level. At the same time, the system can influence the patient's task performance and gaze-related index in response to the virtual companion's emotions [8]. Apple's release of Animoji enables the communication between users using 3D animated emesis. The implementation principle is to use facial recognition sensors to detect changes in user expressions and record user speech with a microphone to generate cartoonist 3D animated demos that can be shared by users with each other via iMessage [9]. The release of this feature has reaped a strong response from the community, stimulating interest and raising the interest of researchers in related fields. This topic will combine virtual reality and instant messaging, with the client driving 3D virtual models by parsing key point positioning data of faces to achieve virtual animated real-time communication [10].

In recent years, research in face key point localization has become increasingly abundant and mature, while research in deep learning has also made many breakthroughs, bringing better innovative approaches and more opportunities for other related research fields. The key point localization is the foundation of face recognition and other research, and the application scenarios are very broad [11]. As shown in Figure 1, the researchers have proposed many algorithms for face key point localization and achieved good results in related fields, but in practical applications, faces are often affected by various internal and external factors such as expression, posture, lighting, and occlusion, making it very difficult to achieve accurate face key point localization, which remains a great challenge [12]. This topic will address the design and optimization of the face key point positioning model and its application in mobile based on the actual application scenarios.

This topic is based on the XMPP to implement an instant messaging system combined with virtual reality. The system includes a client and a server. The client side realizes to build and maintain an instant communication system: based on the basic instant communication system, it locates the face key points in each frame of the video stream through the face key point positioning model, records, noise reduction, echo cancellation, encoding, and packaging the user's voice in real-time through the client microphone, and transmits the face key point positioning data and voice data to the friend user in real time. After receiving the face key point data, the 3D animation model is driven to enable the virtual model to make real-time accurate expressions, while decoding and playing the received voice data.

2. 3D Virtual Animation Key Point Model Design and Implementation

2.1. 3D Virtual Animation Is Raw Data Preprocessing. The design and implementation of the face key point localization model are developed using TensorFlow, and the model training is completed before exporting the file and porting it to the mobile application [13, 14]. The model design and implementation are divided into three parts: data preprocessing, model building, and model training. The main purpose of data preprocessing is to prepare data for model building, and the effect of modeling is greatly affected by good or bad data preparation; the model building is to research and analyze relevant algorithms, design a model that meets the subject matter, and make improvements in experimental analysis; model training is to learn the preprocessed data through relevant machine learning algorithms to get a model that can be used for predicting new data, which is a process of continuous iteration and exploration.

The data for this project is based on the 300 W challenge data, a well-known benchmark for evaluating key point detection algorithms. The challenge datum is a combination of image compilation from five datasets: LFPW, HELEN, AFW, IBUG, and the private 300 W test set. The last private 300 W dataset was originally used to evaluate eligibility for the competition and was then privately owned by the competition organizers, hence the name. Each image in the dataset has 68 markers with bounding boxes generated by the face detector. In this project, the 300 W challenge data are divided into a training set part and a test set part [15]. The training part includes the AFW dataset, the training subset of the LFPW dataset, and the training subset of the HELEN dataset, containing a total of 3148 images. The test section includes the remaining datasets: the IBUG dataset, the private 300 W test set, the best subset of the LFPW dataset, and the best subset of the HELEN dataset. In all datasets, each image corresponds to a text file with the coordinates of 68 key points of the face [16].

In the data preprocessing of this project, the dataset is firstly expanded: combined with the specific application scenario of this project, the dataset is flipped, panned, rotated, and scaled to improve the coupling between the dataset and the application scenario while increasing the number of the dataset. For each image in the dataset, we first flip the image left and right and then flip the data of 68 coordinate points corresponding to the image, and the dataset becomes twice the original one after the flip is completed. After the flip is completed, each image and its corresponding face key point coordinates are randomly panned, rotated, and scaled, and the final output is $112 * 112$. Since the original image is the input of the convolutional neural network in the model, the size needs to be equal, so here the image size is unified into $112 * 112$, and finally, each image is expanded into 20 images after the above operation. For the translation transformation, the translation is relatively simple, which is a simple coordinate addition and subtraction process:

$$\begin{bmatrix} x' & y' & 1 \end{bmatrix} = \begin{bmatrix} dx & dy & 1 \end{bmatrix}. \quad (1)$$

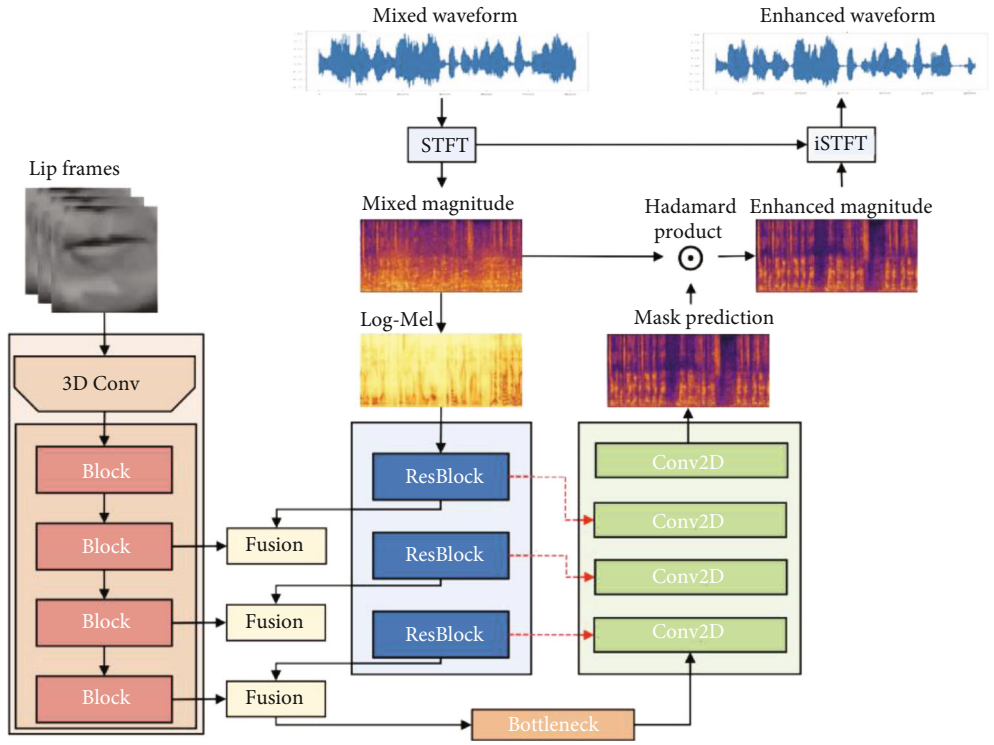


Figure 1: Framework diagram of combined virtual reality instant messaging system.

Calculation with the transformation matrix is as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \begin{bmatrix} dx \\ dy \\ 0 \end{bmatrix}^{\mathrm{T}} = x'dx + y'dy. \tag{2}$$

For the rotational transformation, the formula is calculated to obtain the following relation:

$$\begin{bmatrix} x' \\ 1 \\ y' \end{bmatrix} \begin{bmatrix} dx \\ 1 \\ dy \end{bmatrix} = \begin{bmatrix} \cos\theta & 1 & 0 \\ x & \tan\theta & 1 \\ \tan\theta & y & \sin\theta \end{bmatrix}. \tag{3}$$

For the scaling transformation, if the original coordinates are scaled directly, the center point of the image will be shifted.

$$\begin{bmatrix} x' \\ 0 \\ y' \end{bmatrix} = \begin{bmatrix} S_x & 1 & 0 \\ x & \tan\theta & dy \\ S_y & dx & \sin\theta \end{bmatrix} \begin{bmatrix} x \\ 1 \\ y \end{bmatrix}. \tag{4}$$

If you want to keep the center point unchanged, you can first translate the center point to the origin, then scale it, and then translate it back, i.e.,

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} S_x & 1 & 0 \\ x & 1 & dy \\ S_y & 0 & dy \end{bmatrix} \begin{bmatrix} x - x_0 \\ 1 \\ y - y_0 \end{bmatrix} \begin{bmatrix} x \\ 1 \\ y \end{bmatrix}. \tag{5}$$

2.2. 3D Virtual Animation Key Point Model Construction Design. The face key point localization model is based on a cascaded convolutional neural network, which mainly includes three parts (convolutional neural network, connection layer, and point coordinate similar inversion), and the original image is passed through two convolutional neural networks connected by the connection layer, and then, point coordinate similar inversion is performed to obtain the final 68 face key point coordinate results.

As shown in Figure 2, the convolutional neural network used in this model includes a total of four convolutional pooling operations and two fully connected layers; each convolutional pooling operation contains two convolutional layers and one pooling layer, the pooling layer uses the maximum pooling method, and the fully connected layer uses ReLU as the activation function. The key point coordinates initialized by the key point offset are summed up after the original image passes through the convolutional neural network to obtain the key point coordinates of the face in the first layer of the network.

The main purpose of the connection layer is to similarly transform the original image and its corresponding face key point coordinates, so that the image head pose changes to the standard position to reduce the impact of the different image head poses on the model; meanwhile, the connection layer generates the key point heat map according to the face

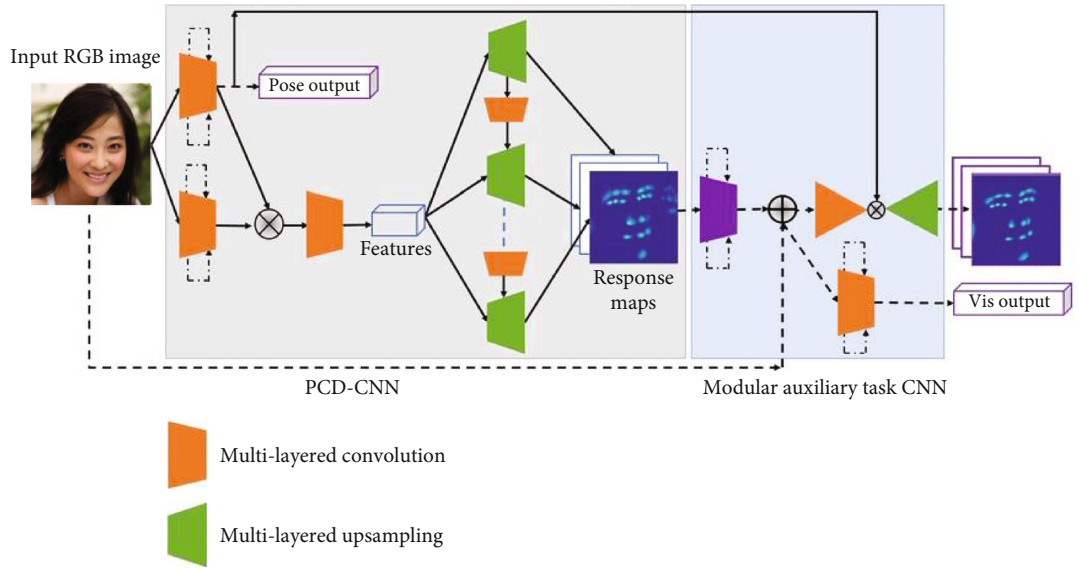


Figure 2: Block diagram of face key point localization model.

key point coordinates as the input of the second layer of the convolutional neural network; the key point heat map can better provide the face key point information and reduce the connection layer which mainly consists of four parts: similar transformation parameter calculation, image similar transformation, heat map generation, and feature map generation [17]. Among them, the similar transformation parameter calculation part calculates the parameters to similarly transform the image and key points to the standard position based on the initialized key point coordinates, the output value of the first layer, and the formula introduced in the data preprocessing and obtains the parameters of the similar transformation inverse transformation [18–20]. The heat map generation section firstly transforms the key point coordinates of the face by similar transformation and then generates the key point heat map according to the transformed key point coordinates as the second input of the second layer network. The feature map generation section takes the output of the first fully connected layer of the convolutional neural network as the input and passes through a fully connected layer to obtain the feature map as the third input of the second layer network, which is used to add feature information to the heat map.

The three outputs of the connection layer are used as the input of the second layer network, and after the second layer convolutional neural network, the key point offset of the second layer network is obtained, and the result of the similar transformation with the key point coordinate result of the first layer is added to obtain the face key point coordinate, which is the result of the similar transformation; therefore, it is necessary to use the inverse transformation parameters obtained from the connection layer to perform the similar inverse transformation of the point coordinates to obtain the result of the second layer network, which is the final face key point coordinate result.

For model training, the two layers of networks are trained separately, and after the training of the first layer is completed, the training of the second layer of networks is continued. The error rate is calculated as the Euclidean distance between the model prediction result and the real 68 key points divided by the distance between the two pupils of the face of that image [21]. The training results are shown in Figure 3. The error rate of the first layer network is higher and harder to converge because the first layer network only contains a convolutional neural network, and the separate convolutional neural network is difficult to achieve better results for the face key point localization problem, and the input of this model is the whole picture, which increases the difficulty of key point localization; therefore, the first layer network is less effective. The results of the second layer network training show that the second layer network converges faster and has a lower error rate due to the introduction of the face heat map, which again verifies the effectiveness of the cascade structure in the model and the advantage of the face heat map.

2.3. 3D Virtual Animation Is a Basic Instant Module Design. Functional requirements are the most basic conditions that a system can meet, but a system designed only based on functional requirements usually cannot be used directly by users, because in the process of use, users usually require the system to meet certain performance standards. The unit learning theme is the structured part of "association and structure". The goal of unit learning is to grasp the essence of knowledge. The unit learning activity is a part of activity and experience, transfer, and application. Therefore, the theme of unit learning is to move from "knowledge units" to "learning units", based on students' learning and development, and to organize the "learning" units in a big conceptual way. The logic of the subject reflects a rich, three-dimensional activity and openness. In the past, subjects were usually closed, but now we want to turn them into something open and unfinished, with unfinishedness and openness, providing space for students to explore and rediscover. We randomized the data according to an 8 : 2 distribution. Performance requirements, that is, nonfunctional requirements, refer to the design of the

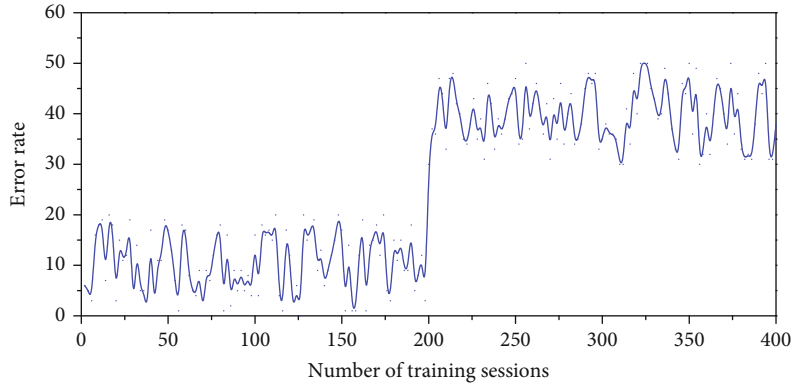


Figure 3: 3D virtual animation key point model training results.

system should meet some principles and functional minimum restrictions, performance indicators to be greater than these minimum requirements, and constraints to meet the needs of users. The good or bad performance of a system determines the user experience of this system, that is, the success or otherwise, so we should consider all aspects of the system's performance as much as possible during the preparation of the system. For performance requirements, the system considers a total of five major aspects, which are the simplicity of the system, the response time of the system, the compatibility of the system, the maintainability and expandability of the system, and the ease of use of the system.

For the basic instant messaging module, it mainly needs to implement a user system, buddy system, and communication system, specifically user registration, user login, buddy management, and real-time communication with the client. For user registration, the client fills in the user's name information and password information and transmits them to the server side for processing. The server side processes the registration request, generates a unique identifier on the server side upon success, and writes this user information to the database. When a user logs in, username and password information are transmitted to the server side for processing, and the server side processes the log-in request. If the user account or password is entered incorrectly, the prompt "Account or password error" is returned; otherwise, the user is redirected to the client communication homepage.

Friendly management mainly includes adding friends and friend lists [22]. After the user logs in successfully, clicking View Buddy List will request the buddy data from the server, and when it succeeds, all the buddies will be returned, and the client will display them in the form of a list. When you click Add Buddy, the requested data is saved in the database, and when the other party accepts the request, the server saves the buddy relationship and updates the data in the client's buddy list. In real-time communication, the client clicks on a buddy for real-time communication and sends a message, which is transmitted to the other buddy by the server in real time, and similarly, if a buddy message is received, the client receives it in real time and displays it in the conversation interface [23]. The client saves each chat log to the local database and loads the chat log every time it opens a conversation with a friend. If a friend sends a message to this client when this client is not online, the message data is saved in the server's offline message database and the offline message is loaded immediately when this client logs in. The design of the basic instant messaging module is shown in Figure 4.

With the sound module, it mainly needs to realize voice recording, voice coding and decoding, and sound playback. For voice recording, the client collects the voice data in real time through the microphone and decomposes the voice data into multiple voice packets of the same size according to the specific size. Since the actual application scenario of this topic requires simultaneous recording and playback of voice and the use of voice playback, it generates large noise interference and echoes interference, so the recorded voice needs to be coded noise reduction, echo cancellation, and other processing. In the voice coding and decoding, for the recorded voice, each recorded voice packet needs to be coded and rewritten, then saved as a file and transmitted; for the received voice, each received voice packet needs to be decoded and played back. When the voice is played, the voice packets need to be cached and processed, and the voice needs to be played in order, and at the same time, the synchronization process with the animation model driver needs to be performed when the voice is played.

3. 3D Virtual Animation Instant Network Communication System Design and Test Analysis

3.1. Instant Messaging System Design and Development. The main process of user registration is first to initialize XMPPStream, then connect to the server, and then send password authorization after connecting to the service successfully. In this project, the client's registration interface is designed with four input boxes for entering a username, nickname, password, and password confirmation, and two button controls for canceling registration and initiating a registration request. To protect the sensitive data of passwords, the password and password confirmation input boxes are controlled in the interface in a way that can be hidden.

The main process of user login is first to initialize XMPPStream, then connect to the server, and after a

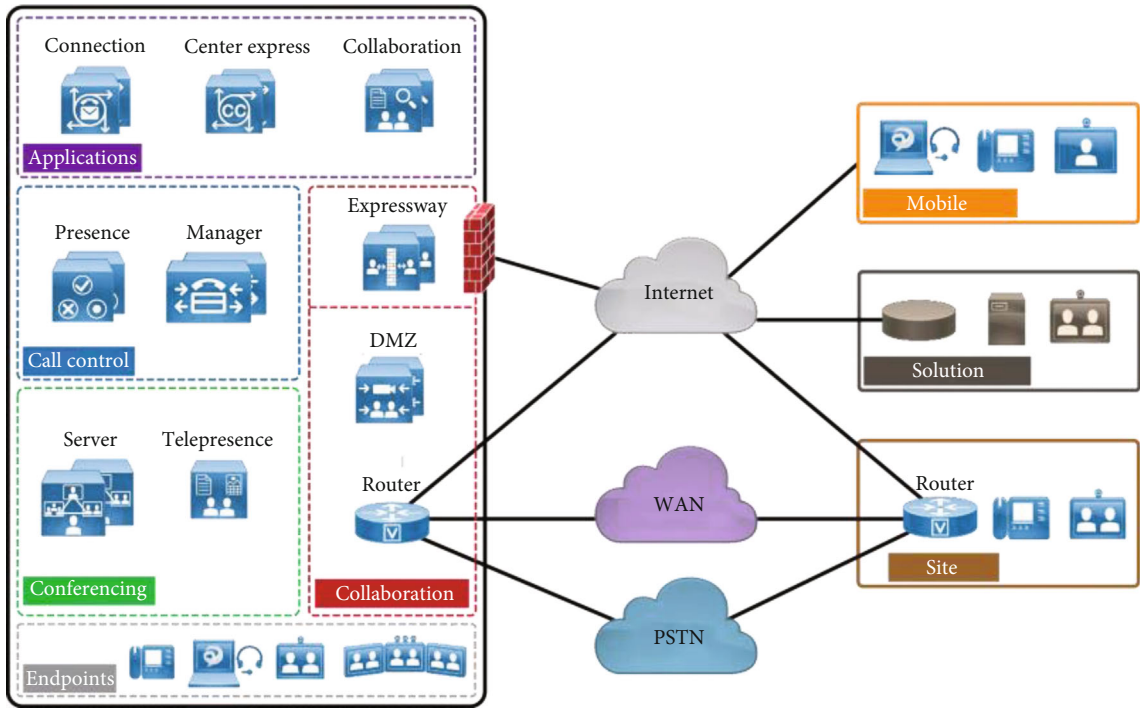


Figure 4: Basic instant messaging module flow chart.

successful connection to the service, send the password authorization, and after successful authorization, send the "online" message. After the connection is successfully established, the client will request offline message data and buddy data from the server, and then jump to the chat history interface and buddy list interface. In this project, the client's buddy management interface adopts the interface design like the buddy list, click the expand button in the upper right corner of the interface to select the add buddy menu, click the pop-up input box to add buddy, enter the user's name, and click the confirmation to send the request to the server to add buddy. The middle part of the friendly interface is designed with a fragment, which contains a list control that sorts the list of friends by their first alphabetical order, and implements a quick search by clicking on the letters on the right side. When you click on a friend in the buddy list, a pop-up window appears to indicate whether to initiate a chat, and when you click to confirm, you jump to the real-time communication interface with that friend, and when you click to cancel, the pop-up window disappears [24]. At the bottom of the friends, the screen is two buttons that can be clicked to switch between the friend's screen and the chat screen.

The main process of real-time communication is that client which sends messages to the server, which forwards messages to client B. At the same time, client B can also reply to messages to the server, which are forwarded to client A. When client B is not online, the messages sent by client A to client B are temporarily saved in the offline message database of the server, and then when the client blogs in, it will read the offline data from the server when the client blogs in. The chat data loading does not go to the server side to read the data in real time but synchronizes the offline data to the client local after each successful user login (each time synchronizing the data that is not available locally), and then, the interface layer of the client single goes to the local (SQLite) to read the data. As in Figure 5, in an instant messaging system with hand-free enabled, the microphone captures the echo of the other parts played by the amplifier, and the remote user receives it and hears his voice, creating an echo effect. The echo canceling can eliminate the echo before it is transmitted back to the other party, thus improving the quality of the sound received by the other user.

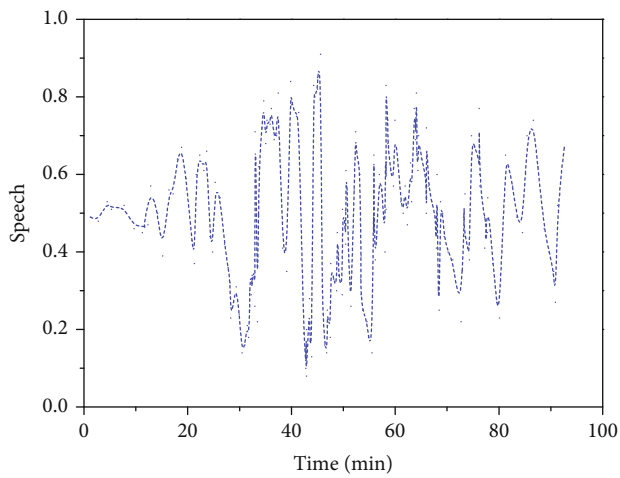


Figure 5: Effect of echo canceled on speech quality.

3.2. Design and Analysis of Facial Key Point Capture and Processing Functions. This project uses FaceTracker, an OpenCV-based facial tracking library. Firstly, FaceTracker needs to be integrated into the project and then tuned

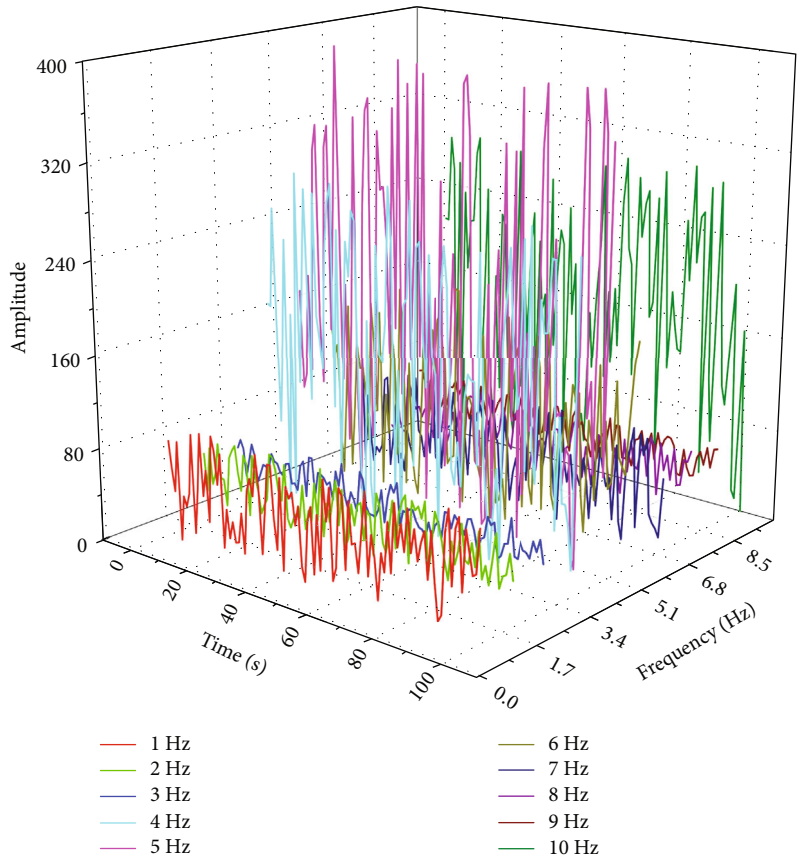


Figure 6: Face positioning timing diagram.

according to the actual situation. After the key facial recognition, a set of data is thrown every other frame, and the structure of the thrown data is shown in the key facial data structure table. After receiving the fatal key point data, it is forwarded to the key facial smoothing module for smoothing. The SG smoothing filter has a good smoothing and denoting effect on two-dimensional points, combined with the advantages of SG smoothing filter and the application scenario of the real-time nature of this topic; the five-point three-time smoothing algorithm of SG filter is used for smoothing, taking five nearby points, and determining a three-time curve, so that the square sum of the difference between the points on the curve and the original points' vertical coordinates is minimized. The coefficients of the cubic curve are determined, and by continuously adjusting the coefficients of the cubic curve, the curve is made smoother and closer to the actual situation.

In this project, different gaps are set for different facial key points to ensure that the movements of the key facial driving the model are effective, to achieve the role of dejittering [25]. The 5 key points show the threshold values corresponding to the facial key points, and the thresholds need to be constant of the adjustment by experiment, due to the large range of motion of the corner of the mouth, so the threshold is larger; taking the left corner of the mouth as an example, the threshold is 0.2. As shown in Figure 6, this paper compares whether the range of activity of a certain frame is greater than 0.2; the displacement is valid and enters the driving module; if not, this displacement is regarded as an invalid displacement. The movement coefficient is required to be adjusted continuously during the driving process of facial key points, and it plays the role of adjustment and optimization in the feature extraction and rereduction algorithm. After the blendShape is calculated by the feature extraction and rereduction algorithm, the driving effect is optimized by continuously adjusting the movement coefficient to drive the 3D virtual model accurately.

3.3. Design and Test Analysis of Driven 3D Virtual Animation Models. In this paper, we calculate the current degree of mouth opening of the user. Then, depending on the size of the user's open mouth, the maximum degree of mouth opening is multiplied by the movement parameter. This weight is assigned to blendShape, and we successfully control the opening of the character. Similarly, there are key points around the mouth, the left eyebrow in the corner of the mouth, the middle of the eyebrow, the opening and closing of the eyes, and the sliding calculation, each with a different calculation algorithm. This algorithm traverses all resources in the resource package. If the type of resource is a game object type, then the bone is instantiated into the scene. The blendShape is a masked mesh rendering property attached to the bone; so when the bone is loaded, it is also loaded. Then, iterating through all the bones and taking the blendShape name can be easily invoked and a face model is loaded into the scene.

Table 1: Virtual portrait key point expression reproduction degree.

Frame rate	Corner of the mouth	Corner of the eye	Eyes	Eyebrows
1	92.47	38.66	17.63	17.54
2	84.43	37.83	16.09	16.52
3	76.35	39.21	15.54	17.68
4	85.33	42.01	17.35	17.89

The encoding part gets data from the sound module and the expression recognition module. Then, key facial packets and voice packets are packaged into the basic data structure synchronously through the key facial queue and sound queue in the expression sound synchronization module. Since the communication of this topic is based on XMPP, which does not support the direct transmission of packet structures and needs to be converted into string format first, this part mainly implements this serialization function. The first thing to consider is to separate the fields into space and write them as strings. As shown in Table 1, the approach is feasible, but the compression rate is too low because the characters used are too low compared to all character sets, and if floating-point numbers are used to express key points, only the numeric part of the string can be used, and the alphabetic part is wasted and the effective utilization is too low.

To improve the compression ratio, we use the following method: First, the grouped sequences are converted into binary data, and the binary strings are converted into strings. This topic is implemented using the Base64 algorithm. Base64 algorithm is one of the most common encoding methods to convert arbitrary invisible binary data into visible text data. Usually, the algorithm is used to transmit long identification information, such as URLs. The input to the Base64 encoding algorithm is a stream of 8-bit bytes, i.e., byte arrays and strings. The advantage of the Base64 algorithm is that the encoding method is simple and easy to understand and the compression ratio is greatly improved compared to the direct conversion method of numeric strings. Its disadvantage is that the encoded strings are not readable, which can cause a certain degree of inconvenience during debugging.

In the process of open-mouth check and shut-mouth check, the user faces the camera and the facial recognition module locates the key to the face by clicking on OpenMouthCheck, the red dot in the upper left corner of the picture. After positioning, the user opens his mouth wide and checks it through the button. Through this check, it is equivalent to preprocessing the user's face. As shown in Figure 7, the parameters of each part are from the beginning of the facial key point identification and positioning, then to the smoothing module to denoise, smoothing, and then to the drive module of blendShape smoothing and feature extraction and then restore algorithm drive, etc.; the 3D virtual model can restore the user's expression more accurately.

3.4. 3D Virtual Animation Instant Network Communication System Performance Test. Functional testing refers to the testing of a product's features and operable behaviors to determine that they meet the design requirements based on

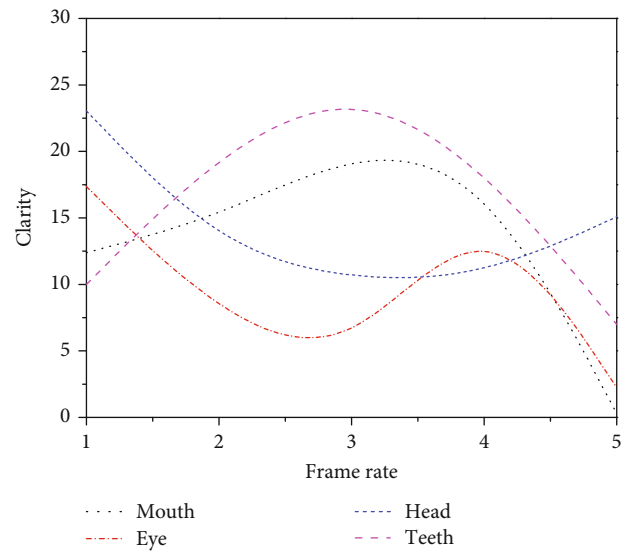


Figure 7: Comparison of facial features before and after virtual animation processing.

product features, operational descriptions, and user scenarios. The ultimate goal of functional testing is to ensure that the program operates in the desired manner, so the software should be tested according to functional requirements, by testing all features and functions of a system to ensure compliance with requirements and specifications. Functional testing can be divided into manual testing and automated testing, and this system uses a manual approach to test the main functions of the system. Since the system is divided into five modules (user management, friend management, instant messaging management, video management, and file storage management, and among them), the file storage management module can be tested by other modules.

Performance testing is the testing of performance indicators to improve the user experience when the functional requirements of the system have been met. The purpose of performance testing is to verify whether the software system can meet the performance indicators proposed by the user and to discover the performance bottlenecks in the software system, optimize the software, and finally play the purpose of optimizing the system. Performance testing mainly includes compatibility, security, stress testing, and load testing.

Security testing refers to the software system in the late stages of development to undergo rigorous security testing to meet the security needs of users using the process. Only products that pass the test can be put online and run. Security testing is to protect the system and users in case of malicious

Table 2: Priority queue for communication network packet processing.

ID	Sending time (s)	Acceptance time (s)	Play time (s)	Loss of condition	Hold time (s)
265	10.45	15.32	10.65	No	0
266	10.66	11.32	11.32	No	0
267	11.32	12.69	5.67	No	0.22
268	11.21	13.42	5.45	No	0.12

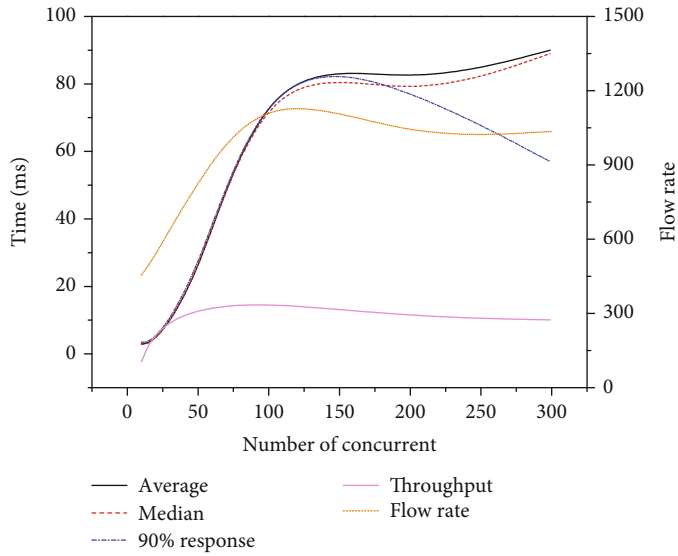


Figure 8: Stress test results of instant network communication system.

and wrong input. Network information security includes two layers of meaning: one is the security of external data exchange and the other is the security of internal LAN.

Web itself has good security monitoring measures; compared to traditional web real-time communication, users need to install plug-ins or applications in the browser; the process itself has potential risks, such as plug-ins and applications with their own Trojan horse virus, video transmission when the unscripted video stream is intercepted, or plug-ins and applications themselves which have the function of extracting user information, while the emergence of WebRTC shield exists. As shown in Table 2, in the video transmission process, the web uses secure communication protocols such as DTLS and SRTP to provide data security to both sides of the peer-to-peer video, which can prevent the leakage of video data on the web and video encryption at the sender and receiver side to encrypt video data; the key is negotiated by the video; both sides negotiate. The overall web components are encrypted, especially the codec, audio, and video transmission module; eliminating the video transmission process was intercepted.

As shown in Figure 8, the results of the stress test show that the average response time of the system can be kept within 100 ms, which meets the performance requirements, while when the number of concurrencies is high, a sudden increase in error rate can be found, so the current number of concurrency that the system can handle is around 200, while the throughput is low compared to large websites. Overall, the system can meet the current user requirements and can provide stable instant messaging and video communication.

The instant messaging management function includes seven main functions: text message, picture message, expression message, voice message, video clip, history message management, and text color change. The seven functions are highly independent, so test cases are designed for each of the seven functions to meet the testing requirements of the instant messaging management module. The instant messaging management module is the main module of the system, and the communication methods such as text, picture, expression, voice, and small video realized by the module greatly facilitate the users to communicate in real time, and each function is crossed and combined in the process of use, which can express the users' wishes completely and clearly and meet their needs. At the same time in the public chat room and private chat process, instant messaging functions are reused, reducing the workload of testing. Our method improves the accuracy by about 10.5% and the efficiency by about 14% compared to other studies.

4. Conclusion

This project combines the development status of cell phone operating systems and instant messaging software, based on the results of end-to-end instant messaging system requirements analysis, and combined with virtual reality technology,

basically realizes an instant messaging system based on XMPP for virtual video chatting, using 3D virtual animation models instead of real faces in the chat, and expression changes are reflected in the model animation, making the chat interaction more interesting and it also makes the transmission process lighter, reduces the burden on the server, and improves the fluency. In the model-driven, the subject first conducts in-depth research on algorithms related to face key point localization, compares several classical algorithms with better results, and implements a face key point localization model based on deep learning, and in the process of model implementation, through in-depth analysis of the actual application scenarios of the subject, proposes and implements improvements to the model, including optimization of data preprocessing and improvement of model algorithms, to improve the robustness of the model is improved. After the training of the model is completed, the migration of the model to run on mobile applications is realized, and each frame of the face video stream captured by the camera is processed in real-time in the subject project to output the relevant face key point localization data, thus driving the subject to further drive the 3D animation model. After acquiring face key point localization data, the system performs denoting and smoothing on the data, drives the 3D animation model by blendShape weight, and continuously tunes the parameters to make the model perform more smoothly and naturally. In this system, real-time voice calls are implemented for aiding virtual video chat. In voice processing, due to the specificity of the subject, the phone needs to open the external playback function for voice recording and playback, so the voice needs to be related to noise reduction, echo cancellation, and codec processing; the subject uses the third-party speed library to realize the voice noise reduction processing, echo cancellation, and voice codec. In the problem of voice delay, by constantly balancing the transmission time and voice recording time as well as the smoothness of playback, the voice delay is minimized. In the future, we will optimize and improve the model based on this.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Ministry of Education Industry University Cooperation Collaborative Education Project: Research on innovation and entrepreneurship education of digital art design major based on competition project construction (the second batch in 2020) and the Ministry of Education Industry University Cooperation Collaborative Education Project: Research on the construction of online advertising course in the new media environment (the second batch in 2020).

References

[1] M. G. Goodman, M. Alvarez, and S. B. Halstead, "Secondary infection as a risk factor for dengue hemorrhagic fever/dengue shock syndrome: an historical perspective and role of antibody-dependent enhancement of infection," *Archives of Virology*, vol. 158, no. 7, pp. 1445–1459, 2013.

[2] P. Jackson and M. T. Raiji, "Evaluation and management of intestinal obstruction," *American Family Physician*, vol. 83, no. 2, pp. 159–165, 2011.

[3] J. Park and Y. Cho, "Design and implementation of automated steganography image-detection system for the KakaoTalk instant messenger," *Computers*, vol. 9, no. 4, pp. 103–110, 2020.

[4] X. Liu, T. Zhang, N. Hu, P. Zhang, and Y. Zhang, "The method of Internet of Things access and network communication based on MQTT," *Computer Communications*, vol. 153, pp. 169–176, 2020.

[5] K. Morris, O. Sugiyama, G. Yamamoto et al., "Towards a medical oriented social network service: analysis of instant messaging communication among emergency physicians," *Advanced Biomedical Engineering*, vol. 9, pp. 35–42, 2020.

[6] F. H. Chen and S. Y. Yang, "A balance interface design and instant image-based traffic assistant agent based on GPS and linked open data technology," *Symmetry*, vol. 12, no. 1, pp. 1–10, 2020.

[7] T. Koonen, K. A. Mekonnen, F. Huijskens, N. Q. Pham, Z. Cao, and E. Tangdiongga, "Fully passive user localization for beam-steered high-capacity optical wireless communication system," *Journal of Lightwave Technology*, vol. 38, no. 10, pp. 2842–2848, 2020.

[8] F. Lamberti, G. Paravati, V. Gatteschi, A. Cannavo, and P. Montuschi, "Virtual character animation based on affordable motion capture and reconfigurable tangible interfaces," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 5, pp. 1742–1755, 2018.

[9] Y. Qiu, J. Xie, H. Lv et al., "FULL-KV: flexible and ultra-low-latency in-memory key-value store system design on CPU-FPGA," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 8, 2020.

[10] G. Liu, Y. Huang, N. Li et al., "Vision, requirements and network architecture of 6G mobile network beyond 2030," *China Communications*, vol. 17, no. 9, pp. 92–104, 2020.

[11] C. M. Chen, M. C. Li, and Y. L. Huang, "Developing an instant semantic analysis and feedback system to facilitate learning performance of online discussion," *Interactive Learning Environments*, vol. 3, no. 5, pp. 1–19, 2020.

[12] F. Li, X. Wang, Z. Wang et al., "A local communication system over Wi-Fi direct: implementation and performance evaluation," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5140–5158, 2020.

[13] H. Wang, Y. Wang, G. Zhuang, and J. Lu, "Asynchronous passive dynamic event-triggered controller design for singular Markov jump systems with general transition rates under stochastic cyber-attacks," *IET Control Theory & Applications*, vol. 14, no. 16, pp. 2291–2302, 2020.

[14] H. Song, J. Thiagarajan, P. Sattigeri, and A. Spanias, "Optimizing kernel machines using deep learning," *IEEE Transactions*

on Neural Networks and Learning Systems, vol. 29, no. 11, pp. 5528–5540, 2018.

[15] S. De, L. Bruzzone, A. Bhattacharya, F. Bovolo, and S. Chaudhuri, "A novel technique based on deep learning and a synthetic target database for classification of urban areas in PolSAR data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 1, pp. 154–170, 2017.

[16] L. Jiang, L. Yan, Y. Xia, Q. Guo, M. Fu, and K. Lu, "Asynchronous multirate multisensor data fusion over unreliable measurements with correlated noise," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 5, pp. 2427–2437, 2017.

[17] H. Wu, Z. Zhang, C. Jiao, C. Li, and T. Q. S. Quek, "Learn to sense: a meta-learning-based sensing and fusion framework for wireless sensor networks," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8215–8227, 2019.

[18] H. Zhang, X. Zhou, Z. Wang, H. Yan, and J. Sun, "Adaptive consensus-based distributed target tracking with dynamic cluster in sensor networks," *IEEE Transactions on Cybernetics*, vol. 49, no. 5, pp. 1580–1591, 2019.

[19] X. Yuan and Y. Pu, "Parallel lensless compressive imaging via deep convolutional neural networks," *Optics Express*, vol. 26, no. 2, pp. 1962–1977, 2018.

[20] D. Nada, M. Bousbia-Salah, and M. Bettayeb, "Multi-sensor data fusion for wheelchair position estimation with unscented Kalman filter," *International Journal of Automation and Computing*, vol. 15, no. 2, pp. 207–217, 2018.

[21] V. Rahu, C. Tong, S. Bhattacharya et al., "Multimodal deep learning for activity and context recognition," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 1–27, 2018.

[22] Q. Zhou and Y. Zheng, "Long link wireless sensor routing optimization based on improved adaptive ant colony algorithm," *International Journal of Wireless Information Networks*, vol. 27, no. 2, pp. 241–252, 2020.

[23] J. Hülsmann, J. Traub, and V. Markl, "Demand-based sensor data gathering with multi-query optimization," *Proceedings of the VLDB Endowment*, vol. 13, no. 12, pp. 2801–2804, 2020.

[24] P. Ghamisi, R. Gloaguen, P. M. Atkinson et al., "Multisource and multitemporal data fusion in remote sensing: a comprehensive review of the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 1, pp. 6–39, 2019.

[25] H. A. Pierson and M. S. Gashler, "Deep learning in robotics: a review of recent research," *Advanced Robotics*, vol. 31, no. 16, pp. 821–835, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9934585, 9 pages
https://doi.org/10.1155/2021/9934585

Research Article

Vehicle-Mounted Self-Organizing Network Routing Algorithm Based on Deep Reinforcement Learning

Shitong Ye,[1] **Lijuan Xu**,[1] **and Xiaomin Li**[2]

[1]*Department of Data Science, Guangzhou Huashang College, Guangzhou 511300, China*
[2]*Department of Mechanical and Electrical Engineering, Zhongkai University of Agriculture and Engineering, Guangzhou 511300, China*

Correspondence should be addressed to Shitong Ye; yst888_0@126.com

Received 6 March 2021; Revised 22 May 2021; Accepted 9 June 2021; Published 1 July 2021

Academic Editor: Wei Wang

Through the research on the vehicle-mounted self-organizing network, in view of the current routing technical problems of the vehicle-mounted self-organizing network under the condition of no roadside auxiliary communication unit cooperation, this paper proposes a vehicle network routing algorithm based on deep reinforcement learning. For the problems of massive vehicle nodes and multiple performance evaluation indexes in vehicular ad hoc network, this paper proposes a time prediction model of vehicle communication to reduce the probability of communication interruption and proposes the routing technology of vehicle network by studying the deep reinforcement learning method. This technology can quickly select routing nodes and plan the optimal route according to the required performance evaluation indicators.

1. Introduction

With people's continuous concern about the intelligent and safety of vehicles, the current artificial intelligence technology related to vehicle-driverless technology has emerged. Driverless technology can replace the human's own operation of the vehicle [1–3], so that the human can liberate their hands when using the vehicle for walking. Driverless technology makes the car react faster when encountering danger in the process of driving, to ensure the safety of passengers [4, 5]. Although the main research on driverless technology of major technology companies is to solve the problem of automatic control of vehicles, it does not mean that driverless technology only has the application field of intelligent vehicle control [6–9]. Vehicles with driverless technology can also be equipped with sensors and communication systems to cruise according to the set route in the military theater. While building the vehicular ad hoc network, due to the use of driverless technology, it can also prevent the enemy from causing casualties when attacking our vehicular system [10–12]. Using driverless technology to build a vehicle self-organizing network, it can also be deployed to extreme environments such as desert and disaster areas to perform tasks such as environmental data, animal and plant data collection, emergency rescue, and disaster relief information contact, etc. And through the vehicle network, the relevant data will be transmitted to the terminal, and the processing and decision-making can be carried out at the terminal, which can reduce the manpower input and avoid the safety problems of personnel in the execution of tasks. When unmanned vehicles are used to build communication networks in war zones or extreme environments with limited communication, technical problems are not only the automatic cruise technology of unmanned vehicles but also the deployment of vehicle networks and the realization of communication routes [13–15].

Vehicular ad hoc network, due to the fast-moving speed of vehicle nodes, the topology of the link will change dramatically in the process of network deployment, the network connection state of some vehicle nodes in the network will continue to change, the communication link will be interrupted, and it is unable to build an effective communication route. Therefore, routing technology is a major technical problem of vehicular ad hoc network. Moreover, vehicles carrying fixed fuel in theater or extreme environment will face

the problem of insufficient fuel when performing tasks and may not be able to continuously supply energy for on-board communication nodes. When the energy is insufficient, node death will cause topology change, communication link interruption, and data transmission interruption. Therefore, when performing data transmission tasks, it is necessary to consider the energy problem of the transmitting node. When data is transmitted through an effective routing node, according to the transmission rate of the node, the energy of the routing node can be maintained for a period of time until the data is successfully transmitted [16–18]. To solve the abovementioned vehicle self-organizing network routing technology problem, this paper proposes an adaptive routing technology in the vehicle self-organizing network based on the research of deep neural network technology, which can consider energy consumption and data transmission efficiency indicators, and automatically construct node data transmission route to complete the successful transmission of task data.

2. Related Work

In the research field of the routing technology of the vehicle ad hoc network, many researchers have done a lot of research work on the transmission efficiency and performance optimization of the multihop routing of the vehicle network. Zhang et al. proposed a link duration model based on duration, which can evaluate link reliability and use it as a key parameter for designing a new routing protocol. The new routing protocol can dynamically adjust the routing path through interaction with the surrounding environment [19]. Li et al. proposed an optimization model for the low-carbon vehicle routing problem under multigraph time-varying networks. The researchers started from the study of the multipath attributes of the real road network and designed a time division method that conforms to the time-varying network carbon emission calculations. The impact of driving speed changes and vehicle load on emissions, a low-carbon vehicle path optimization model under a multichannel time-varying network, is established [20]. Researchers such as Ahmed et al. proposed a highly secure QoS-aware routing algorithm that uses an optimal trust management scheme. Multihop clusters are implemented through an improved whale optimization algorithm, and the trust value is used to complete intercluster routing. This method can quickly discover routes and reduce the packet loss rate [21]. Silva proposes a new routing protocol called adaptive, which considers routing performance indicators such as transmission rate, average delay, and average number of hops. The protocol is based on a predictable connection concept and uses the history of meeting nodes to determine the best way to route and discard data packets on the network, and this method can effectively improve the transmission efficiency of routing data [22]. David and Vanathi proposed a vehicle-mounted self-organizing network clustering model that reduces data packet loss and uses clustering algorithms to cope with frequent topology changes and high mobility of vehicle networks, manage vehicles in an effective way, and provide intervehicle uninterrupted communication [23]. Meng et al. designed a control strategy for official vehicles in the traffic road network to improve the K-means algorithm to make the nodes of the official vehicle network adapt to the route, increase the weight of the backpressure strategy according to traffic pressure conditions, and improve the parameters through optimization the adaptability of official vehicles in the traffic road network [24]. Beirigo et al. proposed a dual-mode vehicle routing in hybrid autonomous and nonautonomous regional networks and introduced a new mathematical programming model in the routing to achieve coordinated routing planning for autonomous and conventional vehicles [25].

3. Time Prediction Model for Vehicular Ad Hoc Networks

For the vehicle-mounted ad hoc network without the cooperation of the roadside auxiliary communication unit, the biggest problem is that the vehicle communication link may be interrupted at any time due to the change of the vehicle speed, which greatly affects the communication quality of the vehicle ad hoc network. Therefore, in the research of vehicle network routing technology, we must first consider the problem of link interruption in vehicle self-organizing network [26, 27]. According to the sensing range of on-board sensors, the distance of end-to-end vehicle nodes, vehicle speed, vehicle speed variation range, task packet size, and network transmission rate, this paper proposes a time prediction model, which is used for the current vehicle. The node predicts a batch of candidate next hop nodes [28, 29]. The candidate nodes selected by the time prediction model meet certain conditions, that is, when the current vehicle node selects these candidate nodes as the relay point of the next hop, the probability of link interruption is small [30, 31]. Therefore, the time prediction model needs to comprehensively consider the transmission time of task packets, sensing range, vehicle spacing, vehicle speed, and possible speed mutation. As shown in Figure 1, the model will select candidate nodes that meet the time prediction function. It can be seen from the schematic diagram in Figure 1 that car A is the current node, and it is assumed that the range that car A's on-board sensor can sense is within the range of the black circle. Among them, car B, car C, and car D are the candidate nodes selected by car A. Although other nodes are also in the sensing range of car A, they do not meet the requirements of time prediction function, so they cannot be used as candidate nodes.

Assuming that the vehicle car A needs to select the next hop node to forward the data packet, the sensing radius of car A is R, and the initial interval between it and a certain vehicle (assuming car B) is $L_0(L_0 < R)$. The communication time for car A to transmit complete data is t. The time prediction model is used to judge whether car B can be used as a candidate node for the next hop: let all vehicles drive at a constant speed, acceleration, or deceleration, and the maximum acceleration is $\alpha_{\max}$. The conditions for car B to be a candidate node for the next hop are as follows: if car A runs at a constant speed with the minimum speed v_0, car B runs at a constant speed with the initial speed v_0 and the maximum

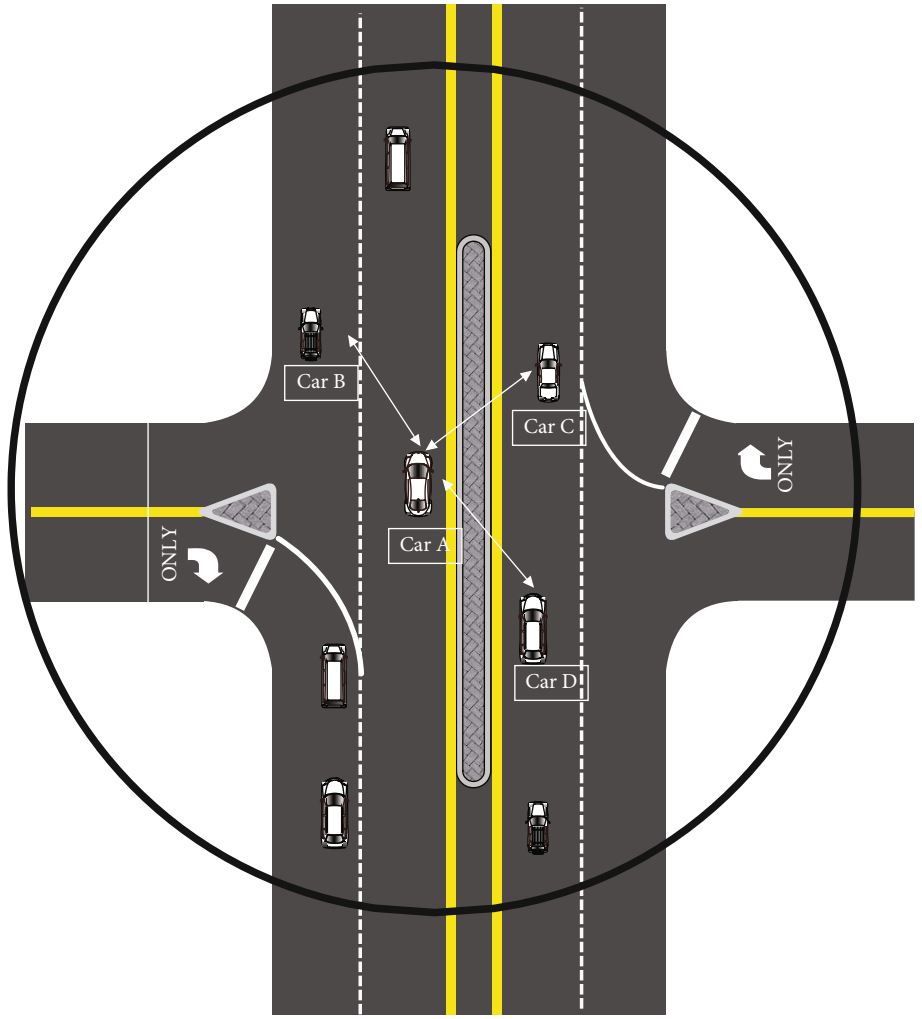


Figure 1: Schematic diagram of current node selection candidate node.

acceleration $\alpha_{\max}$, and after t time, car B is still in the sensing radius R of car A, then car B can be a candidate node for car A's next hop.

First, the relationship between driving distance and time of car A and car B is calculated:

$$\begin{cases} S_A(t) = \int_0^t v_0 dx, \\ S_B(t) = \int_0^t v_0 + a_{\max} dx. \end{cases} \tag{1}$$

After t seconds, the distance relationship between car A and car B satisfies the following relationship, then car B can be the next hop candidate node of car A:

$$S_B(t) - S_A(t) + L_0 \le R. \tag{2}$$

4. Utility Functions of Candidate Vehicle Routing Nodes

For the current vehicle node $x1$, in order to successfully transmit the data packet to $y1$ through the relay node, as shown in Figure 2, it is assumed that $x1$ calculates the candidate nodes of its next hop through the time prediction model, and these nodes from a candidate node set as N_A, $N_A = \{n_1, n_2, \cdots, n_a\}$. After selecting the next hop node $n_i (n_i \in N_A)$ of vehicle node $x1$ from the candidate node set N_A, n_a continues to calculate its next hop candidate node through the time prediction model, assuming that the candidate node set is M_B, $M_B = \{m_1, m_2, \cdots, m_b\}$. Select the next hop node $m_j (m_j \in M_B)$ from the set M_B until the data is finally transmitted to $y1$.

Selecting the best relay node from the candidate node set of each hop can constitute the multihop optimal routing of vehicle network. In this paper, deep reinforcement learning is used to select nodes from each candidate node set. In order to train in the deep reinforcement learning algorithm, it is necessary to determine the reward (or penalty) of the current system under the node selection behavior. In this section, we use the comprehensive utility function to evaluate the reward (or punishment) of node selection behavior.

The comprehensive utility function mainly considers the energy loss and transmission rate of data transmission between nodes. Suppose that in the vehicular network, the vehicular sensor nodes all adopt wireless communication

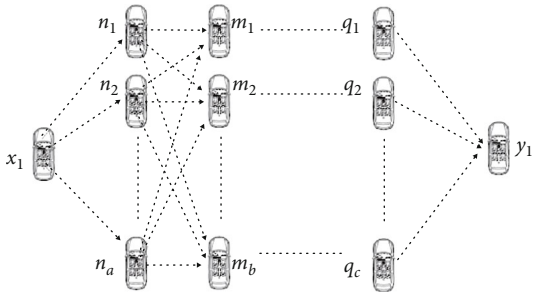


Figure 2: Schematic diagram of vehicle network packet transmission route.

mode, the distance between the current node and the next hop node is L, and the loss of link energy is represented by Loss(L). Considering the attenuation of transmission link, the Loss(L) is as follows:

$$\text{Loss}(L) = 40 \log L - 10(\log G_l + \log G_r) - 20(\log h_l + \log h_r), \tag{3}$$

where G_l is the antenna gain of the vehicle sensor, G_r is the antenna gain of the receiver, h_l is the antenna height of the transmitter, and h_l is the antenna height of the receiver.

Assuming that the transmitting power of the current node is h_l and the unit noise power is P_c, the direct data transmission rate between the current node and the next hop node is

$$v = \log_2\left(1 + \frac{P_T}{P_c n_0}\right), \tag{4}$$

where n_0 is interference noise.

The comprehensive utility function of the current node and the next hop node is expressed as:

$$\text{RE} = \frac{v}{P_T + \text{Loss}(L)}. \tag{5}$$

The higher the transmission rate and the lower the link loss between the current node and the next hop node, the larger the comprehensive utility function.

5. Vehicle Routing Based on Deep Reinforcement Learning

In vehicle network, the system state only considers the channel state of nodes. In Figure 2, the vehicular sensor node $x1$ transmits data to $y1$ through the relay node, and the node switches to the next state by selecting the relay node. We model the relay node selection scheduling problem as a Markov decision process.

In the Markov decision function, the next state S_{t+1} of the system is only related to the current state S_t

$$P_{ss'}^{a} = P(S_{t+1} = s' \mid S_t = s, A_t = a). \tag{6}$$

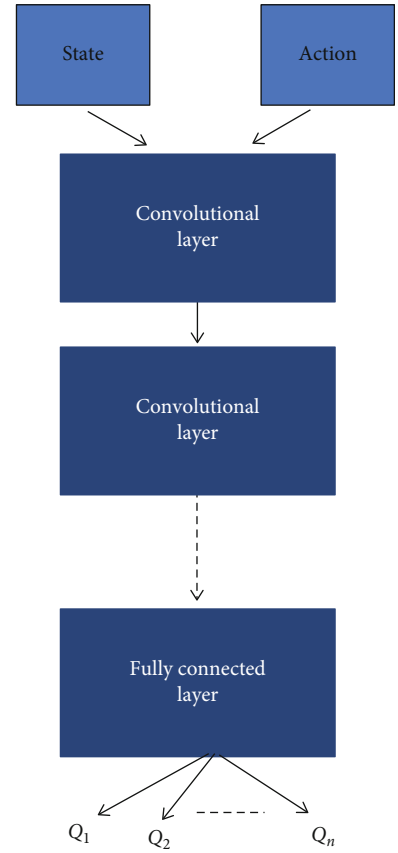


Figure 3: Obtaining Q value through deep network.

Let $A_{x1} = \{a_{t,1}, a_{t,2}, \cdots, a_{t,n}\}$ be the next hop candidate node set of the starting node $x1$ in time slot t, that is, A_{x1} is also the action set in time slot. $a_{t,i}(a_{t,i} \in A_{x1})$ means that the i-th node is selected as the relay node from the candidate node set of $x1$ in time slot t.

The current system performs node selection in state S_t. In order to successfully transmit data from the original node $x1$ to $y1$, assuming that k relay nodes are needed, the profit of the future k steps is obtained through the state value function:

$$V_\pi(s) = E_\pi\left[\sum_{t=0}^{k} \gamma^t \text{RE}_{t+1} | S_t = s\right]. \tag{7}$$

Among them, γ^t is the discount factor at step γ^t, and RE_{t+1} represents the comprehensive utility function value of the current node and the selected next hop node.

In order to evaluate the state and behavior of the system, the $Q_\pi(s, a)$ value is used to represent the action value function:

$$Q_\pi(s, a) = E_\pi\left[\sum_{t=0}^{k} \gamma^t Q_\pi(s_{t+1}, \acute{a}) | S_t = s, A_t = a\right]. \tag{8}$$

Combined with the state value function, the following action value function is used to evaluate the system's profitability:

Select multi-hop routing for vehicle network based on deep reinforcement learning
Suppose the source node is,$Node_{source}$ the target node is, $Node_{aim}$ there are N nodes between the source node and the target node, and is the node set.

Step 1: According to formulas (1.1) and (1.2), calculate all possible multi-hop routes from the source node $Node_{source}$ to the destination node $Node_{aim}$ to form a candidate multi-hop route set R_N.

Step 2:Calculate the comprehensive utility value of all candidate multi-hop routes in R_N according to formula (1.5).

Step 3:The comprehensive utility value is used as a reward, and according to formulas (1.6)-(1.12), deep reinforcement learning is used to adaptively select the best multi-hop route.

Algorithm 1: Algorithm implementation steps.

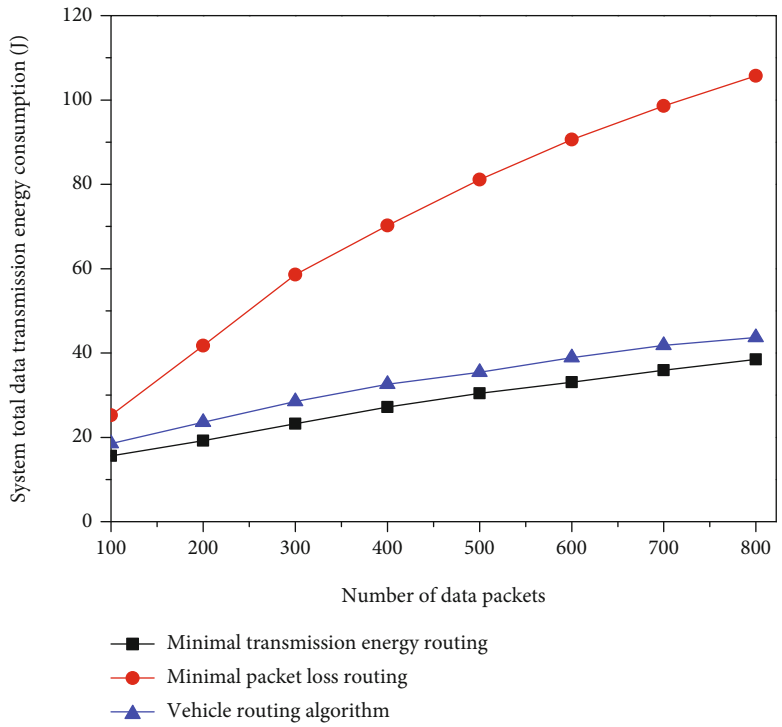


Figure 4: Total energy consumption of system transmission.

$$Q_\pi(s,a) = R_s^a + \gamma \sum_{s' \in S} P_{ss'^a} V_\pi(s'). \tag{9}$$

Among them, R_s^a represents the sum of rewards accumulated in all states after performing a set of actions.

In order to obtain better node selection behavior data in the network, iterative update is required, and iterative formulas are used to achieve optimized learning of action value functions:

$$\begin{cases} Q_\pi(s_t,a_t) = Q_\pi(s_t,a_t) + \lambda B_\pi, \\ B_\pi = R_s^a + \gamma Q_\pi(s_{t+1},a_{t+1}) - Q_\pi(s_t,a_t). \end{cases} \tag{10}$$

Among them, λ represents the learning rate, and γ represents the impact of future returns on current behavior.

Because the traditional Q-learning reinforcement learning method is based on the past state, statistics, and iterative Q value. Therefore, the state and action space applicable to Q -learning is very small, and if a state never appears, Q -learning cannot handle it. Therefore, here we use a deep reinforcement learning algorithm to replace the Q table with a neural network to obtain the Q value corresponding to the state and action. According to the state and node selection behavior, the Q value of each node selection action is output through the convolutional layer and the fully connected layer, as shown in Figure 3.

The optimization objective of deep reinforcement learning is to minimize the loss

$$\text{Loss} = E\left[\left(R_s^a + \gamma Q_\pi(s_{t+1},a_{t+1}) - Q_\pi(s_t,a_t)\right)^2\right]. \tag{11}$$

The gradient descent method is used to update the weight

$$\frac{\partial \text{Loss}(w)}{\partial w} = E\left[\left(R_s^a + \gamma Q_\pi(s_{t+1},a_{t+1}) - Q_\pi(s_t,a_t)\right)\frac{\partial Q_\pi(s_t,a_t,w)}{\partial w}\right]. \tag{12}$$

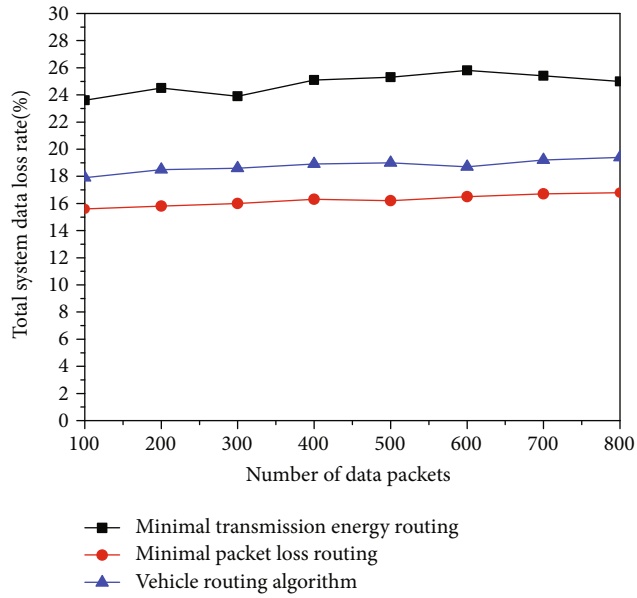


Figure 5: System transmission packet loss rate.

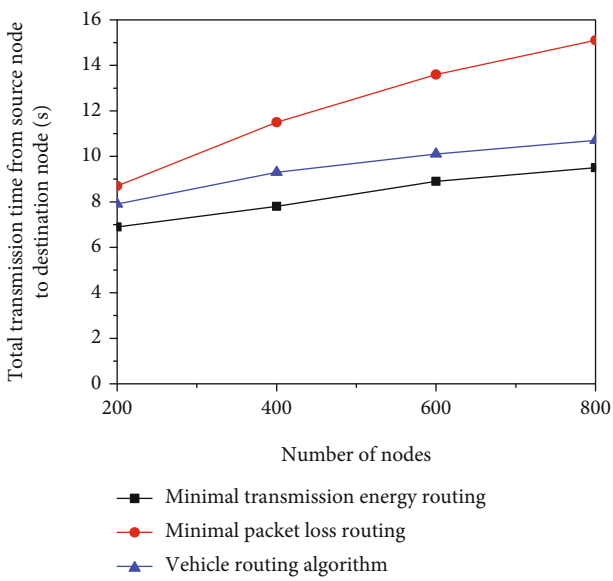


Figure 6: The total time for the source node to transmit data to the destination node.

The deep reinforcement learning method is used to update the Q value iteratively, so that the vehicle network can adaptively construct the optimal multihop routing from the source node to the destination node according to the comprehensive utility value.

The algorithm implementation process is shown in Algorithm 1:

6. Simulation Results

In the simulation experiment, there are 200 vehicular nodes in the simulated vehicular ad hoc network scenario. The nodes use wireless communication and data relay mode. Assuming that the computing power of the nodes is sufficient, the computing time is ignored, the noise power is set as $P_c = 1 \times 10^{-2} W$, the discount factor is $\gamma = 0.6$, and the training error of the deep network is lower than Loss = 1×10^{-4}. In the process of simulation, the simulated road is a straight passage, the width of the road is 16 meters, the width of the vehicle node is 2 meters, the set minimum speed of the vehicle is 20 km/h, and the maximum speed is 110 km/h.

In order to verify the performance of this algorithm, we set up two contrast algorithms in the experiment, one is to select the routing node according to the method of minimum transmission energy consumption, and the other is to select the routing node according to the method of minimum transmission packet loss rate. The two algorithms are tested under the same conditions as the algorithm in this paper.

In order to verify the performance of the proposed algorithm in terms of transmission energy consumption and packet loss rate, we conducted 100 simulation experiments and obtained the average results. In the simulation experiment of vehicular ad hoc network, we calculate the total transmission energy consumption and packet loss rate of the algorithm when the network system transmits 100~800 packets from random source node to random destination node.

According to the experimental statistical chart of the total energy consumption of the system transmission in Figure 4 and the experimental statistical chart of the packet loss rate of the system transmission in Figure 5, the more packets are transmitted, the greater the total energy consumption of the system, and the packet loss rate of the system has no obvious change trend with the increase of the packets. Among them, the minimum energy routing method can better save the system transmission energy consumption, but the system packet loss rate is higher, because the routing node selected by this method only considers the transmission link distance between the current node and the next hop node, so although it can better save the transmission energy consumption of the vehicle network, it will cause higher data packet loss.

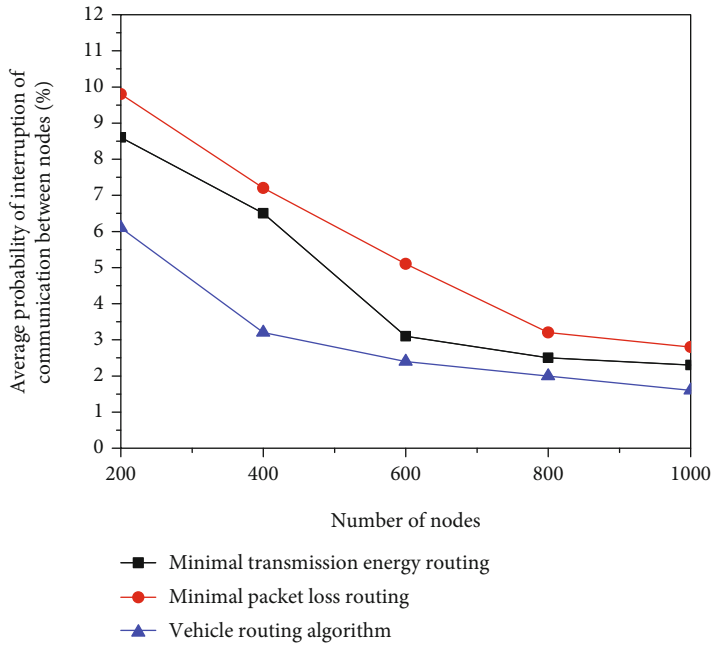


Figure 7: Average probability of communication interruption between nodes.

It can be seen from Figure 5 that the minimum packet loss routing can reduce the total packet loss rate of the system, but from the result of Figure 4, the total energy consumption of the minimum packet loss routing is more. However, the vehicle routing method based on deep reinforcement learning proposed in this paper selects the candidate next hop node according to the time prediction model, although there may be packet loss when the next hop node cannot be selected. However, from the experimental results in Figure 5, this method can maintain a low packet loss rate on the whole, and by using the comprehensive utility as the reward and punishment factor of deep reinforcement learning, the vehicle routing will further consider the transmission energy consumption when selecting nodes. From the results of Figure 5, this method can keep low transmission loss. Therefore, in general, the vehicle routing algorithm based on deep reinforcement learning proposed in this paper, although it cannot achieve the lowest transmission energy consumption and the lowest packet loss rate, is the best in the overall performance of the system transmission energy consumption and packet loss rate.

In order to verify the efficiency of the vehicle network routing algorithm proposed in this paper in data transmission, in the experiment, by increasing the number of experimental nodes, the time required for the vehicle network to transmit data from the random source node to the random destination node is counted. It can be seen from the experimental statistical results in Figure 6 that with the increase of the number of nodes, the time required to transmit data from the source node to the destination node increases gradually. Among them, the minimum energy routing method needs the least data transmission time. This is because the method adopts the shortest path data transmission method, which can quickly transmit the data to the destination node without considering the packet loss. The method in this paper uses the time prediction model, so it takes into account the risk of link interruption, and the transmission time is shorter than the minimum packet loss routing algorithm, which has obvious advantages.

In order to verify the performance of this algorithm in reducing the probability of communication interruption between nodes, in the experiment, we increase the total number of nodes tested and count the average probability of communication interruption between nodes in the vehicle network, as shown in Figure 7. From the figure, we can see that with the increase of vehicle nodes, the average outage probability of communication between nodes will continue to decrease, because the increase of vehicle nodes means that more nodes can act as relay nodes, reducing the possibility of link outage. It can be seen from the comparison in Figure 7 that the algorithm in this paper has a better effect on reducing the average probability of communication interruption between nodes. Because the method in this paper adopts the time prediction model, it can select the better next hop node to complete the data transmission task.

7. Conclusion

In order to reduce the probability of link interruption and improve the energy consumption and transmission efficiency of vehicular ad hoc network, a routing algorithm based on deep reinforcement learning is proposed in this paper. In this algorithm, a time prediction model of vehicular ad hoc network is proposed, which can effectively reduce the probability of communication interruption between vehicle nodes. The algorithm also uses the deep reinforcement learning method to select multihop routing, which can reduce the transmission loss of vehicle network routing and provide transmission efficiency.

Data Availability

The data used to support the results of this study needs to be obtained with the consent of the corresponding author.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

We declare that there is no conflict of interest.

Acknowledgments

This article is supported by Characteristic Innovation Project (Natural Science) in Colleges and Universities in Guangdong Province (No.: 2019KTSCX237) and Innovation Team Project of Humanities and Social Sciences in Colleges and Universities of Guangdong Province (No.: 2020WCXTD008).

References

[1] S. Liu, "Novel dynamic source routing protocol (DSR) based on genetic algorithm-bacterial foraging optimization (GA-BFO)," *International Journal of Communication Systems*, vol. 31, no. 18, article e3824, 2018.

[2] D. G. Zhang, G. Li, K. Zheng, X. Ming, and Z. H. Pan, "An energy-balanced routing method based on forward-aware factor for wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 766–773, 2014.

[3] D. G. Zhang, S. Liu, T. Zhang, and Z. Liang, "Novel unequal clustering routing protocol considering energy balancing based on network partition & distance for mobile education," *Journal of Network and Computer Applications*, vol. 88, no. 15, pp. 1–9, 2017.

[4] D. Zhang, X. Wang, X. Song, and D. Zhao, "A novel approach to mapped correlation of ID for RFID anti-collision," *IEEE Transactions on Services Computing*, vol. 7, no. 4, pp. 741–748, 2014.

[5] J. Yang, M. Ding, G. Mao, Z. Lin, D. G. Zhang, and T. H. Luan, "Optimal base station antenna downtilt in downlink cellular networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1779–1791, 2019.

[6] T. G. Dias, "Driverless cars - another piece of the puzzle: comments on "driverless cars will make passenger rail obsolete," by Yair Wiseman [Opinion]," *IEEE Technology and Society Magazine*, vol. 38, no. 2, pp. 36–38, 2019.

[7] E. Psyllou and J. Pawlak, "Congestion, safety, economic, and environmental challenges of vehicle automation in transport systems: comment on "driverless cars will make passenger rail obsolete," by Yair Wiseman [Opinion]," *IEEE Technology and Society Magazine*, vol. 38, no. 2, pp. 28–35, 2019.

[8] A. Zarkeshev and C. Csiszár, "Patients' willingness to ride on a driverless ambulance: a case study in Hungary," *Transportation Research Procedia*, vol. 44, pp. 8–14, 2020.

[9] L. Dong, D. Sun, G. Han, X. Li, Q. Hu, and L. Shu, "Velocity-free localization of autonomous driverless vehicles in underground intelligent mines," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9292–9303, 2020.

[10] A. Sharif, J. Guo, J. Ouyang et al., "Compact base station antenna based on image theory for UWB/5G RTLS embraced smart parking of driverless cars," *IEEE Access*, vol. 7, pp. 180898–180909, 2019.

[11] S. Liu, K. Hu, W. Ni, Z. Xu, F. Wang, and Z. Wan, "A cognitive relay network throughput optimization algorithm based on deep reinforcement learning," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 2731485, 8 pages, 2019.

[12] M. F. Baumann, C. Brandle, and S. Zimmer-Merkle, "Dis/orientations on driverless driving and autonomous vehicles?," *Applied Mobilities*, vol. 4, no. 2, pp. 251–255, 2019.

[13] R. Mumford, "Driverless cars technology receives 20 million boost in UK," *Microwave Journal*, vol. 59, no. 3, pp. 49-50, 2016.

[14] Z. Wang, Y. Guo, C. Fang, M. Li, Y. Sun, and Y. Zhang, "Optimal control strategy design with minimum energy consumption for connected vehicle systems," in *2019 IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, HI, USA, 2020.

[15] P. Si, Y. He, H. Yao, R. Yang, and Y. Zhang, "DaVe: offloading delay-tolerant data traffic to connected vehicle networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 3941–3953, 2016.

[16] A. Garcia-Santiago, J. Castaneda-Camacho, J. F. Guerrero-Castellanos, and G. Mino-Aguilar, "Evaluation of AODV and DSDV routing protocols for a FANET: further results towards robotic vehicle networks," in *2018 IEEE 9th Latin American symposium on Circuits & Systems (LASCAS)*, Puerto Vallarta, Mexico, 2018.

[17] T.-Y. Huang, C.-J. Chang, C.-W. Lin, S. Roy, and T.-Y. Ho, "Delay-bounded intravehicle network routing algorithm for minimization of wiring weight and wireless transmit power," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 36, no. 4, pp. 551–561, 2017.

[18] X. Chen, Z. Wei, H. Tian, and Z. Feng, "Space-time spectrum sharing for unmanned aerial vehicle networks," in *2018 IEEE/CIC International Conference on Communications in China (ICCC Workshops)*, Beijing, China, 2018.

[19] T. Zhang, T. Zhang, and X. Liu, "Novel self-adaptive routing service algorithm for application in VANET," *Applied Intelligence*, vol. 49, no. 5, pp. 1866–1879, 2019.

[20] L. Shunyong, D. Bin, and G. Xianlong, "Optimization model and algorithm of low carbon vehicle routing problem under multi-graph time-varying network," *Computer Integrated Manufacturing System*, vol. 25, no. 2, pp. 454–468, 2019.

[21] S. Ahmed, N. V. K. Ramesh, and B. N. K. Reddy, "A highly secured QoS aware routing algorithm for software defined vehicle ad-hoc networks using optimal trust management scheme," *Wireless Personal Communications*, vol. 113, no. 4, pp. 1807–1821, 2020.

[22] R. F. Silva, "Adaptive: an adaptive routing protocol for vehicle delay-tolerant networks," *IEEE Latin America Transactions*, vol. 18, no. 2, pp. 223–231, 2020.

[23] S. David and P. T. Vanathi, "Middle-order vehicle-based clustering model for reducing packet loss in vehicular ad-hoc networks," *Journal of Circuits, Systems and Computers*, vol. 29, no. 11, pp. 7–15, 2020.

[24] X. Meng, J. Lv, and S. Ma, "Applying improved K-means algorithm into official service vehicle networking environment and research," *Soft Computing*, vol. 24, no. 11, pp. 8355–8363, 2020.

[25] B. Beirigo, F. Schulte, and R. Negenborn, "Dual-mode vehicle routing in mixed autonomous and non-autonomous zone networks," in *2018 IEEE International Conference on Intelligent Transportation Systems (ITSC)*, Maui, HI, USA, 2018.

[26] S. Radley, C. J. Sybi, and K. Premkumar, "Multi information amount movement aware- routing in FANET: flying ad-hoc networks," *Mobile Networks and Applications*, vol. 25, no. 2, pp. 596–608, 2020.

[27] R. A. Nazib and S. Moh, "Routing protocols for unmanned aerial vehicle-aided vehicular ad hoc networks: a survey," *IEEE Access*, vol. 8, pp. 77535–77560, 2020.

[28] Z. Wan, Z. Pan, W. Ni et al., "A vehicle mobile internet of things coverage enhancement algorithm based on communication duration probability analysis," *IEEE Access*, vol. 7, pp. 98356–98365, 2019.

[29] A. Rawat, M. Khodari, M. Asplund, and A. Gurtov, "Decentralized firmware attestation for in-vehicle networks," *ACM Transactions on Cyber-Physical Systems*, vol. 5, no. 1, pp. 1–23, 2020.

[30] G. A. Klunder, H. Taale, L. Kester, and S. Hoogendoorn, "Improvement of network performance by in-vehicle routing using floating car data," *Journal of Advanced Transportation*, vol. 2017, Article ID 8483750, 16 pages, 2017.

[31] M. W. Levin, M. Duell, and S. T. Waller, "Effect of road grade on networkwide vehicle energy consumption and ecorouting," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2427, no. 1, pp. 26–33, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9861829, 1 page
https://doi.org/10.1155/2023/9861829

Retraction

Retracted: Combined GNSS/SLAM-Based High Accuracy Indoor Dynamic Target Localization Method

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] X. Zhang, S. Zhang, Q. Li, X. Wang, and S. Sun, "Combined GNSS/SLAM-Based High Accuracy Indoor Dynamic Target Localization Method," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 8380869, 12 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 8380869, 12 pages
https://doi.org/10.1155/2021/8380869

Research Article

Combined GNSS/SLAM-Based High Accuracy Indoor Dynamic Target Localization Method

Xiaona Zhang,[1,2] **Shufang Zhang**,[1] **Qiaosong Li**,[3] **Xin Wang**,[4] **and Shujing Sun**[5]

[1]*Information Science and Technology College, Dalian Maritime University, Dalian 116026, China*
[2]*School of Media, Tonghua Normal University, Tonghua, Jilin 134002, China*
[3]*State Grid Jilin Power Supply Company, Jilin 132000, China*
[4]*School of Business Administration, Wonkwang University, Iksan 570749, Republic of Korea*
[5]*Communication Engineering Construction and Maintenance of Wired Communication Department, Air Force Communication NCO Academy, Dalian 116026, China*

Correspondence should be addressed to Xiaona Zhang; zxn0201@dlmu.edu.cn

Received 29 April 2021; Accepted 15 June 2021; Published 28 June 2021

Academic Editor: Wei Wang

As the indoor dynamic target localization does not detect and repair the circumferential jump value in time, which leads to large position and attitude errors and low-velocity stability, a combined Global Navigation Satellite System/Simultaneous Localization and Mapping- (GNSS/SLAM-) based high-precision indoor dynamic target localization method is proposed. The method uses Empirical Mode Decomposition (EMD) to decompose the noisy signal, obtains the noise energy as the dominant mode from the decomposed components, extracts the useful signal energy as the main dividing point, removes the high-frequency signal, constructs the low-frequency component to realize low-pass filtering and denoising, selects a suitable threshold processing function to make the high-frequency signal component retain the detailed signal effectively to realize high-frequency component denoising, detects and fixes the circumferential jump of the observed data, and detects and fixes the circumferential jump of each frequency. The indoor dynamic target positioning method is established by combining GNSS/SLAM to achieve high accuracy target positioning. The experimental results show that the position and attitude errors are small, and the velocity is stable, which indicates that the position information is closer to the dynamic target, i.e., the target positioning is more accurate. To address the problems of scale drift and frequent initialization due to environmental factors in monocular vision SLAM, an Ultra Wideband (UWB)/vision fusion localization model that takes into account the scale factor is proposed, which can give full play to the complementary characteristics between UWB and vision; solve the problems of visual initialization, scale ambiguity, and absolute spatial reference; and improve UWB localization accuracy and localization frequency as well as reducing the number of base stations. The model can reliably achieve 0.2 m accuracy in indoor environments with sparse texture or frequent light changes.

1. Introduction

In recent years, with the growing market for public-facing location services, it has become extremely important to research an efficient, accurate quotient, and low consumption indoor mapping technology, so SLAM technology is gradually becoming a hot topic of research and development in the field of robotics. SLAM problem refers to placing a robot in an unknown environment, where the robot incrementally creates a continuous map of the unknown environment while determining its position in the map. It has been studied in the field of robotics for decades [1]. The cross-fertilization of mapping, navigation guidance and control, and remote sensing science has given rise to a combination of different sensors for positioning and mapping instead of a single navigation system. Multisensor combination navigation systems complement each other's strengths and advantages and can obtain higher navigation performance than any single navigation system. In outdoor environments, the most common technique is to use a combination of Global

Navigation Satellite System (GNSS) and Inertial Navigation System (IGS) for positioning, which can provide centimeter-level positioning accuracy when GNSS signal conditions are good. In general, as an important research direction in the field of navigation, seamless indoor-outdoor positioning technology has received extensive attention and research from governments and scholars [2]. At this stage, GNSS-based outdoor high-precision positioning technology is developing rapidly, and indoor positioning technologies such as UWB and INS have also made great progress, which can meet the positioning needs in most scenarios [3]. However, at this stage, it is still impossible to solve the problem of seamless indoor and outdoor positioning by a single sensor, and the method based on multisensor fusion has become the mainstream technical route [4]. Especially in the face of urban fires, earthquakes, and other emergency rescue, there are fires, smoke, noisy sound, power interruption, house collapse, road obstruction, rescue under the rubble, and other complex situations, intensifying the difficulty and challenge of rapid and high-precision positioning of indoor and outdoor rescue personnel and vehicles [5].

However, in indoor environments, GNSS signals are highly attenuated to the extent that they cannot meet the needs of indoor navigation and positioning. In indoor SLAM, the sensors used are generally divided into two categories: environment-aware sensors and motion-aware sensors. Environment-aware sensors generally include laser range finders, vision sensors, sonar, and optical flow sensors, which are mainly used to achieve localization based on the characteristics of the surrounding environment [6]. Motion-aware sensors include INS and odometer, which can achieve autonomous navigation without relying on the external environment. LiDAR is widely used in indoor SLAM for its advantages such as high localization accuracy and high positioning accuracy, unlike vision sensors which are affected by ambient light and other factors and are more stable, but because LiDAR matching is highly dependent on the features of the surrounding environment, it performs poorly in environments with low features, such as promenades. The advantage of GPS positioning is convenience and low cost because it is done through cell phones. Theoretically, the location of the cell phone can be determined by calculating the signal difference of three base stations, and the terminal held by the user only needs a cell phone. Therefore, if the user's cell phone has a signal, location positioning can be performed at any time without being affected by weather, tall buildings, location, etc. While inertial navigation can achieve autonomous navigation without relying on external environmental information, the sensor errors contained in it will gradually accumulate with the long working time, and the errors will fail after accumulating to a certain degree [7]. Therefore, the effective combination of the two sensors to complement each other's shortcomings and to create highly accurate indoor maps with high efficiency has become an important research direction for scholars at home and abroad.

To solve the existing problems, the GNSS/SLAM combined high-precision indoor dynamic target positioning method is proposed. By preprocessing the indoor dynamic observation data, detecting the weekly jump values of different frequencies, repairing both large and small weekly jumps, and obtaining high-quality data and accurate positioning, GNSS/SLAM combined navigation has the advantages of high positioning accuracy and stability. The combined GNSS/SLAM navigation has the advantages of high positioning accuracy and stability and has a broad application prospect for the development of indoor positioning and wireless communication technology. The original postprocessing algorithm is improved, and a near real-time improvement algorithm is proposed, which greatly reduces the output delay of the navigation results without reducing the mapping accuracy, and paves the way for transplanting the algorithm to the embedded platform, which can make real-time mapping of the drawn map in the computer while collecting information in the surrounding environment while walking on the mobile machine platform in real-time. As can be seen, sensor map-based positioning methods require preestablished maps, so in addition to the study of positioning technology, map-building technology is also an important research direction. The methods of map building can be generally divided into two categories: mapping and map building and simultaneous positioning and map building, in which the former process of map building requires other high-precision positioning systems to provide positioning information, while the latter estimates the location and the map of the surrounding environment simultaneously through the information provided by the sensors themselves. The mapping approach is more costly and difficult to use in GNSS signal occlusion areas, while the SLAM mapping approach is more flexible but less accurate in comparison.

2. Current Status of Research

The visual SLAM process consists of two parts: the front-end and the back-end. In the front-end part, the feature point method or the direct method is used to obtain data associations of visual information from between consecutive or interrupted frames [8]. In contrast, the direct method does not extract feature descriptors and directly uses pixel grayscale values to obtain image associations. Compared with the feature point method, the direct method can estimate denser maps and is more robust to scenes with insignificant textures, but the direct method is more sensitive to illumination changes and has difficulty handling closed-loop detection [9]. Among them, the feature point method is the more commonly used method, while the direct method has only been proposed in recent years [10]. In the back-end part, SLAM methods usually transform the whole problem into a maximizing a posteriori probability estimation (MAP) problem, using the data association provided by the front-end to estimate both the positional and surrounding environment maps [11]. The back-end estimation methods are generally divided into two types: filtering and optimization. The filtering approach is mainly based on the extended Kalman filter (EKF), while the optimization approach iteratively solves the problem by constructing a nonlinear least-squares problem using the Gauss-Newton method or the Levenberg-Marquardt method. The filtering method is more time-efficient, but its accuracy is poor, and it is difficult to deal with

closed-loop problems, while the optimization method is more accurate and can deal with closed-loop detection and optimization problems, with the disadvantage that it is less time-efficient and in rare cases may fall into the local minimum problem and affect the robustness of the estimation [12]. Due to the improvement of hardware computing power and research on matrix sparsity in recent years, optimization-based back-end methods can now run in real-time on common computing platforms, so optimization methods are now more frequently used in visual SLAM research.

Soatti et al. uses SURF as a feature descriptor to solve the closed-loop detection problem by training a bag-of-words model (BOM) of the feature descriptor, and this method is called FAB-MAP [13]. Where closed-loop detection specifically refers to determining whether the location of a frame appears in an established map, thus establishing a link between this frame and the map, and is used to optimize the map that improves the drift problem of the positional estimation. Later, Wen et al. used binary descriptors such as Brief or ORB to build bag-of-words models so that closed-loop detection can be done in a shorter time to meet the needs of real-time applications, and this method is called DBoW [14]. Hoang et al. proposed a visual SLAM method based on the direct method and the LSD-SLAM, which replaces the minimized reprojection error commonly used in the feature point method with minimized photometric error, proves that the direct method can also obtain the same performance as the feature point method and can build denser maps, and is a representative method of direct method visual SLAM [15]. Singh et al. introduce a tightly coupled EKF-based VIO method, the MSCKF, which uses a filtering approach to fuse visual and inertial guidance information, extending the back-end estimation to multiple frames of information in a sliding window instead of the then common optimization of two adjacent frames [16]. Based on this, MSCKF proposes a linearization point in the fixed Jacobi matrix computation to avoid the observable direction anomaly problem and introduces an online estimation method for the external parameters of the camera-IMU system [17].

Aiming at the problem of missing absolute localization reference for visual SLAM, visual odometry, and other visual indoor localization technologies, we carry out the research of a single-image spatial-visual localization model based on coded graphics assistance. Firstly, the algorithm model and technical process of the coded image, automatic detection and recognition of coded image, and automatic extraction of coded image and object coordinates are studied; secondly, the coded image assisted single image space rear rendezvous positioning model is constructed, and Tukey antidifference weighting factor is introduced to weaken or suppress the influence of the residuals of coded image coordinate observation on the positioning results and achieve high accuracy positioning.

3. GNSS/SLAM Combined High-Precision Indoor Dynamic Target Positioning Analysis

3.1. GNSS/SLAM Combination Design. GNSS dynamic positioning in urban environments is limited by the observation environment, with high signal interference noise, serious multipath effects, and frequent signal interruptions by occlusion, which seriously affects positioning accuracy and continuity [18]. Currently, in GNSS positioning, the GNSS carrier-phase double-difference positioning technique can be used to remove highly correlated error sources in space and time, such as ephemeris errors, to improve the positioning accuracy. Usually, the carrier-phase double-difference observation model can be described as:

$$\begin{cases} \nabla\Delta p = \nabla\Delta p_0 - \nabla\Delta T - \nabla\Delta I - \nabla\Delta M + \nabla\Delta\varepsilon_p, \\ \lambda\nabla\Delta\phi = \nabla\Delta\phi_0 - \nabla\Delta T - \nabla\Delta I - (1+\lambda)\nabla\Delta N + \nabla\Delta m + \nabla\Delta\varepsilon_\phi, \end{cases} \tag{1}$$

where the $\nabla\Delta$ symbol refers to double differentiation. p_0 and ϕ_0 denote pseudo-range and carrier phase observations, the geometric distance between the receiver and the satellite is p_0, T denotes tropospheric delay, I denotes ionospheric delay, λ denotes carrier phase wavelength, M denotes pseudo-range multipath error, m denotes carrier phase multipath error, and ε_p and ε_ϕ denote pseudo-range noise and carrier phase noise. For GNSS RTK with GPS and BDS data, the equation for estimating the unknown parameters can be written as:

$$\begin{bmatrix} H^G & A^G \\ H^C & A^C \end{bmatrix} \begin{bmatrix} \varepsilon_p \\ \nabla\Delta N^C \end{bmatrix} = \begin{bmatrix} H^G\varepsilon_p + A^G\nabla\Delta N^C \\ H^C\varepsilon_p + A^C\nabla\Delta N^C \end{bmatrix} = \begin{bmatrix} \nabla\Delta L^G \\ \nabla\Delta L^C \end{bmatrix}, \tag{2}$$

where the superscripts G and C denote GPS and BDS, respectively; H is a geometric matrix containing satellite geometry information; ε_p denotes the baseline increment vector; A is the ambiguity coefficient matrix, and $\nabla\Delta L$ is the observation minus the computation vector. Inertial devices mainly include three-axis accelerometers and three-axis gyroscopes. Limited by technical methods and process levels, inertial devices inevitably contain a variety of systematic and random errors, and at the same time, the influence of environmental factors makes these errors more complicated. Usually, inertial devices have been factory calibrated in a series of laboratory procedures; most of the system errors are compensated; the residual part needs to be estimated online through modeling; in addition, inertial devices also contain a variety of power-law spectrum of random noise, whether system errors or random errors will eventually be reflected by the accelerometer and gyroscope observations; this subsection focuses on the vehicle tactical grade. This subsection focuses on the analysis of SPAN-FSAS, an inertial guide, and discusses the effects of the selection of the systematic error term and the setting of the random error parameter on the combined GNSS/SINS results.

Before introducing the method of visual map building, it is necessary to introduce the form of the composition of the visual map used in this paper. Whether the map is built with high precision GNSS or SLAM, the form of the map is the same, which ensures the consistency of the positioning

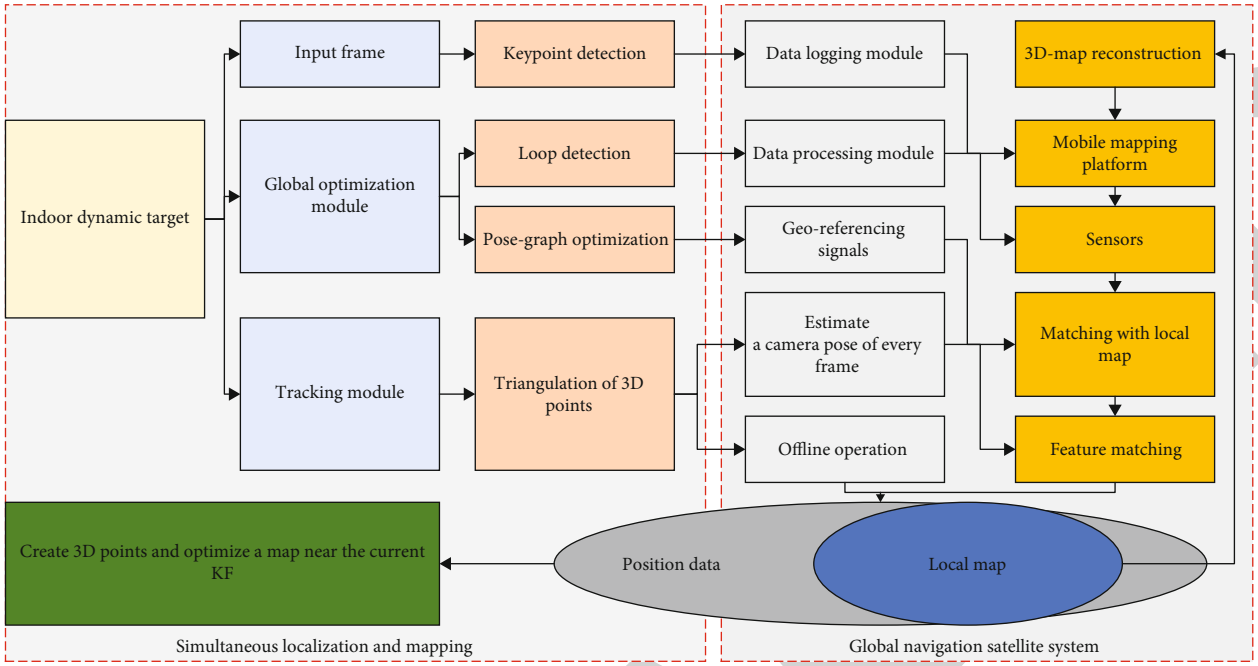


Figure 1: Combined GNSS/SLAM framework.

map. When positioning, there is no need to consider what the building method is, and the same information for positioning is obtained [19]. The map data in this paper is organized in keyframes, which are some of the frames selected with certain rules among all the image frames in the construction map. The main rule is that there is a certain distance interval between adjacent keyframes, to avoid redundancy of data. And each keyframe contains the information extracted from the sensor information, which is enough for positioning, and the complete map is composed of this keyframe, as shown in Figure 1.

It should be noted that the vision sensor in this paper is a monocular camera, and as analyzed in the previous paper, the visual feature map for localization needs the 3D position information of the features, so the focus of the high-precision GNSS-based visual map building method is the accurate estimation of the 3D position of the features. The simplest 3D position estimation method is to use the matching information of the features and the localization information for triangulation. Suppose that for the same feature point P, the normalized coordinates on the two frames obtained by feature detection and matching are x_1 and x_2, and the two frames obtained by the positioning information are $[R_1, p_1]$ and $[R_2, p_2]$. The depths of the high-precision pose GNSS are s_1 and s_2, which are also the depths of the feature point P in the two frames. There is the following constraint relation.

$$s_1 x_1 = \frac{s_2 R x_2 + t}{s_2 R x_2 - t}. \tag{3}$$

In determining whether the current frame is a keyframe and added to the map, the most important judgment criterion is the average parallax between the feature matching pairs of the current frame and the previous keyframe, i.e., the distance between the positions of the matching point pairs under the image coordinates. This is because when the parallax is relatively small, it can lead to a large error in the depth calculated by triangulation due to the camera's measurement unit of one pixel, which has a considerable impact on the matching localization. In addition, the matching accuracy of different feature point pairs may be different, and some of them are even mismatched, and the depths obtained by direct triangulation will become unreliable.

For these reasons, it is not reasonable to simply perform triangulation, so in this paper, after triangulation of the feature points, a sliding window optimization is added to estimate the position of each feature point when building a map based on high accuracy GNSS. In this optimization problem, the only state quantity to be estimated is the depth of all feature points within the window, as in equation (4).

$$X = [\lambda_0, \lambda_1, \cdots, \lambda_m]. \tag{4}$$

And the optimization problem takes the form of

$$\max \left\{ \lim_{l,k \longrightarrow \infty} \sum_{(l,k) \in C} e^{\left\| r_A\left(z_l^{c_k}, X\right) \right\|^2_{p_l^{c_k}}} \right\}. \tag{5}$$

Since the window includes multiple keyframes with a certain distance interval, the features within the window generally have enough parallax, and the estimation of feature depth is more reliable after optimization. After the optimization of the feature depth, the reprojection error of each feature is also recalculated, and if the error is larger than a present threshold, the feature can be considered as a wrong match, and the depth is invalid. When the keyframe with

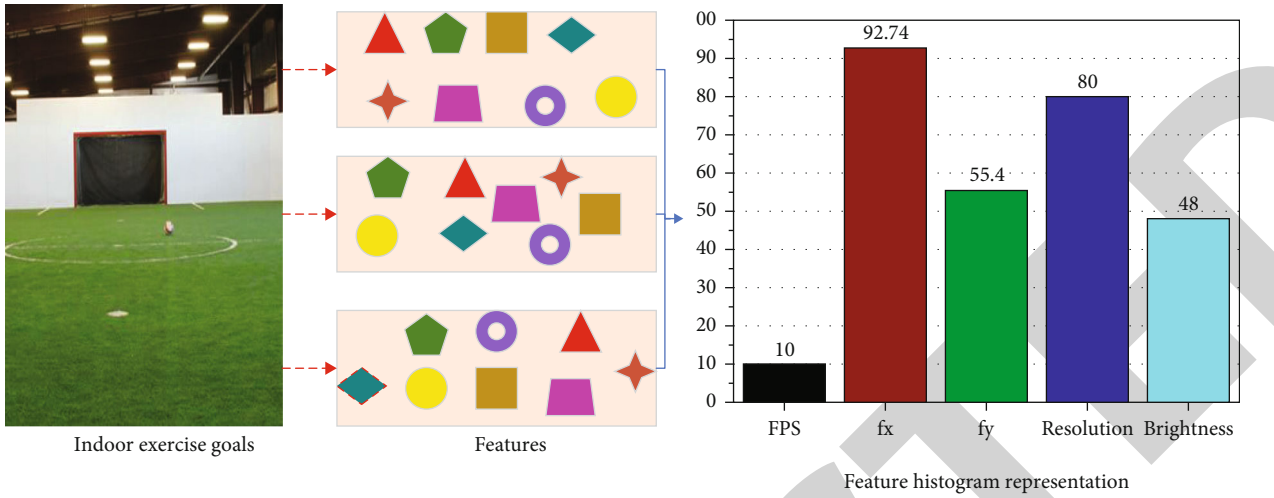


Figure 2: Schematic diagram of the bag-of-words model.

the earliest timestamp in the window is removed, that is, when the feature depth information on the image is no longer changed by the optimization, it is ready to be joined to the map. The keyframe information described in the previous subsection needs to be computed before joining the map; that is, the feature descriptor is computed based on the position of the feature point on the original image, and then, the bag-of-words vector for this keyframe is computed based on the feature descriptor. After completing the computation of this information, this keyframe is added to the map incrementally.

Closed-loop detection is a common concept in SLAM, which is essentially a method of information retrieval. The main purpose is to determine whether the current location has already appeared on the map, and when the information of the current location is found to be like some of the information in the map, it is considered that a closed-loop is found, and the map within the closed-loop can be optimized by establishing constraints on the matching relationship between the closed-loop frames. And in the localization process, the first step of global localization can be considered as the closed-loop detection process, which is to find similar locations in the map. Specifically, for the visual feature map in this paper, it is to find keyframes that are like the scene of the current image frame in the map composed of keyframes, and these keyframes will be used as candidate frames for feature matching with the current frame to get the global localization information. This is equivalent to the coarse localization process before the fine localization, which can greatly reduce the time complexity of global localization [20]. This is because the current frame does not need to feature match with every frame in the map but only needs to match with the map frames that are determined to be similar. The main solution for closed-loop detection in visual SLAM is to use the bag-of-words model. Bag-of-words (BOW) model is originally a document representation method in the field of information retrieval. In determining the similarity of two documents, the order of occurrence of words in a document is not considered, but only the frequency of occurrence of different words, i.e., two documents are considered similar when the frequency of occurrence of each word is similar. And the analogy of this concept to image features is to divide image features into a certain number of categories and count the frequency of occurrence of different categories in each image.

As shown in Figure 2, for an image, the histogram statistics of categories can be performed, and the images with similar histogram distribution have higher similarity, which is the information of bag-of-words vectors in keyframes mentioned before, and the distance between the bag-of-words vectors of different images can be considered as the similarity of images.

Our algorithm uses multiple fusion algorithms, which results in more accurate results and more efficient localization. And the classes of these features are called dictionaries, which are used to query which class the features belong to base on the descriptors of the features. The dictionary of feature descriptors in visual SLAM is obtained by clustering the features on many images, which has the advantage that it can cope with a wide variety of scenes and can perform closed-loop detection in unknown scenes. In the localization scenario of this paper, on the other hand, the map is well established, so the scene is known, and a dictionary specifically for this map can be generated by clustering all the feature points appearing on the map. Feature matching refers to determining the correspondence between feature points of different images and determining which points are the same in the 3D space. The similarity between feature points is generally measured by feature descriptors, which are usually expressed in terms of Euclidean distance for floating-point descriptors like SIFT and Hamming distance for binary descriptors like BRIEF. The Hamming distance refers to the number of values that differ between two binary numbers with the same number of bits, and the larger the Hamming distance, the smaller the similarity.

$$d(x,y) = \mathrm{Sup}\left\{\sum_{i=1}^{n} x[i] \otimes y[i]\right\}. \tag{6}$$

With the similarity measure, what needs to be done is to

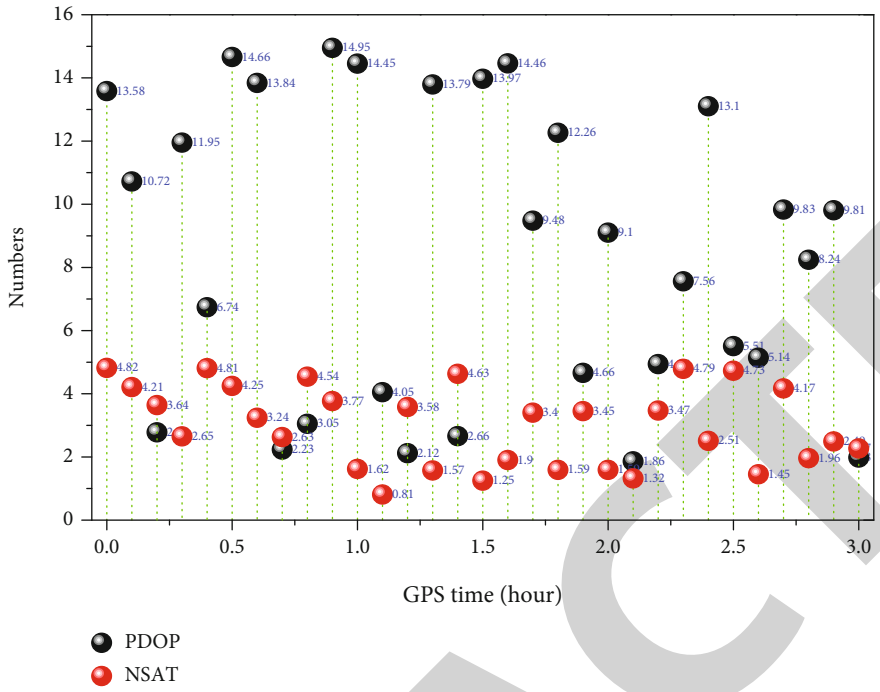


Figure 3: PDOP values and number of available satellites.

find which feature points are like each other. One of the simplest ideas is Brute-Force Matcher, which is the idea of calculating the distance between each feature point and all its possible matches and choosing the closest one as the matching point. This method is very simple, but it works quite well when the number of features is small. The number of features in a single keyframe in the map is generally small between 100 and 200, so using violent matching is a more efficient approach. Based on the sliding window state described above, the form of the optimization problem during fusion localization is somewhat altered from that of local positional estimation, specifically by adding reprojection error term for the features observed in the map frames.

$$\max\left\{\lim_{l,k\longrightarrow\infty} e^{\sum_{(l,k)\in C}\left\|r_A\left(z_l^{c_k},X\right)\right\|^2_{p_l^{c_k}}} + \lim_{l,\beta\longrightarrow\infty} e^{\sum_{(l,\beta)\in C}\left\|r_A\left(z_l^{c_\beta},X\right)\right\|^2_{p_l^{c_k}}}\right\},$$
$$X=[x_0,x_1,\cdots,x_m,\lambda_0,\lambda_1,\cdots,\lambda_m],$$
$$\mathrm{x}_k=\left[p^w_{b_k},\mathrm{v}^w_{b_k},\mathrm{q}^w_{b_k},b_k,b_g\right],$$
$$x^b_c=[p^c_b,\mathrm{q}^c_b]. \tag{7}$$

The objective of this optimization problem is to simultaneously optimize the poses of the image frames in the window and the relative poses derived from PnP in the previous subsection. The final optimized relative poses will be based on the poses of the map frames to obtain the poses of the current frame in the map coordinate system, i.e., the global positioning result. When no closed-loop detection occurs, i.e., when no map frame is observed, the whole problem turns back to the relative pose estimation problem and continues to be derived based on the previous global positioning result. This is the fusion localization method of tightly coupled optimization, which simultaneously performs the estimation of relative positional and global localization, and finally fuses the two, improving the disadvantages of poor robustness and low frequency of global localization information based on feature point matching, giving global localization results with higher accuracy and frequency.

3.2. Indoor Dynamic Target Localization Analysis. SLAM is a synchronous positioning and mapping technology, and GNSS is a superior navigation technology, which is commonly used in various fields [21–23]. In some scenarios, GNSS signals cannot be received due to the complex propagation environment and interference in the propagation channel, and its navigation function fails, so GNSS/SLAM combination can obtain high-precision environmental features and provide accurate positioning information. In the GNSS SLAM combination algorithm, for the attitude angle of the LIDAR coordinate system in the reference coordinate system, the position vector of the sensor and all the observed landmark points are the system state vector at k moments [24–26]. Location recognition is a key element of mobile robotics. In practical usage scenarios, we usually need to estimate the current robot's location information without any a priori information. This operation is also referred to as global localization. In this paper, the data in a 3D spatial coordinate system is first divided into two spheres: an outer sphere and an inner sphere. In this case, the radius of the outer sphere and the radius of the inner sphere can be measured. Based

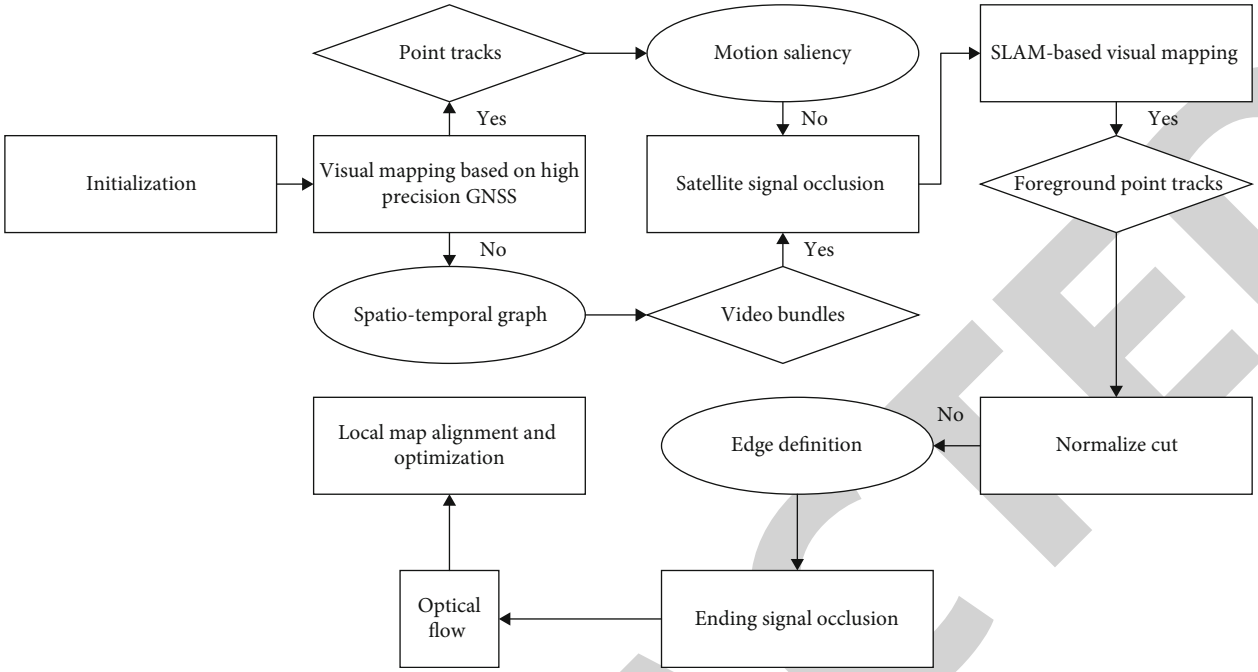


Figure 4: Flow chart of visual map creation.

on the principle of quadrant distribution of the 3D spatial coordinate system, it can be divided into 8 subspaces. Eventually, 16 subspaces can be obtained. The most essential difference between the tight and loose combination is whether SINS provides position a priori information for the GNSS solution. The position information is the link between SINS and GNSS, and the two sensors share the position to achieve the fusion of observation information. Based on this argument, we design a combination method with a semicompact structure.

SINS as the main filter distributes the position a priori information for all sensors, and each subfilter decides by itself whether to use the position a priori information through the judgment of the pretesters. This distributed architecture is essentially a semitight combination, where all sensor information is handshake confirmed and fused in the position domain, while each sensor's internal solution remains independent, leaving only the receiving and sending interface for position information. The advantage of the semicompact combination is that it can robustly fuse data from multiple sensors to obtain a better result; here, only GNSS and SINS sensor fusion is used as an example. The SPAN-FSAS device from Novatel was used to collect a single GPS vehicle data in a suburban area with a duration of about 2.8 hours. To analyze the effect of satellite signal interruption on the semitight combination, we simulated three 60 s segments of partial satellite information loss lock starting from 0.75 hours, and only the first 4 satellites were retained in the first segment according to the satellite serial number size, and only the first 3 satellites were retained in the second segment. According to the size of satellite serial number, only the first 4 satellites are retained in the first segment, only the first 3 satellites are retained in the second segment, and only the first 2 satellites are retained in the third segment, as shown in Figure 3 in the STC results labeled with the three segments of loss of lock.

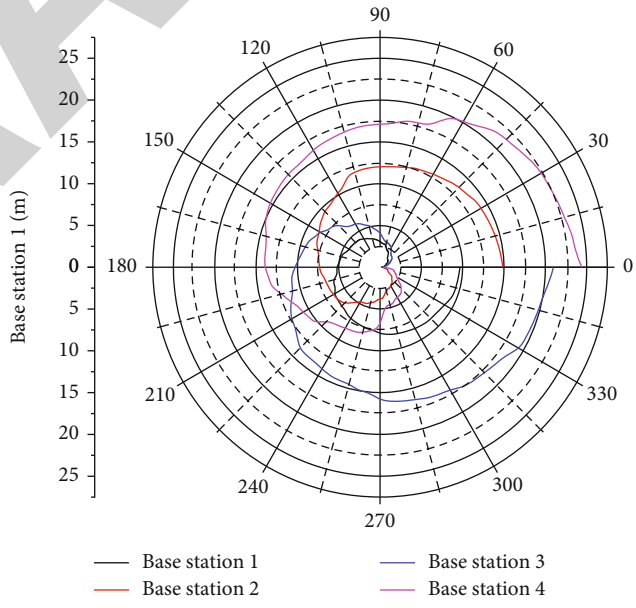


Figure 5: Distance measurement data.

From the error analysis of the positioning experiment, we can see that the lateral and longitudinal errors in most areas are within 20 cm, and the root mean square error of heading angle estimation is 0.420°, which can meet the needs of intelligent vehicle applications in most cases. The maximum lateral error of 84 cm was found in the open intersection area, where the road features on both sides are less constrained, so the lateral error was larger, while the maximum longitudinal error was found in the straight-line area due to the high repetition of the feature points on the straight-line road,

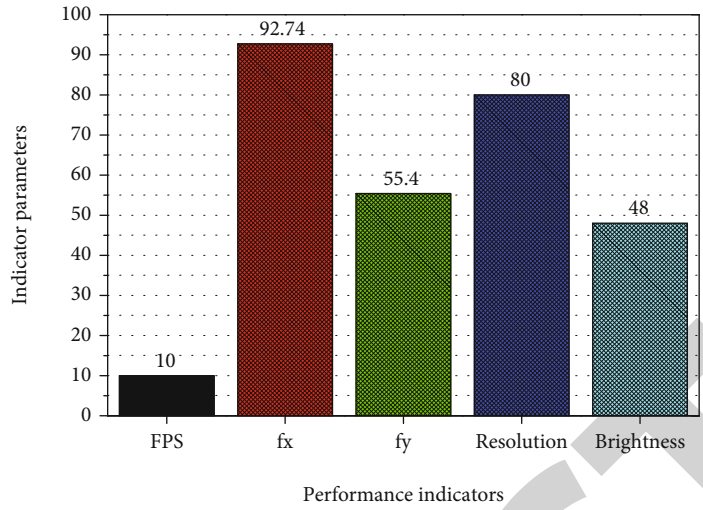


Figure 6: Performance parameters.

which may cause some error in the estimation of the longitudinal positioning. Overall, the localization results are reasonable considering that most of the maps used were created by the VI SLAM approach. When using the maps created by the hybrid approach for localization, it not only shows the performance of the positioning system but also is affected by the accuracy of the maps. Therefore, to further compare and verify the effectiveness of the localization method and the map building method, additional experiments were conducted to localize the maps entirely using high precision GNSS building, as shown in Figure 4.

Local map and map alignment and optimization and the detailed methods for these two parts are described below. It should be noted that due to the lack of robust and reliable criteria for judging whether the satellite signal is obscured, the judgment here is performed by hand, which is easy to implement and reasonable since the map building process is an offline process. After having a more accurate estimate of the positional pose and feature depth, all that must be done is the creation of the local map [27, 28]. It is easy to find that this part is the same as the building steps in the method based on high-precision GNSS mapping; that is, the keyframe information that can be used for localization is calculated from the positional pose and feature position, including the calculation of feature descriptors and the calculation of bag-of-words vectors, which will not be described repeatedly here.

In terms of system implementation, the map building method in this paper is an extension of local bit pose estimation, so the underlying programming platform is the same as the open-source tool, and the time synchronization of GNSS information with other sensor information is achieved through software synchronization, and the software timestamp of GNSS information is accurate enough in this system due to the small amount of data, low transmission delay, and high frequency, and the additional visual image closed-loop used in the experiment detection is implemented using the open-source library DBOW, and the map data is saved using the serialization function of the Boost library. Although the map building process is an offline process after data acquisition, the methods used in this paper are real-time, so all experiments are run in real-time on an ordinary laptop CPU.

4. Analysis of Results

4.1. System Performance Analysis. Figure 5 shows that the UWB ranging data in the rectangular line is partially obscured by the columns and appears discontinuous. Its corresponding Δr_i is shown in Figure 5. The bars of different colors indicate the ranging residual values of the corresponding base stations, respectively. As can be seen from Figure 5, the range values are relatively unstable at several peaks, which is because the range error of the UWB increases relatively as the distance increases. This is because the range error of UWB increases as the distance increases. The corresponding range residuals also increase in Figure 5. The weight of the range residuals is analyzed to dynamically determine the weight of the covariance of the observations at different base stations. From Figure 5, the average range residuals of all four base stations are less than 0.1 m, which indicates that the overall quality of the range values of path 1 is high. The smallest residual value of 0 indicates that the ranging value is obscured at a certain moment, and the residual calculation is not done at this time. Continuous wide coverage scenario is the most basic coverage method for mobile communication, aiming at ensuring users' mobility and service continuity and providing users with seamless high-speed service experience.

Another advantage of the mapping method in this paper is that it does not require route closure in the actual sense to optimize the map. Visual closure is done only once in the whole route, so the accuracy of mapping using visual closure is experimentally lower than that of GNSS closed-loop mapping. Theoretically, both visual and GNSS closed-loop provide the same information, and if the number of closed-loops is the same, the result should be similar. The purpose of this paper is to deal with the visual mapping of the

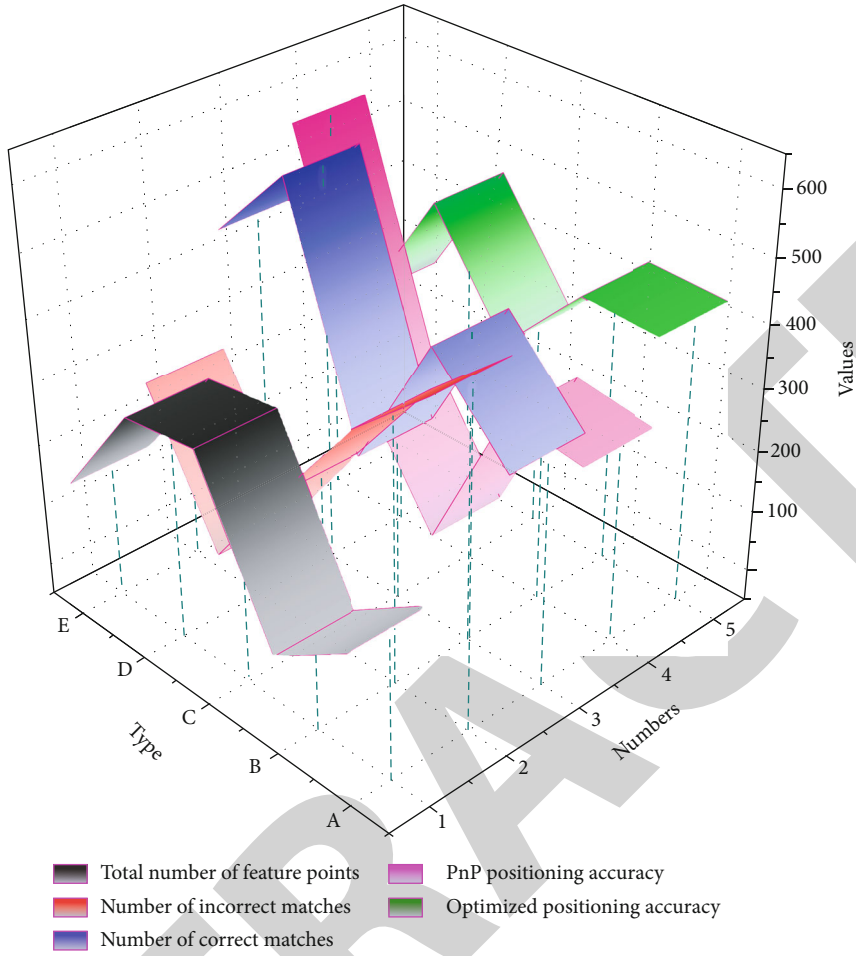


Figure 7: Matching cases and positioning accuracy in five cases.

occluded and unoccluded areas, and the use of visual closed-loop is quite limited. The comparison experiments here illustrate the flexibility of this paper and the effectiveness of the framework.

The experimental process around the figure of the conference table walking in a circle for this camera in the direction of different will automatically adjust the brightness, resulting in the video stream of light and dark changes; serious areas will be too black and too white, basically cannot see the textural information in the figure. Figure 6 is a sparse texture of the experimental environment; you can see that the corridor walls for the brochure, the ground, and the top plate are relatively single tone. In addition, during the experiment, the camera is facing the direction of the corridor, and the wall tilt intersection increasing the texture is not clear. The experimental equipment used is a calibrated camera, as well as a UWB base station and tagging equipment, as shown in Figure 6. The UWB equipment details are consistent with Section 4 of this paper. The camera is a CCD (charge-coupled device) imaging device, supporting up to 1080P resolution video stream acquisition, which can be directly inputted to the computer through USB, and the development of the corresponding interface software to achieve real-time extraction of video streams and keyframes.

During the experiment, to see a larger range of objects, the camera depth of precision is set to infinity. In addition, because the camera is an industrial camera, the camera will automatically adjust the brightness after sensing light. Since a fixed light-sensitive parameter cannot be set, it results in a process of obvious light, and dark changes in the acquired video stream when the light changes. As a streaming technology, the vision sensor not only can obtain rich texture information but also can transmit it in real-time. If it is applied to localization methods, it is ideal for indoor localization as it can not only obtain location information but also understand the surrounding real-world scenery. However, the current vision-based indoor localization methods fail in localization due to sparse textures, too bright or too dark light, and scale ambiguity in monocular ORB-SLAM. To address these problems, this paper proposes a combined UWB/vision-based EKF indoor localization model, which achieves indoor localization accuracy of 0.2 m and solves the problems of visual scale ambiguity and localization failure caused by texture sparsity and light brightness variation.

4.2. Dynamic Target Localization Results. The actual road conditions are highly variable, and only five types of feature matching under typical conditions are given above. It can

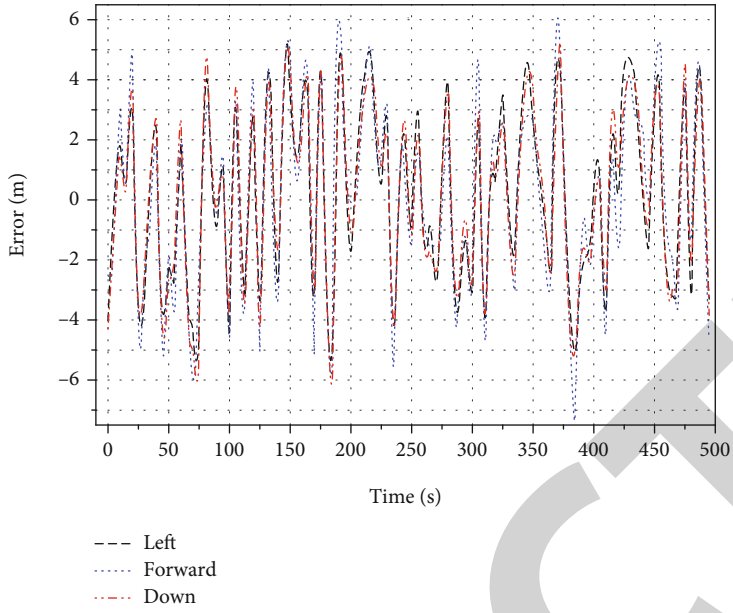


Figure 8: Position accuracy comparison.

be seen that the SIFT operator extracts a lot of feature points, which are distributed in the corner points on the figure and the areas with obvious changes, and the actual number of feature points that can be matched is relatively much less, but they can also be distributed in different areas of the figure so that the geometric configuration strength of the rear rendezvous can be ensured, and the geometric configuration strength through the RANSAC algorithm and antidifference optimization estimation excludes still more error matching points, mainly some road sign points with high feature similarity. The target detection and semantic segmentation are not used when processing the image in real-time here, mainly considering that the difference between vehicle and pedestrian feature descriptors and environmental feature descriptors is relatively large and does not easily cause mutual interference, while the vehicle and pedestrian have been removed when building the map, so the matched feature points in the scene do not fall on the vehicle and pedestrian, but basically on the road, house, and tree. It is better to match the correct point distribution to get a good bit pose solving result.

Figure 7 gives the matching situation and localization accuracy under five road conditions, and we can see that the feature matching rate is only about 10.37%, and most of the extracted feature points cannot find the corresponding road marker points in the visual point cloud map, although the percentage is relatively small, the number of its matching feature points is sufficient, and the ratio of final matching error to correct matching is about 1 : 4.5. Using localization accuracy can also be around cm, and C and D are still slightly better than the optimized localization accuracy, but the optimized results are more stable and have higher accuracy, as shown in Figure 8. It should be noted that the vertical coordinates of the positional results and the positional results of antidifference optimization in the figure are different between the two. In addition, the left and front components of the position are the planes directly under the system, and the lower component is the elevation direction, where the positional positioned under the camera coordinate system, so the heading angle is the roll angle under the ENU system, and the roll angle is the heading angle, and the pitch angle remains the same.

The above conclusion is only for data, and the same set of data is used for map building and positioning. In practice, map building and localization are often separated; for example, if the map built last month is used for localization at this moment, the point cloud map may not be able to accurately represent the actual road information at this moment due to the change of environmental factors. Whether the feature descriptors can still be matched needs to be studied in one step. With the promotion of crowdsourcing technology in the future, we can potentially achieve incremental real-time updates of the visual point cloud map and add more robust multidimensional feature information and semantic information, which can improve the success rate of localization in real scenarios. However, the average time required is 158 ms and 146 ms, and the maximum time required is 693 ms and 541 ms, accounting for the major part of the time consumed. On this basis, the position loosening combination with SINS can achieve continuous high sampling rate localization. Our positioning results are about 20% more efficient and a bit more accurate compared to others.

5. Conclusion

Motion target detection and localization in the dynamic background, the current method cannot be repaired with high precision for different frequency regions in circumferential jump detection and repair, and it is difficult to obtain high-quality data, resulting in large position and attitude errors and unstable speed. That is, the accuracy of indoor dynamic target positioning is low. In this problem, a

GNSS/SLAM combination-based high-precision indoor dynamic target positioning method is proposed to preprocess the indoor dynamic observation data, detect the circumferential jump occurring on each frequency, repair it with high precision, obtain high-quality data, and accurately position the indoor dynamic target, which effectively solves the problems existing in the current method. This method effectively solves the problems in current methods and provides the basis for the realization of high precision indoor dynamic target positioning. In the position error test, the error of method 1 is closer to 0, which proves that the positioning of this method is closer to the dynamic target, and the position is more accurate. The position error of method 2 and method 3 is larger due to the large fluctuation of latitude error, longitude error, and altitude error, which means that the positioning accuracy is lower. Because method 1 detects and fixes the observed data for circumferential hops, regardless of whether it is for large or small circumferential hops, it can detect the circumferential hops occurring at each frequency and can fix them with certain accuracy to obtain high-quality data, which makes the position error closer to 0, i.e., it accurately achieves indoor target positioning.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Foundation No. 61231006, Foundation No. 62071080, and Foundation No. 61501078).

References

[1] Z. Gong, R. Ying, F. Wen, J. Qian, and P. Liu, "Tightly coupled integration of GNSS and vision SLAM using 10-DoF optimization on manifold," *IEEE Sensors Journal*, vol. 19, no. 24, pp. 12105–12117, 2019.

[2] K. W. Chiang, G. J. Tsai, H. J. Chu, and N. el-Sheimy, "Performance enhancement of INS/GNSS/refreshed-SLAM integration for acceptable lane-level navigation accuracy," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 2463–2476, 2020.

[3] T. Shamseldin, A. Manerikar, M. Elbahnasawy, and A. Habib, "SLAM-based pseudo-GNSS/INS localization system for indoor LiDAR mobile mapping systems," in *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 197–208, Monterey, CA, 2018.

[4] R. Jurevičius, V. Marcinkevičius, and J. Šeibokas, "Robust GNSS-denied localization for UAV using particle filter and visual odometry," *Machine Vision and Applications*, vol. 30, no. 7-8, pp. 1181–1190, 2019.

[5] W. Sun, J. Wang, F. Jin, and Y. Yang, "A quality improvement method for 3D laser slam point clouds based on geometric primitives of the scan scene," *International Journal of Remote Sensing*, vol. 42, no. 1, pp. 378–388, 2021.

[6] G. Zhang and L. T. Hsu, "A new path planning algorithm using a GNSS localization error map for UAVs in an urban area," *Journal of Intelligent & Robotic Systems*, vol. 94, no. 1, pp. 219–235, 2019.

[7] G. Sammartano and A. Spanò, "Point clouds by SLAM-based mobile mapping systems: accuracy and geometric content validation in multisensor survey and stand-alone acquisition," *Applied Geomatics*, vol. 10, no. 4, pp. 317–339, 2018.

[8] D. Perea-Strom, A. Morell, J. Toledo, and L. Acosta, "GNSS integration in the localization system of an autonomous vehicle based on particle weighting," *IEEE Sensors Journal*, vol. 20, no. 6, pp. 3314–3323, 2019.

[9] W. Wen, G. Zhang, and L. T. Hsu, "Object-detection-aided GNSS and its integration with Lidar in highly urbanized areas," *IEEE Intelligent Transportation Systems Magazine*, vol. 12, no. 3, pp. 53–69, 2020.

[10] T. Suzuki, "Time-relative RTK-GNSS: GNSS loop closure in pose graph optimization," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4735–4742, 2020.

[11] T. M. Quack, M. Reiter, and D. Abel, "Digital map generation and localization for vehicles in urban intersections using LiDAR and GNSS data," *IFAC-Papers OnLine*, vol. 50, no. 1, pp. 251–257, 2017.

[12] Y. Li, Z. He, Z. Gao, Y. Zhuang, C. Shi, and N. el-Sheimy, "Toward robust crowdsourcing-based localization: a fingerprinting accuracy indicator enhanced wireless/magnetic/inertial integration approach," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3585–3600, 2019.

[13] G. Soatti, M. Nicoli, N. Garcia, B. Denis, R. Raulefs, and H. Wymeersch, "Implicit cooperative positioning in vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 12, pp. 3964–3980, 2018.

[14] C. Wen, Y. Dai, Y. Xia et al., "Toward efficient 3-D colored mapping in GPS-/GNSS-denied environments," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 1, pp. 147–151, 2020.

[15] G. M. Hoang, B. Denis, J. Härri, and D. Slock, "Fusion bayesienne de donnees GNSS, ITS-G5 et IR-UWB pour des applications robustes de localisation vehiculaire cooperative," *Comptes Rendus Physique*, vol. 20, no. 3, pp. 218–227, 2019.

[16] S. K. Singh, S. Raval, and B. Banerjee, "A robust approach to identify roof bolts in 3D point cloud data captured from a mobile laser scanner," *International Journal of Mining Science and Technology*, vol. 31, no. 2, pp. 303–312, 2021.

[17] M. Karimi, M. Oelsch, O. Stengel, E. Babaians, and E. Steinbach, "LoLa-SLAM: low-latency LiDAR SLAM using continuous scan slicing," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2248–2255, 2021.

[18] K. Qiu, T. Liu, and S. Shen, "Model-based global localization for aerial robots using edge alignment," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1256–1263, 2017.

[19] Q. Shi, S. Zhao, X. Cui, M. Lu, and M. Jia, "Anchor self-localization algorithm based on UWB ranging and inertial measurements," *Tsinghua Science and Technology*, vol. 24, no. 6, pp. 728–737, 2019.

[20] R. Karlsson and F. Gustafsson, "The future of automotive localization algorithms: available, reliable, and scalable

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9957067, 11 pages
https://doi.org/10.1155/2021/9957067

Research Article

A Sparse Deep Transfer Learning Model and Its Application for Smart Agriculture

Zhikui Chen,[1,2] **Xu Zhang**,[1] **Shi Chen**,[3] **and Fangming Zhong**[1]

[1]*The School of Software Technology, Dalian University of Technology, Dalian 116620, China*
[2]*The Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian 116620, China*
[3]*Emporia State University, Emporia, Kansas 66801, USA*

Correspondence should be addressed to Zhikui Chen; zkchen@dlut.edu.cn

Received 10 March 2021; Accepted 9 June 2021; Published 23 June 2021

Academic Editor: Keping Yu

The introduction of deep transfer learning (DTL) further reduces the requirement of data and expert knowledge in various uses of applications, helping DNN-based models effectively reuse information. However, it often transfers all parameters from the source network that might be useful to the task. The redundant trainable parameters restrict DTL in low-computing-power devices and edge computing, while small effective networks with fewer parameters have difficulty transferring knowledge due to structural differences in design. For the challenge of how to transfer a simplified model from a complex network, in this paper, an algorithm is proposed to realize a sparse DTL, which only transfers and retains the most necessary structure to reduce the parameters of the final model. Sparse transfer hypothesis is introduced, in which a compressing strategy is designed to construct deep sparse networks that distill useful information in the auxiliary domain, improving the transfer efficiency. The proposed method is evaluated on representative datasets and applied for smart agriculture to train deep identification models that can effectively detect new pests using few data samples.

1. Introduction

Although they have many advantages in performance, deep neural network- (DNN-) based methods often require expert knowledge to label data samples for generating datasets in training. The heavy requirement of labeled data will result in significant training costs, which make it expensive for extension. The deep transfer learning (DTL) can reuse well-trained models for the identification task and transfer knowledge that is learned from laboratory data to help identify in-field data, which alleviates the dependency on labeled datasets to reduce the cost. However, DTL has still not changed the fact that a considerable number of parameters need to be calculated, because they transfer all parameters, and many of the trained DNNs are overparametric. For example, ResNet-18 is the network commonly used as the backbone of DNN-based image recognition, in which up to 11.2 M (million) parameters need to be trained during each epoch [1]. Among these parameters waiting to be transferred, the existence of redundant parameters which are irrelevant to the target task would result in a large amount of unnecessary computation and memory cost. The redundancy in the network seriously affects the quality of transfer, as well as limiting the possibility for popularizing. Meanwhile, the methods of designing small efficient DNNs with fewer parameters will also face difficulty in inheriting knowledge from DTL, due to their unique network structure.

Agriculture is one of the most important basic industries, which covers a wide range of the world. Agricultural production faces many risks, of which pest and disease outbreaks are the most economic threats [2], and early identification of plants in the field is a crucial first step to detect and control the spread of diseases and pests [3]. Traditional in-field plant pest and disease identification methods rely on human's experience of manual observation and evaluation, with relatively low accuracy and efficiency in detection. With the development of intelligent agriculture, more technologies, i.e., remote sensing, IoT devices, computer vision [4], and unmanned aerial vehicle (UAV) [5], are providing new tools for in-field

plant pest and disease detection based on automated image recognition, which can help with large-scale early identification. At the same time, the data is growing. As the amount of data escalates, more identifications of plant pests and diseases by the aid of DNN are proposed. Images collected in the field are used to train deep networks, in which extracted features are used for recognition and classification. In recent years, DNN-based methods have been applied to the pest and disease identification of cash crops and grain crops [6–9].

For further agricultural extension, approaches with general applicability should be able to be used in mobile terminals, smartphones, and other small devices in the edge computing area. To this end, models must balance the performance with applicability and efficiency and adapt to limited processing power on the basis of ensuring detection accuracy. If only the most necessary parts would be transferred in DTL, it is possible to earn simplified models with lower device requirements for image recognition tasks, which could really reduce the computational cost and retain the advantage of inheriting knowledge by transfer learning. The newly proposed *lottery ticket hypothesis* (LTH) provides a theoretical possibility for this [10]. It finds representative sparse subnetworks by pruning the original network, which can be retrained to achieve equivalent or higher performance, but it uses fewer parameters (even only 5%-10% of the original's). It provides a new idea for plant pest and disease identification based on DTL: the network structure of the source domain is firstly pruned to obtain a sparse subnetwork with key knowledge. Then, the subnetwork, instead of the entire network in traditional DTL, will be transferred as the solution of the target task to achieve a sparse DTL. Thus, the requirement of expert knowledge and in-field data samples and the calculation of parameters can be reduced simultaneously.

As indicated above, in this paper, a sparse deep transfer learning method is proposed and applied to solve the problem of plant pest and disease identification based on image recognition. Firstly, a hypothesis is proposed that a transferable sparse subnetwork structure can be found and its portability can be verified. Then, the steps of the method are designed and used in DNN-based plant pest and disease identification, to seek and transfer an optimal sparse subnetwork to the target task to explore the application in practical problems. Finally, simulation experiments are carried out to show that the method can achieve an equivalent (or even higher) recognition accuracy with a more simplified network architecture and fewer parameters, while retaining the advantage of utilizing existing knowledge through transfer learning.

Thus, the main contributions can be concluded in twofold aspects:

(i) To relieve the lack of in-field labeled data and reduce the cost of collecting and labeling data samples by professionals for model training, a DTL-based method is designed, which can moderate the dependence of data in a plant pest and disease identification deep learning model

(ii) To cope with the defect that the DTL-based method cannot reduce high computational complexity and high hardware requirements, a sparse transfer strategy is designed, which transfers the pruned network structure to reduce the parameters that need to be trained in the model, to simplify the network architecture, reduce the volume and computing cost of the model, and thereby provide the possibility of running the model on ordinary office computers, smartphones, and edge computing devices for better agricultural extension

The rest of this paper is organized as follows. Section 2 is the related works about DNN and traditional DTL-based plant identification and LTH. The proposed sparse transfer hypothesis and the sparse deep transfer learning strategy with its steps are given in Section 3. Section 4 proves the proposed sparse transfer hypothesis on benchmark datasets and verifies the performance of the proposed method on the real dataset. Finally, Section 5 concludes the whole work and gives future discussion.

2. Related Works

With the development and popularization of image sensors in DNN-based plant pest and disease identification models, the use of the convolutional neural network (CNN) is becoming an important trend in agriculture. The pests and diseases can be detected and classified by insect individuals, lesions, and representative characteristic changes, which are usually manifested on the leaves of affected plants [11]. Thus, combined with the corresponding agricultural knowledge, images of healthy and diseased leaves can be used as the input of CNN to train the identification model. Methods have been applied to a variety of food and cash crops including but not limited to rice [12, 13], corn [14], tea [15–17], cannabis [18], and apple [19].

In the above deep CNN-based models, there exist two main problems:

(1) *Data requirements*: the high cost of training models from scratch due to the lack of labeled data and the requirement of expert knowledge in labeling them

(2) *Model size*: large-scale network architecture occupies much memory and resources, which is not suitable for low-computing-power devices. It increases the difficulty of storage and transmission, which limits the scope of application

For the contradiction between labeled data requirements and the lack of in-field data in (1), DTL is introduced. Mohanty et al. [20] combine DTL to improve efficiency in training a CNN model to identify 14 crop species and 26 diseases and achieve an accuracy of 99.35% with a hold-out test set. Ramcharan et al. [21] apply DTL to identify 2 types of pests and 3 diseases of cassava images taken in the field of Tanzania, inheriting the knowledge of image recognition from GoogLeNet_Inception v3, and achieve an overall accuracy of 93% for unseen data using 11,670 original images.

Libo et al. [22] use DTL in real-time detection of cole diseases and pests to solve the unbalance classes and false positives generated in training. Thenmozhi and Reddy [23] design a crop pest classification method based on deep CNN and transfer learning. However, DTL usually fails to reduce the requirement of parameters, which results in too large models to deploy on low-computing-power hardware systems.

For (2), the network structure is optimized [24, 25], and particular structures of small efficient CNNs are usually designed with less computation and small volume that is easy to popularize in application. Rahman [26] proposes a two-stage small CNN architecture, which reduced the model size by 99% compared to VGG-16 while remaining an accuracy of 93.3%. Xing et al. [27] develop a weakly dense CNN model for citrus diseases and pest recognition, which is designed from the aspect of parameter efficiency and is simple enough for mobile devices. And regarding designing small efficient neural networks, some state-of-the-art memory-efficient CNN architectures, such as MobileNet [28] and SqueezeNet [29], are usually used as backbones or references. These simple efficient neural networks usually consume less power and take up less memory, which makes it easier to store and deploy them on low-power hardware systems. At the same time, by using fewer parameters, the models require less data for convergence and will be able to avoid dense computing. However, the biggest problem is that because of the distinction in network structure, it is difficult to combine DTL that reuses structures and parameters with these simple CNNs, for commonly used parameters' weights in DTL are usually based on dense networks such as VGG structure or ResNet structure.

To take advantage of DTL in building resource-efficient CNNs, researchers have made a series of efforts [30–32]. In 2019, Frankle and Carbin proposed the LTH [10], which can compress the model by finding representative sparse subnetworks to replace the original dense network. Because the subnetwork is retrainable while preserving the original performance, it may provide a theoretical possibility for generating simple efficient networks from the original dense network and then transfer it to the target task.

To sum up, in the face of these challenges, in this paper, the LTH is modified by using it in DTL to generate transferable sparse structures. Therefore, it transfers only the most necessary knowledge while reducing the volume of the network, to realize the sparse deep transfer learning.

3. Methods

The LTH states that for a feedforward DNN, there is an implicit optimal sparse subnetwork structure which is retrainable to achieve the same accuracy as the dense network within the number of original's iterations. It can succeed in finding the subnetwork to retain knowledge and ability from large-scale datasets such as ImageNet in visual recognition tasks [33]. The possibility that whether the subnetwork is able to transfer in the discussion of LTH has been raised, but whether the transfer can always be realized between tasks has so far been inconclusive.

In this section, on the basis of the studies above, a sparse transfer method named WLTs-SDTL is proposed. It transfers only the most important part of the original network, and LTH is modified for generating sparse subnetworks in DTL. The method is then applied in plant pest and disease identification based on image recognition.

3.1. Sparse Transfer Hypothesis for WLT-Nets

3.1.1. Reviewing the LTH. The particular subnetwork is generated when randomly initialized, and to seek for it is like to find a "winning lottery ticket" in the original network. We named it as a WLT-net. The retrainable WLT-net is able to be found by unstructured pruning according to the following conditions:

Consider a feedforward network whose loss function $\ell = \xi(x)$ and initial parameters are defined by λ_i. After training and optimizing, the network achieves the minimum validation of ℓ with the accuracy rate $\eta\%$ when the number of iterations is n. The exists WLT-net $\xi(x; M(\lambda_i))$ from $\xi(x)$ with iteration $n' \le n$ and accuracy rate $\eta'\% \ge \eta\%$, in which $M(\lambda_i) = m \odot \lambda_i$ is the Mask function; a mask $m \in \{0,1\}^{|\lambda|}$ is used for determining and marking which weights will be retained after pruning.

Then, give the definition of the domain and task in DTL: the source domain is denoted as D_S with task T_S for providing transferable knowledge and the target domain is denoted as D_T with task T_T, in which $D = \{\chi, P(X)\}, X = \{x_1, x_2, \cdots, x_n\} \in \chi$ and $T = \{y, f(x)\}$; y is the label space and $f(x)$ can be regarded as the nonlinear loss function of DNN to map x to y.

Thus, combined with DTL, the sparse transfer hypothesis for WLT-nets is proposed in Assumption 1. The proof procedure shows that when meeting the conditions, a task can be regarded as generated by LTH from a larger dense DNN, while retaining the most necessary backbone architecture and knowledge. Conversely, the proposed hypothesis can be used in designing a transferable WLT-net from an existing arbitrary dense network T_S, to inherit whose knowledge. The required ability is then able to transfer from D_S to D_T through a small efficient WLT-net that is isomorphic to T_T, which completes the process of a WLT-net-based sparse DTL.

Assumption 1. Sparse transfer hypothesis for WLT-nets.

For the task T waiting to be solved, we modeled it as DNN. The network can be regarded as T_T in D_T, where $T_T = \xi(x; M(\lambda_i))$.

According to the reverse reasoning of LT hypothesis, $\exists T_S$, the corresponding dense network in D_S, which makes T_T to be the WLT-net of T_S, only if T_T satisfies the following conditions simultaneously.

When the loss function ℓ achieves minimum validation (compared with T_S):

(1) The number of iterations is $n_T \le n_S$

(2) The percentage of accuracy rate is $\eta_T \ge \eta_S$

(3) $\xi(x; M(\lambda_i))$ is able to be obtained from T_S by unstructured pruning using the Mask function $M(\lambda_i)$

$M(\lambda_i) = m \odot \lambda_i$, and mask $m \in \{0, 1\}^{|\lambda|}$.

Now that T_T is viewed as the WLT-net generated from T_S, which is proven that it can keep the performance in LT hypothesis, it provides a shortcut to transfer knowledge through sparse structures.

When meeting the conditions, a small efficient network can be regarded as the sparse part of a larger dense DNN. In this way, transferable knowledge can be sought from an isomorphic structure in D_S for reusing in D_T, to realize the sparse DTL.

The WLT-net in LTH has been proven to retain the original network's performance with only 5%-10% of parameters left. In particular, when λ_i is used for initializing retraining, the performance is better than that of random initialization. It shows that (1) the main functions of the network are retained in the most important parameters and (2) when the original dense network is initialized, knowledge and skills are simultaneously initialized into the particular WLT-net. Since WLT-net is retrainable, these knowledge and skills are able to transfer within a sparse network backbone as long as the network T_T in the target domain has an identical structure.

3.2. WLT-Net-Based Sparse Deep Transfer Learning. In this section, the specific methods and steps of WLTs-SDTL are proposed. The process is illustrated in Figure 1.

Concretely, the steps of implementation are as follows:

Step 1. Locate and identify D_S in DTL. The structure of T_S will determine the subsequent sparse structure.

(i) When D_S is representative in the corresponding research field, such as ImageNet, which consists of a vast scale of generic data samples, a well-trained T_S using a high-performance hardware device could always be promising

(ii) It should be reasonable for choosing a DNN, which has been trained and used to solve certain problems in the practical application, to be the T_S

(iii) It is also able to train a new DNN from scratch as T_S, if the task is too unique to seek references. Then, the proposed method will compress it with an acceptable additional computation cost to get a simple efficient network that is easy to promote and deploy

Step 2. Prepare to seek WLT-nets in the DNN of T_S.

(i) Based on the proposed *sparse transfer hypothesis for WLT-nets*, an iterative pruning algorithm will be used. The weight of initial parameters in T_S is denoted by λ_i, while after n time iterations, λ_n denotes the weight when the loss function achieves minimum validation. In addition, $\lambda_\alpha (\alpha < n)$ is recorded, preparing for *late reset*, a skill to speed up convergence and improve accuracy

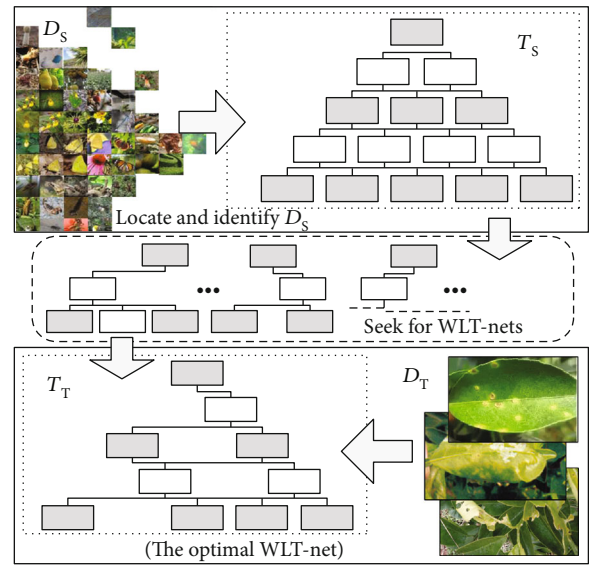


Figure 1: Process of the WLTs-SDTL method.

(ii) μ is defined as the rating standard score to measure the contribution of parameters' weights in DNN, whose definition is

$$\mu(\lambda_i) = \max\left(0, \frac{\lambda_i \lambda_n}{|\lambda_i|}\right). \quad (1)$$

Step 3. Generate WLT-nets by unstructured pruning.

(i) Rank μ by score. The Mask function $M(\lambda_i) = m \odot \lambda_i$ is introduced, using a mask $m \in \{0, 1\}^{|\lambda|}$ to determine whether a parameter should be remained

(ii) Set a pruning ratio p for each epoch. Then, in the layers of DNN, for the weights of parameters whose scores are in the top p%, use the mask $m = 1$ to label them on behalf of retaining. Oppositely, the residual $(100 - p)$% parameters will be pruned, whose masks m are set to 0. The value of p can be defined for each layer separately

Step 4. Parameters are processed according to the value of the mask m obtained in step 3.

(i) For parameters whose mask $m = 1$, reset their weights to the recorded λ_α in step 2

(ii) For parameters whose mask $m = 0$, prune them. Their weights will be frozen in the subsequent training of DNN, which results in a sparse network architecture

(iii) Different from the original LTH, the variation trend of weight is considered. The weights of the parameters being pruned are frozen at 0 only when they tend to 0. When the variation trend of weights in

the training is moving away from 0, they are frozen at λ_i

Step 5. Repeat step 2 to step 4, until the optimal transferable WLT-net ξ is obtained.

Step 6. Use ξ as T_T in the target domain D_T to realize the sparse DTL based on the optimal WLT-net. In the fine-tuned training dataset of D_T, initialize the DNN with λ_α, while the frozen parameters are not trained.

This way, important skills and information can be inherited from D_S through the transfer of a sparse WLT-net architecture. The requirement of data samples and computing parameters in the DNN training are both reduced.

Compared with the original LTH, in the proposed WLTs-SDTL method, we have made optimizations and improvements in the following three aspects:

(i) The LTH is extended and modified for generating sparse networks in DTL. By regarding T_T as a WLT-net of T_S, the correlation between D_S and D_T is established, enabling knowledge to be transferred between tasks. When following the idea in the original LTH, T_T should be initialized using a $\lambda_i{}'$ in D_T, not λ_α, so that the knowledge obtained from D_S will be lost. In the proposed WLTs-SDTL, since we regard T_T as a part generated from T_S, it can then be initialized naturally with λ_α conforming to the LTH, while retaining the performance. Thus, the proposed method can achieve the knowledge transfer using sparse WLT-net effectively

(ii) There is a more reasonable standard to evaluate parameters in pruning. In the pruning process of the original LTH, the rating standard score μ in Mask function M which evaluates the contribution of parameters' weights is defined as $\mu = |\lambda_i|$. Clearly, it failed to consider the sign change of parameters in training, i.e., across the zero axis. As shown in formula (1), the proposed μ emphasizes the role of signs, which leads to a correct expression of the trend in weight changing. Comparative experiments prove that it can improve the final performance of the sparse network

(iii) Further consider the influence of the trend in weight changing during training on the freeze and reset of parameters. In the original LTH, after pruning, the parameters will be reset at λ_i ($m = 1$) or frozen at 0 ($m = 0$). Subsequently, the *late reset* [33] is proposed to use the recorded weights after a period of training iterations for the reset, to make the convergence faster and the final accuracy higher. In the proposed WLTs-SDTL method, we also adopt it that $\lambda_\alpha (\alpha < n)$ is used for resetting parameters instead of λ_i in step 4.

Furthermore, since the *late reset* is effective, considering the frozen ones, the weights of pruned parameters will be frozen at 0 to avoid subsequent training, but why 0?

A reasonable explanation is that these weights contribute less to the network and are not important. However, if it really does not matter, these weights could be set to any value, instead of a particular 0, without affecting the network's performance. In experiments that freeze these parameters to λ_i, it shows that, similar to *late reset*, the validity might depend on whether a specific value can reflect the changing trend of weights in training to some extent. When freezing at 0 correctly, it is equivalent to letting weights whose trends are getting closer to 0 reach to their final value in advance. Thus, in the paper, a parameter will be frozen at 0 only if its trend tends to be 0. When the trend is away from 0, freeze the parameter at its value in λ_i to reduce the impact.

4. Experiment

In this section, experiments are designed to verify the hypothesis and evaluate the performance of the proposed method in actual solutions. Firstly, the proposed *sparse transfer hypothesis for WLT-nets* and *WLTs-SDTL* are verified on the benchmark datasets. Then, the WLTs-SDTL method is used to design a detection model based on image recognition and applied to actual solutions of plant pest and disease identification using open-source lab datasets. Finally, a small scale of real datasets that we collected in Chongqing, China, is used in training a model that realized the citrus greening disease (Haunglongbing) identification. The experiments of training models are run on the server which contains 2 Inter Xeon Silver 4110 8-core CPUs and 2 NVIDIA Tesla M60 GPUs (128 G), while some validations are able to run on ordinary office computers since the sparse models are used.

4.1. Verification of WLT-Net-Based Sparse DTL. The sparse transfer hypothesis for WLT-nets is verified on the benchmark datasets, i.e., CIFAR-10 and SmallNORB. Specifically, define a DNN-based task T_S on the D_S using CIFAR-10, and the proposed method is used to find transferable sparse WLT-net, which will be used as T_T in DTL. The accuracy and computational load of the model after transfer will be compared with those of the fully connected dense DNN structure trained on SmallNORB, the dataset of D_T. Through the above approaches, it validates whether the modified hypothesis is able to realize a sparse DTL to reduce the parameters while maintaining the accuracy in the proposed WLTs-SDTL method.

4.1.1. Datasets. Identification of plant pests and diseases can be modeled as a multiclassification task based on image recognition. Therefore, two of the classical datasets, CIFAR-10 [34] and SmallNORB [35], are chosen for designing a simple experiment to evaluate the feasibility of the hypothesis, which are widely used in identifying ubiquitous objects and regarded as the benchmark to validate various models. The properties of the datasets are shown in Table 1. Since the channel and image size of T_S are different from those of D_S, channel conversion and 4-pixel padding are applied at the training time.

4.1.2. Settings. About the structure of the network, which contains both the original fully connected dense DNN and the sparse WLT-net generated from it, ResNet-18 is chosen for the backbone. As the classical deep residual network's

Table 1: Properties of experimental parameters.

Dataset	CIFAR-10	SmallNORB
Class	10	5
Train	50,000	40,000
Test	10,000	10,000
Image size	$32 \times 32 \times 3$	$28 \times 28 \times 1$
Domain	D_S	D_T
Parameter settings	SGD [$5e^{-3}$, $1e^{-3}$, $1e^{-4}$] with momentum 0.9, weight decay $1e^{-4}$, and learning rate = 0.01	

18-layer version (with 17 convolution layers and 1 fully connected layer, 11.2 M parameters to train), it is also used in the original LTH, so the same configuration is set for the experiments. Settings of experimental parameters are shown in Table 1 too. The pruning rate in finding WLT-net is set to 20% with batch size 128, the maximum number of iterations 30,000, and at most 50 epochs in each iteration. 10 rounds of iterative pruning are performed to find the optimal WLT-net. During the operation, only the convolution layers will be pruned. When retraining the transferred sparse network in T_T for optimizing, by convention, weights of the convolution layers are frozen and only the fully connected layer is fine-tuned.

4.1.3. Verify the Sparse Transfer Hypothesis for WLT-Nets. Firstly, experiments are designed to validate the performance of WLT-net compared with the original dense network on D_S. After training in the source domain CIFAR-10, the original dense ResNet-18 achieved an average accuracy of 89.43% in the test dataset, with 11,173,962 parameters used.

The experimental results are shown in Figure 2, to display the relationship among the pruning level, number of remaining parameters, and average accuracy. When only 10.7% of the original parameters are retained after pruning (1,212,145 used), an average accuracy of 89.24% is still able to be achieved. As illustrated in the figure, it is proven that the pruning method in the proposed sparse transfer hypothesis for WLT-nets can guarantee the accuracy and reduce parameters while generating sparse subnetworks.

4.1.4. Verify WLTs-SDTL. Then, on the basis of the above analysis, to validate whether the sparse WLT-net can inherit knowledge from T_S, more experiments are designed, and the effect of initialization on the results is compared. As for SmallNORB on D_T, when training a dense ResNet-18 from scratch, the average accuracy achieves 89.9%. The optimal WLT-net generated in each round under different pruning levels is transferred to T_T, respectively, and the accuracy is compared. Since the influence of trend in weight change during training has been further considered in this paper, to better compare and prove the effectiveness of the proposed method, four kinds of initialization approaches are used for the sparse network separately: (a) original LTH (using λ_i), (b) random initialization, (c) late reset method (using λ_α), and (d) proposed WLTs-SDTL. The performance of the original dense network before pruning is used as the baseline. The experimental results are shown in Figure 3.

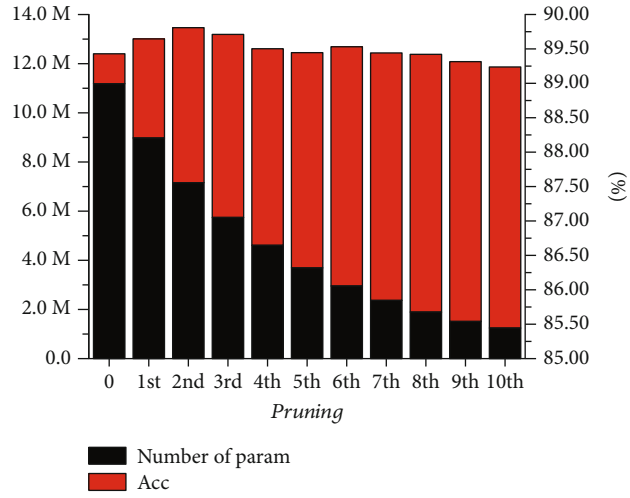


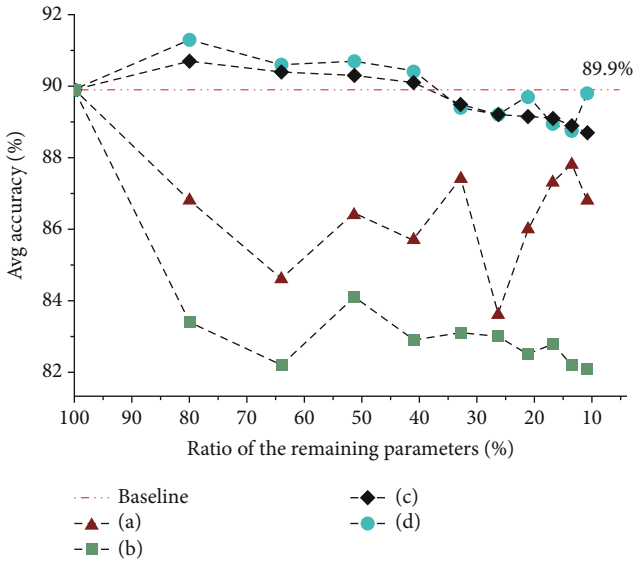
Figure 2: Experimental results: verify the sparse transfer hypothesis for WLT-nets.

As the experimental data shows, a better performance can be obtained than that of training dense DNN directly when proper pruning is carried out. It proves that the WLTs-SDTL can transfer the necessary ability from T_S in D_S to T_T in D_T. The sparse transferable structure is able to greatly save the cost of training parameters, while the overall accuracy can be kept, which realizes a sparse DTL. Compared with the initialization approaches (a), (b), and (c), the proposed WLTs-SDTL is more effective on the whole. At the same time, according to the results, when running a deep-level pruning on the original model, in which only 10% of the parameters remains (1,212,145 compared with the original 11,173,962), the precision loss is acceptable sometimes, which corresponds to the original LTH. Thus, when the target task T_T can accept a slight performance loss in exchange for generalization ability, it offers the possibility of using DNN-based deep computing methods on devices with low computational power such as mobile terminals, smartphones, or edge computing devices.

In summary, experiments on benchmark datasets have verified the feasibility of the proposed *sparse transfer hypothesis for WLT-nets* and *WLTs-SDTL*.

4.2. Identification of Pests and Diseases Based on WLTs-SDTL. In this section, the proposed WLTs-SDTL is used to train a sparse network from a dense detection model based on image recognition and applied in actual solutions of plant pest and disease identification. Specifically, the common diseases of tomato leaves are identified, inheriting the ability from ImageNet and using open-source lab datasets for T_T's fine-tuning training.

4.2.1. Datasets. The ResNet-18 network pretrained on ImageNet is used as D_S to provide the necessary knowledge from weight of parameters. As for the T_T in identifying pests and diseases on tomato leaves, the PlantVillage dataset is chosen. The PlantVillage [36] is an open-source image dataset of



Param. remains with best Acc				
80%	64%	51.2%	40.96%	32.77%
91.3% (d)	90.6% (d)	90.7% (d)	90.4% (d)	89.5% (c)
26.21%	20.97%	16.77%	13.42%	10.74%
89.2% (c,d)	89.7% (d)	89.1% (c)	88.9% (c)	89.8% (d)

Figure 3: Experimental results: using WLTs-SDTL on benchmark datasets.

leaves from 14 crops, which contains 26 categories of plant pests and diseases and corresponding healthy leaves.

Since the samples of different categories in the original dataset are uneven, the crop tomato with sufficient samples is selected, in which categories with fewer samples and images with poor quality are eliminated. Then, data enhancement methods such as horizontal flip are used to adjust the sample size of each category to the same. Finally, a total of 8 categories of pests and diseases/1 healthy leaf are defined; meanwhile, the image size is adjusted to 64×64 uniformly. The specific properties of the dataset are shown in Table 2.

4.2.2. Settings. Deeper-level pruning is carried out. 15 rounds of iterative pruning are performed to find the optimal WLT-net, with at least 3.6% of the parameters being retained (406,495 compared with 11,173,962 in the original dense net). Other experimental settings are the same as them in the previous section.

4.2.3. Identify Pests and Diseases of Tomato Leaves. The experimental results are shown in Figure 4. The identification model using dense ResNet-18 can achieve 96.44% accuracy after training in D_T, and a series of sparse networks with different volumes are obtained, respectively, using WLTs-SDTL.

The highest accuracy is up to 97.69% in pruning level 5, when 67% of the parameters are removed. By and large, WLTs-SDTL can guarantee the accuracy in plant identification while reducing parameter computation. When pruning properly, the accuracy can be higher than that of the original

Table 2: Properties of the PlantVillage dataset.

Category	Sample set	Training set
Tomato healthy	1,592	1,500
Tomato bacterial spot	2,127	1,500
Tomato early blight	1,000	1,500
Tomato late blight	1,910	1,500
Tomato Septoria leaf spot	1,771	1,500
Tomato spider mites	1,653	1,500
Tomato mosaic virus	373	Unused
Tomato leaf mold	952	1,500
Tomato target spot	1,404	1,500
Tomato TYLCV	5,357	1,500

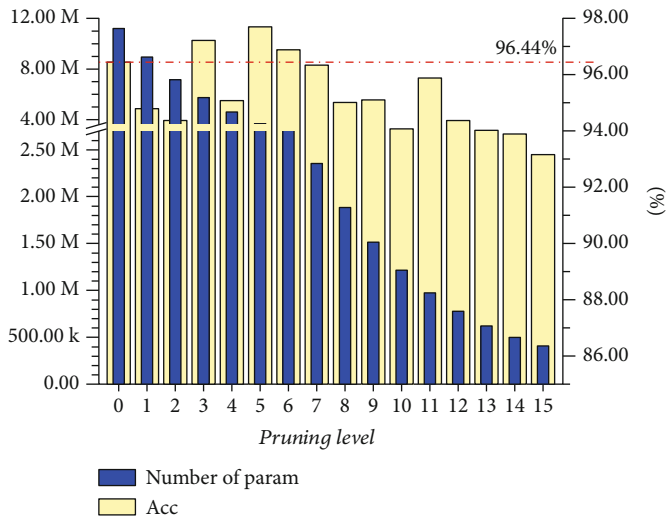


Figure 4: Experimental results of WLTs-SDTL on the PlantVillage dataset.

dense net. If the optimal performance is required, fine-grained pruning can be gradually carried out between levels near the best accuracy, which is between 30% and 50% in this set of experiments. For example, we can set the pruning rate to 10% or less in each round, to find a balance between performance and volume of the model.

When only 3.6% of the parameters are retained, the sparse network is still able to achieve an accuracy of 93.16%, with no other than 406,495 parameters to be trained. Considering that the accuracy is acceptable in daily identification tasks, it verifies that the proposed WLTs-SDTL can generate a small efficient network suitable for mobile terminals or edge computing devices with low computational power in the practical application of pest and disease identification.

4.3. Using WLTs-SDTL for Real Collected Data Identification. In this section, a small scale of datasets collected in Chongqing, China, is used in training the identification model to detect the citrus greening disease (Haunglongbing), combined with lab data. Citrus greening (Haunglongbing) is a

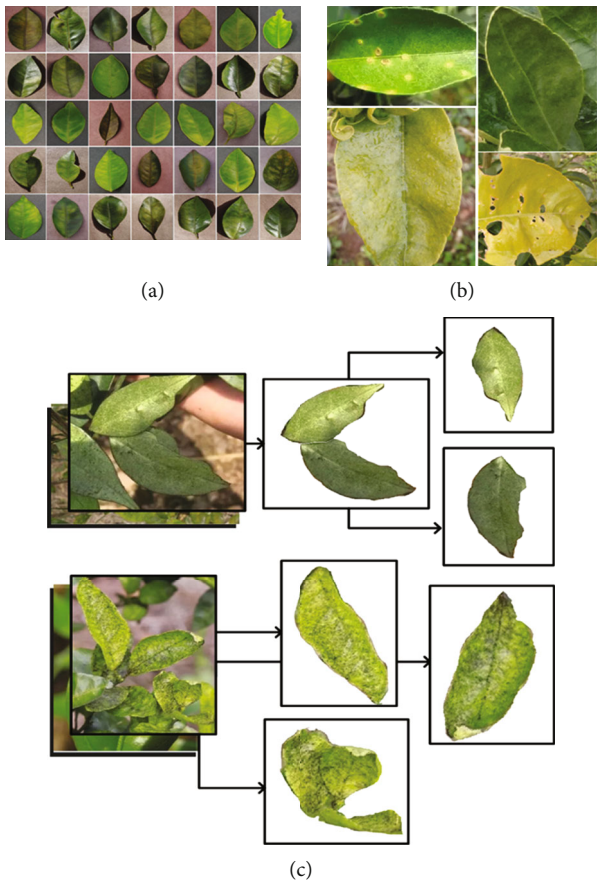


Figure 5: The composition of the dataset: (a) samples in PlantVillage; (b) collected samples; (c) process of clipping.

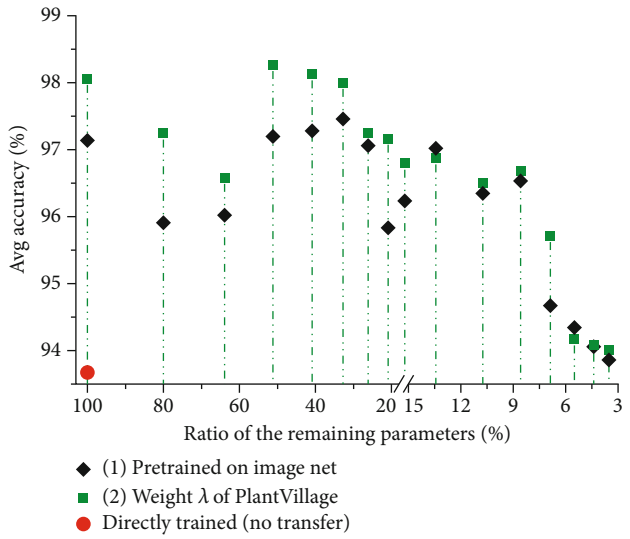


Figure 6: Experimental results of using WLTs-SDTL in identifying Haunglongbing.

devastating disease in the world citrus production, which seriously restricts the development of the citrus industry [37]. The etiolation of leaves can be used for its in-field identification.

4.3.1. Datasets. We have collected 1,266 images from Chongqing, China, as well as from Internet and monographs, in which 238 samples are Haunglongbing. In the process of photographing samples in the field, there may be more than one leaf in a photo, and more training samples can be obtained through clipping, as illustrated in Figure 5(c). Haunglongbing can also be identified by fruit; however, due to the lack of samples, images of diseased fruit are screened out. After the clipping, filtering, and image intensification operations, 1,500 samples are generated for training. Meanwhile, in the category orange haunglongbing of the PlantVillage dataset used in Section 4.2.1, which contains 5,507 images of leaves, 5,000 images are randomly selected (Figure 5(a)). Finally, the training dataset of 6,500 samples is created.

4.3.2. Settings. 15 rounds of deep-level iterative pruning are performed, while the other experimental settings are the same as them in the previous section. About the initial weight used for late reset, two kinds of options are chosen: (1) the weight of ResNet-18 was pretrained on ImageNet, which has been used as D_S in Section 4.2.1 and (2) the final weight λ_B of identification task T_T in Section 4.2.3 has been trained on PlantVillage datasets. The contrast experiment is designed to compare whether the proposed method can inherit better ability from a more *similar domain*.

4.3.3. Identifying Haunglongbing. After training on the original dense network, when performing a traditional intensive DTL (no parameters pruned), the initial weight (1) achieves 97.14% accuracy while the final weight (2) achieves 98.06%. Then, the proposed WLTs-SDTL is used, respectively; the relationship of accuracy and remaining parameters is shown in Figure 6. The highest accuracy is not even able to achieve 94% when training a dense ResNet-18 from scratch, and pruning the network will lead to the loss of overall accuracy; its performance on simplified WLT-net can be no longer discussed in this section (because the performance will not be better than that using DTL).

The experimental results show that, compared with training directly, DTL can achieve better initial performance in identifying citrus Haunglongbing disease under the help of collected data. And regarding the proposed sparse DTL, the following are obtained:

(1) As the parameters decrease, the overall accuracy declines. However, it is still within an acceptable range and higher than no-transfer dense network. When the weight from a similar identification task is used in (2) shown in Section 4.3.2, a higher initial accuracy can be obtained, and the trend of accuracy is more stable in subsequent sparsification than (1) shown in Section 4.3.2. Thus, a similar task, which has been well trained for identifying other pests and diseases, might be a better initial choice in practical application

(2) Although fewer parameters are used, the model can achieve higher accuracy in some pruning levels: in (1) shown in Section 4.3.2, among the 3rd to the 5th levels, the average accuracy, respectively, achieves 97.20%, 97.28%, and 97.46% when 51.2%, 40.96%, and 32.77% of the parameters remain higher than the original 97.14%. And in (2) shown in Section 4.3.2, in the 3rd and 4th pruning level, it achieves 98.26% and 98.13% compared with the original 98.06%, while it is 97.99% in the 5th level. The reason is that sparsification reduces the redundancy of parameters, and the negative feedback of low-contribution parameters is inhibited

(3) Fine-grained pruning in the optimal range can be proceeded for the best performance. And in the experimental results of this paper, note that the range is always between 50% and 30%. Therefore, we speculate that in the proposed WLTs-SDTL, the priority could be given to these pruning levels when models' performance is preferred. And when the volume of the model needs to be compressed as much for widely deploying, it shows that the sparse model can use about 10% of the original parameters (accuracy 96.67% when using 8.59%) to maintain an acceptable performance close to that of the original. The limiting small model with only 3.6% of the original parameters is also taken into account, whose accuracy is able to achieve 94.01%, still higher than that of the dense net without transfer, and is more likely to be used for low-computing-power devices or edge computing devices

To sum up, when the proposed WLTs-SDTL is used in an actual solution of identifying diseases of plants, the sparsification of the network can be realized through pruning to save the computational overhead of parameters while maintaining or even improving performance. Thus, the balance between performance and model size can be dynamically adjusted, and the deployment possibility of low-computational-power equipment is provided.

5. Conclusion

In this paper, a sparse deep transfer learning model is proposed. The method is aimed at modeling the identification of plant pest and disease with limited collected data in the field, and a sparse DTL strategy is designed to transfer only the most important architecture and optimize models' size.

Specifically, (1) the sparse transfer hypothesis is proven, which succeeds in modifying LTH to reduce the parameter computation in DTL by generating sparse transferable WLT-nets. (2) The sparse transfer method named WLTs-SDTL is formally proposed, in which the compressing strategy is designed to construct a deep sparse network, distills useful information from the auxiliary domains, and improves the transfer efficiency. (3) The proposed method is applied to detect pests and diseases with few data samples in training deep identification models. The hypothesis is verified by the benchmark dataset; meanwhile, the proposed method is evaluated on the representative datasets.

Experimental results show that when the proposed method is used in actual solutions, the sparsification of the network can save the cost of computing parameters while maintaining or sometimes improving the performance, thereby dynamically adjusting the balance between the model's accuracy and size, providing the deployment possibility in low-computational-power devices.

Moreover, the sparse strategy can be promoted in identifying new pests and diseases of the plant with few data and even widely used in other tasks based on image recognition and lack of data. On that occasion, depending on specific tasks, it is supposed to wisely choose the suitable network architecture and balance the accuracy and volume of models.

In the future, the proposed method will be studied on more other domains to overcome the scarcity of data and the redundancy of model parameters, improving the effectiveness of sparse deep transfer learning.

Data Availability

The open-source datasets used in this paper, such as IMGnet, are freely available in various deep learning frameworks such as PyTorch and TensorFlow.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Nos. 61672123 and 62076047) and the Fundamental Research Funds for the Central Universities (Nos. DUT20LAB136 and DUT20TD107).

References

[1] N. Shazeer, A. Mirhoseini, K. Maziarz et al., "Outrageously large neural networks: the sparsely-gated mixture-of-experts layer," 2017, https://arxiv.org/abs/1701.06538/.

[2] J. Zhang, Y. Huang, R. Pu et al., "Monitoring plant diseases and pests through remote sensing technology: a review," *Computers and Electronics in Agriculture*, vol. 165, article 104943, 2019.

[3] A. S. Gurle, S. N. Barathe, R. S. Gangule, S. D. Jagtap, and T. Patankar, "Survey paper on tomato crop disease detection and pest management," *International Journal of Applied Evolutionary Computation*, vol. 10, no. 3, pp. 10–18, 2019.

[4] Y. Li, J. Zhou, X. Zheng, J. Tian, and Y. Y. Tang, "Robust sub space clustering with independent and piecewise identically distributed noise modeling," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pp. 8720–8729, Long Beach, CA, USA, 2019.

[5] D. Gao, Q. Sun, B. Hu, and S. Zhang, "A framework for agricultural pest and disease monitoring based on Internet-of-things and unmanned aerial vehicles," *Sensors*, vol. 20, no. 5, article 1487, 2020.

[6] E. Mwebaze and G. Owomugisha, "Machine learning for plant disease incidence and severity measurements from leaf images," in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 158–163, Anaheim, CA, USA, 2016.

[7] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, article 2022, 2017.

[8] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: classification and symptoms visualization," *Applied Artificial Intelligence*, vol. 31, no. 4, pp. 299–315, 2017.

[9] A. A. Alfarisy, Q. Chen, and M. Guo, "Deep learning based classification for paddy pests & diseases recognition," in *Proceedings of 2018 International Conference on Mathematics and Artificial Intelligence*, pp. 21–25, Chengdu, China, 2018.

[10] J. Frankle and M. Carbin, "The lottery ticket hypothesis: finding sparse, trainable neural networks," 2019, https://arxiv.org/abs/1803.03635/.

[11] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.

[12] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, 2017.

[13] D. Li, R. Wang, C. Xie et al., "A recognition method for rice plant diseases and pests video detection based on deep convolutional neural network," *Sensors*, vol. 20, no. 3, p. 578, 2020.

[14] P. Bhatt, S. Sarangi, A. Shivhare, D. Singh, and S. Pappula, "Identification of diseases in corn leaves using convolutional neural networks and boosting," *ICPRAM*, 2019, pp. 894–899, SciTePress, 2019.

[15] X. Sun, Y. X. ShaominMu, Z. Cao, and T. Su, "Image recognition of tea leaf diseases based on convolutional neural network," in *2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, pp. 304–309, Jinan, China, 2018, 10.1109/SPAC46244.2018.8965555.

[16] G. Hu, X. Yang, Y. Zhang, and M. Wan, "Identification of tea leaf diseases by using an improved deep convolutional neural network," *Sustainable Computing: Informatics and Systems*, vol. 24, article 100353, 2019.

[17] J. Chen, Q. Liu, and L. Gao, "Visual tea leaf disease recognition using a convolutional neural network model," *Symmetry*, vol. 11, no. 3, p. 343, 2019.

[18] K. P. Ferentinos, M. Barda, and D. Damer, "An image-based deep learning model for cannabis diseases, nutrient deficiencies and pests identification," in *Progress in Artificial Intelligence: 19th EPIA Conference on Artificial Intelligence, EPIA 2019*, Vila Real, Portugal, 2019.

[19] K. Zhang, Y. Guo, X. Wang, J. Yuan, and Q. Ding, "Multiple feature reweight DenseNet for image classification," *IEEE Access*, vol. 7, pp. 9872–9880, 2019.

[20] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," 2016, https://arxiv.org/pdf/1604.03169/.

[21] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," *Frontiers in Plant Science*, vol. 8, p. 1852, 2017.

[22] Z. Libo, H. Tian, G. Chunyun, and M. Elhoseny, "Real-time detection of cole diseases and insect pests in wireless sensor networks," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 3, pp. 3513–3524, 2019.

[23] K. Thenmozhi and U. S. Reddy, "Crop pest classification based on deep convolutional neural network and transfer learning," *Computers and Electronics in Agriculture*, vol. 164, article 104906, 2019.

[24] Y. Li, J. Zhou, J. Tian, X. Zheng, and Y. Y. Tang, "Weighted error entropy based information theoretic learning for robust subspace representation," *IEEE Transactions on Neural Networks and Learning Systems (T-NNLS)*, pp. 1–15, 2021.

[25] Y. Li, R. Liang, W. Wei, W. Wang, J. Zhou, and X. Li, "Temporal pyramid network with spatial-temporal attention for pedestrian trajectory prediction," *IEEE Transactions on Network Science and Engineering (T-NSE)*, 2021.

[26] C. R. Rahman, "Identification and recognition of rice diseases and pests using deep convolutional neural networks," 2018, https://arxiv.org/abs/1812.01043.

[27] S. Xing, M. Lee, and K. K. Lee, "Citrus pests and diseases recognition model using weakly dense connected convolution network," *Sensors*, vol. 19, no. 14, article 3195, 2019.

[28] G. Andrew, "MobileNets: efficient convolutional neural networks for mobile vision applications," 2017, https://arxiv.org/abs/1704.04861/.

[29] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1mb model size," 2016, https://arxiv.org/abs/1602.07360/.

[30] P. Molchanov, S. Tyree, T. Karras, T. Aila, and J. Kautz, "Pruning convolutional neural networks for resource efficient transfer learning," 2016, https://arxiv.org/abs/1611.06440/.

[31] J. Liu, Y. Wang, and Q. Yu, "Sparse deep transfer learning for convolutional neural network," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31no. 1, pp. 2245–2251, San Francisco, California, USA, 2017.

[32] C. Reinhold and M. Roisenberg, "Filter pruning for efficient transfer learning in deep convolutional neural networks," in *18th International Conference on Artificial Intelligence and Soft Computing*, pp. 191–202, Zakopane, Poland, 2019.

[33] J. Frankle, G. K. Dziugaite, D. M. Roy, and M. Carbin, "The lottery ticket hypothesis at scale," 2019, https://arxiv.org/abs/1903.01611v1/.

[34] R. C. Çalik and M. F. Demirci, "Cifar-10 image classification with convolutional neural networks for embedded systems," in *2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, pp. 1-2, Aqaba, Jordan, 2018.

[35] Y. LeCun, F. J. Huang, and L. Bottou, "Learning methods for generic object recognition with invariance to pose and lighting," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*, vol. 2, pp. II–104, Washington, DC, USA, 2004.

[36] D. P. Hughes and M. Salath'e, "An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing," 2015, https://arxiv.org/abs/1511.08060/.

[37] Y. Lan, Z. Huang, X. Deng et al., "Comparison of machine learning methods for citrus greening detection on UAV multispectral images," *Computers and Electronics in Agriculture*, vol. 171, article 105234, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9988651, 11 pages
https://doi.org/10.1155/2021/9988651

Research Article

Communication Signal Modulation Mechanism Based on Artificial Feature Engineering Deep Neural Network Modulation Identifier

Fei Lu,[1] **Zhenjiang Shi**,[1] **and Rijian Su**[2]

[1]*Department of Engineering Technology, Open University of Guangdong (Guangdong Polytechnic Institute), Guangzhou, Guangdong 510091, China*
[2]*School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, Henan 450000, China*

Correspondence should be addressed to Rijian Su; srj@zzuli.edu.cn

Received 29 March 2021; Revised 24 May 2021; Accepted 31 May 2021; Published 18 June 2021

Academic Editor: Wei Wang

Based on the characteristics of time domain and frequency domain recognition theory, a recognition scheme is designed to complete the modulation identification of communication signals including 16 analog and digital modulations, involving 10 different eigenvalues in total. In the in-class recognition of FSK signal, feature extraction in frequency domain is carried out, and a statistical algorithm of spectral peak number is proposed. This paper presents a method to calculate the rotation degree of constellation image. By calculating the rotation degree and modifying the clustering radius, the recognition rate of QAM signal is improved significantly. Another commonly used method for calculating the rotation of constellations is based on Radon transform. Compared with the proposed algorithm, the proposed algorithm has lower computational complexity and higher accuracy under certain SNR conditions. In the modulation discriminator of the deep neural network, the spectral features and cumulative features are extracted as inputs, the modified linear elements are used as neuron activation functions, and the cross-entropy is used as loss functions. In the modulation recognitor of deep neural network, deep neural network and cyclic neural network are constructed for modulation recognition of communication signals. The neural network automatic modulation recognizer is implemented on CPU and GPU, which verifies the recognition accuracy of communication signal modulation recognizer based on neural network. The experimental results show that the communication signal modulation recognizer based on artificial neural network has good classification accuracy in both the training set and the test set.

1. Introduction

Modulation recognition of communication signals has a wide range of application requirements in modern wireless communications [1]. In noncooperative communication, since the receiver does not know the communication parameters of the sender, especially the modulation mode, it cannot effectively perform subsequent operations such as demodulation [2]. Communication signal modulation recognition is widely used in noncooperative communication scenarios such as electronic countermeasures, communication reconnaissance, signal recognition, and electronic supervision. (2) In the cognitive radio system, the communication receiver hopes to realize a universal receiver to accurately receive the information sent by the sender [3]. In the design of universal receiver, signal modulation identification is a key technology. Only when the modulation mode of the signal is accurately identified can the carrier frequency bandwidth and other information of the signal be estimated more accurately, so as to facilitate the subsequent demodulation and decoding of the signal. Therefore, it is of great significance to study the modulation recognition of communication signals. In today's increasingly complex wireless communication environment, the electromagnetic signal space is more and more complex, the transmission of information is more and more large, and the signal changes more and more quickly. Therefore, the research on automatic and high-speed modulation pattern recognition of multiclass communication modulation signals has a very high application value.

The method of wavelet transform [4] is proposed to extract three different features for the recognition of MPSK and MFSK. A simple method based on high-order cumulants [5] is proposed for the classification of digital modulation schemes. It is proposed that ten different types of digital modulation can be identified by using the multilayer perceptron method [6] in combination with the spectrum and statistical accumulative characteristic sets of signals. In recent years, the modulation recognition algorithms are becoming ever more diverse. Random forests, decision trees, support vector machines, and other models are widely used [7, 8]. SVM occupies a dominant position in the field of machine learning due to its complete theoretical basis and good practical effects. Modulation recognizers based on the SVM model are emerging endlessly [9, 10]. The essence is to slightly improve the SVM model based on the extracted communication signal feature vector, such as improving the multiclassification strategy of SVM, improving the kernel function of SVM, and improving the penalty term of SVM [11, 12]. According to the different priori information obtained from the signal, the classical signal detection algorithm can be mainly divided into three categories: the signal detection algorithm based on matched filtering, the signal detection algorithm based on feature, and the signal detection algorithm based on energy detection. Matched filtering [13] algorithm is the best signal detection method in theory, but because it needs to obtain the priori information of the target signal in a comprehensive way, the application scene is strictly required in the actual application process. On the other hand, as the bandwidth of the target spectrum increases, when there are multiple signal forms in the detected frequency band, it is necessary to design a matched filter for each signal and compare it one by one, which is of great computational complexity. The feature-based signal detection algorithm requires less signal prior information and can obtain better detection performance under certain prior conditions. The method based on correlation detection is the most common one among feature detection algorithms. The main idea is to design correlation rules by using the inherent correlation characteristics of signals [14]. However, the performance of correlation detection depends heavily on the statistical characteristics of the signal to be detected, and the detection performance varies greatly with the difference of feature selection. The signal detection algorithm based on cyclic features [15] is also an algorithm with excellent performance and good antinoise performance. However, the cycle feature detection is not suitable for systems that require high real-time performance due to the large amount of calculation. In general, the feature-based signal detection method can obtain good detection accuracy in some cases, but the detection performance is also limited by prior information, and the performance is still limited under noncooperative conditions. Energy detection method [16, 17] takes the energy of the signal in the detection domain as the evaluation standard and can be divided into time domain detection and transformation domain detection according to different detection domains. Among them, the time domain detection method has poor antinoise ability and poor real-time performance, and its performance needs to be improved, and most of them are based on narrow-band detection tasks. In the transformation domain detection algorithm [18, 19], the power spectrum method is widely used in practical engineering due to its advantages of clear and simple physical meaning, low computational complexity, and direct application in the field of multisignal detection. However, due to the influence of complex electromagnetic environment and partial impedance mismatch of receiver analog front end, the noise base of the obtained broadband signal may have a large fluctuation, and this fluctuation phenomenon will lead to a serious decline in detection performance. For this kind of situation, the power spectrum detection algorithm based on the correction of noise base has been widely studied [20, 21]. Its basic idea and technical route are as follows: estimate the noise base, use the estimated base to correct the power spectrum, and finally build the detection model based on the modified Gaussian white noise. The signal detection performance of this method depends heavily on the accuracy of the estimation of the noise base, and the robustness is poor in complex electromagnetic environment. Moreover, the detection algorithm based on power spectrum is only applicable to stationary signals. For nonstationary signals, the signal cannot be comprehensively described on the power spectrum because the spectrum structure may change with time [22, 23]. Therefore, signal detection based on time-frequency domain method is also widely studied [24–29]. Due to the diversity of modulation methods, it is difficult to find a universal algorithm that can recognize all modulation methods so far. Different algorithms can be applied to different signal types. Therefore, in practice, it may be necessary to combine a variety of different algorithms to meet the expected requirements. The modulation recognition algorithms mainly fall into two categories: one is the likelihood ratio method based on hypothesis testing, which takes each modulation mode as a possible alternative hypothesis, and determines the modulation mode of the signal by constructing the likelihood function and obtaining the maximum likelihood; the second is the pattern recognition method based on feature extraction. The idea is to extract some feature information from the signal or extract some information and then carry out some operations to obtain some characteristic values, and through these information classifications, identify the modulation type. Past research communication signal modulation style intelligent recognition technology has a lot of research results; in some application scenarios, end-to-end algorithm is less than human characteristic modulation recognition algorithm based on artificial features, but the selection of model does not have definite theory support, and the degree of matching between input data and the model is much less than the reliability of the model which is often restricted to the adequacy of the data. The training of a large number of signal samples will inevitably lead to a large amount of time and resource consumption. The application of artificial features in the field of signal modulation recognition needs further research.

Based on deep neural network and artificial feature engineering, a communication signal modulation recognizer is constructed. In this paper, a concrete design scheme of neural network is proposed firstly, and then, a communication signal modulation recognitor is implemented on CPU and GPU, and the experimental performance is given. In addition, on the basis of studying the different features, a kind of quasirecognition signal set recognition scheme is designed.

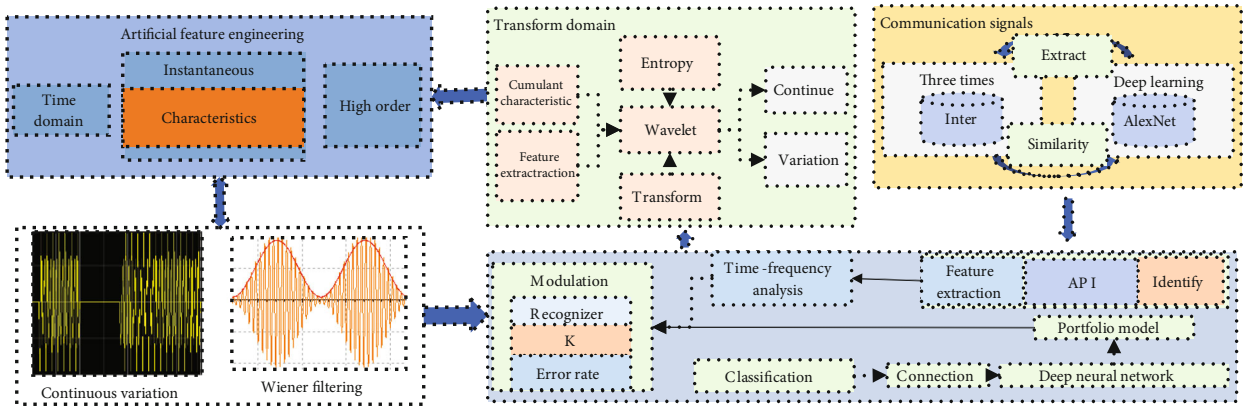


Figure 1: General framework of modulation recognition of artificial feature engineering deep neural network.

A statistical algorithm of spectral peak number for FS and K class recognition is proposed. In order to solve the problem of constellation rotation caused by phase mismatch of QAM signal during downconversion, an improvement was made. Finally, the whole recognition scheme is verified by simulation experiment.

2. Modulation Mode and Artificial Feature Engineering Deep Neural Network Modulation Recognition for Communication Signals

2.1. General Framework of Modulation Recognition for Artificial Feature Engineering Deep Neural Network. In order to realize the combination of artificial feature and deep neural network, the 20-dimensional artificial feature vector was constructed by designing and selecting the appropriate feature parameters, and the four-layer fully connected BP neural network was used to classify the artificial feature vector. Then, time-frequency analysis of the signals is carried out and appropriate preprocessing methods are selected. The modulation recognition algorithm based on deep learning is studied by using the AlexNet model and the Inception-ResNet-V2 model, respectively, in order to realize the end-to-end modulation recognition. Finally, based on the Indulge-ResNet-V2 model with better recognition effect, the existing model is improved by designing the combined model and the triple network model, and new signals outside the preset range are identified to expand the application range of the existing model, as shown in Figure 1.

In the construction of artificial feature vectors, extracting characteristic parameters with characterization ability from the received signals is the first step and also a key link in the realization of modulation recognition. In the recognition of MASK signal, MFSK signal, MPSK signal, and MQAM signal, three new transform domain features are designed, and on this basis, a 20-dimensional artificial feature vector is constructed, which contains six temporal instantaneous features, five high-order cumulant features, and nine transform domain features.

Different characteristic parameters can represent different signal modulation modes from different perspectives. The difference of transient characteristics in time domain provides the intensity and stability of signal temporal waveform richness and instantaneous amplitude distribution, while high-order cumulative features can suppress the influence of Gaussian white noise and enhance the recognition effect at low SNR. The characteristics of the transform domain provide the information of discrete spectral line distribution, normalized spectral density, and instantaneous amplitude spectral density of the signal. The mind mapping of all feature parameters in the artificial feature vector is shown in Figure 2.

MASK, MQAM, MPSK, and MFSK signals vary in amplitude to different degrees. The envelope of MPSK signal is slightly undulating, while the envelope of MFSK signal is not undulating. These signals are different in terms of time domain waveform richness, intensity of instantaneous amplitude distribution and stability of amplitude information, etc. Six time domain instantaneous features are adopted, and the specific extraction algorithm of each characteristic parameter is as follows:

(1) Using the zero center of the signal to normalize the instantaneous amplitude, the characteristic parameters are obtained

Firstly, the zero center normalized instantaneous amplitude of the signal is calculated by using the instantaneous amplitude of the signal.

$$\beta(n) = \frac{\beta(n)}{\sum_{n=1}^{N} \beta(n)} - 1. \tag{1}$$

Then, a digital low-pass filter is used to smooth the filter by removing some burrs and amplitude mutations, and then, the parameters can be obtained by calculating the mean value and variance. If the filtered number sequence is $\beta(n)$, then the calculation expression of characteristic parameters is as follows:

$$M_{\text{mean}} = \frac{1}{N \sum_{n} \beta(n)}, \tag{2}$$

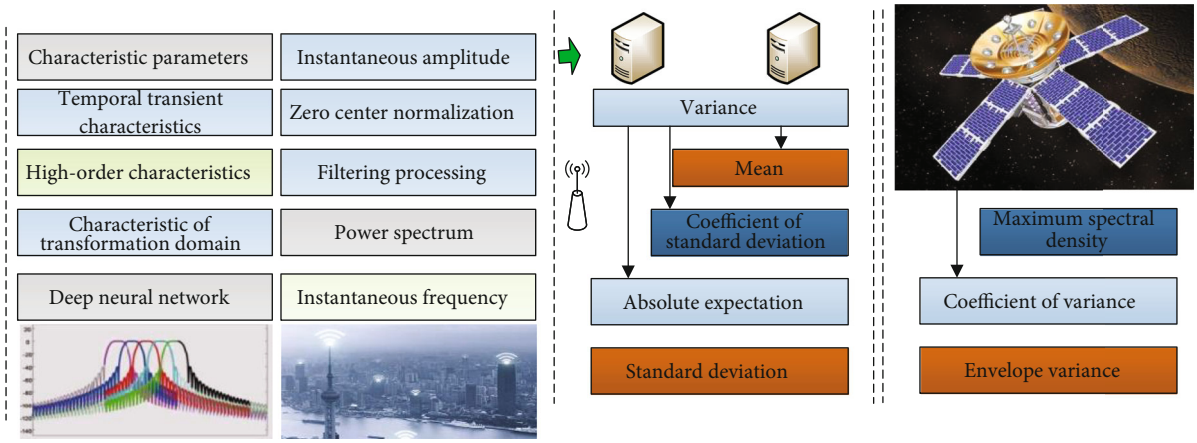


Figure 2: Modulation identification diagram of deep neural network for all feature parameters in artificial feature vectors.

$$M_2 = \frac{1}{M}\sum_{n}^{N}\left(\beta(m) - \overline{\beta(n)}\right)^2. \tag{3}$$

(2) Calculate the standard deviation coefficient and the expectation of absolute value of the two normalized instantaneous amplitudes of the zero center of the signal. The calculation expression is as follows

$$M_\beta = E\|\beta_N(\beta_n(m))\|. \tag{4}$$

(3) Calculate the standard deviation of the absolute value of the normalized instantaneous amplitude of the zero center of the differential signal, and obtain the characteristic parameters

$$ta_{ac} = \sqrt{\frac{1}{M}\left\|\sum_{N}\beta^2(n)\right\|}. \tag{5}$$

(4) The characteristic parameters can be obtained by calculating the variance coefficient of the envelope of the differential signal. Its calculation expression is as follows

$$R = \frac{\beta^2}{m^2}. \tag{6}$$

The analysis of signal in time domain and transformation domain only describes the same existence from different perspectives. The transformation of perspectives often brings new information. Some signals that are difficult to see the features in time domain will be easier to see the features after being converted to transformation domain. Therefore, the characteristic parameters of the transformation domain also play an important role in modulation identification and can provide information on discrete spectral line distribution, normalized spectral density, instantaneous amplitude spectral density, and other aspects of the signal.

2.2. Research on Recognition of Modulation Mode of Communication Signal. Many studies use high-order cumulants or clustering algorithms to identify QAM signals in class. For QAM signal, using clustering recognition is the most direct method, but the cluster computing complexity is much higher than the level of higher-order cumulant, and higher-order cumulants also have some disadvantages; FS cumulant characteristic value K signal exactly makes impossible using higher-order cumulant characteristic value of FSK signal recognition, so the combination of the two methods is in this paper. Since high-order cumulants have lower operational complexity, it is preferred to use high-order cumulant features for in-class recognition of QAM signals. When the recognition cannot be completed by using high-order cumulant features, clustering is carried out. The theoretical values of high-order cumulants of the four QAM signals of 8QAM, 16QAM, 32QAM, and 64QAM are shown in Table 1:

The simulation shows that 16QAM and 64QAM cannot be distinguished effectively even with eighth-order cumulants. Therefore, after the signal set {16QAM, 64QAM} is obtained by using Ki and K classifications, the clustering algorithm is used to separate the two. It is assumed that the clustering point is M, and the threshold value of the clustering point is $t(N1)$, which is determined to be 64QAM when $M > t(M)$; otherwise, it is 16QAM.

In the case of phase asynchronization during mixing, as the constellation diagram rotates, the range of constellation points sampled at regular intervals is smaller than that in the case of phase synchronization, so the radius of clustering should be reduced accordingly. However, as the radius of subtraction clustering is set as a fixed value, the recognition rate of signals is reduced.

As can be seen in Figure 3, when the phase is out of sync, the constellation diagram rotates, and the clustering radius value should have been smaller. However, since its value

Table 1: Theoretical value of higher-order cumulants of QAM signals.

Signal	8QAM	16QAM	32QAM	64QAM
IC201	0	0	0	0
IC211	0.668*E*	*E*	*E*	*E*
IC401	*E*	0.68*E*2	0.18*E*2	0.61*E*2
IC411	*E*2	0	0	0.31

has been fixed, for 64QAM signals, the clustering radius is relatively large, which leads to too few clustering points, thus leading to low recognition rate.

2.3. Modulation Pattern Recognition of Communication Signal Based on Modulation Recognizer of Deep Neural Network. With the improvement of computer processing speed and storage capacity, the design and implementation of CNN has gradually become a trend. Communication signal modulation pattern recognition method based on deep neural network is used to recognize modulation signal. Firstly, the received modulation signal is normalized and the time-frequency feature image is preprocessed to generate the training set and test set required for network training. Secondly, a classifier for communication signal modulation pattern recognition, namely, CNN, is designed and built. The training set is input into CNN for training, and the CNN network model is obtained. Finally, the identified modulation signal is preprocessed to generate a test set in the dataset, and the training set is input into the CNN network model to identify the modulation mode of the communication signal. This method takes the time-frequency domain graph as input and the signal modulation mode as output. The specific algorithm flow chart is shown in Figure 4.

Time-frequency analysis represents nonstationary signals as two-dimensional functions are related to time and frequency, which can be analyzed and processed intuitively. It is a kind of important method for processing nonstationary signals. CWD is selected for time-frequency analysis of communication signal modulation mode, and then, the grayscale feature image generation algorithm is used to generate the training set and test set required by the network. Training set and testing set include 2ASK, 2FSK, 2PSK, five kinds of AM and FM modulation signal time-frequency characteristics of the image, SNR range for [−SdB, SdB], interval of 1 dB, each signal-to-noise ratio under each modulation signal sample for 200, including 60 for the test set and 140 as the training set, the training set including a total of 7700 samples, and a test set containing a total of 3300 samples.

After the design and construction of CNN's network structure is completed, some network super parameters, such as learning rate and weight attenuation coefficient, need to be set. The setting of super parameters in this network is shown in Table 2.

After setting the super parameters, the following network model training can be carried out. During training, 30% of the training set is used as the validation set to verify the performance of the network. Through the training of the network, Figure 5 shows the loss value of the training set and the relationship between the accuracy and the number of iterations.

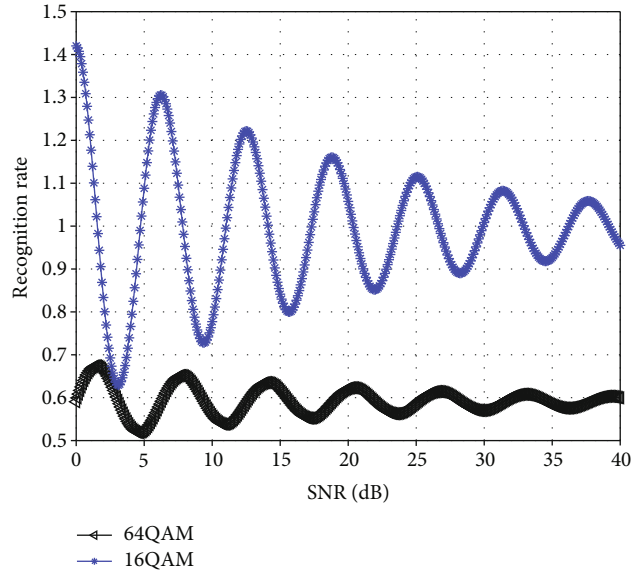


Figure 3: Recognition rate of subtraction clustering when the mixing phase difference is 45 degrees.

As can be seen from Figure 5, after 40 iterations of training, the accuracy of training set (ACC) is basically stable at about 0.99, while the loss value of training set is also stable at about 0.15. The accuracy of the verification set fluctuates around 0.85, while the loss value of the verification set fluctuates greatly and shows an oscillating trend.

The first step is data preparation: in the deep neural network modulation identifier, the input is directly the original communication signal sampling data, and there is no need for feature extraction and preprocessing, but the appropriate dimension transformation must be carried out to adapt to the deep network input. Note that our three deep networks are input dimensions that are all (None, 1024, 1), which is a three-dimensional tensor: the first dimension None represents the number of samples and does not need to be specified. The second dimension, 1024, means the length of our communication signal is 1024. Dimension 1 means there is only one channel. As for the label calibration of modulation mode, label unique thermal vectorization, and the division of training set and test set, it is completely consistent with the description in the implementation of modulation recognizer of deep neural network.

The second step is to build the model: based on the Keras and TensorFlow libraries, three kinds of deep neural networks are cooperatively built on the CPU and GPU computing platform according to the above network architecture description.

The third step is training model and test: with good training, dataset is divided into three deep neural network training, respectively, calculated separately on each training wheels on deep modulation recognizer network training set and testing set of accuracy and losses, after the network convergence, to test the three networks in different SNR test data accuracies.

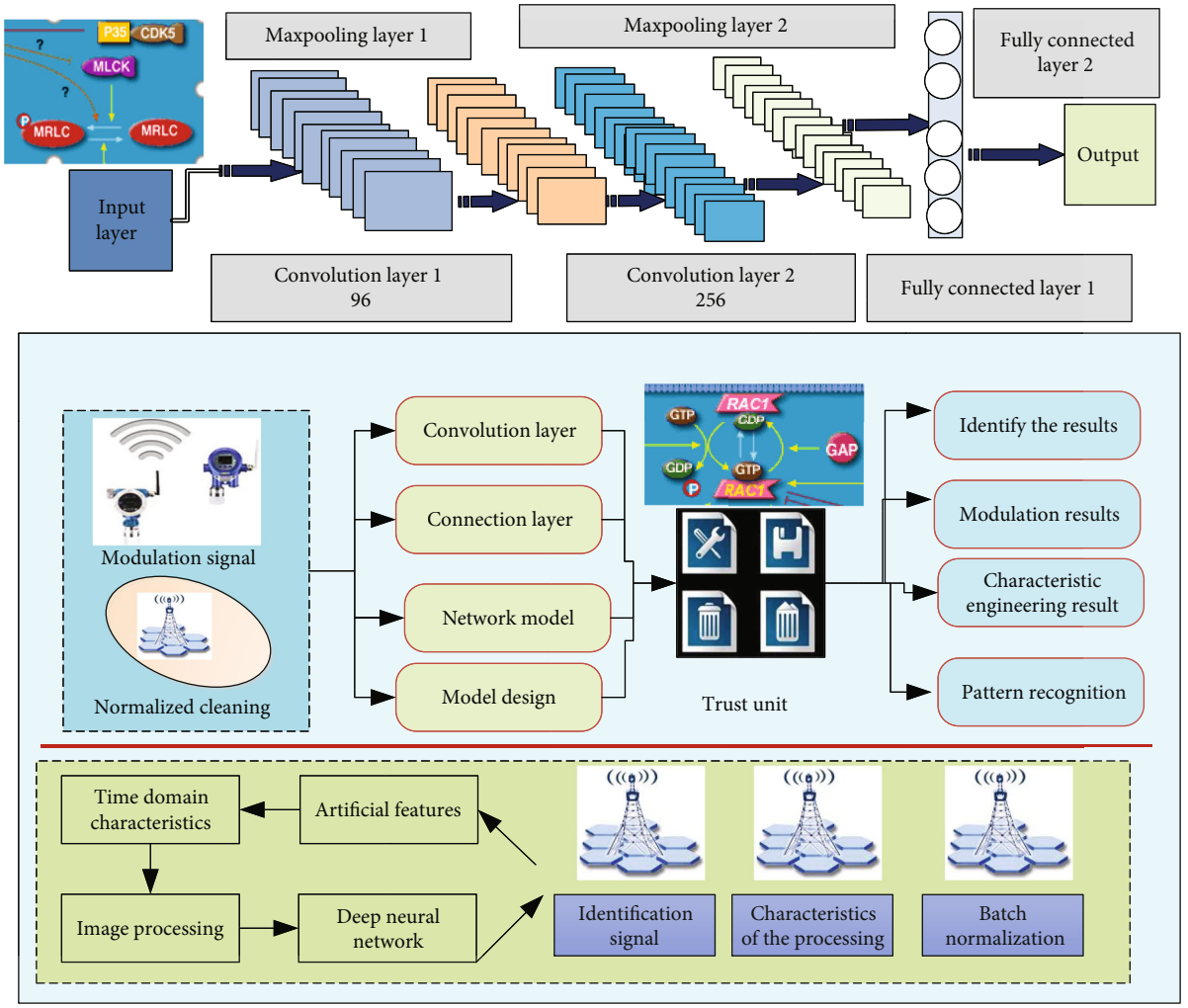


Figure 4: Flow chart of communication signal modulation pattern recognition based on deep neural network.

Table 2: Network super parameter settings.

Super parameter	Set the value
Vector	0.0001
Attenuation coefficient of weight	0.00006
Number of iterations	60
Batch size	129

The number of neurons in the input layer and output layer of the neural network can be clearly obtained by considering the characteristic vector dimension of the signal and the modulation mode types of the signal to be recognized. The signal features we used in the analysis included 5 spectral features and 12 cumulant features, so each signal could be described by a 17-dimensional feature vector, so we set the number of neurons in the input layer as 17. Since there are seven types of modulation signals we need to recognize, the number of neurons in the output layer is set as 7. For the setting of the number of hidden layer neurons, we might as well consider setting a little larger. Because when the greater the number of hidden layer neurons, the greater the capacity of the neural network model, it is more easy to fitting the training data; that is to say, the performance of the neural network on the training set will be better, but it also means that the model had the greater risk than the fitting; we can have a very good avoidance over fitting strategy that is regularization. We can not only apply L2 regularization to the weight of the network but also randomly inactivate the neurons in the two hidden layers. Based on this, we can set the number of neurons in the hidden layer as 40 and 32 and set the random inactivation probability of neurons in the hidden layer as $P = 0.6$. According to the above description, the concrete structure of our modulation recognizer for deep neural network can be shown in Table 3.

The first is the ReLU activation function, which is used for the activation of hidden layer neurons. ReLU activation function is not only simple to calculate but also can well alleviate the problem of gradient disappearance, which can greatly accelerate the convergence speed of neural network and improve the performance of neural network. Although the "dead zone" phenomenon exists in the ReLU function, it can be effectively alleviated by appropriate parameter initialization strategy.

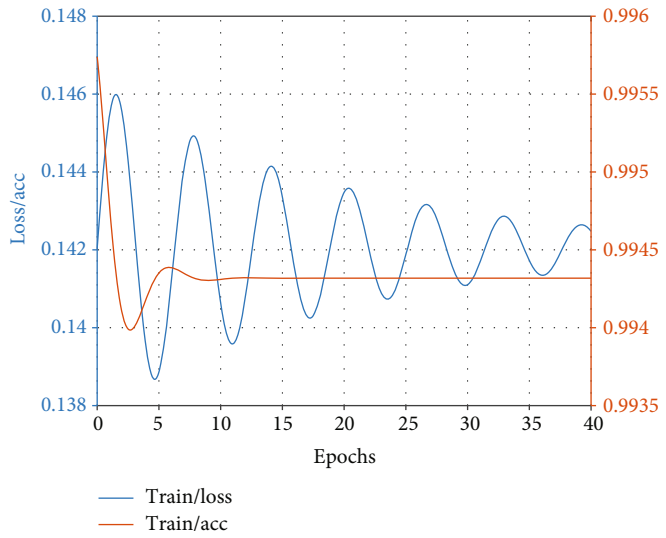


Figure 5: The relation between the loss value of validation set and the accuracy and the number of iterations.

Table 3: Specific structure of feedforward neural network.

Network layer	Input dimensions	Output dimensions	Number of parameters
Densel	(None, 18)	(None, 43)	725
Dropout (0.8)	(None, 41)	(None, 42)	2
Den set	(None, 41)	(None, 31)	1332
Dropout (0.7)	(None, 31)	(None, 31)	1
Dense3	(None, 30)	(None, 6)	234

Table 4: Signal modulation mode label and unique heat vector corresponding table.

Modulation mode label	Unique heat coding vector
0	(1, 0, 0, 0, 0, 0, 0)
1	(0, 0, 0, 1, 0, 0, 0)
2	(1, 0, 0, 0, 1, 0, 0)
3	(1, 0, 0, 0, 0, 1, 0)
4	(1, 1, 0, 0, 0, 0, 0)
5	(1, 0, 1, 0, 0, 0, 0)
6	(1, 0, 0, 0, 0, 0, 1)

The second is the softmax activation function, which acts on the output nerve layer neurons. Softmax function can map the output to the probability of the communication signal corresponding to each modulation mode, so that the output of the neural network can form a probability distribution vector, which can be used to train the neural network modulation recognizer for the classification task with the cross-entropy loss function.

As for the selection of the loss function, considering that the modulation pattern recognition of communication signals is essentially a multiclassification problem, the loss function can be explicitly designed as the cross-entropy loss function. Because the cross-entropy loss function is used as a measure of the difference between two probability distributions, the output layer using softmax activation function can ensure the output is a probability distribution vector dimensions for the output category number, so for real label samples, it must also be converted to a probability distribution of the same dimension vector to calculate cross office according to the loss. Hot codinq can achieve this goal. Therefore, we can consider the real taqs data of hot codinq and convert it into multiple types of multidimensional vector modulation modes. In this hot vector alone, the modulated signal corresponds to the model. One dimension of the corresponding vector is 1, and the other dimensions are 0. The specific corresponding relationship is shown in Table 4.

2.4. Collaborative Implementation of Deep Neural Network Communication Signal Modulation Identifier on CPU and GPU Platform. The first step is data preparation: in the deep neural network modulation identifier, its input is the feature vector extracted from the original communication signal sampling data, and the feature vector is preprocessed by the maximum and minimum normalization. The label calibration of the modulation method, the unique thermal vectorization of the label, and the division of the training set and the test set are analyzed.

The second step is to build the model: based on the Keras and TensorFlow libraries, this four-layer deep neural network is built on the computing platform of CPU and GPU according to the network structure described.

The third step is to train the model and test: the divided training dataset is used for training the neural network, and then, the test dataset is used to test the performance of the network.

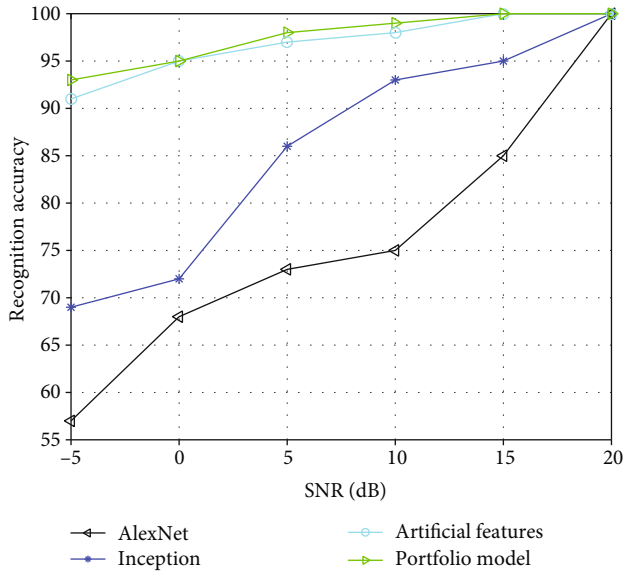


Figure 6: Recognition accuracy of different models for 18 kinds of simulation signals.

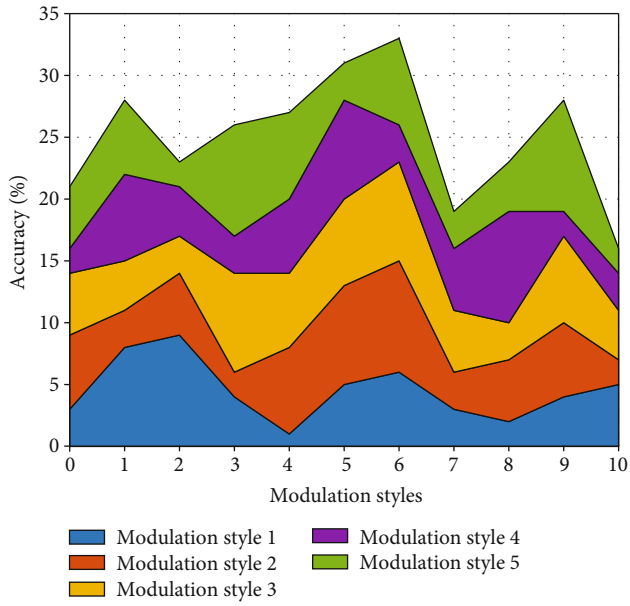


Figure 7: Confusion matrix of the combined model for 18 kinds of simulation signals at 11 dB.

2.5. Experimental Verification. In the feature extraction part, a signal sampling sequence consisting of 40960 sampling points is intercepted. After feature extraction, 20-dimensional feature vectors can be obtained. Each simulation signal ranges from 1 SdB to 20 dB. The features are extracted with different signal-to-noise ratios. It can be found that comparing the two single model and modulation recognition model based on ergonomic 20-dimensional features, the combined model has achieved better results. At -5 dB, Inception-ResNet significantly improves the recognition effect of the V2 model by 93% under the condition of low signal-to-noise ratio. And the recognition effect is slightly better than modulation recognition based on 20-dimensional human body characteristics. 1000 samples are generated every 2 dB, with a ratio of 3 : 1 : 1 By machine extraction, training set, verification set, and test set are added, respectively.

The combined model is based on the migrated Inception-ResNet-V2 model. Figure 6 compares the recognition accuracy of the combined model with the three existing models in the previous two chapters for 18 kinds of simulation signals under different signal-to-noise ratios. Can be found, and the deep study of two single model and the modulation recognition model based on 20-D characteristics of human engineering, combined model has achieved better effect, when -5 dB evidently improve recognition accuracy rate of 93% combined the Inception-ResNet- before the V2 model recognition effect under the condition of low SNR, and the recognition effect is slightly better than for modulation recognition based on 20-D human characteristics.

It can be seen that the addition of 8 artificial feature parameters enables the combined model to provide a more effective feature expression for the signal, and it is effective to use the model to assemble the features online in the modulation identification. Combined model is compatible with the artificial features and advantages; based on time-frequency diagram of the automatic feature in the case of low signal-to-noise ratio (SNR), artificial features play an important role, and when the signal-to-noise ratio is not less than 7 dB, using the migrated Inception-ResNet-V2 model can be the signal of time-frequency diagram that obtained the recognition accuracy higher than 95%; accordingly, sectional model has achieved above 98% accuracy.

Table 5: Recognition rate of offline test.

Modulation method	Recognition rate (SNR = 16 dB)	Recognition rate (SNR = 15 dB)
4ASK	97.4%	99.7%
2FSK	94.3%	98.3%
4FSK	91.5%	98.5%
8FSK	89.1%	93.2%
8PSK	95.7%	98.8%
8QAM	92.2%	96.3%
16QAM	6.9%	97.1%
32QAM	94.5%	98.2%
64QAM	91.3%	95.6%
AM	91.6%	95.3%

Figure 7 shows the confusion matrix of the combined model for 11 kinds of simulation signals at 11 dB, where labels 0-10 represent signals with different modulation styles. It can be seen that the combined model can realize the accurate recognition of the 11 simulation signals at 11 dB.

The data of 4ASK, 2FSK, 4FSK, 8FSK, 8PSK, 8QAM, 16QAM, 32QAM, and 64QAM signals are generated by MATLAB. The generation method is as follows: 258 random baseband codes are generated, the codes are formed by the rising cosine filter, the roll drop coefficient is 0.6, the corresponding modulation is carried out according to the selected modulation type, and the carrier frequency is 100 MHz. After

Table 6: Test recognition rate of 16 kinds of communication signals.

Modulation method	Test way	Signal source	Recognition rate
AM	Offline test	MATLAB to produce	95.3%
FM	Online test	FM radio	92.6%
1JSB	Online test	Signal generator	98.6%
LSB	Online test	Signal generator	99.1%
2ASK	Online test	Signal generator	96.7%
4ASK	Offline test	MATLAB to produce	99.3%
2PSK	Online test	Signal generator	98.4%
4PSK	Online test	Signal generator	92.2%
8PSK	Offline test	MATLAB to produce	98.5%
2FSK	Offline test	MATLAB to produce	98.1%
4FSK	Offline test	MATLAB to produce	98.4%
8FSK	Offline test	MATLAB to produce	93.7%
8QAM	Offline test	MATLAB to produce	96.3%
16QAM	Offline test	MATLAB to produce	97.3%
32QAM	Offline test	MATLAB to produce	98.2%
64QAM	Offline test	MATLAB to produce	95.7%

AM signal is generated, a low-pass filter with a cut-off frequency of 8 kHz is used for filtering. After filtering, 100 MHz carrier is used for modulation, and the depth of modulation is random value in the range of [0.4,1]. After modulation, noise is superimposed, and then, orthogonal down conversion and low-pass filtering are carried out. The filter bandwidth is 300 kHz, and the filtered IQ data is stored as a file.

The code that calls modulation recognition reads the IQ data file and performs modulation recognition, using 8195 data points for each identification. If it is identified as QAM signal, the entire IQ data file is read for identification. The recognition rate of each signal is shown in Table 5:

Test methods and test results of 16 signals are shown in Table 6. For the modulation method that uses signal generator to generate signals for testing, the recognition rate is measured when the output power of signal source is -82 dBm; for the modulation method that uses offline testing, the signal-to-noise ratio of the recognition rate is 15 dB.

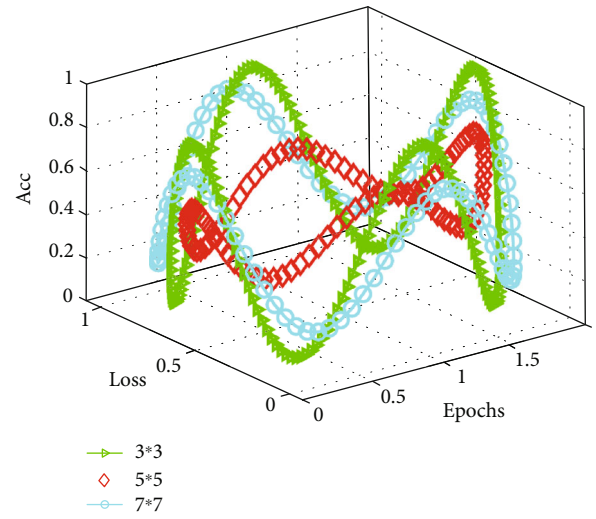


Figure 8: Relation between the accuracy of verification set and the number of training iterations for different deep cores.

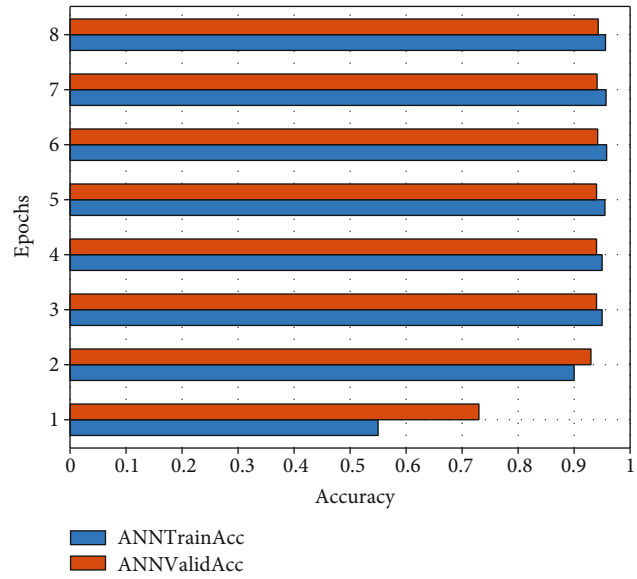


Figure 9: Accuracy performance curve of deep neural network on training set and verification set.

Deep nucleus size is usually odd, commonly $3 * 3$, $5 * 5$, etc., but there is no theoretical basis. The size of the deep core is usually determined by a large number of practical tests. And then there is a special deep core of 1×1, and deep cores of that size are usually used for dimension-reduction or linear transformations. The setting of the network hyperparameter is the same as that of the network hyperparameter in 4.3.2. After network training, network models with different deep core sizes were used to verify the accuracy of the set in the training process to analyze the influence of different deep core sizes on network performance. When the deep core size is different, the accuracy of the verification set changes with the increase of the number of training iterations, as shown in Figure 8:

As can be seen from Figure 8, when the size of deep core is $3 * 3$, the accuracy of verification set is the highest, and with the increase of the number of training iterations, the fluctuation range of the accuracy of verification set is the smallest, and $3 * 3$ deep core will be used for deep calculation.

The training process of the deep neural network communication signal modulation recognitor is shown in Figure 9. The curve in Figure 9 represents the accuracy curve of the neural network on the training set and the verification set.

With the increase of the number of network training rounds, the accuracy of the deep neural network in the training set and the test set gradually improved and tended to be stable, and the modulation recognizer in the training set and the test set reached a high level of accuracy, both of which exceeded 98%. This means that our neural network has been fully and effectively trained and no fitting phenomenon has occurred.

Through experimental verification, the online splicing features take into account the dual advantages of time-frequency analysis and artificial features, and the combined model design with better recognition effect is completed. Then, after selecting an appropriate feature extraction subnetwork, a triplet network model is built, and the similarity of the input time-frequency graph is calculated by using the Euclidean distance between the feature vectors, and a reasonable sample decision method is designed, which not only improves the matching degree between the input time-frequency graph samples and the Inception-ResNet-V2 model but also improves the matching degree between the input time-frequency graph samples and the Inception-ResNet-V2 model. It also realizes the recognition of the newly modulated signal and expands the application range of the existing model. Finally, the algorithm test and performance analysis based on the above two models are completed by using the real signal and simulation signal under different SNR.

3. Conclusion

Based on Keras and TensorFlow libraries, the deep neural network communication signal modulation recognizer is implemented on CPU and GPU hardware platform, and the training and testing tasks are completed. Experiments show that the deep neural network has good performance on both the training set and the test set, and can accurately identify the modulation mode of the communication signal with low SNR, which shows the robustness of the neural network. The deep neural network in this chapter had better reached the accuracy level of 97%, and we can see that the modulation recognizer based on deep neural network can achieve a better accuracy level. Modulation recognizers for communication signals are constructed based on two kinds of mainstream deep neural network structures: three kinds of modulation recognizers are constructed based on deep neural network architecture, and their performance differences are compared; based on the cyclic neural network architecture, we construct two modulation identifiers and compare their performance differences. The experimental results show that the modulation recognizer based on deep neural network has a good accuracy performance on the simulation dataset. For the modulation signal identification problem adopted in this paper, we try to collect more signal modulation modes to research more signal characteristics, so that the algorithm can identify more communication signal modulation modes.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

We declare that there is no conflict of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China: the research on ferromagnetic resonance temperature imaging based on superparamagnetic nanoparticles (Grant 61773018), research on low complexity coding model and method based on 3D-HEVC (Grant 61771432), research on information acquisition and processing method of temperature field in superparamagnetic nanoparticle targeted (Grant 61374014), and a real-time and accurate method for measuring 2D temperature distribution in magnetic nanoparticle-mediated hyperthermia (Grant 61803346).

References

[1] M. Richterova, "Signal modulation recognizer based on method of artificial neural networks," *PIERS Online*, vol. 1, no. 5, pp. 575–578, 2015.

[2] S. Wu, "Communication modulation recognition algorithm based on STFT mechanism in combination with unsupervised feature-learning network," *Peer-to-Peer Networking and Applications*, vol. 12, no. 3, pp. 1615–1623, 2019.

[3] Q. Guan and Y. Zhang, "Cyclic cumulant based on communication signal multilayer neural network modulation pattern recognition," *IOP Conference Series Materials Science and Engineering*, vol. 677, pp. 42113–42125, 2019.

[4] S. Peng, H. Jiang, and H. Wang, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 6, no. 3, pp. 28506–28523, 2018.

[5] X. Bi, "A method for modulation recognition of maritime search and rescue communication signal based on neural network," *Journal of Coastal Research*, vol. 103, no. sp1, pp. 847–862, 2020.

[6] P. Wu, B. Sun, and S. Su, "Automatic modulation classification based on deep learning for software-defined radio," *Mathematical Problems in Engineering*, vol. 2020, Article ID 2678310, 13 pages, 2020.

[7] Z. Y. Li and Y. D. Zhang, "DenseNet-ResNet-LSTM model for modulation recognition of communication signal," *Journal of*

Physics: Conference Series, vol. 1693, no. 1, pp. 12150–12165, 2020.

[8] H. Hosseinzadeh, F. Razzazi, and A. Haghbin, "A self training approach to automatic modulation classification based on semi-supervised online passive aggressive algorithm," *Wireless Personal Communications*, vol. 82, no. 3, pp. 1303–1319, 2015.

[9] R. M. Al-Makhlasawy, A. A. Hefnawy, and E. MMA, "Modulation classification in the presence of adjacent channel interference using convolutional neural networks," *International Journal of Communication Systems*, vol. 33, no. 13, pp. 4295–4306, 2020.

[10] K. Li, W. Pan, and Y. Li, "A method to detect sleep apnea based on deep neural network and hidden Markov model using single-lead ECG signal," *Neurocomputing*, vol. 294, no. 7, pp. 94–101, 2018.

[11] L. Sun, J. Chen, and K. Xie, "Deep and shallow features fusion based on deep convolutional neural network for speech emotion recognition," *International Journal of Speech Technology*, vol. 21, no. 4, pp. 931–940, 2018.

[12] A. Lozano-Diez, R. Zazo, D. T. Toledano, and J. Gonzalez-Rodriguez, "An analysis of the influence of deep neural network (DNN) topology in bottleneck feature based language recognition," *PLoS ONE*, vol. 12, no. 8, pp. e0182580–e0182592, 2017.

[13] W. Shirui, Z. Yuelun, and L. Shubin, "Performance of deep neural network-based artificial intelligence method in diabetic retinopathy screening: a systematic review and meta-analysis of diagnostic test accuracy," *European Journal of Endocrinology*, vol. 2020, no. 183, pp. 41–49, 2020.

[14] Y. Yang, M. Chen, and Y. Wang, "Digital signal modulation classification using data conversion method based on convolutional neural network," *Journal of Physics: Conference Series*, vol. 1693, no. 1, pp. 12039–12050, 2020.

[15] R. Arunkumar and P. Karthigaikumar, "Multi-retinal disease classification by reduced deep learning features," *Neural Computing & Applications*, vol. 28, no. 2, pp. 329–334, 2015.

[16] N. Teimouri, M. Omid, K. Mollazade, and A. Rajabipour, "An artificial neural network-based method to identify five classes of almond according to visual features," *Journal of Food Process Engineering*, vol. 39, no. 6, pp. 625–635, 2016.

[17] Y. Qian, N. Chen, and K. Yu, "Deep features for automatic spoofing detection," *Speech Communication*, vol. 85, pp. 43–52, 2016.

[18] J. Li and G. Jian, "A new feature extraction algorithm based on entropy cloud characteristics of communication signals," *Mathematical Problems in Engineering*, vol. 2015, Article ID 891731, 8 pages, 2015.

[19] Y. Yao, L. Yu, and Y. Chen, "Feature extraction method of radiation source in deep learning based on square integral bispectrum," *Journal of Physics: Conference Series*, vol. 1678, no. 1, pp. 12074–12080, 2020.

[20] Y. He, Y. Zhu, and W. Lin, "HTTP tunnel Trojan detection model based on deep learning," *Journal of Physics: Conference Series*, vol. 1187, no. 4, pp. 42055–42066, 2019.

[21] Y. Huang, W. Jin, and Z. Yu, "Radar emitter signal recognition based on deep learning and ensemble learning," *Systems Engineering and Electronics*, vol. 40, no. 11, pp. 2420–2425, 2018.

[22] M. Li, W. Chen, and Z. Tao, "A novel seizure diagnostic model based on kernel density estimation and least squares support vector machine," *Biomedical Signal Processing and Control*, vol. 41, no. 3, pp. 233–241, 2018.

[23] M. Sawalkar and V. V. Yerigeri, "Classification of fruit family based on features extraction using Penn classification," *IOSR Journal of Electronics and Communication Engineering*, vol. 13, no. 6, pp. 48–54, 2018.

[24] P. Heum, "Extraction of major features for major adverse cardiac events from KAMIR dataset based on deep learning algorithm," *International Conference on Future Information & Communication Engineering*, vol. 9, no. 1, pp. 69–72, 2017.

[25] J. K. Kim, B. D. Kim, and D. W. Yoon, "Deep neural network-based automatic modulation classification technique," *The Journal of Korean Institute of Information Technology*, vol. 14, no. 12, pp. 107–123, 2016.

[26] J. Han, N. Lin, J. Ruan, X. Wang, W. Wei, and H. Lu, "A model for joint planning of production and distribution of fresh produce in agricultural Internet of Things," *IEEE Internet of Things Journal*, 2020.

[27] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the Internet of Things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

[28] A. Zielonka, A. Sikora, M. Woźniak, W. Wei, Q. Ke, and Z. Bai, "Intelligent Internet of Things system for smart home optimal convection," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4308–4317, 2021.

[29] S. H. Ahmed, V. H. C. de Albuquerque, and W. Wei, "Guest editorial: special section on advanced deep learning algorithms for industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2764–2766, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9804643, 1 page
https://doi.org/10.1155/2023/9804643

Retraction

Retracted: Marketing Method and System Optimization Based on the Financial Blockchain of the Internet of Things

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] C. Yan, J. Zhu, Y. Ouyang, and X. Zeng, "Marketing Method and System Optimization Based on the Financial Blockchain of the Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 9354569, 11 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9354569, 11 pages
https://doi.org/10.1155/2021/9354569

Research Article

Marketing Method and System Optimization Based on the Financial Blockchain of the Internet of Things

Chaosong Yan,[1] **Jun Zhu**,[2] **Yinglong Ouyang**,[3] **and Xingyu Zeng**[4]

[1]*Olin Business School, Washington University in St.Louis, St. Louis, Missouri 63130, USA*
[2]*Jiangsu Jari Information Technology Co., Ltd., Lanyungang, Jangsu 222000, China*
[3]*China Minsheng Banking Corp., Ltd., Beijing 100031, China*
[4]*School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing 100876, China*

Correspondence should be addressed to Chaosong Yan; chaosong.y@wustl.edu

Received 30 April 2021; Accepted 25 May 2021; Published 9 June 2021

Academic Editor: Wei Wang

This article makes relevant research and analysis from theory and practice, respectively. At the same time, with reference to the current state of the Internet, relative analysis was used to focus on the state of commercial banks. The analysis mainly focuses on the problems encountered in the current Internet development in various forms such as the bank's sales status and customer products. In addition, it really made a reasonable opinion about the relevant sales status of this bank. It systematically studied the development history of marketing products of financial blockchain in China's financial blockchain and analyzed the process of gradual improvement of the functions and characteristics of marketing products of financial blockchain in China. The characteristics and shortcomings of the marketing products of the financial blockchain were discussed in detail, and from the perspective of technological innovation, the application of emerging technologies in the marketing of China's financial blockchain was analyzed. This article analyzes the many challenges and opportunities faced by security firms in the development of the Internet. Under the new situation, what kind of development model to adopt and how to transform and upgrade is the strategic proposition that security firms must think about. Based on Internet finance, this article finds that the equity crowdfunding model can be selected in the future development of security companies. On the one hand, the use of equity crowdfunding by security firms can enhance their direct financing capabilities, which will benefit the security industry, especially Internet security firms. On the other hand, through equity crowdfunding, security companies can expand direct financing channels for small, medium, and micro enterprises to promote entrepreneurship. The paper promotes the development of Internet finance, thereby improving the ability of the capital market to serve the real economy.

1. Introduction

In the course of these years of development, the development of the Internet has gradually accelerated, and it has played a significant role in improving the pattern of microfinance and other aspects [1]. In the later period, some Internet companies gradually launched some lending platforms or used some new models of crowdfunding to integrate financial and other aspects of funds. Therefore, we can see that the overall development status of the Internet has had a great impact on the financial industry [2]. For the development of modern banks, how to solve the current problems encountered by banks in the establishment of banks and some sales methods is the most concerned issue of many researchers [3]. In this article, the emphasis is on the analysis of Bank A, from both theoretical and practical perspectives, to realize reasonable research on the various problems that this bank has produced. At the same time, some more effective marketing strategies and methods have been proposed for the development of this bank. Blockchain provides a decentralized trust establishment model and provides new development concepts and development opportunities for the future of

finance, Internet of Things, and other fields. With regard to these opinions, it is hoped that the bank can have a more stable development direction during the development process.

The overall development method and speed of Internet finance are very advantageous. Therefore, for the previous commercial banks, many related businesses gradually began to lose their advantages [4]. Commercial banks must make reasonable innovations so that future developments will not be replaced. Whether it is from its own competitiveness or from some emerging products, it must have its own unique form. In the current Internet finance, the third-party payment method has become a relatively common payment method that people use now. The ways that many companies use to raise funds have also changed significantly [5]. It can be seen that Internet finance has made some previous banking services no longer available for special use by banks. There have also been significant changes in the overall profitability of the banks. In this article, the focus is on the analysis of the various forms of influence that Internet finance has commercial banks. Although the application prospect of blockchain is very wide, its theoretical research is not perfect, and there are still some security problems that need to be solved urgently. Whether an effective solution can be proposed will directly affect the implementation and development of blockchain applications. On the basis of research, it is hoped that it can bring better creativity and development capabilities to commercial banks. During the operation of a bank, it must also be able to improve its overall capabilities from the perspective of services and products. With the continuous rise and development of new technologies such as cloud computing and the Internet of Things, market competition in the IT industry has become increasingly fierce [6]. Only by accurately grasping the market demand points, providing products that meet customer needs, and precision marketing, IT companies can achieve rapid development.

From the perspective of customer needs, it is guided by marketing theories such as 4Ps and 4Cs in marketing theory. For using literature research methods, empirical research methods, questionnaire survey methods, and fully combining the actual situation of the L Group, it is aimed at the marketing management of the corporate call center. In terms of problems such as scattered customer management, imperfect sales mechanism, asymmetry in distribution information, single promotion method, and insufficient assessment and incentives, they research the marketing management strategy of its call center. Through research, it is found that the steady growth of the IT industry has intensified industry competition; as an important link to connect with customers, call centers play a huge market role in both telesales and customer service; call centers must establish a set of appropriate marketing management strategy combination; through customer management, business opportunity management, network marketing, assessment and incentives, and other measures, they improve call center performance and management, so as to better support the company's performance development and expansion. This article proposes a KDM secure Bitcoin encrypted wallet solution, which solves the problem of capital loss caused by local file loss by introducing cloud storage to back up encrypted files. The research in this article proposes a set of executable solutions for the development of marketing management strategies of the L Group's call center, which solves the current problems of the L Group's call center, promotes the improvement of the management of the L Group's call center, and thus plays a greater role for the L Group. This article also provides reference for the marketing management of the call center for IT-type peer companies or companies (B2B type) that require transactions.

2. Related Work

According to the combining of more than 1,300 financial blockchain marketing startups or InsurTech companies, Jesus et al. [7] found that the marketing innovation of financial blockchain is divided into four major types, namely, product innovation and financial blockchain marketing innovation, enterprise operation innovation, marketing data, and intelligent platform application of financial blockchain. Yu et al. [8] suggested that the marketing-related service industry of financial blockchain should look for subdivision opportunities in the marketing industry chain of financial blockchain and use technological innovation to solve the inefficiency dilemma in the existing business model and break the original. For the current monopoly of traffic, channels, pricing, etc., we can use information technology to open up the innovation outlets in the marketing industry chain of the next financial blockchain. Maiti and Ghosh [9] believe that with regard to the changes in the marketing industry of financial blockchains, financial blockchain marketing companies should focus on the following aspects: first, when customers need it, provide customers with seamless digital channels, valuable insights, or sales; second, use big data to price risks and form insights; third, develop strategic cooperation with a new generation of digital intermediaries or commercial entities that represent customer needs; fourth, offset premiums through scale effects and operational efficiency loss or create new value.

In terms of the application of big data, Singh and Singh [10] divided the development stage of the marketing industry big data application of the financial blockchain into three stages: internal circulation, extension expansion, and national application. At the same time, through the marketing of financial blockchain research on the status quo of data application, which puts forward the potential breakthroughs in the application of big data in the marketing industry of financial blockchain, they are expanding the scope of underwriting, achieving personalized pricing, optimizing underwriting claims, improving antifraud performance, improving operational efficiency, and helping risk monitoring. Iftekhar et al. [11] believe that in the era of big data, by collecting information such as social networking platforms, data mining is performed on consumers, and people's behavior patterns and consumption habits can be analyzed, so as to provide different consumers with personalized and customized services and carry out precise risk control. Therefore, individualization and customization are the goals to be achieved by the marketing development of the financial blockchain in the future. In terms of blockchain,

Viriyasitavat et al. [12] in the marketing blockchain project team of financial blockchain believe that the marketing of financial blockchain and blockchain have genetic similarities and connections, and sociality is the common attribute of the two. Realizing the collection and collaboration of individuals is a common appeal, and rebuilding trust is the core value. Therefore, the marketing of financial blockchain is a typical scenario for blockchain applications. The blockchain application of financial blockchain marketing should not only focus on talent training but also focus on the integration and innovation of business and technology and focus on platform construction; it should not only focus on the construction of corporate and industry capabilities but also focus on external relations; particularly, it is the integration of blockchain technology [13]. Through the mosaic analysis of the marketing business of blockchain technology and financial blockchain, it is believed that the current research and development or alliance of blockchain technology by financial blockchain marketing companies in the country are still in the initial stage of concept or experimentation of the existing industrial applications. It is still possible to be replaced or improved in the process of future technology changes. The key to the adoption of blockchain technology in the future depends on solving the pain points of industry development and then achieving the promotion of the marketing industry value of financial blockchain [14]. The application of blockchain in the marketing industry of financial blockchain needs to break through several technical bottlenecks, including network scalability to be strengthened, security still to be improved, and industry standards and systems still being explored [15]. In terms of artificial intelligence, the application of artificial intelligence technology in the marketing industry of the country's financial blockchain is prospected [16]. It believes that with the development of genetic testing technology, intelligent testing technology may change the underwriting model and pricing model of financial blockchain marketing companies. Financial blockchain marketing companies should prepare in advance, fully understand, and adopt new reasonable actuarial methods, new systems, etc., to avoid risks and serve customers [17–19]. At the same time, intelligent control is the product of the combination of control theory, artificial intelligence, and computer science. In the future, intelligent control and automatic planning technology will also be more mature, and the application will be more extensive [20, 21]. In terms of the Internet of Things, the "marketing of the financial blockchain of the Internet of Things" is defined as the marketing company of the financial blockchain with the Internet of Things technology as the core to provide customers with smarter financial blockchain marketing products and claim services [19]. At the same time, it takes property insurance as an example to analyze the necessity and feasibility of the marketing of the financial blockchain of the Internet of Things [22–26]. Necessities include that the Internet of Things can help solve the current problem of unfairness caused by the static marketing and pricing of the financial blockchain, the inaccurate settlement of claims in the marketing industry of the financial blockchain, and the huge demand for the marketing of the country's financial blockchain. The feasibility mainly includes the advantages of the Internet of Things itself in terms of technology, timeliness, relevance, and accuracy, as well as relevant policy support given by the state [27]. The innovative significance of the marketing products of the Internet of Things financial blockchain is summarized: from the customer's point of view, its premium collection is more scientific, the efficiency of claim settlement is improved, and the probability of risk occurrence is reduced. From the perspective of the marketing company of the financial blockchain, the marketing products of the Internet of Things financial blockchain innovate profit models and create new benefits, innovate risk control models and reduce costs, improve customer service levels, and increase insurance renewal rates [28–30].

3. Construction of a Marketing Model Based on the Financial Blockchain of the Internet of Things

3.1. Hierarchical Distribution of Internet of Things Finance. The 4P theory was born in the United States in the 1960s, and it emerged with the introduction of the marketing mix theory. 4P, respectively, corresponds to Product, Price, Place location or channel, Promotion. This theory believes that if a marketing mix includes the right products, the right price, the right distribution strategy, and the right promotion strategy, then it will be a successful marketing mix, and the company's marketing goals can also be achieved by this. In the application of the Internet of Vehicles based on the blockchain, users are most concerned about the security of private data, including identity privacy and content privacy in the communication process. Marketing managers regard 4P mobile banking and marketing-related basic theory overview theory as a classic of marketing theory, which is mainly due to the conciseness and wide application scope of the theory.

$$u(x) = \int \left(at^2 - bt - c\right) dt,$$
$$v(x) = \int \left(at^2 + bt - c\right) dt. \tag{1}$$

Target market analysis refers to segmenting the market according to a specific standard, taking one or more of the segmented markets as the target market, and positioning its products based on this standard. Choosing the target market is after segmenting the market, in order to provide consumers with better services and products; the company chooses the corresponding market as the target. It is not that the target market must be selected because the submarket has advantages. Rather, because enterprises cannot cover the entire market due to restrictions on human resources and capital in the development process, submarkets are needed to play a role, and corresponding strategies can be formulated according to the requirements of consumers to pursue economic benefits.

$$f(x) = \begin{cases} \dfrac{v(x) - u(x)}{v(x) + u(x)}, \\ \dfrac{ax + b}{ax - b}. \end{cases} \quad (2)$$

Traditional marketing methods are mainly used for manual sales via telephone or agents. Such methods are costly, low in efficiency, and poorly targeted. In particular, telemarketing is often used by customers as harassing calls and triggers complaints. In terms of marketing, traditional models are still adopted, such as magazines, newspapers, soft articles, positions, EDM, and Baidu keyword placement. These methods have high input and low output, and they are still in their infancy in new media marketing.

$$G(n) = \{g(i)\}, \quad i = 1, 2, \cdots, n. \quad (3)$$

PEST refers to Political, Economic, Social, and Technological, and it analyzes the macro environment of enterprises based on this model. In the macro environment, how to operate and develop enterprises needs to learn from this model. The external environment is very important to the development of an enterprise, and it promotes or restricts the development of the enterprise, but the enterprise cannot control the external environment.

$$y(s, t) = f(s, t) \cdot \max\left(ax^2 + bx + c, s, t\right). \quad (4)$$

In the early 1980s, the concept of SWOT was first proposed by Werick, a professor of management at the University of San Francisco in the United States. SWOT refers to Strength, Weakness, Opportunity, and Threat. The purpose is to find and determine the competitive advantages, developmental disadvantages, opportunities, and threats that the company has. The scientific analysis method organically combines the company's strategic deployment, the company's internal resources, and the external development environment.

$$p = \frac{\sum f(x(i)) \cdot x(i) + \sum a \cdot y(s, t)}{\sum f(x(i)) \cdot x(i) + \sum a \cdot y(s, t) - c}. \quad (5)$$

3.2. Blockchain Marketing Algorithm. Compared with traditional consumer groups in the new era of the Internet, consumer groups have the three major attributes of mobility, localization, and socialization. In the consumption process, they also present three major characteristics of all-weather, personalized, and multichannel. The corporate marketing system in the new era is to establish a consumer-centric C2B system, supported by a big data platform.

$$q[x] = \frac{1 + (f(x) + n)^{-p}}{1 - (f(x) + n)^{-p}}. \quad (6)$$

(1) In the Internet era, companies have been forced to reconstruct the entire operating process from the perspective of customers. First, they must truly understand customer needs and adjust existing marketing models from the perspective of customers. A major advantage of the WeChat platform is that it can quickly collect and sort out customer information, acquire and create user needs, so as to formulate product development and service strategies, and conduct precise marketing for target customer groups. (2) When formulating marketing strategies, it is necessary to consider delivering the various values of the product (values are generated by functions, features, brands, etc.) to customers. WeChat is an important platform for communicating with users. It can directly transmit product features, marketing activities, brand value, and other information to customers through message push.

$$s \cdot (p \cdot f(m) + q \cdot y) \bmod m = k \cdot p. \quad (7)$$

(3) With the rise of technology, especially the emergence of social networks, the direct dialogue between enterprises and consumers is no longer confused, but point-to-point static communication mechanism gradually evolved into a many-to-many, three-dimensional dynamic communication mechanism. The WeChat platform can provide consumers with $7 * 24$ hours of uninterrupted communication channels. At the same time, combined with the o2o mode, it forms a closed-loop online and offline, integrates multichannel communication, collects customer feedback, and achieves real-time response and comprehensive coverage. The KDM secure symmetric encryption algorithm is used for encryption, which avoids the problem of key entropy leakage and improves the security of encrypted wallets. The data collected from customers has become the basis for future product and service improvements. (4) The extensive use of data, social tools, and LBS technology makes people's "life trajectory" recorded on the Internet, forming the so-called "big data" to locate customer portraits, analyze consumer behavior characteristics, and conduct risks. Prevention and control provide a data basis. The biggest feature of WeChat marketing is that it can quickly acquire a large number of customers and achieve the accumulation of large amounts of data, so as to achieve precise positioning and personalized marketing.

P theory: product (Product), price (Price), channel (Place), and promotion (Promotion). When an enterprise conducts marketing planning in the market, it is not only necessary to fully understand and analyze the external environment in which the enterprise is located but also necessary to establish a set of marketing strategy portfolios by analyzing the market from Figure 1. With this tool, the enterprise can effectively obtain from the market. Only by executing the marketing strategy plan, conforming to the current environment, and satisfying the market demand, the enterprise tasks can be achieved. On the basis of the original 4P combination theory, two factors P are added, namely, Power and Public Relations, and a modern big marketing strategy, referred to as 6Ps, is proposed. 6Ps can better target the market with trade barriers that existed at that time. Subsequently, in addition to the 6Ps of big marketing, several factors including Probe, Partition, Precedence, Position, and People were added to arrive at the P marketing theory combination, which made the original 4P theories finally perfected.

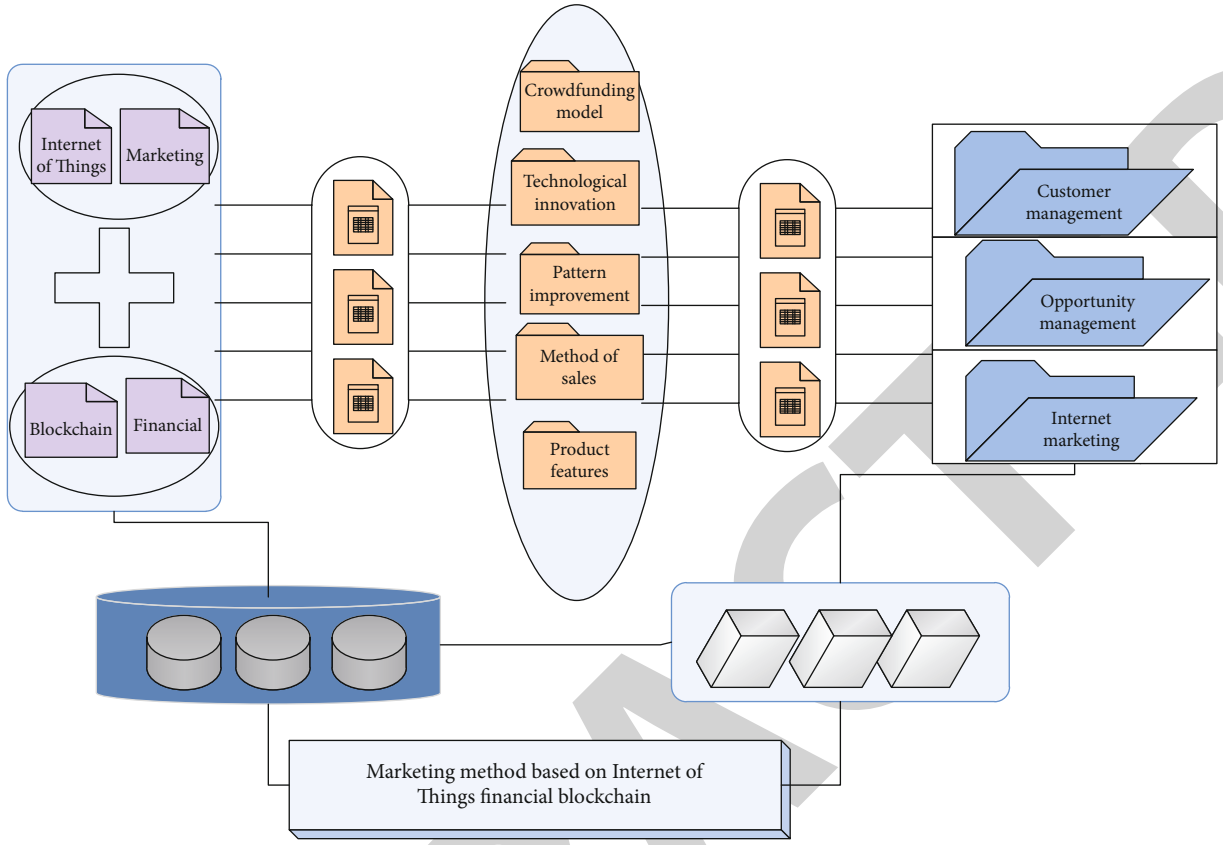


Figure 1: Hierarchical distribution of the financial blockchain of the Internet of Things.

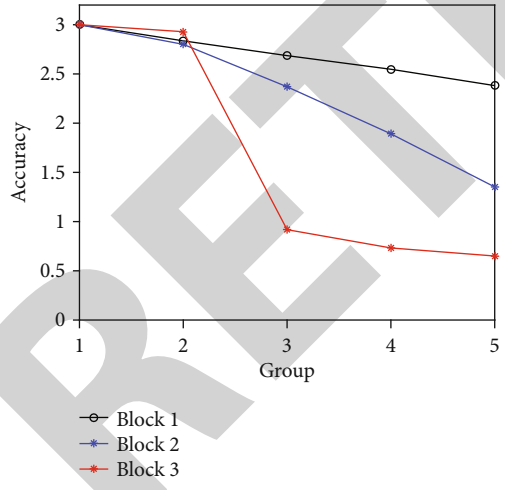


Figure 2: Comparison line chart of the accuracy of IoT marketing models.

Based on the above WeChat marketing model foundation in Figure 2, they formulate five WeChat marketing strategies for CS finance's product development strategy, sales channel strategy, customer experience strategy, precision marketing strategy, and brand building strategy and combine big data to achieve and optimize these marketing strategies. At the same time, they build and optimize the WeChat public platform structure according to these five marketing strategies, make full use of the advantages of WeChat's huge user base, integrate resources, shorten marketing channels, and make it a marketing service integrating product sales, marketing, customer service, and brand promotion, realizing the innovation of marketing models of financial blockchain marketing companies, and helping insurance companies break through the bottleneck of transformation and development. (3) Implement WeChat marketing strategy in all management links of CS finance. The expansion layer encapsulates programmable smart contracts, various script codes, and algorithms. It provides an interface for upper-level applications through smart contracts, making the blockchain have programmable properties. They implement WeChat marketing strategy in the daily operation and management of the company's actuarial, business, service, and marketing departments, formulate relevant regulations, and improve the company's product development, marketing promotion, customer service experience and brand building, and other aspects of efficiency and quality, to provide customers with a full-process service from insurance, preservation to claim settlement. At the same time, each department is interlocking and complementary to each other, forming a closed marketing loop.

3.3. Optimization of Marketing Model Parameters. The marketing industry of traditional financial blockchain has developed so far. Although many aspects have been partially improved, for most financial blockchain marketing companies, there are still many development bottlenecks that need

to be broken through. Mainly reflected in the following: (1) Business model: although traditional financial blockchain marketing companies pay more attention to online platforms, their actual marketing strategies habitually rely on offline channels; the degree of industry informatization is low; the insurance process is lengthy and cumbersome; products and businesses homogeneity of the model; insufficient interaction with customers. (2) Actuarial pricing: except for industry-leading companies and innovative companies, most insurers' pricing factors are slow to update; pricing methods continue to stay in the GLM model; pricing systems are closed and opaque, and lack communication with the outside world. (3) Claim settlement: manual intervention dominates; most companies have low technical content in claim settlement; lack of linkage with the underwriting link. (4) Organizational structure: the traditional financial blockchain marketing company has many levels of organizational structure, and the departments are complicated and not flexible enough; the front and back offices are not smoothly connected, and the disconnection is serious; the labor-intensive development path of human sea tactics is still adopted. (5) Expenses and costs: increasing business volume brings huge business data, and it is more difficult for traditional insurance companies to follow-up information processing capabilities, and operation and maintenance costs increase; in addition, some insurance companies have vicious competition and set fee rates in violation of regulations, giving them financial services. In addition, we have carried out a detailed security certification for the solution, which proves that the cloud storage will not obtain any valid information about the wallet private key during the interaction with the user. The behaviors of blockchain marketers' noncontractual benefits make it difficult for traditional financial blockchain marketing companies, especially small and medium property insurance companies, to dilute their expenses.

(1) Product strategy: it refers to the content of decision-making such as product development, product planning, product design, and delivery date from Figure 3. Its influencing factors include product characteristics, quality, appearance, accessories, brands, trademarks, packaging, guarantees, and service. (2) Price strategy: it refers to the principles and objectives of price setting. The influencing factors include payment methods, credit conditions, basic prices, discounts, wholesale prices, and retail prices. (3) Promotion strategy: it refers to the strategy of how to encourage consumers to purchase and increase sales through various promotional methods. Its influencing factors include advertising, personnel sales, publicity, business promotion, and public relations. (4) Distribution strategy: it refers to the channel strategy of how products reach the end consumers through various distribution channels. Its influencing factors include distribution channels, regional distribution, types of intermediaries, transportation methods, and storage conditions. The above four types of marketing combinations are collectively referred to as marketing mix strategies. The basic idea of marketing mix strategy is as follows: first, formulate product strategy; secondly, formulate price, promotion, and distribution channel strategies to form the overall strategy, so that products that meet customer needs can meet market prices and attractive promotional methods; finally, arrive the consumer process. The success or failure of a company's marketing strategy is directly related to the performance level of the company's sales market. A set of effective and in line with the market's marketing strategy is the basis for success in today's market competition.

4. Application and Analysis of Marketing Model Based on Internet of Things Financial Blockchain

4.1. Simulation of Marketing Strategy Module. The L Group is the leading cloud computing overall solution provider and cloud service provider. It has formed an overall solution service capability covering three levels of laaS, PaaS, and SaaS, relying on high-end servers, mass storage, cloud operating systems, and information security technology. They create a leading cloud computing infrastructure platform for customers, based on government, enterprise, and industry informatization software, terminal products, and solutions, to fully support the construction of smart government, enterprise cloud, and vertical industry cloud. The L Group's turnover has increased rapidly for four consecutive years, with operating income of 45.1 billion. The L Group ranked 10th among the top 100 electronic information industries in China, and its comprehensive strength ranked among the top two Chinese IT companies, the first Chinese independent brand software vendors, and China independent brand IT service providers. The second place, the third place is among the top 500 competitiveness of China's large enterprise groups. The sales volume of L Group's servers ranks fifth in the world and No. 1 in China: the L Group storage has ranked first in sales of state-owned brands for 10 consecutive years; the L Group has the largest market share of management software for 11 consecutive years. The core products of the L Group are server and ERP business, and the competitive situation it faces can be analyzed for these two types of products. It adopts multiple technologies such as automatic photo shunting, dereflecting, component recognition, and image angle correction. Through a large amount of data accumulation and model iterative learning, the entire process from case reception, damage determination to maintenance plan decision-making has been formed. The application layer encapsulates the actual application requirements and implementation cases of blockchains similar to programmable digital currency and programmable finance. According to the official statement of Ant Financial, the fixed loss insurance can save about 2 billion in the cost of the financial blockchain marketing industry every year.

The development of an intelligent quotation system to ensure that customers can get the most accurate quotation in the fastest way is the key to the promotion of insurance policies. The company chose an intelligent quotation system for life insurance quotation in Figure 4. The customer needs to answer seven basic questions. If the customer answers "yes," the system will proceed to the next step under the guidance of the logic tree. If the answer is "no," the system will immediately generate an insurance policy quotation. After

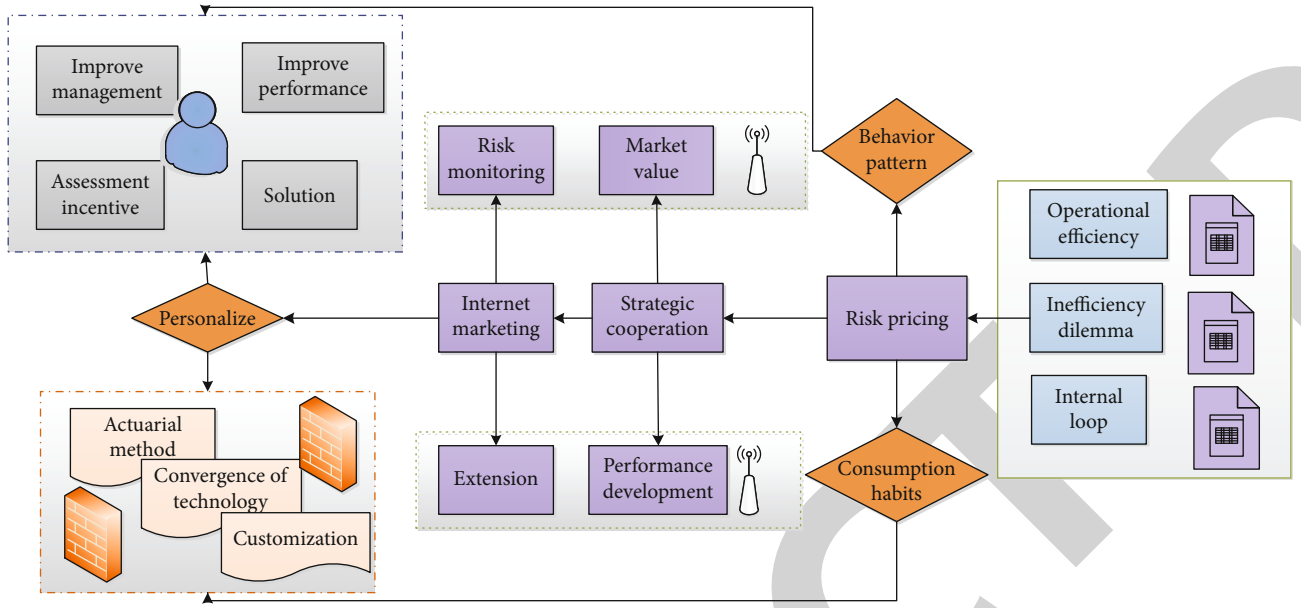


Figure 3: Blockchain marketing algorithm flowchart.

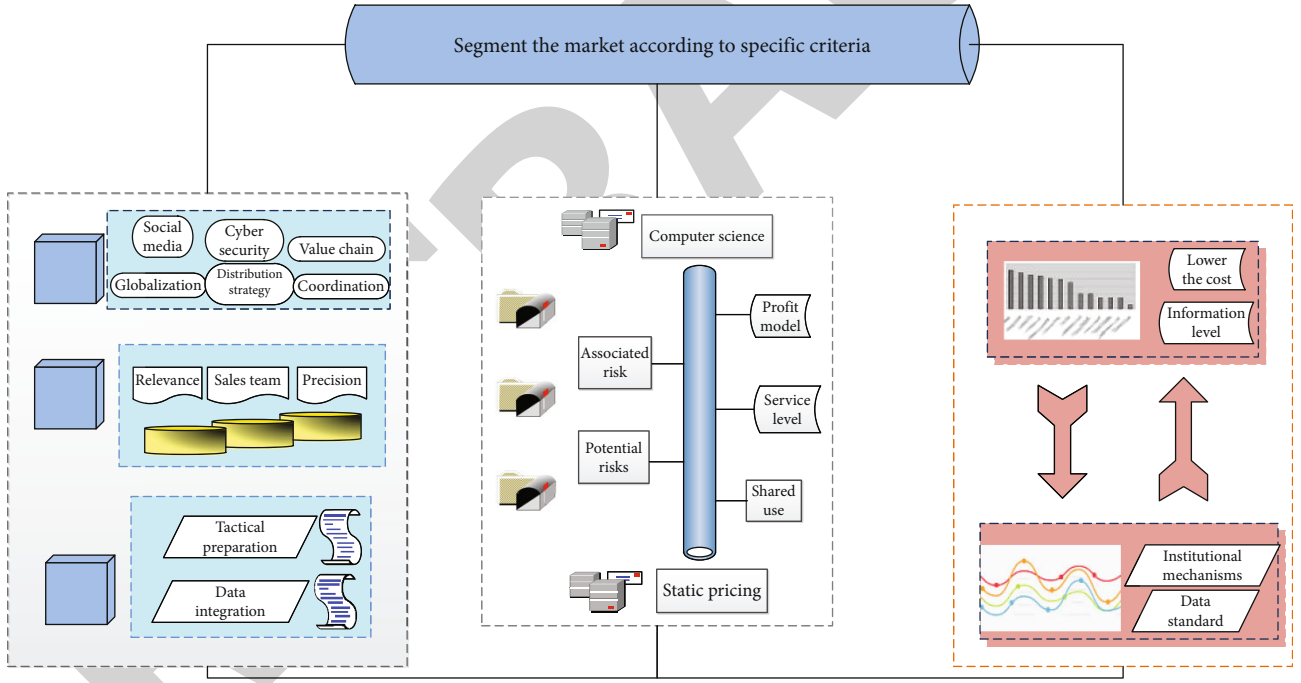


Figure 4: Framework diagram of the marketing model based on the financial blockchain of the Internet of Things.

the customer pays successfully, the policy will take effect immediately. The entire purchase process only takes 10 to 15 minutes, which is very fast. According to the company's internal staff, about 90% of customers can quickly apply for insurance through self-service.

The remaining 10% of customers are likely to have a more complicated medical history, at which time manual services will be involved. From the perspective of improving customer experience in Figure 5, marketing innovation can effectively improve work efficiency and reduce business costs. It is one of the most direct innovation paths for all financial blockchain marketing companies to use technology to improve customer satisfaction.

4.2. Statistical Analysis of Model Reliability. Specifically, for a specific financial blockchain marketing customer and financial blockchain marketing business scenario, a three-dimensional customer relationship management matrix can be built, as shown in Figure 6. The 4C theory focuses corporate marketing on consumers and redefines the four elements of marketing: Consumer, Cost, Convenience, and Communication. It is pointed out that the first task of an enterprise

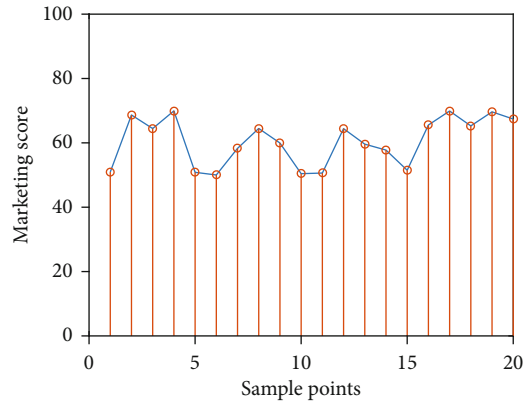


Figure 5: The performance of the IoT marketing model.

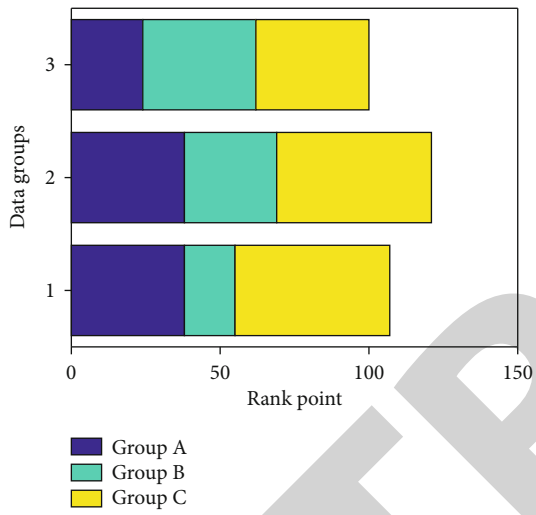


Figure 6: Stacked histogram of the evaluation scores of different groups of the model.

should be to pursue consumer satisfaction and then to strive to reduce the customer's purchase cost and continue to pay attention to and improve the convenience of customers when purchasing products. In the system architecture of the blockchain, many innovative technologies are covered. Among them, the nontamperable chain data structure, the consensus mechanism of the distributed network, and the flexible and operable smart contract are the representative key technologies. Group A is the classification of customers, which can be divided by value contribution, family structure, and age group; Group B is the customer's life cycle, and major nodes include the actual age of the customer and major family events, such as marriage, childbirth, and children's marriage; Group C is the life cycle of an insurance policy, including the entire period from the customer's contact with the financial blockchain marketing to the consideration of whether to insure, to the payment of premiums in the subsequent period after the policy becomes effective, changes in preservation, processing of claims, survival benefits, and surrenders. All actions that may occur during the life cycle of the policy are based on the policy.

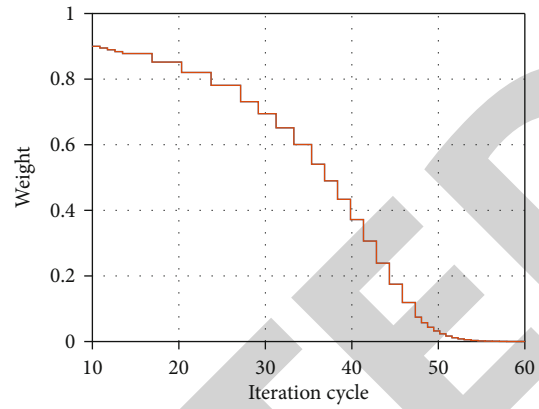


Figure 7: Ladder diagram of model weight changing with the number of iterations.

Under this preferential policy environment, the national online retail sales increased significantly in Figure 7, a year-on-year increase of 33.3 percentage points, and the total amount reached 3.9 trillion; of which, the online retail sales of physical and nonphysical goods were 82% and 18%, respectively, with a year-on-year increase of 31. 6% and 42.4%. Among the total retail sales of consumer goods, online retail sales of physical goods accounted for 10.8%. All these indicate that the consumption mode of Chinese residents has shifted from traditional consumption to online shopping, in which e-commerce plays an important role. In addition, the shortage of technical talents is also a major technical risk faced by the marketing technology innovation of financial blockchain. According to a survey by PricewaterhouseCoopers (PwC), 87% of financial blockchain marketing industry respondents feel that they have difficulties in hiring and retaining talents with corresponding innovative skills.

Among them, L server shipments in the four-way and eight-way markets reached 23% and 54%. The market targeted by ERP products is mainly the corporate market, and the core target customer group is group companies. Due to the impact of the financial crisis in recent years, the corporate market has been in a downturn. ERP vendors, especially their competitors, have suffered losses in the past two years. With the advantages of the group's brand, management, and products, L International maintains an annual growth rate of 30%. In it, China's software business revenue was 475.1 billion yuan, a year-on-year increase of 29.6%. In the software ranking, L software business revenue ranked fifth in China. The software business of the top four companies is not in competition with the software business of the L Group, while UFIDA ranks 39. Some of financial blockchain marketing industry respondents chose "very difficult," compared to only 28% of banking industry respondents. In the past, it took as long as one year to develop a product from conception to completion. The nodes in the network are dynamically divided into several clusters, and each cluster is composed of a cluster head node and several common cluster member nodes. Nowadays, by implementing development strategies to determine product positioning and quickly locate customer portraits, the development cycle can be shortened to

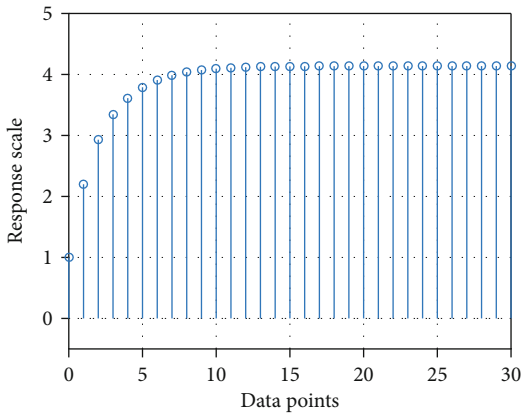


Figure 8: Match stick diagram of model response rate.

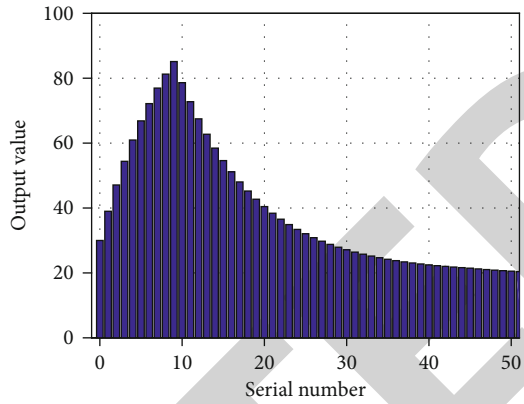


Figure 9: Distribution of the output value of the marketing model with the number of sequences.

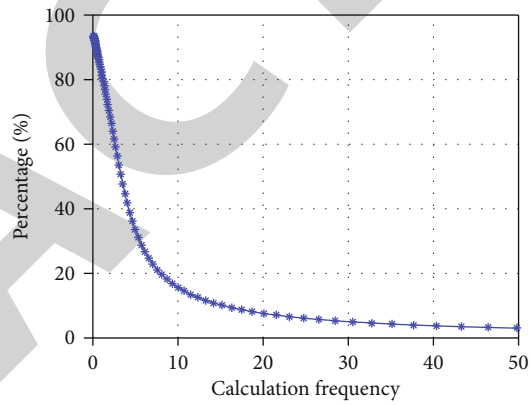


Figure 10: The dependence curve of the proportion of the blockchain with the calculation frequency.

three months. The increase in the express delivery of individual product development means an increase in the number of product developments throughout the year. This plays a vital role in the Internet era where rapid product iteration is emphasized.

At the same time, the experience of developing customized products through cooperation with third-party channels is applied to the development of its own products in Figure 8, making the products more in line with customer needs and avoiding the disconnection between its own products and market demand. Offline traditional insurance usually takes 5-7 days, and it also takes a lot of time for the agent to communicate, but using WeChat to apply for insurance can reduce this time to 2 minutes, and the claim settlement time from the longest 10 days to the shortest 30 minutes, the processing time for security services was shortened to 5 minutes, 75% of claims cases were closed on the same day, and the average time limit for claim services was 2 days. And the online insurance model can effectively avoid misleading sales.

4.3. Example Application and Analysis. The current customer base of L server is mainly the government market, and the enterprise market has been included in the official market mission since 2019. The market research company IDC released China's X86 server market data. The data showed that the total sales of China's X86 servers was 1,474,828 units, a year-on-year increase of 12.6%. Domestic manufacturers surpassed foreign countries with a 61% market share for the third consecutive year. It can be seen from Figure 9 that the share of international IT vendors in the X86 server market has gradually decreased for four consecutive years, from 53% to 39%. The market share of domestic manufacturers has increased year by year. It was 47% and 61%, indicating that the competitiveness of domestic manufacturers is increasing year by year. Domestic servers are gradually expanding their share in the Chinese server market. Among them, the cluster head node records the information of the members in the cluster and other related information of the cluster head. When searching for the target node for communication, it is first to search in the cluster. If it is not found in the cluster, it will search in other clusters through the cluster head. Market research company Gartner announced the server market research report for the third quarter. According to data, global server shipments in the third quarter increased by 1% year-on-year to 2.53 million units, and the growth was sluggish; in the global regional market, only the shipment volume of the Chinese market maintained growth, with a year-on-year increase of 15.63%. In the third quarter, China became the only bright spot in the global service provider market and the only driving force for the growth of global server market shipments.

It can be seen from Figure 10 that in the third quarter, the growth rate of local brands in the Chinese server market was higher than that of foreign manufacturers, and the advantages of domestic manufacturers have further emerged. In the third quarter, in the top five server markets, the three domestic manufacturers L Group, Huawei, and Lenovo ranked second, third, and fourth, respectively. Dell maintained its position in shipments with a slight advantage, while HP ranked fifth. In the third quarter, local companies further encroached on foreign companies.

At present, in the domestic group management software market, the concentration of the Chinese group management software market continues to increase, with domestic companies occupying a dominant position. It can be seen that

domestic management software companies represented by L Group and UFIDA have won the trust of group corporate customers with their high-quality services and advanced products. The top ten vendors occupy 79.3% of the market. To sum up, although the L Group has separate competitive positions in the server and ERP markets, as the only domestic IT vendor that integrates software and hardware, it is beyond the reach of other domestic companies in terms of overall solution services. It should be noted that the cluster head is variable. According to the dynamic changes of the network, such as the new node joining, the original node exiting, and the contribution ability and other factors, the roles of the cluster head and members in the cluster will change accordingly. As customers pay more attention to overall solutions, the overall competitive advantage of the L Group has become more and more obvious. As a link between the company and customers, the L Group call center plays an irreplaceable role in sales and service performance. The business opportunity signing performance and customer service department service revenue generated by the L Group's call center server product line accounted for 15% of server sales. The business opportunity signing business generated by the ERP product line and the service revenue generated by the online support department accounted for 22% of the ERP product sales. Therefore, with the rapid progress of L Group's performance and the vigorous expansion of the market, the call center is bound to keep up with the development of the group's market.

5. Conclusion

This article starts from the application practice of blockchain, artificial intelligence, Internet of Things, big data, cloud computing, drones, and other modern technology elements in the marketing industry of financial blockchain. On the basis of the existing literature, the connotation of marketing technology of financial blockchain will be redefined. It explored the process, innovation logic, and basic principles of the marketing technology innovation and development of international and domestic financial blockchain under the guidance of innovative ideas. Following the innovation logic and basic principles, the paper focuses on the application path of financial blockchain marketing technology innovation, expounds the new risks that financial blockchain marketing technology innovation may face, and puts forward several corresponding risk control recommendations. The research in this article has two practical significances: one is to propose a set of practical solutions for the improvement of marketing management strategies of the L Group call center. In the context of the continuous growth of the IT market, by strengthening the marketing management of the L Group's call center, the L Group's ability to serve customers and the ability to accurately grasp customer needs are promoted, and it improves marketing capabilities and profits, so that the call center can play for the L Group greater market value; in other words, the L Group is a typical representative of the domestic IT industry. The research in this article can provide reference for the marketing management of call centers for IT-type peer companies or B2B companies that trade between companies.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by Washington University in St. Louis support.

References

[1] D. Knezevic, "Impact of blockchain technology platform in changing the financial sector and other industries," *Montenegrin Journal of Economics*, vol. 14, no. 1, pp. 109–120, 2018.

[2] H. Yao, T. Mai, J. Wang, Z. Ji, C. Jiang, and Y. Qian, "Resource trading in blockchain-based industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3602–3609, 2019.

[3] Y. Zhang and J. Wen, "The IoT electric business model: using blockchain technology for the internet of things," *Peer-to-Peer Networking and Applications*, vol. 10, no. 4, pp. 983–994, 2017.

[4] S. Zhao, S. Li, and Y. Yao, "Blockchain enabled industrial Internet of Things technology," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1442–1453, 2019.

[5] Z. Zheng, S. Xie, H. N. Dai, X. Chen, and H. Wang, "Blockchain challenges and opportunities: a survey," *International Journal of Web and Grid Services*, vol. 14, no. 4, pp. 352–375, 2018.

[6] S. Zehir and M. Zehir, "Internet of things in blockchain ecosystem from organizational and business management perspectives," *Digital Business Strategies in Blockchain Ecosystems*, vol. 6, pp. 47–62, 2020.

[7] E. F. Jesus, V. R. L. Chicarino, C. V. N. de Albuquerque, and A. A. de A Rocha, "A survey of how to use blockchain to secure internet of things and the stalker attack," *Security and Communication Networks*, vol. 2, 18 pages, 2018.

[8] Y. Yu, Y. Li, J. Tian, and J. Liu, "Blockchain-based solutions to security and privacy issues in the Internet of Things," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 12–18, 2018.

[9] M. Maiti and U. Ghosh, "Next generation Internet of Things in fintech ecosystem," *IEEE Internet of Things Journal*, vol. 4, pp. 22–29, 2021.

[10] S. Singh and N. Singh, "Blockchain: future of financial and cyber security," *Contemporary computing and informatics*, vol. 3, pp. 463–467, 2018.

[11] 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. 2, 20 pages, 2020.

[12] W. Viriyasitavat, L. Da Xu, Z. Bi, and V. Pungpapong, "Blockchain and internet of things for modern business process in

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5006248, 11 pages
https://doi.org/10.1155/2021/5006248

Research Article

Automatic Recognition of Communication Signal Modulation Based on the Multiple-Parallel Complex Convolutional Neural Network

Zhen Huang[1], **Chengkang Li**[1], **Qiang Lv**[1], **Rijian Su**[2] **and Kaibo Zhou**[3]

[1]*School of Electrical and Electronic Engineering, Wuhan Polytechnic University, Wuhan, Hubei 430070, China*
[2]*School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, Henan 450002, China*
[3]*MOE Key Laboratory of Image Processing and Intelligence Control, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China*

Correspondence should be addressed to Rijian Su; srj@zzuli.edu.cn

Received 18 April 2021; Revised 20 May 2021; Accepted 28 May 2021; Published 9 June 2021

Academic Editor: Wei Wang

This paper implements a deep learning-based modulation pattern recognition algorithm for communication signals using a convolutional neural network architecture as a modulation recognizer. In this paper, a multiple-parallel complex convolutional neural network architecture is proposed to meet the demand of complex baseband processing of all-digital communication signals. The architecture learns the structured features of the real and imaginary parts of the baseband signal through parallel branches and fuses them at the output according to certain rules to obtain the final output, which realizes the fitting process to the complex numerical mapping. By comparing and analyzing several commonly used time-frequency analysis methods, a time-frequency analysis method that can well highlight the differences between different signal modulation patterns is selected to convert the time-frequency map into a digital image that can be processed by a deep network. In order to fully extract the spatial and temporal characteristics of the signal, the CLP algorithm of the CNN network and LSTM network in parallel is proposed. The CNN network and LSTM network are used to extract the spatial features and temporal features of the signal, respectively, and the fusion of the two features as well as the classification is performed. Finally, the optimal model and parameters are obtained through the design of the modulation recognizer based on the convolutional neural network and the performance analysis of the convolutional neural network model. The simulation experimental results show that the improved convolutional neural network can produce certain performance gains in radio signal modulation style recognition. This promotes the application of machine learning algorithms in the field of radio signal modulation pattern recognition.

1. Introduction

Secure and efficient transmission of information is the basic requirement of wireless communication. In the actual communication system, the baseband signal cannot be transmitted directly due to the channel spectrum characteristics and the modulation is usually used to load the code element information carried by the baseband signal to the digital characteristics of the sinusoidal signal and then transmit it through the antenna [1]. Depending on the specific channel type and occupancy, the choice of modulation type can vary, and in turn, the signal after different modulations can exhibit different structural and statistical characteristics in the time-frequency domain. Modulation recognition refers to the use of mathematical models such as machine learning to select the correct modulation type for a received signal, from a given number of modulation systems, after supervised training. MR is an important application of pattern recognition in the field of signal processing and has significant practical importance in both collaborative and noncollaborative communication [2]. With deep learning networks being widely used in speech recognition, image feature learning, etc. and

achieving many practical results, it has led researchers to introduce deep learning into the process of modulation pattern recognition of communication signals, hoping to make communication devices capable of self-learning and self-renewal, so that they can better cope with the problems and challenges brought about by the complex electromagnetic environment and the increase of modulation patterns in the future [3].

Modulation pattern recognition is one of the key technologies for software radio, communication countermeasures, and illegal spectrum monitoring, which has important military and civil values. In the military field, the future war is an information-driven war and electronic countermeasures are an important part of information warfare and modulation pattern recognition is one of the functions that must be considered in the receiver design process, where the receiver automatically demodulates the intercepted signals through modulation pattern recognition technology. In electronic reconnaissance and jamming, modulation technology is the key to implement precise jamming, which in turn disrupts enemy communications. In civil applications, modulation pattern recognition technology is a key technique to identify the illegal use of frequency bands to prevent spectrum abuse and interference with normal communications [4]. Traditional modulation pattern identification methods include maximum likelihood hypothesis testing methods and statistical pattern recognition-based methods, the former with high complexity and poor robustness to models and parameters and the latter with performance strongly correlated with human-selected feature parameters. In recent years, machine learning (ML) technology is hot and ML is one of the most important branches of artificial intelligence, which can classify and predict data intelligently and has excellent recognition performance, and the research of modulation pattern recognition applying ML is receiving more and more attention and interest. Therefore, this paper studies deep learning-based algorithms for digital communication signal modulation and recognition.

This thesis focuses on the deep learning-based modulation pattern recognition method for communication signals. The scheme adopts a multibranch CNN architecture to realize the convolutional mapping of the input signal in the complex domain and complete the preprocessing work of signal denoising and channel equalization to improve the input for modulation recognition; it investigates the impact of abstract features learned by CNN and artificially designed expert features, multiple machine learning classification models on modulation recognition, and modulation recognition algorithms; with the help of a general-purpose software radio platform, a variety of modulation signal sequences are collected with the help of a general-purpose software radio platform, the data sets used for training and testing are established, and the algorithms are designed and coded and validated by a deep learning framework and software platform. The first chapter is the introduction part of the thesis, which introduces the background and significance of the thesis and finally gives the research content and structural arrangement of the thesis. Chapter 2 is the related work section, which systematically describes the research status and analyzes the advantages and disadvantages of domestic and foreign technologies in modulation identification, signal denoising, and channel equalization. The third chapter analyzes and studies the communication signal feature processing and explains the specific implementation of the algorithm, and finally, the design study of the modulation identifier is carried out. Chapter 4 is the analysis of the results. By analyzing the performance of the algorithm proposed in this paper and simulation tests, the method can identify the modulation patterns of communication signals well under low signal-to-noise ratio, which proves the feasibility and effectiveness of the method. Chapter 5 summarizes the full text of the work and provides an outlook.

2. Related Work

The maximum likelihood hypothesis testing method based on decision theory theoretically ensures that its decision results are optimal under the Bayesian least-cost criterion and can guarantee the performance of the method in a certain low signal-to-noise environment. However, the main drawbacks of the method are as follows: more a priori knowledge is required, such as signal-to-noise ratio, carrier frequency, symbol rate, oversampling multiplier, and other parameters, and secondly, the existence of unknown parameters leads to a complex computational push-to process and high computational complexity, which is difficult to implement in practical production [5]. The classification function is relatively single, and all common modulation types cannot be identified by one framework; in the actual channel, the noise is not necessarily Gaussian white noise and there are multipath fading, time delay, Doppler effects, and other effects [6]. The maximum likelihood hypothesis testing method based on decision theory is more susceptible to these influences and less robust [7]. Mishra et al. implemented deep learning-based denoting by improving the deep network structure. In the literature [8], the authors used ResNet to remove rain noise from photographs and the experimental results showed that the model has superior denoting performance [9]. Gupta et al. utilized a single-layer, stacked LSTM network to identify seven signals such as 2ASK, 4ASK, BFSK, 4FSK, BPSK, QPSK, and 16QAM [10]. Patnaik et al. used CNN and DBN to identify CPFSK, BPSK, and 16QAM signals [8]. The cascade structure is poorly parallelized; the model is too complex, and the training is time-consuming. In addition, the temporal signal loses some timing features after the convolutional layer [11].

In order to take full advantage of the temporal and structural features of the signal, this paper studies CNN and LSTM parallel modulation-style recognition algorithms. In order to further improve the signal recognition performance, the integrated learning algorithm of the heterogeneous basis classifier is also studied [12]. Since deep learning can automatically extract features from large-scale data and perform learning, it can be quickly developed and widely used [13]. Although the deep learning parameter design still lacks some theoretical proofs and has some shortcomings, practice has proved that applying deep learning of the modulation recognition problem

is a very effective method and it is worthy of in-depth research and exploration.

3. Machine Learning Decision Theory-Based Communication Signal Modulation Pattern Recognition

3.1. Communication Signal Feature Processing. With the increase in computer processing speed and storage capacity, the design and implementation of CNN has gradually become an exhibition trend [14]. In this chapter, a convolutional neural network-based communication signal modulation pattern recognition method is used to determine the modulated signal [15]. Firstly, the received modulated signal is preprocessed by normalization and time-frequency feature image generation to generate the test set of the training set required for network training; secondly, the classifier for modulated pattern recognition of communication signal, CNN, is designed and built and the training set is input to CNN for training to obtain the CNN network model; finally, the modulated signal to be recognized is preprocessed to generate the test set of the dataset and the CNN network model is obtained. Finally, the modulated signal to be identified is preprocessed to generate a test set in the dataset and the training set is input into the CNN network model to identify the modulation pattern of the communication signal. The method takes the time-frequency domain map as the input and the modulation pattern of the signal as the output, and the specific algorithm flow chart is shown in Figure 1.

The digital signal (−1, +1) to be transmitted is mapped to N carriers (IFFT) and superimposed together and then sent. For example, there is a bandwidth of 1000 M and the size of the IFFT is 1024. In this case, the bandwidth of 1000 M is divided into 1024 parts. The purpose is to reduce the bandwidth and thereby reduce the ISI. Then, the frequency of the first carrier is 1000 M/1024, the frequency of the second carrier is 2*1000 M/1024, and so on. The function of this part is to make N carriers orthogonal to each other to eliminate interference between carriers.

The effect of noise and the channel makes the communication signal amplitude vary very much. If the acquired time domain signal is used directly for time-frequency analysis, it will cause some difficulties in processing the time-frequency map after time-frequency analysis. Therefore, the signal has to be normalized [16]. In this paper, the acquired time domain signal will be normalized by the min-max normalization method, which is indicated below. The amplitude of the modulated signal is simply scaled by the min-max normalization method, and the modulated signal amplitude is taken to be compressed between the interval [0, 1], so that the amplitude values of the modulated signal are relatively close to each other and the effect of channel fading on the signal amplitude is reduced.

$$G\left(\overline{\overline{\inf}}\right) = \frac{\inf - \min\left(\inf\right)}{\max\left(\mathrm{INF}\right) - \min\left(\mathrm{INF}\right)}. \tag{1}$$

A feature is an abstract representation of an object or a class of objects. Relative to objects, features use a set of low-dimensional tensors to express the focal properties of the original object and are the key to distinguishing multiple objects. The features together with the training set data determine the theoretical upper limit of the machine learning task, and the models and algorithms are intended to approximate this limit as closely as possible. Therefore, the selection of features should be cantered on the task [17]. For the modulation recognition task, the selection of features can be divided into two stages: one is based on manually designed export features; the other is based on CNN self-learning abstract features. This section will elaborate and analyze these two kinds of features [18].

$$\left\{ \begin{array}{l} h(x) = \dfrac{|h(x)|}{M\{|h(x)|\}} - \beta \\ y_n(x) = y(x) - M\{|h(n)|\} \end{array} \right\}. \tag{2}$$

In the convolutional layer, each convolutional kernel can be considered as a linear system for extracting a certain feature, but before the training of the network, the operational parameters of the whole system are unknown, so the weight parameters of the convolutional kernels are randomly initialized, and the parameters of this system can only be updated by the BP algorithm to continuously optimize by reducing the value of the objective function, and when the training is completed, the system can be used to extract the input features. The cascade of convolutional layers allows the input signal to be mapped by layers of abstraction to obtain the feature vector needed by the classifier.

$$B_{ij} = M\left\{ |h(n)|^{i-j} \mid \underset{i+j}{\longrightarrow} h(n)^j \right\}. \tag{3}$$

In the encoder stage, it is downsampled by the pooling layer to compress the size of the output feature map, after the signal is mapped by the convolution layer of CAE, and after several identical operations, it reaches the bottom convolution layer and the output of the bottom layer can be considered as the abstraction of the original input signal $h(n)$; in the decoder stage, these abstract features are continuously up sampled by the deconvolution layer and finally reach the top output layer. The goal of CAE is to make the output consistent with the input, and the computational process can be characterized as "compression before reconstruction" [19].

The result obtained after the time-frequency analysis of the modulated signal is a representation of the modulated signal in the time-frequency plane, which cannot be precisely input into the deep neural network model for processing. Therefore, it is necessary to convert the time-frequency map of the signal to generate a digital image first and then use the deep learning algorithm to identify the modulation of the signal [20]. Usually, the color image can be greyed out by using the maximum value method or the component method or the average method. Among them, the component method uses one of the three components of R, G, and B in the color image as the gravy value; the maximum value

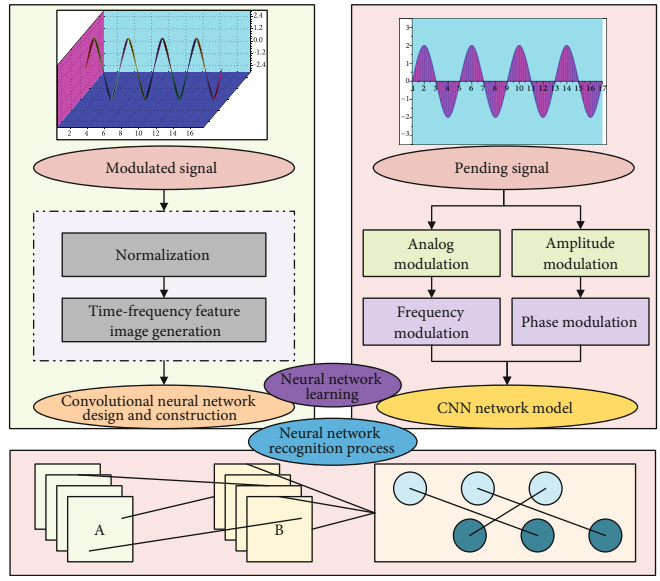


Figure 1: Flowchart of modulation pattern recognition of communication signal based on the convolutional neural network.

method uses the maximum value as the gravy value; the average value method uses the average of the three components of R, G, and B in the color image as the gravy value. Different gray scale processing methods will produce different gray scale feature images [21]. In this paper, we select the gray scale processing method of the average method and its gray scale value calculation formula is

$$\text{Gray}(x) = \frac{R+G+B}{3} * \beta + \mu. \tag{4}$$

By generating gray scale feature images of modulated signals with different signal-to-noise ratios, it is found that when the signal-to-noise ratio is low, the modulated signals are more affected by the background noise, which makes them appear more disordered in the gray scale feature images and the feature information becomes somewhat blurred [22]. When the signal-to-noise ratio is high, the modulated signals are less affected by the background noise and their gray scale feature images are clearer and more regular. By comparing the binary feature image and gray scale feature image, it is found that the medium gray scale feature image can retain almost all the original feature information and has certain noise immunity.

For noise, Gaussian white noise can be used, because the noise of general signals is mainly divided into two categories—one is external noise and the other is internal noise of the receiver. The general noise characteristics of the collected signal are very similar to Gaussian white noise, so it is appropriate to replace the internal noise of the receiver with Gaussian white noise. For the amplitude of the noise level, in order to ensure a certain detection probability, the signal-to-noise ratio is required to be greater than 10 db.

3.2. Research on Modulation Recognition Algorithm Based on the Neural Network. Decision trees can be divided into classification trees and regression trees based on the nature of the data labels. When the data labels are continuous values, we call the decision tree in a regression tree; when the data labels are a series of discrete values, it is referred to as a classification tree. Each leaf node of the classification tree represents a classification result, and the branches of the tree are equivalent to the features on which the classification is founded. Since the digital signal modulation identification in this paper is a classification problem, the decision trees discussed refer to classification trees without separate emphasis. The creation of a classification tree can be summarized as follows: training data is input to the decision tree model and new branches are derived from the root node to the leaf nodes using a recursive approach based on the direction of data flow determined by the judgment conditions in the internal nodes until a leaf node is generated. Classification trees can be generated using a variety of algorithms [23].

Classification trees are generated by discriminating branches according to internal node conditions, which are essentially a feature selection process. In algorithms such as ID3, information entropy is used to perform feature selection [24]. The so-called information entropy can be understood as a quantitative indicator of data uncertainty. For example, for the training data set S, the information entropy is calculated using the following formula:

$$F(s) = \lambda * \sum_{n=0}^{255} h(n) \log_2 h(n). \tag{5}$$

In the CART algorithm, the concept of information first is not continued but replaced by the GINI value [25]. If there is a data set S, the actual set of categories to which the data set belongs is L (N):

$$L(N) = \lambda * \sum_{i}^{j} G(i \mid j) h(i \mid N) h(j \mid N). \tag{6}$$

CNN consists of a series of convolutional layers, pooling layers, and fully connected layers, of which the convolutional layer is the core of CNN [26]. Convolution is a mathematical operation performed on two functions, a linear operation that satisfies the exchange law of addition and the union of multiplication, and other operational properties can be divided into the following notation.

$$U(x) = V(x) * W(x). \tag{7}$$

The above equation is a convolution operation on a one-dimensional continuous time system, where V (x) represents the input to the model; $W(x)$ is called the kernel function, and $U(x)$ is the model output. The two-dimensional discrete convolution can be expressed as follows:

$$U(i,j) = M(i,j) * N(i,j) = \sum_{i,j}^{m,n} \frac{M(m,n)}{N(i-m, j-n)}. \tag{8}$$

In the convolution operation of the input, a large number of feature maps will be obtained; if it was directly input to the next layer, it will make the input signal of the next layer too large. Generally, before the feature map is input to the next layer, the output feature map pooling operation, on the one hand, can reduce the number of parameters and training time; on the other hand, it can reduce redundancy and enhance generalization. The pooling function generally uses the overall statistical features of the neighbouring outputs at a specified location to replace the network's output at that location. Before the final output, CNN changes the obtained feature map into a one-dimensional form before the classification output [27].

Communication signals not only have temporal characteristics but also have different constellation maps for different modulation signals. This suggests that communication signals have strong spatial characteristics, so this paper explores the CNN-based modulation pattern recognition method [28]. For the received baseband I and Q signals, they are concentrated to form a 2*256 size matrix, which is denoted by I_{2*256}, and the elements in I_{2*256} are denoted by $I(m,n)$. For signals, different processing of I and Q information will produce distinct recognition effects, and specifically for CNN networks, different kernel sizes will directly affect the classification performance.

After completing the training, the algorithm can integrate the base classifiers by linear weighting. In the first step, the initial value of the weight is determined here. It should be pointed out that $H(X)$ is called the weight of the base classifier, which essentially means the importance of each classifier. The specific definition is as follows.

$$H(X) = \log_2 \sum \ln \frac{1 - E_x}{E_x} * \beta. \tag{9}$$

Next, DX is updated, and finally, the base classifiers of X iterations are summed to form the final classifier H. In the process of updating the parameters of the network iteratively, the stochastic gradient descent method is used because the descent gradient of the optimized objective function does not depend on a single sample for calculation during the network training process but each iteration samples a random portion of data from the training data set and passes it into the neural in the network training process. The gradient of the decline of the optimization objective function does not depend on a single sample for calculation. The specific flow of the modulation recognition algorithm in this article is shown in Figure 2.

3.3. Modulation Identifier Design Study. Simulation analysis shows that the time-frequency map of the signal can describe the modulation pattern characteristics of the communication signal well. In this paper, the time-frequency map is used as the input of the CNN and the size of the time-frequency image is generally 32*32, 64*64, and 128*128. First, this paper will compare the performance of neural networks with a different number of convolutional layers; then, it will compare the neural networks with different input sizes; finally, it will compare neural networks with different convolutional kernel sizes to get an optimal network model.

Considering the characteristics of the modulation pattern of the communication signal and the size of the time-frequency map, the CNN contains a convolutional layer and a pooling layer to extract the effective feature vector; the nonlinearity of the network model is provided by the activation function Re LU used after each convolutional layer, and using the Re LU function as the activation function can suppress the gradient disappearance or explosion that occurs during the training of the network; the last layer is the final layer and is a fully connected layer used to integrate local features to obtain global features of the input data; finally, the global features are classified and identified by using the SOFTMAX activation function [29–33]. The following will illustrate the effects of different neural network structures on the classification recognition of modulated signals from three aspects: the number of nonconvolutional layers, the input size, and the convolutional kernel size.

In this paper, convolutional kernels of size lama are used extensively, which is equivalent to convolutional kernels to map the real and imaginary parts of the input signal $h(n)$ separately. The $h(n)$ can only characterize the amplitude information of the signal, while the frequency and phase information need to be jointly characterized by the real and imaginary parts. However, according to the convolutional layer algorithm, if a convolutional kernel of size $2 \times$ m is used, the size of the output feature map will be reduced to 1 $\times$ 1024, so only one convolutional layer of $2 \times$ m is used in RESTNET. In order to obtain the resonant feature mapping of the convolutional layer for the frequency and phase of the input signal, this paper proposes using MR-Net in series with the traditional sequential structure to build a CNN as a modulation recognition network. Where the input signal $s(n)$ is the recovered signal from the received signal $y(n)$ input to the pretrained Recover-Net, finally, the classification probability vector is obtained using the SOFTMAX layer output.

This paper proposes to verify the communication modulation awareness algorithm in a real environment by building

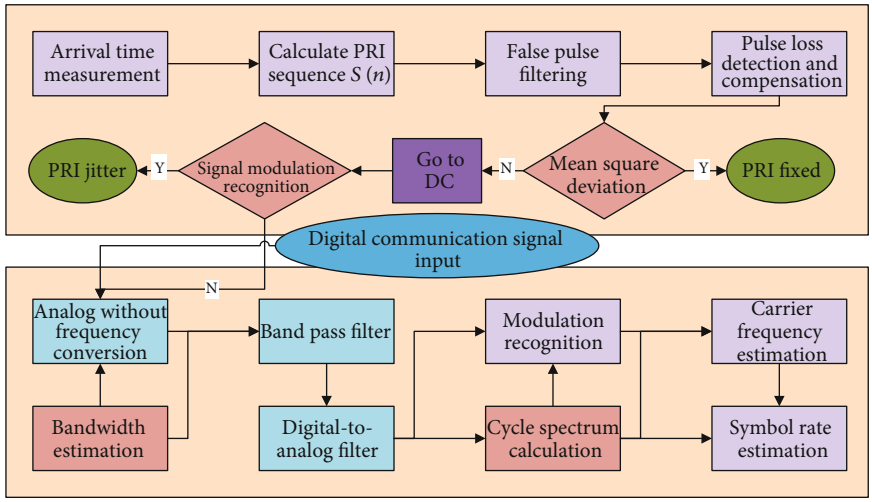


Figure 2: Modulation recognition algorithm flow.

a wireless communication transceiver platform. NI-USRP is based on the public version of the software radio platform USRP Radio by NI Instruments, and some of the external circuitry is modified. NI-USPR hardware has a common software-defined radio (SDR) architecture, and in its FPGA digital signal processing logic, the communication transmitter modulates user data into digital baseband data and the output becomes analogy baseband signal I/Q by high-speed digital-to-analogy converter DAC, and after filtering out high harmonics and spurious by a low-pass filter, the analogy baseband signal is up converted to transmit carrier frequency 915 MHz by analogy RF quadrature, and the communication transmission is completed by power amplification and antenna; at the communication receiver, the antenna receives the faint at the receiving end, the weak wireless communication signal received by the antenna is adjusted by low-noise amplification, then, the analogy RF quadrature downconversion is completed, and the I/Q baseband signal is output to the high-speed analogy-to-digital converter ADC, and the quantized IQ data stream is sent to the FPGA digital signal processing logic. The FPGA digital signal processing logic is used to implement the digital downconversion (DDC) at the receiver side and digital upconversion (DUC) at the transmitter side. In the figure, the upsampling and downsampling logic adjusts the sampling rate of the complex baseband signal to achieve efficient processing of the nonbroadband communication signal and to facilitate further processing via Ethernet to the host computer.

4. Results and Analysis

4.1. Algorithm Performance Analysis. The test set of 10,000 data samples in the radio signal data set was tested on the classic convolutional neural network and the improved convolutional neural network, and the final prediction results were obtained. The modulation pattern recognition accuracy results of different algorithms are shown in Figure 3. It can be seen in Figure 3 that the improved convolutional neural network has a 3.5% higher modulation pattern recognition accuracy than the classic convolutional neural network. The results show that the improved convolutional neural network and classical convolutional neural network proposed in this paper have advantages in the accuracy of modulation pattern recognition.

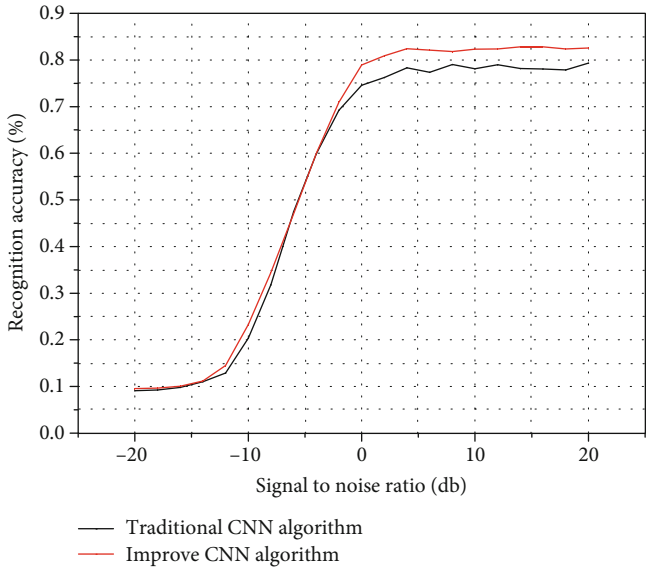


Figure 3: Modulation pattern recognition accuracy of different algorithms.

The running time of the CNN algorithm with different numbers of convolutional layers is given in Figure 4. It can be seen that the running time of the CNN algorithm increases with the increase of the number of convolutional layers. Considering the ACRP and the running time of CNN algorithm, this paper determines to use two convolutional layers, with the first layer having 14 dimensions and the second layer having 5 dimensions.

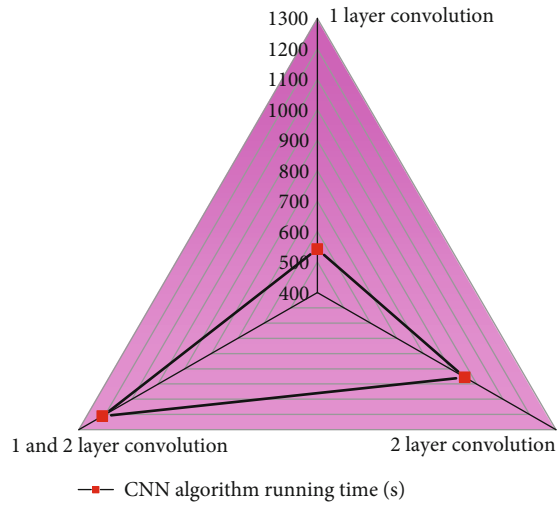


Figure 4: Running time of algorithms with different numbers of convolution layers.

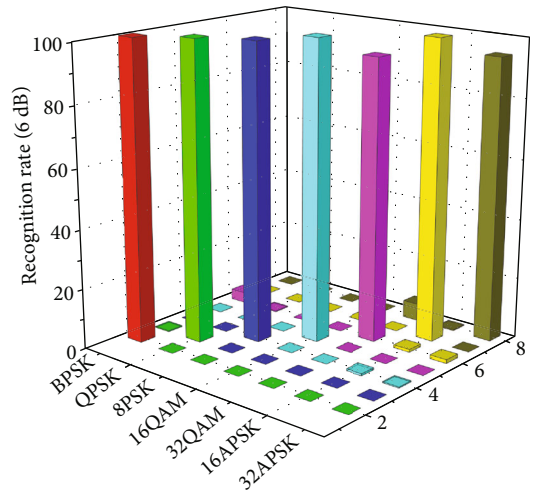


Figure 6: Single signal recognition rate of the CNN algorithm with a 6 dB signal-to-noise ratio.

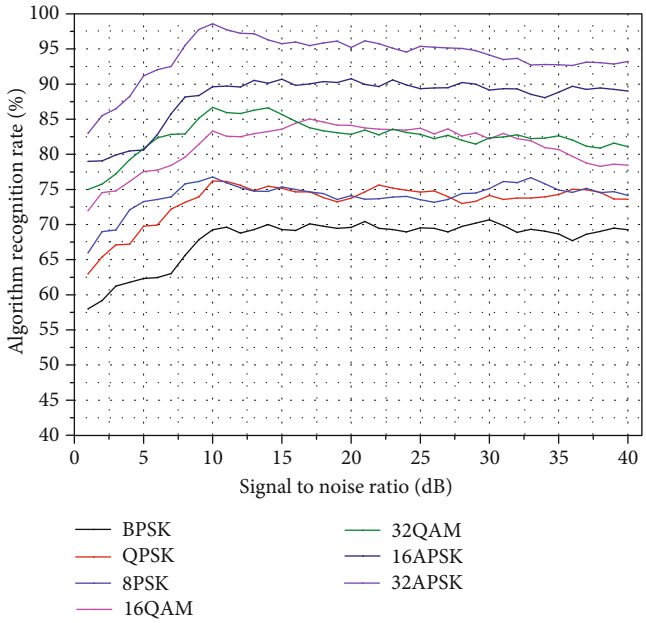


Figure 5: CNN algorithm recognition performance.

Figure 5 shows the correct recognition rates of the proposed CNN algorithm for BPSK, QPSK, 8PSK, 16APSK, 16QAM, 32QAM, and 32APSK signals. As can be seen from the figure, the correct recognition rate of all the seven signals is higher than 92.13% when the signal-to-noise ratio is greater than 2 dB; the CNN algorithm has the highest recognition rate for 16QAM signals, followed by QPSK signals; the CNN algorithm has the worst recognition performance for 32QAM signals, followed by 8PSK signals.

The individual signal recognition rates of the CNN algorithm show the seven signals when SNR = 6 dB are shown in Figure 6. It can be seen that at the SNR of 4 dB, the CNN algorithm has good recognition performance and only 40 samples are misconceived overall and the ACRP reaches 98%; the recognition rates of 16APSK, 16QAM, and QPSK signals all reach 100%, and the recognition rates of 32APSK and BPSK signals reach. The recognition rates of 16APSK, 16QAM, and QPSK signals are 99.99%, 32APSK and BPSK signals are 99.7% and 98.91%, respectively, and the recognition rates of 32QAM and 8PSK signals are 93.9%.

4.2. Communication Signal Simulation Test Analysis. For the simulation test, a test set containing 5 modulated signals of 2ASK, 2FSK, 2PSK, AM, and FM was used, of which the number of samples in the test set was 3000. The average recognition accuracy of the five types of signals from −10 dB to 10 dB is tested, and the test results are presented in Figure 7(a). Recognition accuracy of each signal-to-noise ratio from −10 dB to 10 dB for each of the five types of signals is shown in Figure 7(b). As the S/N ratio increases, the average recognition rate of each type of signal also increases gradually, and finally, the average recognition rate of all five types of signals at 10 dB is greater than 97.1%. The recognition rates of 2FSK and AM are 99.99%; the recognition rate of FM increases rapidly from 99.3% to 99.76% at the beginning and then remains unchanged; while the recognition rates of 2ASK and 2PAK are relatively low at −10 dB, 84.3% and 90%, respectively. With the increase of SNR, the recognition rate of 2FSK and the recognition rates of 2FSK and FM signals also increase as the S/N ratio increases. However, when the signal-to-noise ratio is small, the energy of the noise is higher than the energy of the signal, so that 2ASK and 2PSK are drowned in the background noise, leading to a decrease in the recognition rates of 2ASK and 2PAK. Finally, the recognition rates of both 2ASK and 2PAK reach 97.2% at 10 dB. It can be concluded that the higher the signal-to-noise ratio, the better the recognition effect of the convolutional neural network, and at a lower signal-to-noise ratio, the recognition effect slightly decreases but it can still perform classification recognition; the convolutional neural network

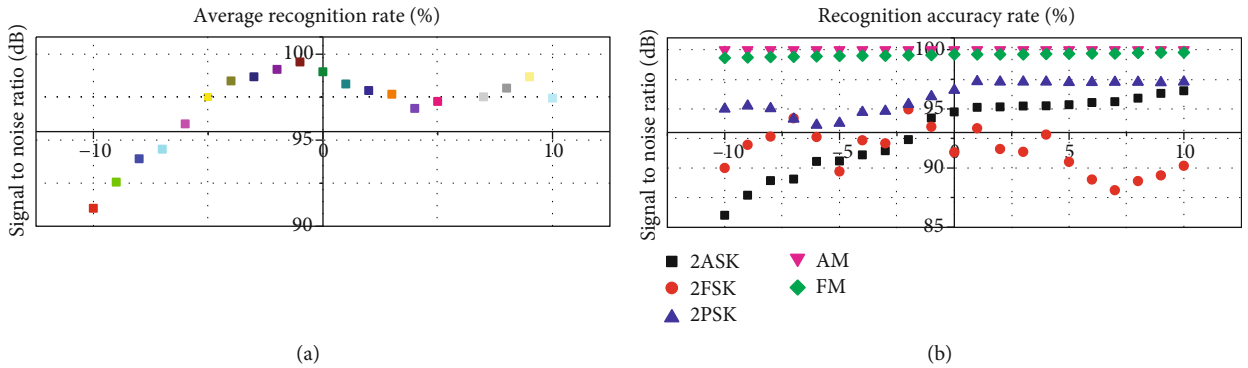


Figure 7: Comparison of average recognition accuracy and recognition accuracy.

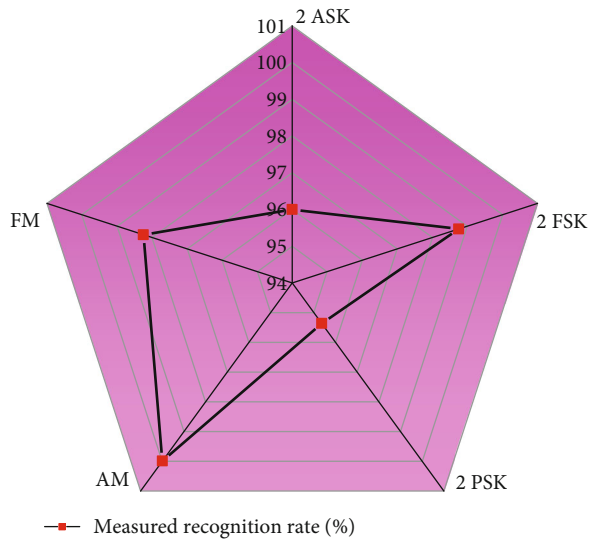


Figure 8: Measured signal test results.

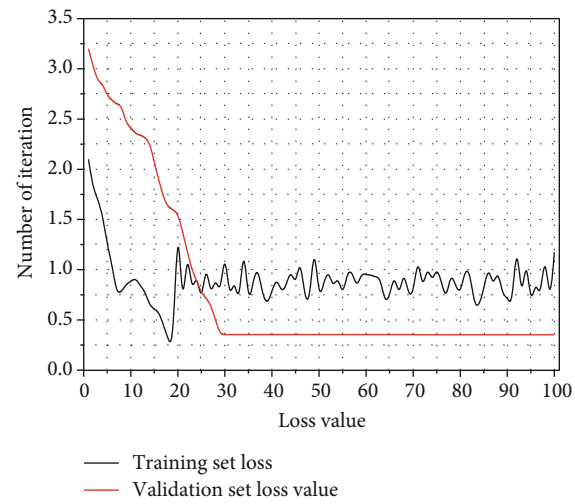


Figure 9: The relationship between the training set loss value and the validation set loss value and the number of iterations.

has good performance in modulation recognition and also has certain antinoise performance.

The database required for the network is produced by preprocessing the measured signals with the same number of samples as in the simulation test, which is 3000, and then tested. The test results are shown in Figure 8. In Figure 8, it can be seen that the recognition rate of the measured signals is higher than 95.36% after the classification and recognition by the modulated recognizer designed in this paper. It also verifies the conclusion that the convolutional neural network has good recognition performance in modulation recognition.

The accuracy of the training set and the accuracy of the validation set change with the number of iterations, and the loss value of the training set and the loss value of the validation set change with the number of iterations. Figure 9 shows the loss of the training set. The relationship between the value and the loss of the validation set and the number of iterations are shown. Through comparison, it is found that the transformation trends of the training set accuracy and the validation set accuracy are basically the same, while the transformation trends of the training set loss value and the validation set loss value are not consistent. It can be seen in Figure 9 that after 10 iterations of training, the loss value of the training set continues to decrease, while the loss value of the validation set shows a trend of oscillating transformation. It shows that the network has an overfitting phenomenon and the complexity of the network needs to be reduced.

For continuous phase modulation signal, this paper considers the GMSK signal, which is a commonly used communication modulation signal with high spectrum utilization and high noise and channel interference immunity. Unlike modulation types such as MPSK and MQAM, GMSK signals do not have fixed theoretical constellation points in the complex plane, as shown in Figure 10, where we demonstrate the effect of Recover-Net on GMSK signals before and after recovery from the time domain waveform; the channel is still an indoor channel simulated by MATLAB with SNR = 10 dB. It can be observed that the amplitude and phase recovery of the time domain waveform by Recover-Net is good and there is only a small deviation in the details at some peaks, with

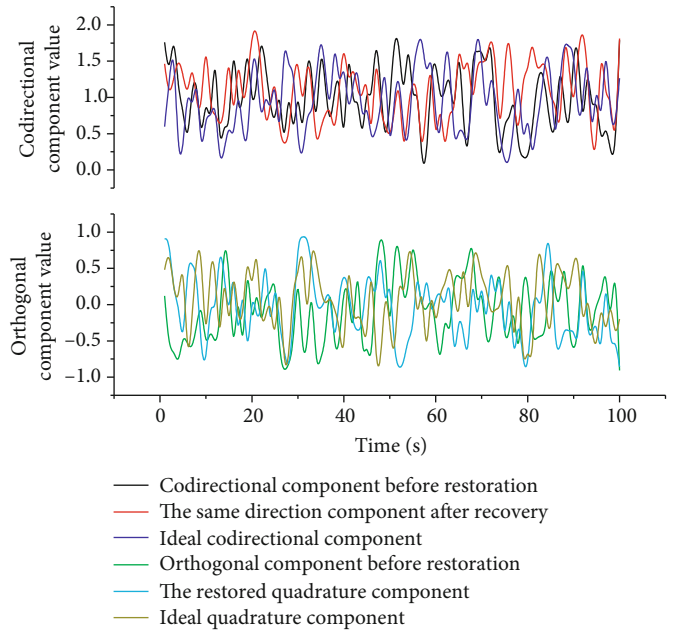


Figure 10: Example of the time domain before and after the recovery of the GMSK signal.

EVM = 204.17%, MER = −6.26 dB for the prerecovery signal and EVM = 12.14%, MER = 18.14 dB, and SER = 0.038%.

5. Conclusion

In this paper, the problem of communication signal modulation pattern recognition based on deep learning is studied. Firstly, the mechanism of communication signal generation and the related theory of deep learning are introduced to provide the theoretical basis for the identification of modulated signals. Secondly, from the time-frequency domain of the signal, multiple time-frequency analysis methods are compared to select the time-frequency analysis method that better characterizes the modulated signal and the feature image generation algorithm is used to convert the time-frequency image into a database that can be used by the neural network. Finally, through the design of the modulation recognizer based on the convolutional neural network and the analysis of the performance of the convolutional neural network, a convolutional neural network model is established to realize the fast and accurate recognition of modulation patterns of communication signals in a complex electromagnetic environment and it can have good recognition effect under the condition of low signal-to-noise ratio. Compared with classical convolutional neural network modulation pattern recognition, recognition accuracy is improved; the convolutional neural network can be improved by changing the network layer structure, using a sequential convolutional module structure or using small convolutional kernel to extract the fine details in the radio signal; the characteristics of the radio signal will be clearer, compared with classical convolutional neural network modulation pattern recognition; the improved convolutional neural network not only shortens in training time but also improves the recognition accuracy. Theoretical analysis and computer simulation results show that the proposed algorithm based on the machine learning decision theory for communication signal modulation and recognition is practical, effective, and easy to implement and has the value of application in practical engineering.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

We declare that there is no conflict of interest.

Acknowledgments

The National Natural Science Foundation of China, for the research on ferromagnetic resonance temperature imaging based on superparamagnetic nanoparticles, Grant 61773018; research on low complexity coding model and method based on 3D-HEVC, Grant 61771432; research on information acquisition and processing method of temperature field in superparamagnetic nanoparticle targeted, Grant 61374014; and research real-time and accurate method for measuring 2D temperature distribution in magnetic nanoparticle-mediated hyperthermia, Grant 61803346, supported the study.

References

[1] Z. Lv, A. K. Singh, and J. Li, "Deep learning for security problems in 5G heterogeneous networks," *IEEE Network*, vol. 35, no. 2, pp. 67–73, 2021.

[2] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4074–4077, 2019.

[3] P. Bai, H. Xu, and L. Sun, "A recognition algorithm for modulation schemes by convolution neural network and spectrum texture," *Xibei Gongye Daxue Xuebao/Journal of Northwestern Polytechnical University*, vol. 37, no. 4, pp. 816–823, 2019.

[4] Y. Chen, W. Liu, Z. Niu, Z. Feng, Q. Hu, and T. Jiang, "Pervasive intelligent endogenous 6G wireless systems: prospects, theories and key technologies," *Digital Communications and Networks*, vol. 6, no. 3, pp. 312–320, 2020.

[5] S. Jayakumar and N. S, "A review on resource allocation techniques in D2D communication for 5G and B5G technology," *Peer-to-Peer Networking and Applications*, vol. 14, no. 1, pp. 243–269, 2021.

[6] Y. He, I. Ahmad, L. Shi, and K. Chang, "SVM-based drone sound recognition using the combination of HLA and WPT techniques in practical noisy environment," *TIIS*, vol. 13, no. 10, pp. 5078–5094, 2019.

[7] T. Long, T. Zeng, C. Hu et al., "High resolution radar real-time signal and information processing," *China Communications*, vol. 16, no. 2, pp. 105–133, 2019.

[8] M. Patnaik, V. Kamakoti, V. Matyas, and V. Rehak, "PROLE-Mus: a proactive learning-based MAC protocol against PUEA and SSDF attacks in energy constrained cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 2, pp. 400–412, 2019.

[9] N. Mishra, S. Srivastava, and S. N. Sharan, "Countermeasures for primary user emulation attack: a comprehensive review," *Wireless Personal Communications*, vol. 115, no. 1, pp. 827–858, 2020.

[10] V. Gupta, S. Tripathi, and S. De, "Green sensing and communication: a step towards sustainable IoT systems," *Journal of the Indian Institute of Science*, vol. 100, no. 2, pp. 383–398, 2020.

[11] M. Sheraz, M. Ahmed, X. Hou et al., "Artificial intelligence for wireless caching: schemes, performance, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 631–661, 2021.

[12] V. Kumar, R. K. Jha, and S. Jain, "NB-IoT security: a survey," *Wireless Personal Communications*, vol. 113, no. 4, pp. 2661–2708, 2020.

[13] X. Hu, Y. Zhang, X. Liao, Z. Liu, W. Wang, and F. M. Ghannouchi, "Dynamic beam hopping method based on multi-objective deep reinforcement learning for next generation satellite broadband systems," *IEEE Transactions on Broadcasting*, vol. 66, no. 3, pp. 630–646, 2020.

[14] Y. E. L. Morabit, F. Mrabti, and E. H. Abarkan, "Survey of artificial intelligence approaches in cognitive radio networks," *Journal of Information and Communication Convergence Engineering*, vol. 17, no. 1, pp. 21–40, 2019.

[15] O. Maraqa, A. S. Rajasekaran, S. al-Ahmadi, H. Yanikomeroglu, and S. M. Sait, "A survey of rate-optimal power domain NOMA with enabling technologies of future wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2192–2235, 2020.

[16] T. M. Getu, W. Ajib, R. Landry, and G. Kaddoum, "Toward overcoming a hidden terminal problem arising in MIMO cognitive radio networks: a tensor-based spectrum sensing algorithm," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9833–9847, 2019.

[17] G. J. Sutton, J. Zeng, R. P. Liu et al., "Enabling technologies for ultra-reliable and low latency communications: from PHY and MAC layer perspectives," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2488–2524, 2019.

[18] M. A. Qureshi and C. Tekin, "Fast learning for dynamic resource allocation in AI-enabled radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 95–110, 2019.

[19] A. Bohloulzadeh and M. Rajaei, "A survey on congestion control protocols in wireless sensor networks," *International Journal of Wireless Information Networks*, vol. 27, no. 3, pp. 365–384, 2020.

[20] A. Mukhopadhyay, P. K. Singh, R. Sarkar, and M. Nasipuri, "Handwritten Indic script recognition based on the Dempster–Shafer theory of evidence," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 264–282, 2020.

[21] Y. Zhao, Z. Li, B. Hao, and J. Shi, "Sensor selection for TDOA-based localization in wireless sensor networks with non-line-of-sight condition," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9935–9950, 2019.

[22] S. Vimal, Y. H. Robinson, S. Kadry, H. V. Long, and Y. Nam, "IoT based smart health monitoring with CNN using edge computing," *Journal of Internet Technology*, vol. 22, no. 1, pp. 173–185, 2021.

[23] R. V. Priya, "Emotion recognition from geometric fuzzy membership functions," *Multimedia Tools and Applications*, vol. 78, no. 13, pp. 17847–17878, 2019.

[24] R. C. Shit, S. Sharma, D. Puthal et al., "Ubiquitous localization (UbiLoc): a survey and taxonomy on device free localization for smart world," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3532–3564, 2019.

[25] T. Akhtar, C. Tselios, and I. Politis, "Radio resource management: approaches and implementations from 4G to 5G and beyond," *Wireless Networks*, vol. 27, no. 1, pp. 693–734, 2021.

[26] H. Tongal, "Comparison of local and global approximators in multivariate chaotic forecasting of daily streamflow," *Hydrological Sciences Journal*, vol. 65, no. 7, pp. 1129–1144, 2020.

[27] P. Kumari, H. P. Gupta, and T. Dutta, "An incentive mechanism-based Stackelberg game for scheduling of LoRa spreading factors," *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 2598–2609, 2020.

[28] S. Shibu and V. Saminadan, "Clustering-based resource allocation scheme for dense femtocells (cradf) to improve the performance of user elements," *Wireless Personal Communications*, vol. 113, no. 2, pp. 1183–1200, 2020.

[29] M. al-Sharman, D. Murdoch, D. Cao et al., "A sensorless state estimation for a safety-oriented cyber-physical system in urban driving: deep learning approach," *IEEE/CAA Journal of AutomaticaSinica*, vol. 8, no. 1, pp. 169–178, 2021.

[30] W. Wang, Z. Gong, J. Ren, F. Xia, Z. Lv, and W. Wei, "Venue topic model–enhanced joint graph modelling for citation recommendation in scholarly big data," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 1, pp. 1–15, 2021.

[31] W. Wei, Q. Ke, J. Nowak, M. Korytkowski, R. Scherer, and M. Woźniak, "Accurate and fast URL phishing detector: a convolutional neural network approach," *Computer Networks*, vol. 178, article 107275, 2020.

[32] S. H. Ahmed, V. H. C. de Albuquerque, and W. Wei, "Guest editorial: special section on advanced deep learning algorithms for industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2764–2766, 2021.

[33] H. Hu, B. Tang, X. Gong, W. Wei, and H. Wang, "Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 2106–2116, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6661926, 12 pages
https://doi.org/10.1155/2021/6661926

Research Article

Distributed Path Planning of Unmanned Aerial Vehicle Communication Chain Based on Dual Decomposition

Xiaohua Wei and Jianliang Xu

School of Mechanical and Electrical Engineering, Quzhou College of Technology, Quzhou, Zhejiang 324000, China

Correspondence should be addressed to Xiaohua Wei; 377036471@qq.com

Received 29 December 2020; Revised 20 May 2021; Accepted 25 May 2021; Published 7 June 2021

Academic Editor: Wei Wang

Limited by the insufficiency of single UAV's load and flight time capabilities, the multi-UAV (unmanned aerial vehicle) collaboration to improve mission efficiency and expand mission functions has become the focus of current UAV theory and application research. In this paper, the research on UAV global path planning is carried out using the ant colony algorithm, and an indoor UAV path planning model based on the ant colony algorithm is constructed. In order to improve the efficiency of the algorithm, enhance the adaptability and robustness of the algorithm, a distributed path planning algorithm based on the dual decomposition UAV communication chain is proposed. This algorithm improves the basic ant colony algorithm from the aspects of path selection, pheromone update, and rollback strategy in view of the inherent shortcomings of the ant colony algorithm. In order to achieve the best performance of the algorithm, this paper analyzes each parameter in the ant colony algorithm in depth and obtains the optimal combination of parameters. The construction method of the Voronoi diagram was improved, and the method was simulated to verify that the method can obtain a Voronoi diagram path that is safer than the original method under certain time conditions. Through the principle analysis and simulation verification of the Dijkstra algorithm and the dual decomposition ant colony algorithm, it is concluded that the dual decomposition ant colony algorithm is more efficient in pathfinding. Finally, through simulation, it was verified that the dual decomposition ant colony algorithm can plan a safe and reasonable flight path for multiple UAV formation flights in an offline state and achieve offline global obstacle avoidance for multiple UAVs.

1. Introduction

UAV (unmanned aerial vehicle) is widely used in military warfare and people's livelihood due to its low cost, ease of operation, fearlessness, and danger [1]. Militarily, UAV can perform tasks such as battlefield reconnaissance and surveillance, ship escort, and air early warning [2]. It can also be used as bait to attract enemy firepower, or as a target aircraft for army training. In civil use, UAV can be used for aerial photography, disaster monitoring, traffic patrol, and security monitoring [3].

Algorithms are the core content of path planning problems. Excellent path planning algorithms can not only ensure the optimal global path but also accelerate the speed of path replanning when encountering unexpected situations, thereby improving the overall efficiency [4–6]. Commonly used path planning algorithms include genetic algorithm, probability map method, particle swarm algorithm, and artificial potential field method [7]. Among them, the genetic algorithm draws on the genetic laws of survival of the fittest in nature, through selection, crossover, mutation, and other means, and finally achieves an optimal solution [8]. The algorithm is simple in process, has many advantages, can perform fast and random search, and does not require high details of the problem. It has strong robustness and is easy to be combined with other algorithms. In recent years, great progress has been made in the research of UAV path planning based on genetic algorithm, but there are still some defects, such as slow convergence speed and premature fall into local optimality [9]. At the same time, because the genetic algorithm involves the selection of multiple parameters, the parameter selection is different, and the convergence effect of the

algorithm is different. If these parameters are not properly selected, the optimal solution will not be obtained [10]. For example, the selection of fitness function not only needs to consider the algebraic value of the flight path but also needs to consider the path feasibility [11]. Therefore, if it is not selected properly, it may not get the global optimal and converge to the local optimal.

Related scholars have proposed a double genetic algorithm mechanism [12]. The method achieves two evolutionary goals based on different fitness functions and plans paths for static and dynamic threats in the environment, which can find the optimal path while avoiding threats [13]. The researchers applied genetic algorithms to the solution of path planning problems and used polar coordinates to describe the location of path points and threats, which improved the efficiency of path planning [14]. Relevant scholars solved the problem of avoiding obstacles for robots by manually setting gravitational fields and repulsive fields [15]. The basic idea of this method is to abstract UAV motion in the mission environment as a kind of motion in a man-made field, construct a gravitational field for the target point, so that it produces "gravitation" for UAV, and the threat area has a repulsive field [16]. It produces "repulsive force" and guides the UAV's flight through the combined force. This method can realize fast control, so it is widely used for real-time motion control. The planned path is safe and smooth, but when the threat is close to the target area, because the repulsive force is greater than the gravitational force, the UAV will not be able to reach the target position, and this method may fall into local optimum [17]. Relevant scholars generate path points, construct a connected graph in free space, and convert the solution space of the path planning problem into a topological space, so that the complexity of the problem is independent of the complexity of the environment and the dimension of the planning space [18]. The disadvantage of this method is that when the obstacles in the planning space are dense or there are narrow passages, the efficiency of the method will become low [19, 20]. In addition, because this method randomly samples the signpost nodes when constructing the connected graph, it is easy to cause the final search path to deviate from the optimal.

Depending on the degree of master of environmental information, path planning can be divided into global path planning and local path planning, environmental information offline completed entirely known cases before the flight path planning is called the global path planning, environmental information is not completely known. Before global path planning, it is necessary to master comprehensive environmental information and carry out path planning according to all information in the map. Local path planning does not require known comprehensive environmental information. Instead, sensors collect environmental information in real time to determine the distribution of obstacles around the aircraft and carry out path planning in the local range. The research objects of path planning include robots, ground vehicles, underwater submarines, and aircraft. The path planning of aircraft is usually called track planning, while in the field of indoor microunmanned aerial vehicle autonomous navigation, the flight planning process is usually called path planning rather than track planning.

In this paper, we use the ant colony algorithm to study UAV global path planning and carry out in-depth and detailed theoretical research and analysis on the key problems in the ant colony algorithm. In order to solve the problem that the ant colony algorithm has a slower initial solution and is easy to fall into local optimization, a UAV global path planning algorithm is proposed. And we optimize the selection of important parameters in the algorithm, so that the algorithm has the best results in terms of optimization ability and algorithm efficiency. On the basis of identifying obstacle modalities in the global environment, the safety circle is delimited for the obstacle, and the safety circle area is set as the dangerous area. In the process of generating the Voronoi diagram path diagram, it focuses on the improved method of finding the delta center of the duality generation method after generating Delaunay triangulation elements, and the method was simulated to verify that the method can obtain a safer Voronoi diagram path. Through simulation verification, it is concluded that the dual decomposition ant colony algorithm in the global path search is more efficient than the Dijkstra algorithm in time and space. For this topic, the dual decomposition ant colony algorithm is selected as the global obstacle avoidance method.

The organizational structure of this paper is as follows: in the second part, the related theories and key methods are introduced. In the third part, the global path planning algorithm of UAV based on the ant colony algorithm is designed. In the fourth part, the methods proposed in this paper are simulated, and the results are analyzed and discussed. Finally, the full text is summarized.

2. Related Theories and Key Methods

2.1. Dual Optimization Theory. The optimization theory in mathematics is to discuss what kind of plan is the best and how to find the best plan among many plans. The dual optimization theory is an important branch of optimization theory. The following is a brief introduction to the dual optimization theory to provide a basis for the research discussion.

The general optimization problem is as follows:

$$\begin{matrix} \min & f_0(x) \\ f_i(x)<,i=1,2,\cdots m & h_i(x)=0, i=1,2,\cdots,p \end{matrix}. \tag{1}$$

The basic idea of Lagrange duality is to extend the objective function by weighting the constraint function. The Lagrange function of the optimization problem is defined as follows:

$$L(x,v,\lambda)=\sum_{i=1}^{p} v_i h_i(x)+\sum_{i=1}^{m} f_i\lambda_i(x)+f_0(x), \tag{2}$$

where λi and vi are both Lagrangian multipliers, and vectors λ and v are called dual variables or Lagrangian multiplier vectors. This process actually gives the general construction

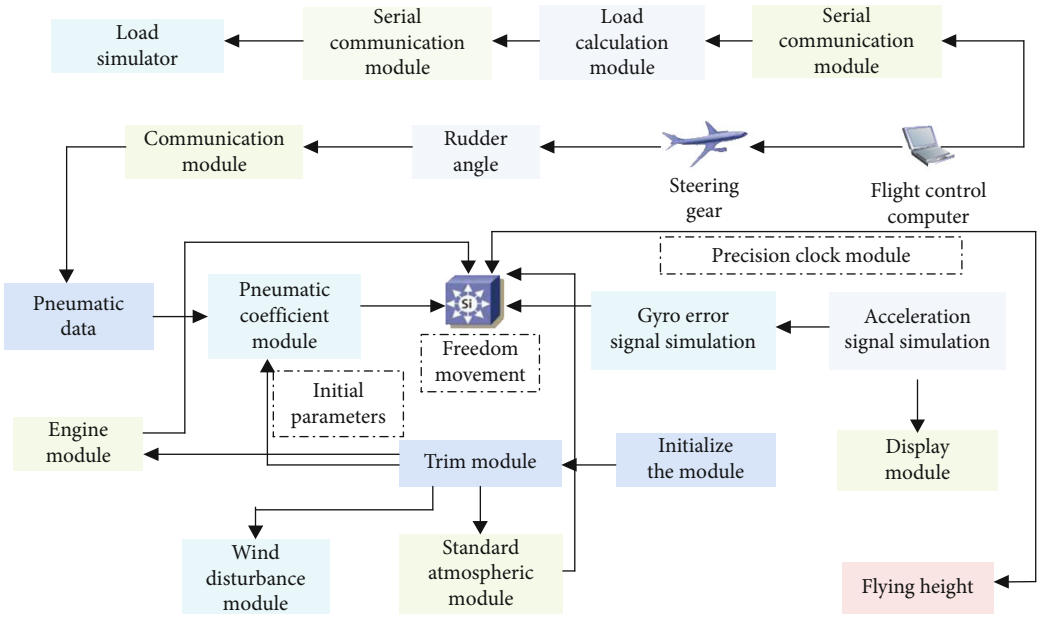


Figure 1: UAV flight simulation system block diagram.

method of the Lagrangian function. Further, the Lagrange dual function needs to be defined as follows:

$$h(v,\lambda) = \inf\left(\sum_{i=1}^{p} v_i h_i(x) + \sum_{i=1}^{m} f_i \lambda_i(x) + f_0(x)\right). \quad (3)$$

The value of x in the Lagrangian function defined above should be within the specified range. If the Lagrange function does not specify a lower bound on x, then the value of the function may include $-\infty$. It can be seen from the form defined above that the dual function is a piecewise continuous lower bound of a set of affine functions with (λ, v) as independent variables, regardless of whether the original optimization problem is a convex function. After transformation, its Lagrangian dual function must be convex.

No matter whether the initial problem is convex or not, the objective function of the Lagrange dual problem is concave, and the constraints are also convex. Therefore, its optimization problem must be a convex optimization problem. According to the theory of convex optimization, the only optimal solution must be obtained, and this result actually guarantees the solvability of the dual optimization problem.

When the optimization of the original problem is transformed into that of its dual problem, the optimization values of the two problems may not be equal. This phenomenon needs to be discussed and analyzed. The next conversion can ensure that the original problem and the dual problem are equal or determine the cause of the difference between the original problem and the dual problem.

2.2. Structural Analysis of UAV-WSN System. The WSN (wireless sensor network) node is usually composed of a microprocessor module, a data storage module, a communication module, a sensor module, and a power supply. The microprocessing module is mainly composed of a bus and an analog-to-digital converter. It is generally used as an IO interface. The storage modules of WSN nodes are generally relatively small, mainly used to store data and programs. The storage module of the WSN node is mainly limited by its size and power consumption. If the WSN node performs read and write operations, the power consumption will increase greatly, so the design of the WSN node must consider the needs of its application. The main modulation methods used by the WSN communication module are amplitude shift keying and frequency shift keying. Because energy consumption is the most prominent problem restricting the use of WSN nodes, when WSN nodes are not communicating, in order to prevent the nodes from wasting too much unnecessary monitoring and achieve energy-saving purposes, WSN nodes are in a dormant state. When the WSN node is activated, it starts to communicate. The WSN sensing module mainly includes temperature sensing, pressure sensing, light sensing, and other methods. The WSN application mainstream sensing module is MEMS (microelectromechanical systems) module. The advantage of MEMS module lies in its small size, low cost, and relatively accurate information about the surrounding environment, which has been widely welcomed by people. The operating system of the WSN node is mainly an embedded operating system, which also has other operating systems. Because the data processed by the WSN node is not very complicated, but its power consumption, size, and storage capacity are strictly limited, so the operating system of the WSN node requires high efficiency and certain fault tolerance. The block diagram of the UAV flight simulation system is shown in Figure 1.

Referring to the TCP/IP network hierarchy model, WSN can also be divided into the physical layer, data link layer, and network layer. Among them, the physical layer and the data link layer are mainly to study how to make the energy more efficient when WSN nodes communicate with each other.

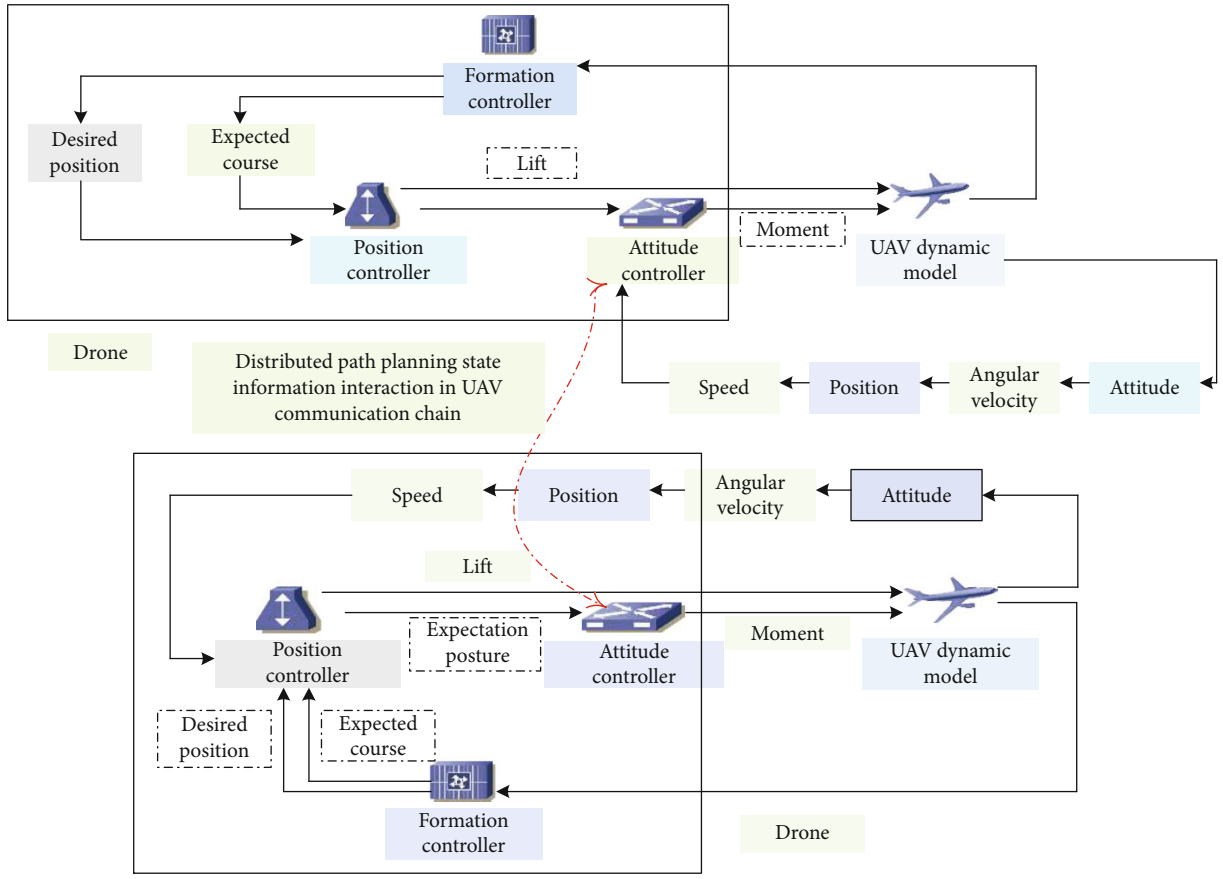


Figure 2: UAV formation control structure diagram.

The research on the network layer of WSN mainly focuses on the Routing method. At the network layer, WSN nodes mainly communicate in a multihop manner, and energy has always been one of the most concerned topics in WSN networks. The main research of the network layer is the overall energy of the entire network. If the WSN nodes are fixed, then when collecting information, the information needs to be stored and forwarded through neighboring nodes, which will generate a lot of energy consumption. If mobile nodes are introduced in this network to collect information, UAV is undoubtedly one of the best choices. With UAV's strong maneuverability and sufficient energy supply, it can act as a powerful node with information collection in the network that can move freely and have unlimited energy, which can effectively reduce the energy consumed by information forwarding in the network and extend the entire network.

WSN nodes are mainly randomly deployed in the monitoring area through manual setting or high-altitude broadcasting. WSN nodes form a wireless sensor network mainly through self-organization. The WSN node has the ability to sense, collect, store, and process data. It can process and analyze the information in the monitoring area and can also return the data to the UAV in a multihop manner. The UAV has a relatively strong emission capability and high energy. At the same time, the UAV can also be equipped with a corresponding processor, which has data storage and processing capabilities, can collect data in the entire area, and can transmit data information to the observer. The observer analyzes the received data through the terminal of the network and draws a conclusion to make corresponding changes to the perceived objects. Therefore, the UAV-WSN architecture should include WSN nodes, UAVs, observers, and perceived objects. The structure diagram of UAV formation control is shown in Figure 2.

2.3. Analysis of the Characteristics of UAV-WSN System. UAV-WSN system is mainly composed of WSN nodes in the network. As a sink node, UAV collects information and sends data back to the observer. WSN nodes can also be used as interrupt nodes to route and forward information from WSN nodes that UAV cannot receive. The WSN nodes are connected peer-to-peer.

(1) Large scale and wide range

In order to monitor the situation in a certain area, it is usually necessary to deploy a large number of wireless sensor nodes in the monitoring area. Sometimes the number of wireless sensor nodes may reach thousands or even more. On the one hand, the geographical area of monitoring is particularly broad, and a large number of nodes need to be deployed for monitoring, and then, UAV is used to collect information. For example, deploying WSN nodes in the original forest for monitoring, it is necessary to deploy a large

number of WSN nodes and, then, use UAV high mobility for information collection. If the space is not very large, but the density requirements of the WSN nodes that need to be deployed are very high, a large number of WSN nodes need to be placed, and UAV can effectively reduce the energy consumed by information forwarding. The advantage of this is that by adjusting the communication distance and angle, the information can have a greater signal-to-noise ratio, and the distributed processing of a large amount of collected information can also improve the accuracy of monitoring and improve the fault tolerance of the entire system. The high dependence on the sensor of a node makes the whole system in a normal working state; the presence of a large number of redundant nodes, even if an individual node fails, will not greatly affect the entire system, which also makes the system very strong fault-tolerant performance; when a large number of nodes cover the monitoring area, the monitoring area is in full coverage, which effectively reduces blind spots and improves the monitoring capacity of the entire system. The schematic diagram of the UAV positioning system is shown in Figure 3.

(2) Dynamic network

The WSN topology is not fixed and may change due to the following factors: first, the sensor node malfunctions or fails due to the harsh environment of the wireless sensor location or the battery exhaustion; second, the changes in environmental conditions may affect the wireless. The mutual communication bandwidth between sensors affects the quality of communication, and when it is disconnected, it makes the topology of the entire WSN network unstable; third, relative to a fixed wireless sensor, if the sensor can be moved, which brings uncertainty to the entire system, then the network structure will also change with the changes of the members in the system; fourth, when new wireless sensor nodes are added in the network, then the existing network must be reorganized to form a new topology structure, which also shows that the wireless sensor network has dynamic system reconfigurability.

Because UAV can move arbitrarily, the network topology is also changing at any time. This change is not only limited to the communication between UAV and WSN but also related to the transmission of information between WSN nodes. Because the communication bandwidth of UAV is fixed, within a certain range, the number of communication received by WSN nodes is also very limited, which requires that the information between WSN nodes needs to be collected and sent to UAV.

3. UAV Global Path Planning Algorithm Based on the Ant Colony Algorithm

3.1. The Ant Colony Algorithm. The process of ants searching for food sources can be described as follows: in the absence of pheromone information, individual ants randomly select one of the paths when they pass the intersection and return to the original path after finding the food source. In this process, a large number of ants will leave more pheromones along the same path. Therefore, the probability of subsequent ants choosing a short path will be relatively large. As more and more ants choose a shorter path, the concentration of pheromones on this path will continue to increase, and because pheromones are a volatile substance, the pheromone on other paths will gradually dissipate with the passage of time, and eventually, all ants will choose the shortest path forward. This collective behavior of the ant colony exhibits a self-catalytic enhancement effect: the more ants choose a path, the more attractive this path will be for subsequent ants, so a positive feedback is formed. Positive feedback ultimately selects the shortest foraging path. Figure 4 shows the process of ants finding the shortest foraging path in an obstacle environment.

In the ant colony algorithm, the concept of artificial ants was abstracted according to the ants in the biological world, and some key ant behavior characteristics in the ant colony foraging process were given to artificial ants. Worker ants have some characteristics that real ants do not have. Each artificial ant is a simple individual satisfying the following behavioral characteristics:

(1) The movement of artificial ants is a transition between two discrete states

(2) The artificial ant moves along the adjacent position. When it is at any position, it selects the next visited position according to a certain probability. This probability is related to the pheromone concentration and distance between the two positions

(3) Artificial ants release pheromones along the way during movement

(4) The artificial ant has a memory function, and it will not select the location that has been visited during a pathfinding process

In the ant volume system, the pheromone increment is inversely proportional to the length of the path, so that more pheromone is obtained on the shorter path, thereby generating greater attraction for subsequent ants.

The ant-week system model releases pheromones after the ants have established a complete trajectory. The pheromone updating formula is as follows:

$$\tau_{ij}(n+t) = \Delta\tau_{ij}(n+t,t) + (1+\rho) \times \tau_{ij}(t), \tag{4}$$

$$\Delta\tau_{ij}(t+n,t) = \sum_{k=0}^{m} \Delta\tau_{ij}^{k}(t+n,t), \tag{5}$$

where p is the pheromone volatilization coefficient. The addition of volatility coefficient can prevent the algorithm from premature convergence to a suboptimal solution to a certain extent, so as to ensure that the algorithm still has a probability to explore new areas in the space after iterating for a period of time. $\Delta\tau_{ij}^{k}(t+1,t)$ represents the amount of pheromone released by the kth ant on the branch (i, j).

The process of the ants searching for the shortest path from the ant nest to the food source is the result of the

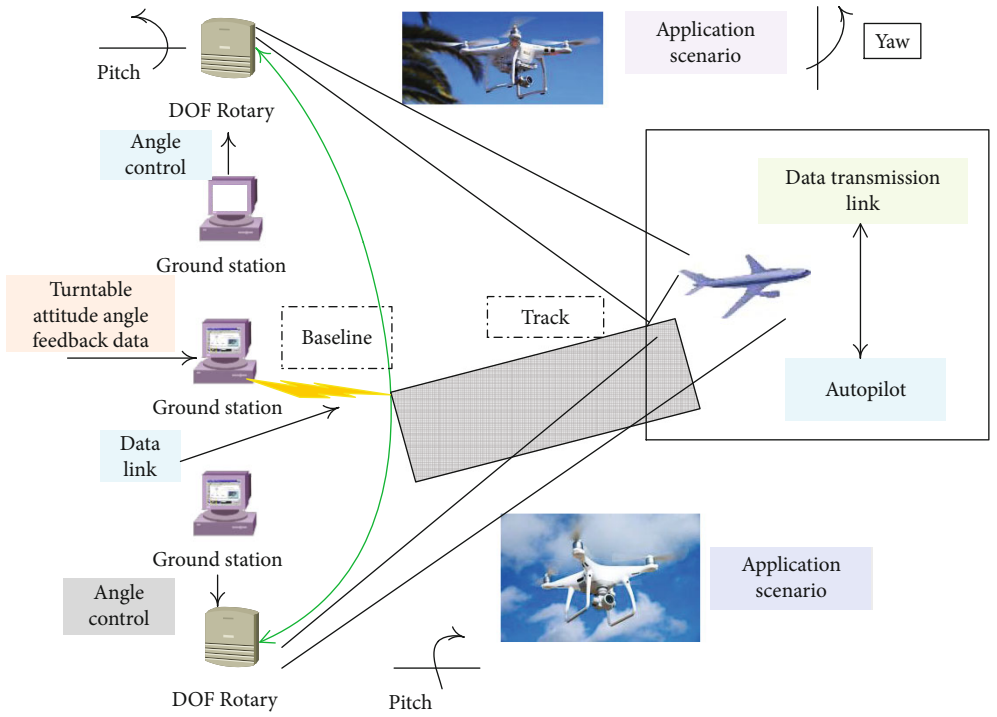


Figure 3: Schematic diagram of UAV positioning system.

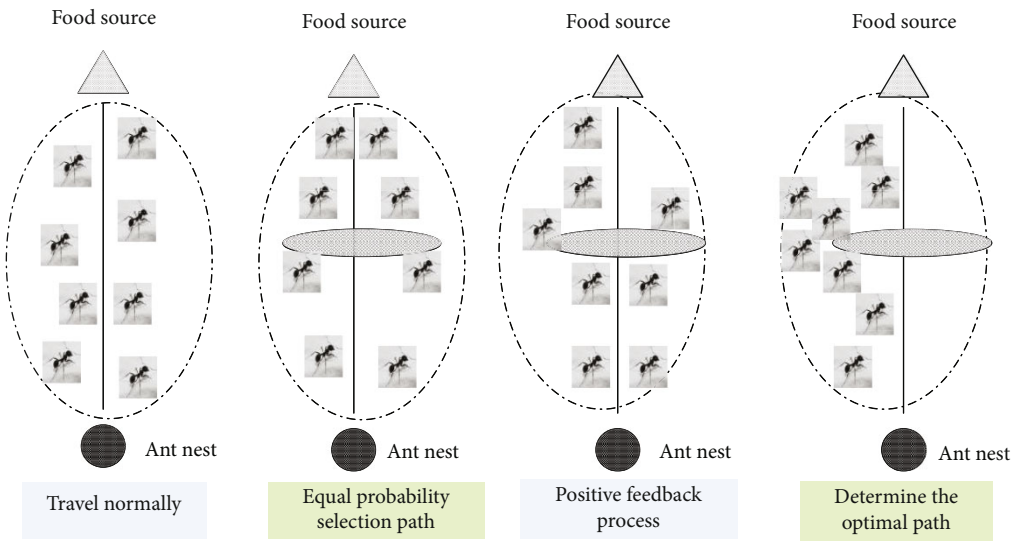


Figure 4: Schematic diagram of ant colony searching for food source.

interconnection, mutual influence, interaction, and mutual cooperation between a large number of ant individuals. The ant colony algorithm uses artificial ants to simulate the path-finding behavior of ants, which is essentially characterized by positive feedback and good parallelism. At the same time, as an intelligent bionic algorithm, the ant colony algorithm has very good optimization capabilities. The performance of the solution of the ant colony algorithm increases with the increase of the number of ants, until it reaches the upper limit, that is, the ideal situation that can always find the optimal solution. In addition, through reasonable improvements in parameter settings, probability calculation heuristics, and pheromone update strategies, the algorithm can also be improved in terms of convergence speed and search capabilities.

However, the ant colony algorithm itself is a probabilistic algorithm, and it is difficult to mathematically prove its correctness and reliability systematically. When the ant colony algorithm solves the optimization problem, the high-level search behavior of the algorithm is generated by the mutual influence of the simple behavior of the underlying ants. The design of the search algorithm for a single ant is simple, but the control of the entire system is not easy. The designer must be able to map the complex optimization behavior of the

entire system to the simple behavior of the individual ants at the bottom and achieve the desired by controlling the bottom search process. High-level behavioral characteristics are more difficult.

3.2. UAV Path Planning Algorithm Design Based on the Ant Colony Algorithm. The ant colony algorithm has the advantages of positive information feedback, good parallelism, and optimization capabilities, but it still has some shortcomings. For example, due to the lack of pheromones in the early stage, the movement of the ant colony lacks guidance information, and the solution speed is slow. With the effect of feedback, the algorithm is prone to fall into local extreme value. In view of the shortcomings of the basic ant colony algorithm and the specific application object of UAV path planning, this article made some adjustments and improvements to the basic ant colony algorithm.

The ant colony algorithm uses the evolutionary process of the colony to search for the optimal solution. A single ant builds a feasible solution and adds it to the feasible solution set after a pathfinding process. The solution constructed by m ants after a cycle forms a subset of the feasible solution set. Therefore, the larger the number of ants, the larger the subset constructed, the stronger the global search ability of the ant colony algorithm, and the better the stability of the algorithm. In the actual application of the algorithm, if all ants are set to release pheromones, when the number of ants is too large, although the randomness of the search is strengthened, the change of pheromones on each edge tends to be average, which weakens the positive feedback effect of information and reduces the convergence speed. In this paper, we have improved the pheromone global update mode to only release pheromones by the iterative optimal ants, and there is no problem that a large number of ants weaken the positive feedback. However, the increase in the number of ant population will inevitably lead to the increase in the time complexity of the algorithm, and the time spent in each iteration will increase. On the contrary, if the number of ants is too small, the subset of the problem constructed is too small, the randomness of the search path is weakened, and the global optimal solution may not be found when the size of the problem is large. Therefore, the determination of the number of ants is a balance between the speed of convergence and the searching ability.

According to the definition of current visibility, the visibility value of a candidate grid is only related to its distance from the current grid. In this way, there is a problem: when the ants are in any grid, the ants will always prefer to select the grid with the smallest distance from the current grid, regardless of the grid and the target point. When the initial pheromone concentration is low and it is impossible to give enough guidance information for the ant's path selection, the ant selects the grids in these directions with almost the same probability to start the search. Due to the lack of information related to the distance of the destination as a guiding factor, the ant's path search is aimless, and it often takes a large circle in the environment map to find the destination, which not only results in a slow initial search speed of the algorithm but also a low search ability. It is difficult to search for the true optimal path.

In response to the above problem, the visibility from grid r to grid s in the transition probability calculation formula is changed to the visibility between the candidate grid and the target grid. After adopting this strategy, the ants no longer blindly conduct path search, but preferentially select the grid closest to the target point in the candidate grid set, so this strategy is called the target attraction strategy.

In the ant colony algorithm, each ant individual uses a greedy heuristic to perform a path search. Each ant can find a feasible solution. The ant individuals in the ant colony have established a large number of solutions at the same time. The higher-quality solutions are rewarded, and the pheromone on its path is more strengthened, so that through continuous positive feedback, the solution of the problem gradually evolves toward the direction of global optimality until the optimal solution is found.

The above is the principle of the optimal solution of the ant colony algorithm in the ideal case. However, in fact, the ant colony algorithm does not always find the global optimal solution. In some cases, the positive feedback of the system may lead to premature convergence of the solution. For example, because there is a local optimal solution in the system, or just because of random oscillations in the early iterations, some nonoptimal individuals in the group affect the entire group, preventing the group from further exploring the global optimal solution.

After adopting the goal attraction strategy and defining the visibility with the reciprocal of the distance between the candidate grid and the target grid, the premature convergence problem will appear in some specific cases: when there is a local optimal solution in the algorithm, this solution is global; when the path length difference of the optimal solution is not large, the path of the ant colony will converge to this local optimal solution.

This article proposes a new global pheromone update rule as follows:

$$\tau_{rs}(t+n) = \Delta\tau_{rs}(t, t+n) \times \rho + (1+\rho) \times \tau_{rs}(t+1). \quad (6)$$

In the optimization process of the ant colony algorithm, with the enhancement of positive feedback, the pheromone on the edge of the optimal solution (it may also be actually a suboptimal solution) continues to accumulate, and the pheromone on the other edge continuously volatilizes. The gap is getting bigger and bigger, and the ants will repeatedly construct the same solution, search stagnation, and no longer explore new solutions. Therefore, slowing down the increasing speed of the pheromone gap between the edge searched by more ants and the edge rarely searched can ensure that the system evolves at a reasonable speed, avoiding the premature convergence phenomenon caused by the positive feedback speed being too fast appear.

In the ant's path search process, sometimes, there is no way to go after the ant has reached a certain position, that is, the candidate grid set allowed is an empty set, the algorithm is stuck in a deadlock state, and the path search is stalled. Special consideration should be given to this

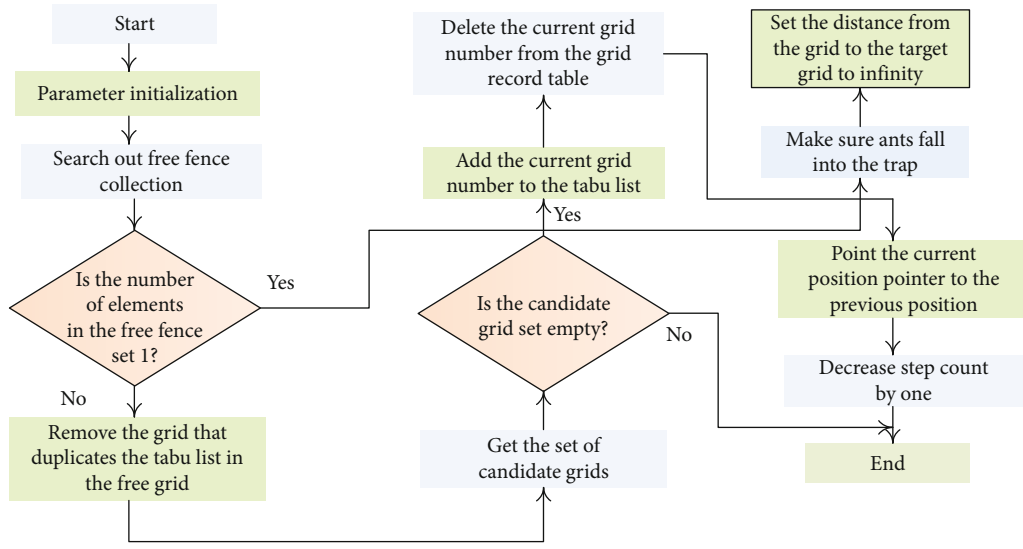


Figure 5: Flow chart of path rollback strategy.

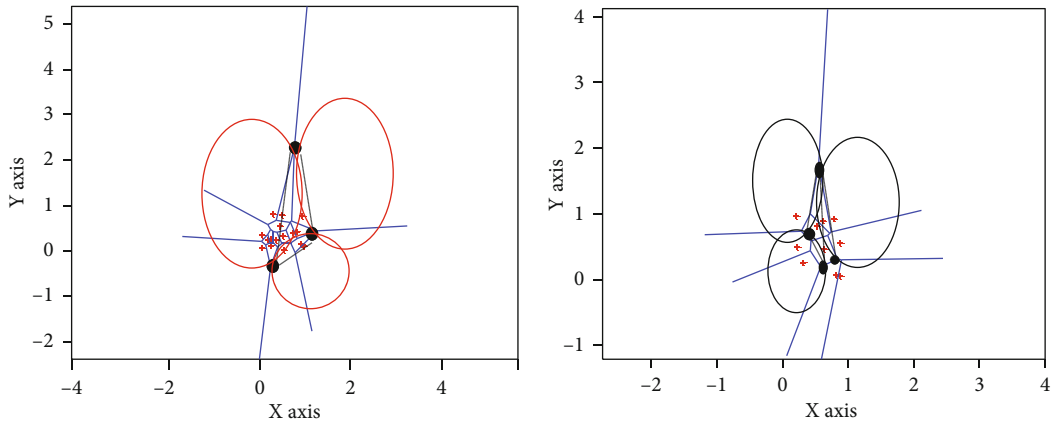


Figure 6: Voronoi diagram construction results. (a) Voronoi diagram construction results of the original dual generation method. (b) Voronoi diagram construction results of the ant colony dual generation method.

situation, and a certain backoff strategy should be set to make the algorithm adapt to various complex environments and improve the robustness of the algorithm. The flow chart of the path rollback strategy is shown in Figure 5.

4. Simulation and Discussion

4.1. Simulation of the Method of Generating V Map. This paper mainly studies the obstacle avoidance of multiquadrotor UAV formation on the basis of formation control constraints. In the obstacle avoidance research, it involves the global flight path planning and local obstacle avoidance. Then, the Dijkstra algorithm and A* algorithm are introduced and compared. Finally, A* algorithm is used to simulate the global path planning of multi-UAV formation.

In order to test whether the ant colony duality generation method can obtain a more secure V-graph path, this section simulates the original duality generation method and the ant colony duality generation method under the abovementioned simulation test conditions and compares them.

Figure 6(a) is the simulation result of the original dual generation method under the given obstacle environment. Among them, the black dot is the center of the obstacle safety circle, the black dotted line is the Delaunay triangulation generated by the duality generation method, the red circle is the circumscribed circle of each triangle in the Delaunay triangulation, and the solid black line is the Voronoi diagram finally obtained path results.

Figure 6(b) shows the simulation results of the ant colony dual generation method in the obstacle environment. Among them, the black dot is the center of the safety circle of the obstacle, the black dotted line is the Delaunay triangulation generated by using the even generation method with the center of the safety circle as the vertex, the black circle is the circle tangent to the three adjacent safety circles, and the solid blue line is the final Voronoi diagram path result.

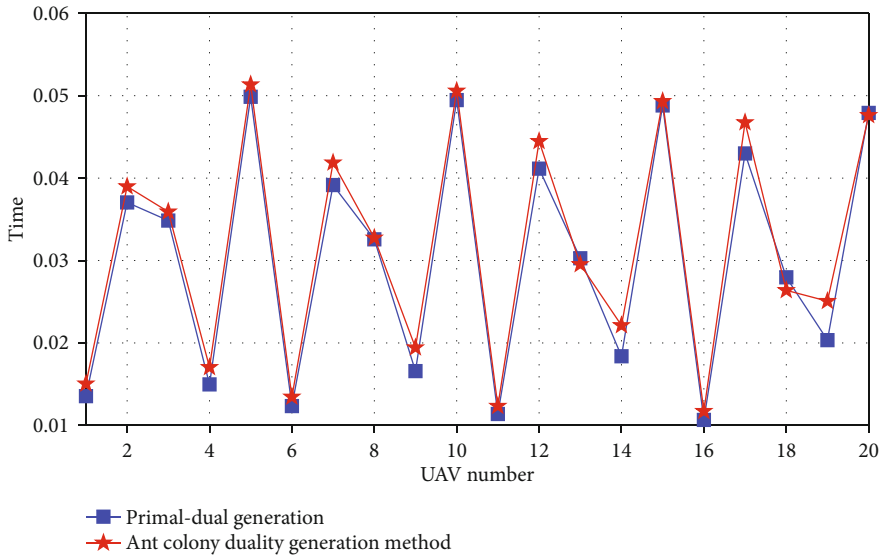


Figure 7: Comparison of time before and after improvement of the duality generation method.

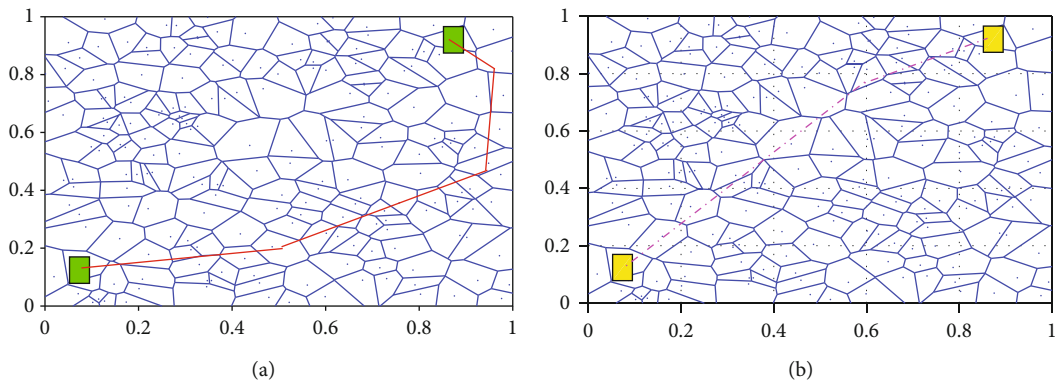


Figure 8: Comparison of simulation results in a multidimensional environment. (a) The Dijkstra algorithm. (b) The dual decomposition ant colony algorithm.

From the experimental results, it can be seen that although the path maps obtained by the two methods are not very different overall, the original dual generation method produced a path that intersects the safety circle of the obstacle. It can be seen from the figure that compared with the original dual generation method, the improved dual generation method generates a Voronoi diagram with an equal safety circle distance to the obstacle, making the obtained Voronoi diagram path safer.

Figure 7 shows the time used by the two methods in the above experiment. It can be seen that the ant colony duality generation method is not as good as the original duality generation method in terms of time performance. But in terms of UAV offline path planning, there is enough time to do this path planning, the time requirement is not very high, and UAV has a certain flying height; at this height, the amount and scale of obstacles are usually not very large, so the improved dual generation method can get a safer Voronoi diagram path under certain time conditions.

4.2. Simulation Comparison between the Dijkstra Algorithm and Dual Decomposition Ant Colony Algorithm. In order to compare the performance of the Dijkstra algorithm and dual decomposition ant colony algorithm in global path planning, this section selects multidimensional simple environment with one obstacle by analyzing the predecessors' setting of the number of obstacles in the path planning simulation. Simulation experiments were carried out on these two environments under the abovementioned simulation experiment conditions.

Because the simulation of directly inputting the Voronoi diagram is difficult to visually observe, so when doing the simulation, this article will rasterize the V diagram, the center of each square is the vertex of the Voronoi diagram, and the connection between the centers is the V diagram. The simulation effect is shown below.

Figure 8(a) is the simulation result of the Dijkstra algorithm, and Figure 8(b) is the simulation result of dual decomposition ant colony algorithm. It can be seen that in this example, the Dijkstra algorithm accesses more nodes than the dual decomposition ant colony algorithm, and in the results obtained, the dual decomposition ant colony algorithm finds more paths than the Dijkstra algorithm finds.

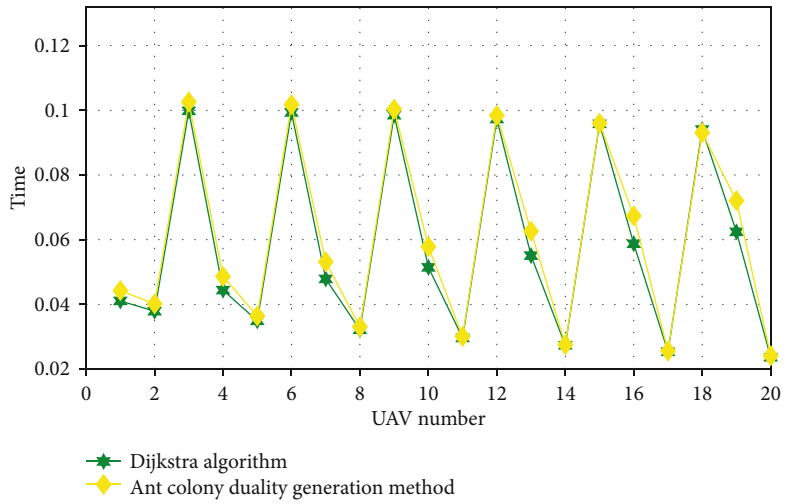


Figure 9: Comparison of time between the dual decomposition ant colony algorithm and Dijkstra algorithm.

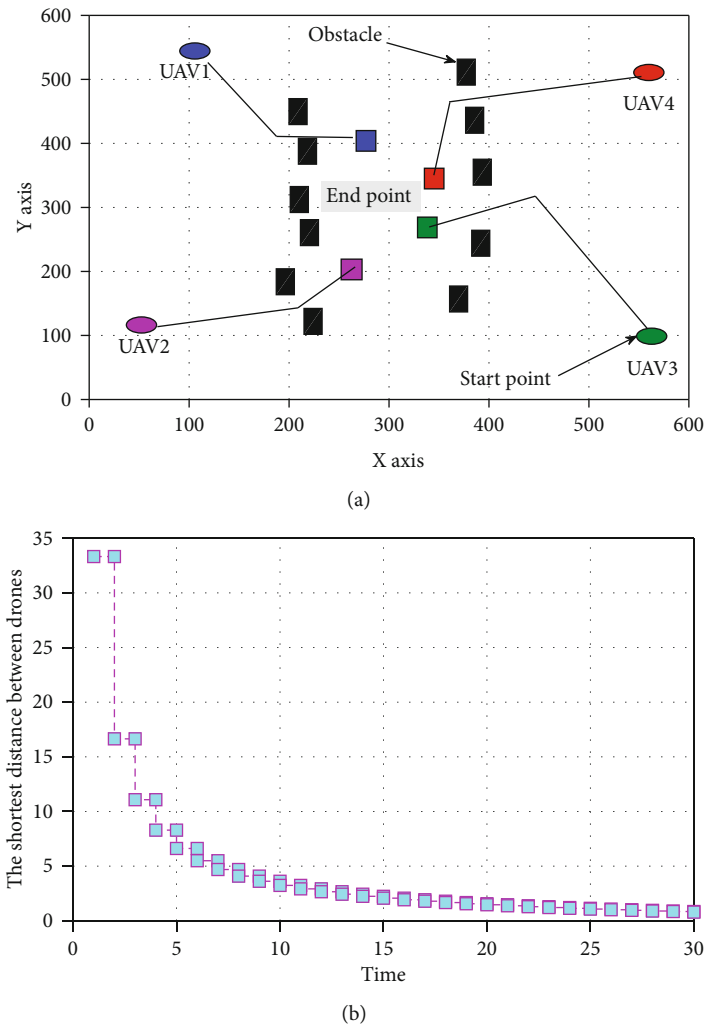


Figure 10: The results of UAV distributed path planning under dual decomposition ant colony algorithm. (a) Planned trajectory results. (b) Curve of the shortest distance between drones.

Figure 9 shows the average program elapsed time calculated by running the Dijkstra algorithm and the dual decomposition ant colony algorithm ten times in this simulation experiment. It can be seen from Figure 9 that the dual decomposition ant colony algorithm is more efficient than the Dijkstra algorithm in terms of its own time.

Through the simulation comparison of the Dijkstra algorithm and dual decomposition ant colony algorithm in a relatively complex obstacle environment of 20 dimensions, it is concluded that the dual decomposition ant colony algorithm is superior to the Dijkstra algorithm in path optimization and time consumption. The ant colony algorithm is more efficient. Therefore, this paper chooses the dual decomposition ant colony algorithm to realize the global path planning for the UAV cluster.

4.3. The Dual Decomposition Ant Colony Algorithm to Find the Optimal Path of the Cluster Simulation. This section will use dual decomposition ant colony algorithm to simulate multi-UAV formation path planning based on application scenarios and formation rules.

Considering the two typical tasks of UAV formation assembly and formation reconstruction, the effectiveness of the sequential convex optimization method to solve the collaborative trajectory planning problem is verified. Formation assembly requires that the drones in the formation arrive at the designated assembly target point from their respective initial positions at the same time and form the desired formation, which provides conditions for the formation to perform subsequent collaborative tasks. Formation reconstruction requires that the UAV formation change its formation through coordinated flight to meet the mission or safety needs.

It can be seen from Figure 10(a) that the trajectory obtained by the dual decomposition ant colony algorithm planning method can realize the transfer from the starting point to the target point and avoid obstacles in the environment. The roundabout approach guarantees the coordination of the entire formation time. According to the variation curve of the shortest distance between drones in Figure 10(b), the shortest distance between formations is always maintained above the safe distance, that is, the planned trajectory of the UAV formations can meet the collision avoidance constraints. Therefore, the dual decomposition ant colony algorithm can plan a feasible cooperative trajectory that satisfies the constraints for the UAV formation.

5. Conclusions

In this paper, we use the popular ant colony algorithm in the path planning field to study the indoor UAV global path planning problem; theoretically analyze the characteristics, advantages, and disadvantages of the ant colony algorithm in detail; and propose a distributed UAV communication chain path planning algorithm. The algorithm improves the basic ant colony algorithm in path selection strategy, pheromone update, fallback strategy, etc., and enhances the algorithm's adaptability to the environment and the algorithm's robustness, making the algorithm's ability to achieve the best condition. The ant colony duality generation method is compared with the original duality generation method, and the improved method can obtain a safer flight path Voronoi diagram under a certain time limit. In a simple environment with one obstacle and a complex environment with multiple obstacles, the simulation comparison between Dijkstra's algorithm and the dual decomposition ant colony algorithm is obtained, and the dual decomposition ant colony algorithm is obtained in the path optimization and time consumption. Both are superior to the Dijkstra algorithm, and the dual decomposition ant colony algorithm is more efficient. To verify the effectiveness of the dual decomposition ant colony algorithm in multi-UAV path planning, in the application scenario proposed in the previous chapter (under the environment of four obstacles and four UAVs), the dual decomposition ant colony algorithm is used. The four-rotor UAV formation flight has been path-planned, and it is concluded that the dual decomposition ant colony algorithm can plan a safe flight path that conforms to the multirotor UAV formation rules. The real-time programming algorithm in this paper is designed based on the assumption that the sensor detection range can only support one-step programming. In future research, an online planning algorithm can be designed according to the actual detection range of the sensor, which can be used for path optimization within a certain range.

Data Availability

No data support is available.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

We declare that there is no conflict of interest.

Acknowledgments

This study was supported by the Scientific Research Project of Zhejiang Provincial Education Department (No: Y201839845), Science and Technology Plan Projects of Quzhou (No: 2018k12).

References

[1] H. Van Nguyen, H. Rezatofighi, B. N. Vo, and D. C. Ranasinghe, "Online UAV path planning for joint detection and tracking of multiple radio-tagged objects," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5365–5379, 2019.

[2] T. I. Zohdi, "The game of drones: rapid agent-based machine-learning models for multi-UAV path planning," *Computational Mechanics*, vol. 65, no. 1, pp. 217–228, 2020.

[3] K. Y. Kok and P. Rajendran, "Enhanced particle swarm optimization for path planning of unmanned aerial vehicles,"

ECTI Transactions on Computer and Information Technology (ECTI-CIT), vol. 14, no. 1, pp. 67–78, 2020.

[4] S. Primatesta, G. Guglieri, and A. Rizzo, "A risk-aware path planning strategy for uavs in urban environments," *Journal of Intelligent & Robotic Systems*, vol. 95, no. 2, pp. 629–643, 2019.

[5] H. Wang, J. Wang, G. Ding, J. Chen, F. Gao, and Z. Han, "Completion time minimization with path planning for fixed-wing UAV communications," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3485–3499, 2019.

[6] X. Peng, Y. Lan, J. Hu, F. Ouyang, K. Xiao, and Z. Gao, "Turning mode and whole region-coverage path planning and optimization of agricultural small UAV," *Journal of South China Agricultural University*, vol. 40, no. 2, pp. 111–117, 2019.

[7] M. d. S. Arantes, C. F. M. Toledo, B. C. Williams, and M. Ono, "Collision-free encoding for chance-constrained nonconvex path planning," *IEEE Transactions on Robotics*, vol. 35, no. 2, pp. 433–448, 2019.

[8] G. Zhang and L. T. Hsu, "A new path planning algorithm using a GNSS localization error map for UAVs in an urban area," *Journal of Intelligent & Robotic Systems*, vol. 94, no. 1, pp. 219–235, 2019.

[9] Y. Ding, B. Xin, and J. Chen, "Precedence-constrained path planning of messenger UAV for air-ground coordination," *Control Theory and Technology*, vol. 17, no. 1, pp. 13–23, 2019.

[10] R. K. Dewangan, A. Shukla, and W. W. Godfrey, "Three dimensional path planning using Grey wolf optimizer for UAVs," *Applied Intelligence*, vol. 49, no. 6, pp. 2201–2217, 2019.

[11] G. A. Thanellas, V. C. Moulianitis, and N. A. Aspragathos, "A spatially wind aware quadcopter (UAV) path planning approach," *IFAC-Papers OnLine*, vol. 52, no. 8, pp. 283–288, 2019.

[12] Y. Qian, F. Wang, J. Li, L. Shi, K. Cai, and F. Shu, "User association and path planning for UAV-aided mobile edge computing with energy restriction," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1312–1315, 2019.

[13] Y. Choi, Y. Choi, S. Briceno, and D. N. Mavris, "Energy-constrained multi-UAV coverage path planning for an aerial imagery mission using column generation," *Journal of Intelligent & Robotic Systems*, vol. 97, no. 1, pp. 125–139, 2020.

[14] S. Cho and D. H. Shim, "Sampling-based visual path planning framework for a multirotor uav," *International Journal of Aeronautical and Space Sciences*, vol. 20, no. 3, pp. 732–760, 2019.

[15] U. Challita, W. Saad, and C. Bettstetter, "Interference management for cellular-connected UAVs: a deep reinforcement learning approach," *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2125–2140, 2019.

[16] V. Jamshidi, V. Nekoukar, and M. H. Refan, "Analysis of parallel genetic algorithm and parallel particle swarm optimization algorithm UAV path planning on controller area network," *Journal of Control, Automation and Electrical Systems*, vol. 31, no. 1, pp. 129–140, 2019.

[17] U. Challita, A. Ferdowsi, M. Chen, and W. Saad, "Machine learning for wireless connectivity and security of cellular-connected UAVs," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 28–35, 2019.

[18] P. Oettershagen, J. Förster, L. Wirth, G. Hitz, R. Siegwart, and J. Ambühl, "Meteorology-aware multi-goal path planning for large-scale inspection missions with solar-powered aircraft," *Journal of Aerospace Information Systems*, vol. 16, no. 10, pp. 390–408, 2017.

[19] S. Kumar Debnath, R. Omar, N. B. Abdul Latip et al., "A review on graph search algorithms for optimal energy efficient path planning for an unmanned air vehicle," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 2, pp. 743–749, 2019.

[20] C. Liu, S. Zhang, and A. Akbar, "Ground feature oriented path planning for unmanned aerial vehicle mapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 4, pp. 1175–1187, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6415136, 11 pages
https://doi.org/10.1155/2021/6415136

Research Article

Optimization of Sports Good Recycling Management System Based on Internet of Things

Panhong Ren,[1] Mengjian Nie,[2] and Hui Ming[3]

[1]*Sanquan College of Xinxiang Medical University, Xinxiang 453003, China*
[2]*Henan Normal University, Xinxiang 453000, China*
[3]*Xinxiang Medical University, Xinxiang 453003, China*

Correspondence should be addressed to Hui Ming; 091049@xxmu.edu.cn

Received 10 April 2021; Revised 18 May 2021; Accepted 20 May 2021; Published 7 June 2021

Academic Editor: Wei Wang

Based on the technology of Internet of Things, this paper proposes a sports good recycling management system and realizes the system optimization. Based on the management of returned good flow based on the Internet of Things, we can comprehensively apply system integration technology, system architecture, cloud computing technology, etc., based on the current situation of surrounding return flow management, combined with the reality of the sports product recycling business process. From two aspects of process information and operation information, based on the information system of each link of the Internet of Things recycling, a recycling logistics information management system platform is constructed. This can solve the problem of the integrated management of the recycling logistics process and the operation control of the recycling logistics management system. In the stage of sports information management, combined with the Internet of Things technology, we should do a good job of sports artificial intelligence management and information knowledge exploration. In the development of sports under the framework of the Internet of Things, we should combine the needs of sports information management and the technology mode of the Internet of Things, so as to realize the scientific and standardized management of sports good recycling under the Internet of Things technology.

1. Introduction

As a green industry, the sports industry is called the sunrise industry and the civilian production industry. It is an important part of the modern industrial system. It has a large market demand and integrates related characteristics. It has become a national strategy in the context of national fitness. The development of the sports industry has also reached a new level, becoming the dominant source of power for the promotion of social and economic value. However, it must be admitted that sports goods involve complex links such as production, processing, circulation and sales, and lack of information correlation and sharing before each link. When quality and safety problems are found, products cannot be tracked and traced effectively to find the key to the problem [1]. Therefore, the sports good supply chain should grasp the great opportunity of developing information economy and promoting intelligent application and accelerate the realization of informatization and visual supervision of each link of the supply chain with the support of the new generation of information technology such as the Internet of Things [2]. With the continuous development of technology, robots and the development of the Internet of Things industry has been applied to all walks of life, in sport management industry, however, did not involve, traditional sports goods still stay in the human management stage, although this way of management directly, but brought great inconvenience to people, the most important thing is given at the end of the sports people to drive a little way to return the sports goods, this caused people tired body and mind pressure, then the traditional sporting good length because management is not caused by the loss of sporting goods, such not only increased the economic operating costs. And this increased the difficulty of work for people, hence the need for a comprehensive and fast sports good recycling management system.

The application of Internet of Things (IoT) technology is mainly to realize effective connection of objects on the basis of agreement, information exchange and communication, and, finally, to realize an intelligent identification process by combining information, information, communication, and sensing. Generally speaking, IoT technology should be mainly used in three layers, that is, the application of the perception layer and the network layer, as well as the application of the processing layer [3]. Based on the Internet of Things technology under a sports information management process, mainly combining with the form of content couplet net perception technology, the sports information acquisition effect rate increases, combined with the high-speed network of a cloud computing platform, a mode of realizing interconnection application, completes the sports information of intelligent control, and makes the realization of real-time monitoring and automated analysis. By combining the form of information collection, we can generate the record and do a good job of the relevant management of the event information, so as to show the effective application of the awareness layer and the network layer in the application of the multisingle bit cooperative supervision. A hardware perception process in the perceptual layer should combine the integration and application of information resource effectiveness, realize a comprehensive analysis of its own activities, combine a body system structure of historical records, and show the effectiveness of data sharing of interconnected body information [4]. The realization of a network layer structure, combined with the function of data backup and data mining, realizes the effective management of sports information. The application process of mass distribution of the Internet of Things shows the root analysis of the intelligent management layer, the effective application of the reweighting technology system, and the integration of the basic situation of the environment, so as to realize the whole exploration process. In a word, sports information management based on the Internet of Things technology, combined with the process of information system design and implementation evaluation, can realize the effective operation and guarantee of the information system. Pay attention to the realization of security management, realize the effective application of organizational information resources, and do a good job in the overall investment and evaluation of information technology. In the process of information system management, combined with the strategic development of organization competition, the management of sports information industry will be promoted comprehensively, and an effective application and management of sports information service industry will be realized.

Based on the above background, this paper firstly analyzes the operation process, network structure, and business links of the recycling logistics. Combined with the characteristics of the Internet of Things technology, it proposes the feasibility of using the Internet of Things technology to optimize the operation process and the management system. Secondly, the actual requirements of information construction of recycling logistics management are analyzed from the perspective of objects involved in recycling logistics, business process, and Internet of Things technology, and the framework of information platform of recycling logistics is designed based on business function requirements. We combined the actual sports to lead the sporting good recycling business. According to the framework developed by the information platform, the platform has achieved good results in the actual recycling business test. This paper proposes a sports good recycling management system based on the Internet of Things technology, which realizes the process informatization and operation informatization by combining the actual sports good recycling business process, integrated application system integration technology, system architecture, and cloud computing technology, etc. This can solve the problem of comprehensive management of the recycling logistics process and operation control of recycling logistics management system, to achieve the scientific and standardized management of sporting good recycling under the Internet of Things technology.

The rest of this paper is organized as follows: the related work is discussed in Section 2. Section 3 elaborates the recycling management mode of sporting good. In Section 4, based on the Internet of Things technology, a sports goods recycling management platform is established. The fifth part carries on the system optimization and the test. Section 6 summarizes the text.

2. Related Work

The close combination of Internet of Things technology and modern logistics technology also significantly improves the efficiency of the sports good logistics management and operation process, which is more convenient and safe, thus attracting the interest of researchers. Hou et al. [5] analyzed the influencing factors of sports product recycling management from the perspective of the value of recycled goods, constructed a mathematical model based on the influencing factors, and provided a reference analysis method for the value of different recycled products by using the analytic hierarchy process. Mahmood and Zubairi [6] in combination with the practical recycling logistics problems of lead-acid sporting goods, considering a fully closed recycling system of each production link, such as purchasing, production, scrap, recycling, and reprocessing, built an integrated management for solving minimum cost and emission pollution management cost and a minimum pollution impact model of a multiobjective fuzzy algorithm solution. Ren et al. [7] developed a returned logistics network model which is constructed to simulate the sports good recycling business in India, through a different stimulus allocation mechanism, with comprehensive consideration of forward and reverse returned logistics operation and design recovery system efficiency factors, through the simulation to predict recovery cost, transportation time, and the relationship between the logistics cost and resource use efficiency. Saha et al. [8] developed a returned logistics network of recycling product flow to determine the premise, through the comprehensive positive outlets and recycling outlets, assuming the outlets to recovery and recovery center, and dealing with the finish H different role infrastructure multilayer two-way integrated logistics network. Angani et al. [9] established the optimization design model of a recycling logistics network based on the

interaction mode of new and old exchange through the trade-in strategy, constructed the 0-1 mixed integer programming model, analyzed the number and location of the set network, and proposed the feasibility analysis of reducing the difficulty of collection by establishing the trade-in method. Zhou et al. [10] comprehensively considered the uncertain elements that caused the recycling demand of sports goods, established a multiperiod and multiobjective model, and solved the problem of logistics operation allocation among recycling outlets. Ingemarsdotter et al. [11] proposed to design a sports good traceability system based on the Internet of Things by starting with the Internet of Things technology, electronic product codes, and other related technologies and first established the corresponding sports good traceability model. Ramly et al. [12] analyzed the overall requirements of the system in depth. Based on the Internet of Things technology, they design each submodule from the aspects of product sales, product quality inspection, and market supervision to ensure that the quality of sporting goods in circulation can meet the market demand.

With the continuous development of technology, the development of the Internet of Things industry has been applied to the sports industry, but in the sports good management, it has not been involved. The traditional sports good management is still in the stage of manual management. Although this management mode is direct, it brings great inconvenience to people. The loss of traditional sporting goods due to poor management not only increases the cost of economic operation, but also increases the difficulty of work for people. Therefore, a comprehensive and fast sporting good management system is needed.

3. Recycling Management Mode of Sporting Goods

3.1. Operation Process. The operation process of recycling logistics mainly refers to a series of logistics activities such as the centralized collection of sporting goods, transportation, distribution, classification, processing, and packaging with the help of logistics nodes and networks, discarding users to identify product transportation methods. Related processing enterprises, after effective processing, extract useful resources, materials, or parts and return to the manufacturer for processing; the specific operation process is shown in Figure 1.

Network nodes participating in the sports product recycling logistics operation can be divided into several different nodes according to their functions: recycling network, recycling center, and processing end point. The recycling network is responsible for completing the business links of sports products entering the recycling system. In the process of recycling, it must act as the main node carrying the functions of concentration, classification, and transportation; carry out logistics operations such as classification, storage, and transportation of sports products according to different flow directions; and finally transport them to the processing center. Sports products are used as raw materials at the end of processing, and renewable resources are obtained through processing.

(1) *Recycling point*: there are two main construction modes of recovery network: one is to set up an independent recovery network in the area where products are concentrated as scrap and the other is to rationalize forward logistics nodes and make them function as recovery logistics nodes at the same time [11]. The operation mode of recovery outlets can be divided into independent operation of production enterprises, joint operation of production industries, and entrusted operation of third-party logistics enterprises according to different responsible subjects.

(2) *Recycling centers*: a recycling center as the intermediate link to undertake logistics tasks; its core function is to connect the recovery valve point and the processing end point, but it also has the functions of sports products storage, classification, and transportation. The operation efficiency of the recycling center plays a crucial role in the production plan of the processing end point [12]. For sports products collected and distributed from the recycling outlets, the recycling center will classify them according to different recycling and processing directions and then carry out storage and transportation operations in different ways until the sports products enter the processing terminal enterprises.

(3) *Processing end point*: the processing center is the final node in the recycling logistics system and the final target link of the recycling logistics system. According to the different processing methods of sports products, the processing end points can be divided into repair processing end points, remanufacturing processing end points, and comprehensive processing end points. However, its core and goal are to realize the recycling of resources through the harmless treatment of sports products.

3.2. Network Structure Analysis. The structure of the recovery logistics network can be expressed as a four-layer network structure of the collection and distribution layer, the recovery layer, and the processing and recycling layer according to the different logistics direction and node operation mode. The collection and distribution layer and the recovery layer, as the core, and layers of the recovery logistics system undertake the function of building the recovery logistics network [13]. The sports product recovery logistics network can be expressed as an upper, middle, and lower three-layer structure according to the order of logistics links. The lower layer is the recovery network to meet the needs of sports product distribution. The middle level is a recycling center, which carries out logistics operations of sports products. The upper layer is the sports product processing enterprises, which can realize the function of resource recycling. Therefore, the recycling logistics network can be expressed as the structure shown in Figure 2.

The underlying network structure of sports product recycling logistics is recycling N points, whose main function is to carry out the preliminary collection of sports products. Therefore, according to the different sports products

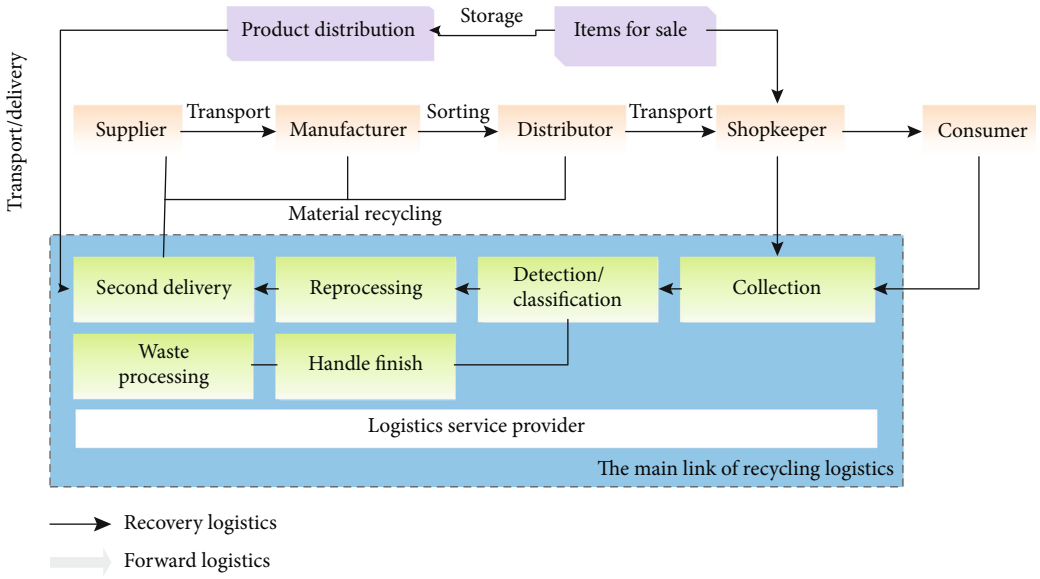


Figure 1: Recycling logistics business flow chart.

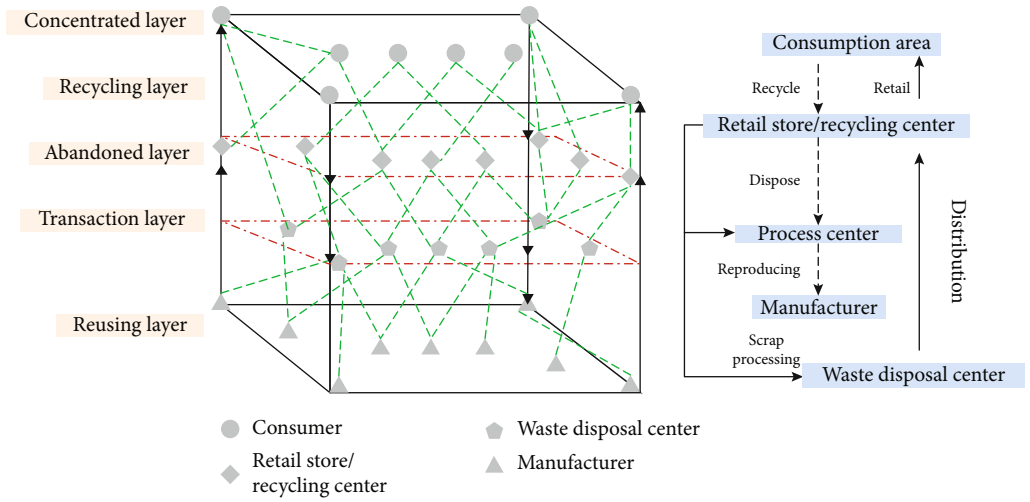


Figure 2: Sports good recycling logistics network structure.

recovered, generally, the recycling points are arranged to produce sports products with large frequency and large quantity. After completing the primary collection and distribution task of sports products in the recovery network, the sports products will continue to be concentrated again to the upper layer of the network structure, namely, the recovery center, through vehicles. The recycling center is an important link and node in the recycling network. By means of logistics operations such as recycling center, storage, classification, special product packaging, and transportation, it continues to transport to the processing center in the recycling network of sports products and finally reaches the processing node of sports products at the top of the network. For the sports product return flow network, after a reasonable layout of network recycling, recycling, and processing, the network nodes in the transportation operation link are also an important link in determining the operating efficiency and logistics cost of the recycling network, especially the focus of the recycling network to the recycling center. Whether it is for reasonable configuration and route planning, or to reduce transportation costs or recover the maximum profit, the recycling network structure is determined under the conditions of the main research content.

3.3. Management of Recycling Logistics by the Internet of Things. Perception technology [14], data processing technology, satellite positioning technology [15], and visual display technology [16] are used to manage and control the operation of the business process of recovery logistics, making the operation of the recovery logistics system more efficient. The technology used in this system is shown in Figure 3.

The Internet of Things technology has outstanding performance and powerful application ability in the field of recycling logistics. Because the Internet of Things application has to be able to have good perception, noncontact processing, good interaction with information, achieve excellent

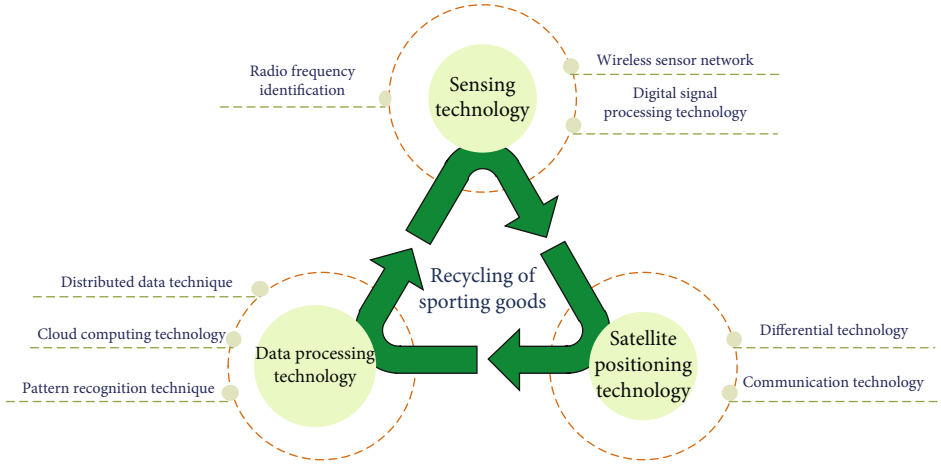


Figure 3: IoT technology applied in recycling systems.

feedback and control of technical characteristics and development of the logistics link mainly around items and objects, and rapid response and feedback of the item, the Internet of Things technology can effectively crack information barriers in the process of logistics operation, accelerate the speed of the logistics system, logistically link operations to reduce costs, and improve logistics efficiency of overall system performance, which embody in the following aspects: (1) improve the recycling logistics business link effective monitoring, (2) improve the efficiency of logistics operation, and (3) realize digital management.

4. Sporting Good Recycling Management Platform

4.1. Target Positioning and Demand Analysis. Based on recycling logistics information demand, combined with the Internet of Things for the positive influence and significance of returned logistics management, in view of the returned logistics system of main business, with the aid of returned logistics building method based on Internet of Things, around the recovery and recycling of sports products returned logistics information platform construction, the main content of the information platform for the design of the overall Japanese standard is summarized as the following aspects: (1) for the returned logistics management information platform to build good foundation for the environment; (2) build a recycling information docking platform to realize the whole-process control of recycling business; (3) integrate information resources of all links of recycling logistics to realize information resource sharing among transportation enterprises, recycling enterprises and sports product processing enterprises; (4) provide information support for logistics recovery and resource recycling operation control; and (5) achieving sustainable utilization of renewable resources and developing a circular economy.

According to the management departments and enterprises involved in the recovery logistics system, the information construction of the recovery logistics system is carried out according to the business and technical requirements, so as to realize the comprehensive management and operation control of the recovery logistics. In addition, the construction of the information platform also involves some nonfunctional requirements of the platform [17]. Therefore, the platform demand analysis can be divided into four aspects: user demand analysis, functional demand analysis, technical demand analysis, and nonfunctional demand analysis, as shown in Table 1.

For the management departments and related enterprises involved in the process of recycling logistics, based on the theory of the Internet of Things, the application of relevant information technology is carried out to construct the information of recycling logistics. Technical requirements can be divided into system integration technology, SOA architecture technology [18], cloud computing technology, data warehouse and data mining technology, and system interface technology.

(1) *Information collection and transmission technology*: perception and identification technology. This technology includes GPS, GIS, and bar code technology [19–23], which is used to realize the positioning, tracking, and monitoring of participating transportation vehicles and recovered goods. Network and communication technology is used to realize the interconnection, high-speed transmission, and sharing of information among monitoring units, enterprises, and individuals in the recovery logistics network, which provides technical support for the intelligent implementation of recovery operation planning and scheduling management of logistics vehicles.

(2) *Transportation operation optimization technology*: logistics optimization technology is an important technology to realize sports product recovery business. This technology mainly includes vehicle scheduling technology, vehicle path optimization technology, and cooperative processing technology of forward logistics operation task.

(3) Data analysis and mining technology is used to establish the virtual data warehouse which is used to solve

Table 1: Recycling management system information construction requirements.

System needs	Items			
User needs	Management department	Related enterprises		
Functional needs	Network management	Warehouse management	Sorting management	Transportation management
Technical needs	System integration technology	SOA architecture technology	Cloud computing technology	Data mining technology
Nonfunctional needs	Platform interface requirements	Hardware and software requirements	Environmental protection requirements	Economic and technical index

the problem of how to read information from the heterogeneous data source skill which is a structured data environment, data warehouse, and data transmission via the Internet data analysis of rationalization, which is the application of different themes and decision-making support to achieve the ultimate goal. A data warehouse can well integrate different data sources and can update the subject data in real time, which can achieve the corresponding goals stably. Platform architecture technology includes system comprehensive integration technology, system interface technology, and SOA architecture technology [24, 25].

4.2. System Framework. Analyzing the return to logistics network characteristics of the Internet of Things, we must consider platform information requirements, logistics product requirements, and commercial transportation, as shown in Figure 4.

(1) *Core business layer*: core business is the core of the platform to implement recovery logistics business management, including recovery network management, transportation management, storage management, and classification management. Recycling network management refers to the unified management of each recycling network connected to the platform to collect and process the recycling information. The transportation management mainly manages the daily recycling transportation plan and the transportation operation process; warehouse management is used to meet the storage needs of the recycled sports products due to temporary storage. Classified management is mainly aimed at determining the flow direction of the recycled sports products according to the categories, so as to facilitate them to better enter the recycling system and transport to the corresponding recycling enterprises.

(2) *Auxiliary business layer*: auxiliary business is the core of the platform and the supporting business of business development, mainly including customer management, vehicle management, and equipment management. Customer management refers to the management of centralized waste production units, recycling enterprises, logistics enterprises, and other platform customers that participate in recycling activities. Through customer evaluation and statistical analysis, it provides high-quality customer recommendation and other services for upstream and downstream customers. Vehicle management is to manage the information of the quantity, type, load, in-transit state, and driving state of the transport vehicle. Equipment management mainly realizes the management of logistics equipment.

(3) *Value-added business layer*: value-added business layer includes four parts: e-commerce, logistics information service, recovery fee settlement, and sports product information service. Through the development of e-commerce, the platform provides product information services for recycling enterprises and centralized waste production units and provides a platform for trading information sharing and collection. Logistics information service refers to the platform that forms a recovery or transportation plan by integrating e-commerce order information, provides logistics enterprises with transportation information of recovered goods, and optimizes and improves the utilization rate of logistics enterprises' resources

4.3. Platform Construction. In the development of recycling logistics information platform for sports lead storage sports goods, the development environment and operation environment are shown in Table 2.

According to the recycling business process of sporting goods, combined with the content of the information platform framework, the recycling platform of sporting goods is developed. With recycling network as the core, the sports good recycling platform emphasizes the comprehensive control of the whole process of recycling, storage, transportation, and production of sports goods. The specific functions include the following: various ways to display and query the recovery dot; use of IoT related technology to manage the outbound and inbound operations of goods and inventory checking process; based on the intelligent management directly oriented to the specific logistics transportation and distribution command and operation level, the vehicle scheduling and operation plan are completed by using the scheduling optimization model, and a variety of advanced technologies are used to intelligentize the recovery and transportation process. In view of the related customers involved in the recovery, the establishment of customer management

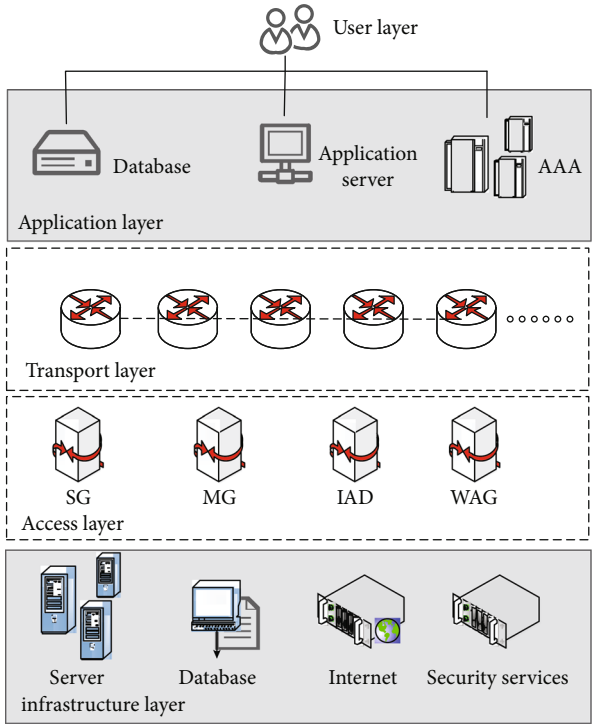


Figure 4: Sporting good recycling system framework.

system, to achieve the classification of customer management and improve the efficiency of customer management, use human resource management system to improve the efficiency of human resource management. Through the information platform, sports good information will be displayed to customers, so that merchants and customers can negotiate, trade, and pay online, and at the same time, the transaction process and payment process of both sides can be managed. At the same time, the emergency support system can be used to give early warning to possible safety problems in the platform and make intelligent judgment and generate an emergency plan.

5. Platform Optimization

In this paper, reducing system energy consumption is regarded as an optimization objective, and the alternating direction multiplier algorithm is applied to solve the problem more efficiently. Using the KNN algorithm to classify sports goods, the cloud model theory is introduced to identify unstable data partitions in unstructured data management systems.

5.1. Optimization of Sports Good Classification Module. The garbage classification module is responsible for inputting the images processed by the image preprocessing module into the machine learning model and outputting the final classification results. In this scheme, the machine learning model based on Lenet-5 is finally adopted. A conventional Lenet-5 has seven layers, each containing multiple parameters, totaling more than 60,000 parameters. The number of parameters of the model increases proportionally with the size and number of convolution cores and exponentially with the number of layers of the model. Therefore, the scheme is simplified to remove one convolution layer and the pooling layer. The overall process is shown in Figure 5. The image extraction process is as follows. First, features are extracted through the convolution unit (each convolution unit includes a convolution layer and a sampling layer), and finally, all local features are summarized by the full connection layer.

The neuron multiplies the input signal with the corresponding weight value, adds the bias term, and finally outputs the final result through the activation function, which can be expressed by the following formula:

$$y_j = f\left[\lim_{n\to\infty}\sum_{i=1}^{n}(x_i + b_j)\right], \tag{1}$$

wherein y_j is the activation function, which is used to determine whether the neuron should be activated under the current input, that is, whether the previous input can affect the next layer. In the training process, due to the inaccurate parameters of the model, the value of the output layer is usually different from the expected value. In this case, we need the back propagation (BP) algorithm to update the parameters of the model.

$$E = e^{-\lim_{n\to\infty}\sum_{i=1}^{n}(y_i - \bar{y}_i)}. \tag{2}$$

E represents the sum of the square deviations of all N sample errors, which is usually used as the objective function. Thus, it can be seen that the training of the model is to find the appropriate weight and bias the term to make the objective function take the minimum value, which is essentially a mathematical optimization problem. A gradient descent algorithm is used to calculate the partial derivative of the loss function. The gradient descent algorithm is used to calculate the partial derivative of the loss function. Take the y-th neuron in the NTH layer as an example again, and the i-th weight update formula of the upper layer is

$$\omega_{ij}^{l-1} = \eta \cdot \frac{\partial E}{\partial \omega_{ij}^{l-1}} + y_j, \tag{3}$$

where η is the learning rate that can be set.

After many experiments, the number and size of the convolution kernel are customized, so that the total number of parameters in this model is reduced to more than 5,000. We set the size of the model as an upper limit of 60 kB, tested the classification accuracy of the model for four kinds of garbage under different parameters, and obtained the results as shown in Figure 6.

After the preprocessing module, the size of the image is $60 * 60$. However, if this size is used as the standard for subsequent processing of the model, the terminal will not be able to fully store the parameters required by the model. Therefore, we halved its resolution, that is, we took the average value of each $2 * 2$ region of the image, converted it into a $30 * 30$ image, and then filled a layer of 0 in its outer ring to obtain a $32 * 32$ image, which is convenient for subsequent processing operations. $5 * 11 * 11$ convolution kernels are

Table 2: Information platform development environment and operation environment.

System development environment		System operating environment	
Operating system	Windows 10	Operating system	Windows XP
Development tools	MyEclipse	Browser	IE 10, etc.
Server	Apache 2.4.10	Configuration	phpStudy 2010
Database	MySQL 5.5.40		
Development platform	phpDesigner 7		

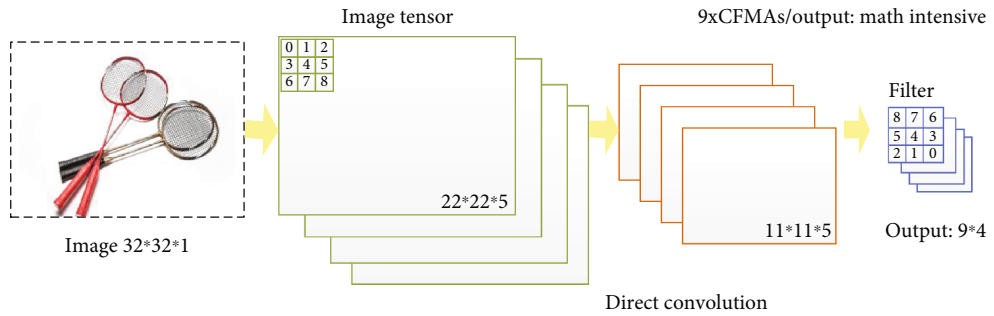


Figure 5: Sporting good classification model.

used to convolve the image to obtain a $22 * 22 * 5$ feature map. In this case, the number of parameters to be trained is $5 * (11 * 11 + 1) = 610$, including each offset term. The original Lenet-5 adopts a Euclidean radial basis function, and each kernel represents a classification result. The category of the value with the smallest output value is the final classification result. In this scheme, considering the specific transplantation, it is changed into the ReLU activation function, and the largest value among the 4 output values is the classification result, with a total of $(8 + 1) * 4 = 36$ parameters.

5.2. Timely Delay Optimization of System Stability. The advantage of the system architecture lies in the rational use of the characteristics of the edge server and cloud server, that is, the edge server has general storage capacity and computing capacity but short network delay and stable communication. The cloud server has strong storage capacity and computing capacity, but the network delay is long and the communication is unstable. For the unstructured data management system, the system carries out cloud modeling for the data partitions in the storage system, analyzes the correlation between the data partitions, identifies the unstable data partitions, and rearranges them, so as to reduce the data transmission cost caused by frequent data movement in the unstructured data management system.

The cloud model theory is introduced to identify unstable data partitions in unstructured data management systems. The transmission time of the passive pull-out data set corresponding to the data partition is used as the criterion to identify the unstable data partition. The total transfer time overhead of the data set d_k stored in the data partition dp_j on node S_i is

$$\cos t\left(s_i, dp_j, d_k\right) = \lim_{n\to\infty}\left(\frac{\sum_{i=1}^{n}\left(\text{size}_k/b_j\right)\cdot r_i^{i+k} - d_k}{\sum_{i=1}^{n}\left(\text{size}_k/b_j\right)\cdot r_i^{i+k} + d_k}\right). \quad (4)$$

A cloud model for each data partition is needed first. The cloud model for constructing data partition is as follows:

$$E_x = \lim_{n\to\infty}\frac{1}{N+1}\sum_{I=1}^{N}u^2 - 1,$$

$$E_n = \lim_{n\to\infty}\sqrt{\left|\frac{1}{N+1}\left(\sum_{I=1}^{N}(u+E_x)^2 - \sum_{I=1}^{N}(u-E_x)^2\right|} - \sqrt{\frac{E_x}{2}}, \quad (5)$$

where N is the number of cloud drops in a cloud model and u is each cloud drop in the cloud model. After the cloud model of each data partition is obtained, the data partition is divided into the Stable Data Partition Group (SG) and the Unstable Data Partition Group (UG) based on the cloud model of each data partition. The ECM calculation formula of cloud model similarity is

$$\text{ECM}(\text{DPC1}, \text{DPC2}) = \frac{E_x}{\sqrt{2/\pi}\left(E_n + E_x\right)^2}. \quad (6)$$

To demonstrate the characteristics of the different servers, this article compares edge servers with cloud servers. The communication delay and communication stability of the two kinds of servers are, respectively, compared from the horizontal and vertical aspects. As can be seen from Figure 7, the communication stability of the edge server is basically 2.5 times that of the cloud server, and the communication delay of the cloud server is 7 times that of the edge server, no matter in which day or in which time of day. The edge server is only one hop away from the device, while the cloud server is many hops away. Obviously, the edge server is more suitable for direct communication with the intelligent terminal and is responsible for some services requiring

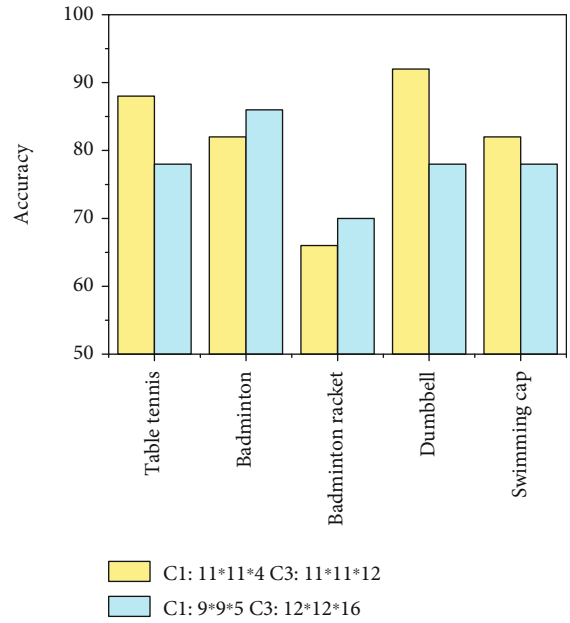


Figure 6: Model classification accuracy under different convolution kernels.

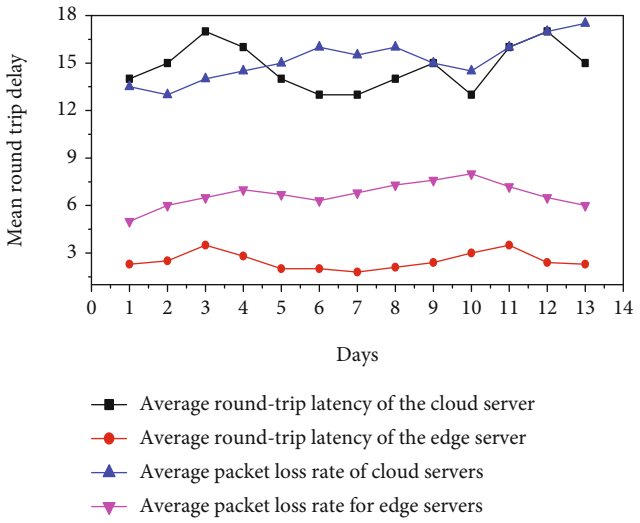


Figure 7: Comparison diagram of the performance of the two server networks.

stability and timeliness, while the cloud server is suitable for some services requiring a lot of computation and insensitive time delay.

5.3. Optimization of Path Planning Algorithm. The path planning algorithm based on genetic algorithm designed in this system comprehensively considers three factors, namely, the degree of recycling bucket filling, the full time, and the time to wait for recycling. We compare it with the distance-first path planning algorithm. The distance-first algorithm only considers the distance between each recycling bucket and sorts the distance between each recycling bucket in order from small to large. According to this order, the final recycling and cleaning path can be obtained. After the path planning algorithm based on the genetic algorithm obtains the final path, the path obtained by the distance-first algorithm is substituted into the fitness function to obtain the fitness values of the two paths, respectively. The path with a smaller fitness value is the better path. The comparison shows that the improved genetic algorithm is superior to the distance first algorithm in any case.

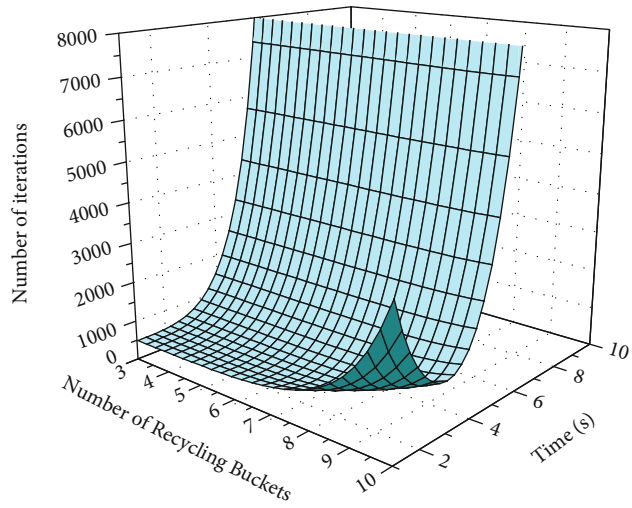


Figure 8: The relationship between the number of iterations, time, and number of recycling barrels.

This system also evaluates the time complexity of the path planning algorithm based on the genetic algorithm, and the influence of different parameters on the algorithm, Figure 8, is the number that shows the recycling barrel and the influence of the number of times of iteration algorithm running time; we can find 20 5000 trash iteration times, with the algorithm of the running time of 5 seconds; it has to do with paltry garbage pickup compared to the time saved. Figure 8 shows the convergence ability of the algorithm under different iteration times. We can find that the fitness value will not change after 5000 times, that is, the algorithm gets the optimal solution, so the number of iterations of the algorithm is also set to 5000.

6. Conclusion

Based on the analysis of the current situation and existing problems of the returned good flow, combined with the characteristics of the returned good flow process and the recycling network structure, this paper proposes to use the powerful intelligence of the Internet of Things and efficient information transmission capabilities and good operation and maintenance capabilities. We combine the return flow business and links, design an information framework platform, and develop and design an information platform framework based on the actual physical recycling business background. Combined with the feedback of the actual operation of the platform, the functions that can be realized in the optimization of the recovery business path are proposed. By establishing the corresponding model, the overall framework of the information platform is proposed from a macro perspective,

with business activities as the main management object and with the help of target positioning. Through the inspection of its operation process, network structure, and management objects, on the basis of the framework, computer technology is used for development and implementation. Through the design, development, and optimization of some functions of the information platform, the positive role of Internet of Things technology in improving the efficiency and reducing the cost of recovery logistics business management has been confirmed, and the ideas for solving the existing problems in recovery logistics management have been provided.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors received the following funding: Henan Provincial Soft Science Research Plan Project: Research on the Governance Path of Social Sports Organizations Participating in Rural Public Sports Service Supply under the Concept of Shared Development (Project No.: 202400410261); the General Project of the "13th Five-Year Plan" of Education Science in Henan Province: Research on the Path of Tailor Poverty Reduction of School Physical Education in Deep Poverty-Stricken Areas under the Concept of Shared Development (Project No.: 2019-JKGHYB-0331).

References

[1] S. Pimenta and S. T. Pinho, "Recycling carbon fibre reinforced polymers for structural applications: technology review and market outlook," *Waste Management*, vol. 31, no. 2, pp. 378–392, 2011.

[2] J. Rifkin, *The Zero Marginal Cost Society: The Internet of Things, the Collaborative Commons, and the Eclipse of Capitalism*, St. Martin's Press, 2014.

[3] E. Manavalan and K. Jayakrishna, "A review of internet of things (IoT) embedded sustainable supply chain for industry 4.0 requirements," *Computers & Industrial Engineering*, vol. 127, pp. 925–953, 2019.

[4] F. K. Shaikh, S. Zeadally, and E. Exposito, "Enabling technologies for green internet of things," *IEEE Systems Journal*, vol. 11, no. 2, pp. 983–994, 2015.

[5] W. Hou, W. Li, L. Guo, Y. Sun, and X. Cai, "Recycling edge devices in sustainable internet of things networks," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1696–1706, 2017.

[6] I. Mahmood and J. A. Zubairi, "Efficient waste transportation and recycling: enabling technologies for smart cities using the internet of things," *IEEE Electrification Magazine*, vol. 7, no. 3, pp. 33–43, 2019.

[7] W. Ren, Y. Sun, D. Zhao et al., "High-performance wearable thermoelectric generator with self-healing, recycling, and Lego-like reconfiguring capabilities," *Science Advances*, vol. 7, no. 7, article eabe0586, 2021.

[8] H. N. Saha, S. Auddy, S. Pal et al., "Waste management using internet of things (iot)," in *2017 8th annual industrial automation and electromechanical engineering conference (IEMECON)*, pp. 359–363, Bangkok, Thailand, 2017.

[9] A. Angani, J. C. Lee, and K. J. Shin, "Vertical recycling aquatic system for internet-of-things-based smart fish farm," *Sensors and Materials*, vol. 31, no. 12, pp. 3987–3998, 2019.

[10] Z. Zhou, Y. Cai, Y. Xiao, X. Chen, and H. Zeng, "The optimization of reverse logistics cost based on value flow analysis – a case study on automobile recycling company in China," *Journal of Intelligent & Fuzzy Systems*, vol. 34, no. 2, pp. 807–818, 2018.

[11] E. Ingemarsdotter, E. Jamsin, G. Kortuem, and R. Balkenende, "Circular strategies enabled by the internet of things—a framework and analysis of current practice," *Sustainability*, vol. 11, no. 20, p. 5689, 2019.

[12] R. Ramly, A. A. Bakar Sajak, and M. Rashid, "IoT recycle management system to support green city initiatives," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 2, pp. 1037–1045, 2019.

[13] O. Elsayed, M. Abouzied, V. Vaidya, K. Ravichandran, and E. Sanchez-Sinencio, "An ultralow-power RF wireless receiver with RF blocker energy recycling for IoT applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 11, pp. 4927–4942, 2018.

[14] F. Benkhelifa, H. ElSawy, J. A. Mccann, and M. S. Alouini, "Recycling cellular energy for self-sustainable IoT networks: a spatiotemporal study," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2699–2712, 2020.

[15] H. Wang, H. Han, T. Liu et al., ""Internet +" recyclable resources: a new recycling mode in China," *Resources, Conservation and Recycling*, vol. 134, pp. 44–47, 2018.

[16] T. J. Sheng, M. S. Islam, N. Misran et al., "An internet of things based smart waste management system using LoRa and tensorflow deep learning model," *IEEE Access*, vol. 8, pp. 148793–148811, 2020.

[17] F. Gu, B. Ma, J. Guo, P. A. Summers, and P. Hall, "Internet of things and big data as potential solutions to the problems in waste electrical and electronic equipment management: an exploratory study," *Waste Management*, vol. 68, pp. 434–448, 2017.

[18] G. C. Nobre and E. Tavares, "Scientific literature analysis on big data and internet of things applications on circular economy: a bibliometric study," *Scientometrics*, vol. 111, no. 1, pp. 463–492, 2017.

[19] M. Maksimovic, "Leveraging internet of things to revolutionize waste management," *International Journal of Agricultural and Environmental Information Systems*, vol. 9, no. 4, pp. 1–13, 2018.

[20] N. Salah and M. Rabee, "Smart recycle bin system based on Wi-Fi and IoT," *International Journal of Computer Applications*, vol. 181, no. 4, pp. 34–37, 2018.

[21] S. Keivanpour and D. A. Kadi, "Internet of things enabled real-time sustainable end-of-life product recovery," *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 796–801, 2019.

[22] J. Han, N. Lin, J. Ruan, X. Wang, W. Wei, and H. Lu, "A model for joint planning of production and distribution of fresh produce in agricultural internet of things," *IEEE Internet of Things Journal*, 2020.

[23] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the internet of things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

[24] W. Wei, M. Guizani, S. H. Ahmed, and C. Zhu, "Guest editorial: special section on integration of big data and artificial intelligence for internet of things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2562–2565, 2020.

[25] S. H. Ahmed, V. H. C. de Albuquerque, and W. Wei, "Guest editorial: special section on advanced deep learning algorithms for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2764–2766, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 7823982, 10 pages
https://doi.org/10.1155/2021/7823982

Research Article

Intelligent City Traffic Scheduling Optimization Based on Internet of Things Communication

Jingwen Jiang

College of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing 210037, China

Correspondence should be addressed to Jingwen Jiang; jiangjw@njfu.edu.cn

Received 13 April 2021; Revised 17 May 2021; Accepted 24 May 2021; Published 7 June 2021

Academic Editor: Wei Wang

The application of Internet of Things technology in the construction of smart city traffic system is the inevitable choice of smart city traffic system to promote the development of road traffic in the process of urban development. First of all, the traffic in the city, on the basis of the research background and significance, research on the Internet of Things, and wisdom of the construction of the urban traffic are summarized and analyzed; secondly, on the basis of requirement analysis and feasibility analysis, combined with the principles and goals of the construction of the system, the establishment of the model of urban traffic based on Internet of Things and the city's architecture based on Internet of Things intelligent urban traffic system and design are put forward. Finally, an improved ant colony algorithm is proposed in the aspect of system optimal path algorithm, and its superiority is verified. The results show that complementary networks can significantly improve the robustness of networks. These conclusions provide a theoretical basis for the construction of smart city public transport network traffic scheduling and provide model support for the operation and management of smart city public transport.

1. Introduction

With the continuous advancement of urbanization and the increasing urban population, the traffic pressure faced by cities is becoming more and more obvious. In order to scientifically relieve the traffic pressure and provide convenience for the emergence of urban residents, it is necessary to timely combine new technologies to promote the transformation and upgrading of the traditional traffic mode. In this development, the application of smart city traffic system is an important breakthrough. Especially with the integration of Internet of Things technology, the construction of new smart city traffic system is of great value to the alleviation of urban traffic pressure.

With wisdom of urban transport system in urban transportation planning in [1] on the intelligence development requirements also rising, under this situation, you need to in the wisdom of city traffic system actively introduce advanced Internet technology, to promote the development of the wisdom of the urban transport system intelligence and provide residents with more effective transportation service. Therefore, integrating the Internet of Things technology into the smart city traffic system is the first need to improve the intellectualization of the traffic system. As a system integrating with advanced technologies such as microelectronic sensors and wireless communication technology, the application scope of Internet of Things technology has been greatly expanded, and it also plays an important role in the construction of smart city traffic system. Through the application of Internet of Things technology, data collection and detection can be carried out more efficiently, thus providing strong technical support for the intelligent development of urban traffic. In terms of the current urban traffic management, it is generally necessary to maintain order through the consciousness of the public. The function of traffic lights and traffic signs can only provide static [2] and limited guidance, and it is difficult to realize the function of optimizing road resources. The smart city traffic system based on the Internet of Things technology can effectively collect all the road and vehicle information in the city and then dynamically calculate the best route planning to promote the orderly development of urban traffic.

Secondly, the application of Internet of Things technology is also the need to improve the level of traffic command

and scheduling. By using the Internet of Things technology to carry out traffic command and dispatch, it can more effectively avoid the occurrence of large area of traffic congestion, so as to provide convenience for the smooth travel of residents. For example, in the process of vehicle running, the smart city traffic system of the Internet of Things can directly perceive, control, and guide road traffic by real-time processing of traffic information and using traffic signal lights, so as to achieve the function of regulating road vehicles. At the same time, with the support of RFID technology [3, 4], vehicle information can also be collected through vehicle electronic tags. This operation can realize information collection and analysis even in the high-speed operation of vehicles, thus realizing dynamic management of vehicles, ensuring accurate prediction of traffic flow, and effectively improving traffic operation efficiency.

It can be seen from the above research that the application of Internet of Things technology can promote driving safety and improve driving comfort. In the intelligent urban traffic system, improve safety and travel comfort is the most basic purpose and, with the help of the Internet of Things technology, can be achieved under the complex traffic environment, intelligent vehicle terminal based on GNSS navigation system, and the real-time data transmission on the Internet, which by the real-time display state, as far as possible to avoid congestion, avoid fault section, and improve the system efficiency.

2. Related Work

The EU has achieved good results in the research direction of Internet of Things technology application. In order to lead the world in the Internet of Things direction, the facility operator actively carry out the technology and application of M2M [5, 6].The Internet of Things application mainly includes the intelligent medical treatment, use of the sequence code product, can timely certification and drug safety, the unsafe drugs, counterfeiting, fraud has played a very good prevention, and play a blow. In addition, the application of intelligent electronic material system can provide real-time consumption information needed by consumers. Power remote monitoring system can be better in real-time and reliable grasp of the safety and use of power; in traditional industries such as logistics and retail, the application of the Internet of Things enables the timely exchange of information on various demands, meets the information needs of different groups, reduces the flow of information, and improves the efficiency of information use. The Internet of Things has been applied to People's Daily life [7, 8]. Recipe can be downloaded through mobile phone, food to be purchased can be viewed remotely through refrigerator lens, and certain activities can be automatically carried out at a certain time set on the network. With the application of various communication facilities in the Internet of Things, people can receive timely information services anytime and anywhere [9]. It is hoped to establish information standards in the new era and create a network society in the era of the Internet of Things, so as to better solve social problems such as the aging of people [10], so as to make information universal and efficient popularization. At the same time, also, actively promote IoT-related industries such as strategic planning, put forward the "by building one of the world's most advanced Internet of Things from the infrastructure, the convergence of multicast communication to build future superfluid strong goal [11], from big to Internet of Things from the infrastructure, the Internet service, Internet technology, Internet of Things environment such as strength, actively carry out relevant research, looking for the Internet of Things market new business opportunities, ensuring the continuous expansion of the Internet of Things industry, and make a superb information communication technology power [12]. With the rapid development of Internet of Things, our timelines of information collection, information transmission security, and standard put forward higher request, in the process of the development of the Internet of Things is faced with many problems: lack of national standard, enterprise technology research level is weak, high cost of RFID tags, industry often lack the talent, the first technology is focused on national security question, and personal privacy is being violated not guaranteed. It can be seen that for the huge demand of IoT-related industries, how to promote the process of IoT is facing certain difficulties.

With the rapid development of science and technology and information technology, urban intelligent urban traffic has become the hot topic today; now, a lot of places are still in the stage of the urban traffic wisdom, and it is the primary stage, but the wisdom of the urban traffic is insufficient, making intelligent urban traffic must in some places to research intelligent urban traffic [13, 14]. The smart city transportation system is guided by the framework of the national smart city transportation system to build a smart city transportation system of "high efficiency, good safety, environmental protection, and comfort." Not only improve the management level and the efficiency of the urban traffic system but also provide traffic information services, convenient, high quality, fast, economical, safe, comfortable intelligent traffic, and transportation services; improve timely, accurate, comprehensive, and sufficient information support and information-based decision support for traffic management departments and related enterprises, so as to make the system decision intelligent [15, 16]. Many scholars are actively deploying the construction of urban smart city transportation. The research on smart city transportation is early, and the technology is advanced. Through intelligent control system, urban signal lights are controlled intelligently to reduce traffic congestion [17]. Department network intercommunication and information sharing, real-time broadcast of traffic information, traffic failure in a timely manner [18]; the use of bus priority control system to ensure the efficiency and safety of bus operation; strengthen the R&D and application of smart city transportation, encourage the cooperation between enterprises and research institutions, etc. [19, 20]. Through the road control station, FIRD technology, advanced monitoring system, and natural traffic flow roadside system, the entering vehicles can be automatically identified, and "road congestion tax" is levied at a specific time, which effectively reduces traffic flow, reduces traffic congestion by 25%, increases road traffic capacity by 80%, and

reduces vehicle exhaust emission by 8%-14%. In addition, through the architecture of smart city transportation, good effects have been achieved in protecting environment and preventing pollution. Through sensors and road probes, traffic jams can be reported in time, and traffic jams in the next period of time can be analyzed and predicted in advance, so as to suggest smooth road driving [21, 22]. Through a variety of sensors, real-time assessment of the traffic bridge can be made to obtain the number, weight, and pollution of vehicles, etc. Once the set limit value is exceeded, the traffic management department will receive an alarm, and then, find and solve the problem [23]. Through the establishment of intelligent urban traffic cloud platform founded information portal, and actively start traffic hotline, travelers can timely understand live traffic electronic bus stop, bus stations, and travel information, such as integrated traffic information query for traffic incidents, road construction, and other special circumstances; the platform can formulate and publish in time travel advice [24] and make people travel more convenient and quick. Urban transportation in the urban traffic information public platform and wisdom cloud platform to build "comprehensive three-dimensional traffic command center", to a more comprehensive, more traffic system planning and management, make "wisdom city traffic" comprehensive traffic development to upgrade, improve the intelligent level of traffic [25–27], and provide more comprehensive traffic information service for the people.

At the present stage, many cities have realized the serious problems brought by urban traffic and are actively promoting the construction of "smart city traffic," and have made certain achievements, but most of them are still in the initial stage. With the in-depth application of technologies such as the Internet of Things and cloud computing, they will surely achieve better development.

3. Smart City Traffic System Function Based on Internet of Things

With the rapid progress of social economy and science and technology, the city scale has been expanded to a certain extent. As a result, the traffic problem is extremely serious. At this time, in order to better solve the traffic problem, we need to do a good job in many aspects of coordination and cooperation and upgrade and optimize the urban traffic system. From the perspective of the government, when solving the problems of urban traffic development, we should not only do a good job in the introduction of new equipment but also need to increase the investment in infrastructure and traffic facilities management and further optimize the management system. The application of smart city transportation system in cities can effectively solve some practical problems and provide convenience for users. Basic functions of smart city traffic system include road condition information, service inquiry, safety guidance, and service guide [28, 29]. Smart city traffic maintains constant development in the process of continuous development. It can not only provide users with accurate traffic information but also facilitate users to query and prompt information. It has a good development prospect in the future development process. Based on the smart city traffic system, users can grasp the traffic conditions in real time and shorten the time wasted on the road. In this paper, the traffic conditions of the existing key sections are recorded and photographed, and a reasonable analysis method is adopted to classify the road levels, so as to facilitate users' travel. In the research framework, users can also query vehicle information in real time through the smart city traffic system, which can not only search vehicle information but also search vehicle information such as annual inspection, so as to provide convenience for users. Compared with relevant studies, in order to enable users to drive safely, the smart urban traffic system can remind users of traffic safety so that users can drive safely. Smart city transportation system service guide includes traffic laws, driving management business, and vehicle business. In case of problems, smart city transportation systems can provide users with legal guidance and accident counseling. On the basis of smart city traffic system, give full play to the technology of the Internet of Things and form a smart city traffic system through the in-depth application of intelligent identification, cloud computing, mobile technology, and data fusion technology. The architecture of smart city traffic system is shown in Figure 1.

Smart city transportation system can be divided into three levels from the overall architecture, namely, perception layer, network layer, and application layer.

3.1. The Perception Layer of the Internet of Things. IoT perception layer, at the end of the architecture, is the basic level in wisdom city traffic system, by using RFID, sensor networks, wireless communication and real-time positioning technology for data acquisition and collection, and timely and accurate monitoring of traffic information, it is for the whole traffic system to collect real-time traffic data for $\_f =$ layer for transmission and application. It consists of an M2M terminal and an infinite sensor network.

3.2. Internet of Things Network Layer. Internet network layer in the architecture of the middle layer has the effect of flow; the main principle is through the mobile network, wireless networks, Internet, satellite communication facilities such as the perception layer collected information processing to the data center, form to real information of the real reaction, through the scientific analysis of information, can provide service of traffic control decision-making.

3.3. Internet of Things Application Layer. The application layer of the Internet of Things is at the top of the architecture. Based on the various data obtained through the perception layer and the network layer, it realizes the parallel processing and optimization of massive information as well as the dynamic configuration and deployment of storage resources under the resources such as intelligent processing, cloud computing, management system, and database. Through information storage and processing system and integrated control system, information is managed, and information is presented to users in various ways for decision-making, service, and business development, including operation support platform and application system two parts.

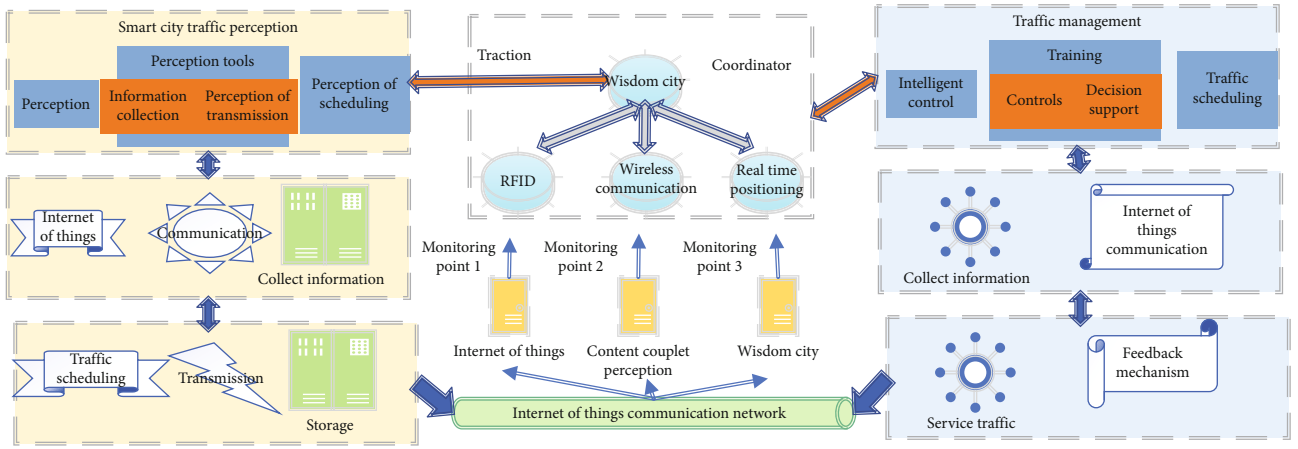


FIGURE 1: Intelligent city transportation system architecture diagram of the Internet of Things.

Wisdom urban traffic system which is built based on the Internet of Things technology in the process of general considerations should pay attention to the following several points: first, to guarantee the traffic system of dynamic, dynamic here mainly refers to during the running of the system, to be able to real-time capture different time through intelligent urban traffic system of traffic information, effective information collection, and through the data analysis response-specific traffic; second, the construction of smart city traffic system should be considered from the perspective of global, that is to say, different traffic scenes and specific application functions should be set under the same system to avoid resource waste. Third, the construction of smart city traffic system should also meet the requirements of intelligence. This construction needs to fully integrate the Internet of Things technology and promote the flexible application of Internet of Things technology in smart city traffic system, so as to master the traffic dynamics in an all-round and dynamic way. In the process of building smart city traffic system, security is the most fundamental construction point. Therefore, data security should be fully paid attention to in the construction of smart city traffic system. In the construction of the corresponding data security management system, the following four dimensions can be generally started (as shown in Table 1).

Selected by some special rule (which will be described briefly), a complete graph is formed by connecting all the possible edges with T-1 existing nodes and new nodes. Next, we will discuss degree distributions that are consistent with different rules.

(1) First, we consider how to select T-1 existing nodes. Here, we use the probability rule that is proportional to the degree R of each node I. Namely, the performance linear preference rule, which means that in the smart city traffic network, when a newly added station chooses to connect to a station, the station with greater performance is more likely to be selected. We can obtain the evolution equation of performance through a quasicontinuous approximation, which is very similar to the BA model. The BA model is that at the beginning, there are M0 nodes in the network, and these nodes are arbitrarily connected, as long as there is at least one link for each node, and the network evolves gradually according to growth and preference connection. Growth adds a new node with m chains to the network at each step

$$\frac{\partial T_i}{\partial t} = \frac{(T+1)\sum T_i}{(T-1)\sum T_j}. \tag{1}$$

The performance distribution is solved as follows:

$$p(T) = \frac{m - t\beta}{m + t\beta} T^{-(1/\beta)}. \tag{2}$$

(2) We consider randomly selecting T-1 existing node rules. Following the BA model, we can write down the degree evolution equation as follows:

$$\frac{\partial T_i}{\partial t} = \frac{m - t - 1}{m + t}(T - 1). \tag{3}$$

The urban traffic dispatching system is used to monitor commuting at important traffic crossings (such as bridges) and control the operating status of traffic lights at corresponding locations. System composition is shown in Figure 2, including bridge stress monitoring and gate access control node, traffic light control node, ZigBee-WiFi gateway, network transmission platform, acquisition service software, and web application software.

The traffic management system includes traffic-integrated information management, signal control system, police command system, and public traffic service system. Integrated traffic information management is to understand the real-time use level of urban traffic on the basis of traffic data collection and provide reasonable travel suggestions for travelers in combination with the theory of traffic planning in reality. Signal control system is through the management of traffic infrastructure, to understand the reality of

Table 1: Basis of safety management system construction.

Security dimension	Object-oriented	Take measures
Application	System users	It can be controlled by access control, data encryption, protocol modification, and other measures
System	Whole smart city traffic	It is necessary to strengthen the protection of basic service software security vulnerabilities, such as database and operating system
Network	All system data	The data generated by the Internet of Things are encrypted and transmitted through the password algorithm and then transmitted through the Internet. After data sharing, the function of data fusion in the intelligent system is achieved
Data	All application data	Through the establishment of smart city traffic dedicated platform to deploy software and store data

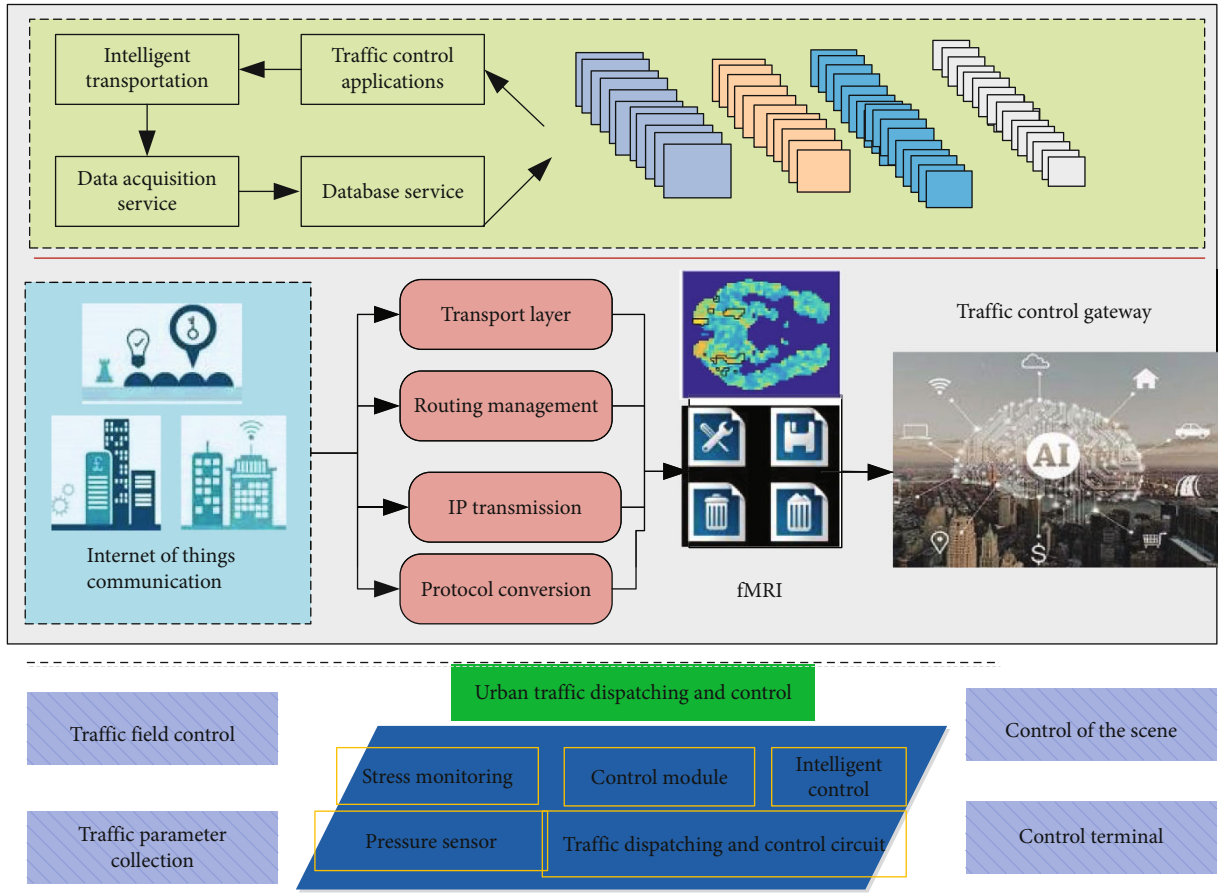


Figure 2: Block diagram of urban traffic dispatching and control system.

urban traffic flow and traffic congestion information and take a reasonable control of the signal indication, so that the traffic flow of the road will reach a reasonable level. The police command system handles emergencies in traffic, obtains road sections and traffic conditions through traffic facilities, releases information in time, and takes corresponding measures to reduce the traffic impact caused by the events. The public transport service system enables travelers to timely understand the travel route, traffic location, and other information; reasonably arrange the transportation mode; and optimize the travel activities.

4. Research on Intelligent City Traffic Scheduling Optimization Algorithm Based on Internet of Things

In the real traffic, the optimal path selection in the traffic network is different from the shortest path problem in theory. One-way driving on road sections, turning restrictions at intersections, as well as optimization standards for the shortest driving distance, the lowest driving cost, and the least traveling time will all have corresponding influences on the optimal path for traffic [27], which shows that people have

different requirements for real-time path in traffic. The relevant influencing factors are mainly divided into the following three aspects.

4.1. Real-Time Traffic Flow. In real traffic, if there are more vehicles on an optimal path from the starting point to the destination, which exceeds the carrying capacity of the path, urban traffic congestion is bound to result. The optimal path planning in order to be more in line with the actual must be according to the traffic flow to quantify the degree of congestion; traffic congestion discriminant is of high maneuverability, average travel speed of the motor vehicle can root for description, and concrete can be divided into smooth, mild crowded, crowded, and severe degree of four.

4.2. Delay Time at the Intersection. In mathematics, path optimization regards each node as an abstract point with zero delay, whose existence does not affect the optimization time and does not play any role. However, in the actual traffic network, the influence of traffic lights and other conditions is very obvious, which inevitably leads to time delay. Its weight in the optimal traffic time path is relatively large, so it can be seen that the delay time generated when vehicles pass the crossing day must be considered in the path optimization algorithm.

4.3. The Influence of One-Way Traffic. Under normal circumstances, the traffic administrative department of the sections in order to optimize certain set up of the road, the intelligent drive from one end to the other side cannot reverse driving, and general path algorithm is assumed that any two nodes are exchanged; in this case, set in accordance with the route is not realistic, so you must consider road affects traffic path.

In the process of solving the optimal traffic path, the basic ant colony algorithm has some problems, such as long searching time and easy to converge to the local optimal solution. In the actual traffic, there exists the problem of steering direction. By introducing the directional guidance information and choosing the preferred traversal path, the efficiency of the algorithm can be improved. The problem of search converging to the local optimal solution is solved by the cooperative characteristic between ant colonies, and the accuracy of search is improved.

Because the basic ant colony algorithm has no directionality when solving the shortest path in the complex traffic network, the directionality has a great weight when solving the shortest path by improving the visibility parameter on the basis of the transition probability.

$$\mu_{jk} = \frac{1 - \omega_{kj} \cdot \beta_{jk}}{1 + \omega_{kj} \cdot \beta_{jk}}. \tag{4}$$

Because of the influence of pheromone on path selection, the pheromone can be divided into the starting point pheromone and the end point pheromone. Each ant starts at the beginning and disperses the most pheromones at the beginning and the end. As the distance goes, the amount of pheromone disperses decrease. Individual ants have only one type of pheromone, which changes when they reach their destination. After N iterations, the path with the most pheromones is the shortest path. According to the collection of the starting point and the end point, the shortest path from the starting point to the end point can be found.

In order to avoid the excessive accumulation of road pheromones and prevent the search from converging to the local optimal solution, it is necessary to ensure the diversity of path choices and explore different paths to find the optimal solution, which requires the local update of pheromones. In the updating process, the pheromone concentration is updated for the road section passed by ants according to the local updating rules. The local update rule is shown in the formula.

$$\lambda(j,k) = \frac{\lambda_{jk} - \Delta\lambda_{jk}^{t}}{2}. \tag{5}$$

Theoretically, if the pheromone concentration of a certain road is too low, the probability of ants' choice will be affected, or even cannot be chosen by ants. If so, the diversity of path choices will be affected, so that the optimal solution may not be the optimal path. Therefore, a minimum standard can be set for the pheromone concentration of the road. When the pheromone concentration of the road section is lower than this standard, it is forced to restore to the minimum standard line to ensure the diversity of route selection. Local pheromone updating can effectively solve the problem of excessive accumulation of path pheromones, prevent the possibility of local convergence, and expand the search scope as much as possible, which is conducive to the selection of the optimal path.

Step 1. Initialize parameters, and the number of iterations is 0.

Step 2. Set the maximum number of iterations, increasing by iteration $N = N + I, A, 1$.

Step 3. Set $A = A + 1$.

Step 4. Ant A moves to the next node k according to the transition probability.

Step 5. Modify taboos and update pheromones locally.

Step 6. If node k is the end point, perform Step 7. Otherwise, jump to Step 4.

Step 7. If A is greater than the total number of ants, perform Step 8; otherwise, skip to Step 3.

Step 8. Update the global pheromone.

Step 9. If the maximum number of iterations is satisfied, perform Step 10; otherwise, perform Step 2.

Step 10. Finish the search and output the optimal result.

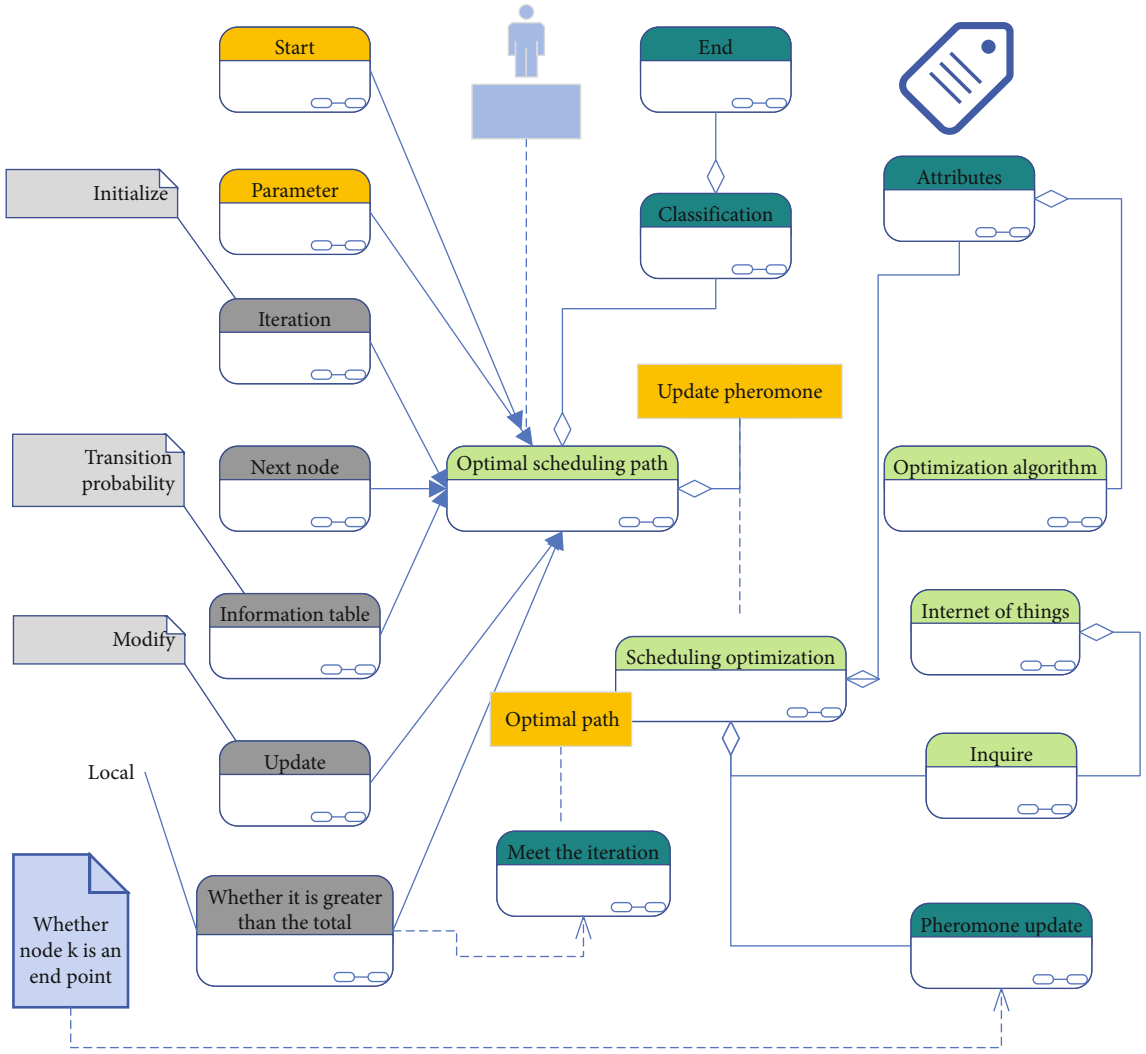


Figure 3: Intelligent city traffic scheduling optimization algorithm for the Internet of Things.

Table 2: Descriptive statistics of main variables.

Variable	Clarification	Mean	Standard deviation	Minimum value	Maximum
TP	First map	22781.48	74785.23	78.02	1132569.21
GDP		59636.5	49872.45	4410.186	506301.2
FDI		0.0033	0.0046	0	0.0564
STE		0.0171	0.0164	0.0003	0.2313
Ind	Second path	50.0139	11.8912	9.73	87.94
EDU		470.4439	409.3027	0	2409.313
POP		0.1026	0.0944	0.0023	1.1448
BUS		119.4722	113.5525	1.9540	1591.47
INN		12.44	55.65	0.006	1061.27
EMP	Inner path	38.75	84.45	1.13	923.2
TECH		2.313	1.894	0.077	12.152

The improved algorithm flow is shown in Figure 3.

The changed algorithm emphasizes the directionality of the transfer, the classification of the pheromone, and local pheromone updates and enables the ants to explore new paths, so it can strengthen the basic ant colony algorithm of pheromone positive feedback, improve the difference of each section of pheromone concentration, and prevent to explore the path of reducing, so as to improve the efficiency of the algorithm. The main variable design is shown in Table 2.

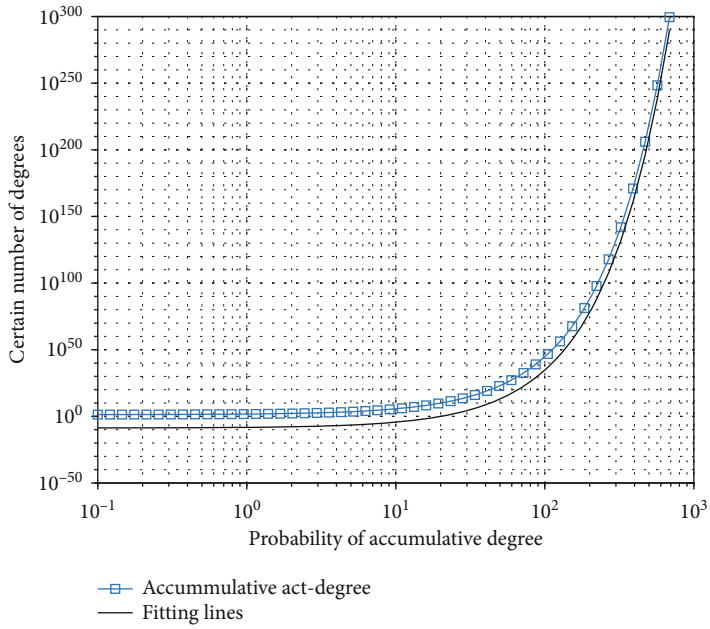


Figure 4: Distribution diagram of traffic scheduling accumulation degree in smart city.

Table 3: Network model characteristic parameter simulation data table.

Network parameters and characteristic quantities (simulation data)	City 1	City 2	City 3	City 4
Number of sites	165	245	123	94
Number of lines	12	13	7	6
The network diameter	5	5	2	2
The average degree	21.65	25.49	21.59	24.58
Mean path length	2.374	2.523	2.182	1.696
Clustering coefficient	0.9069	0.9033	0.8102	0.8335

5. Results Analysis

We has carried on an empirical analysis of a city traffic scheduling in 2017, which includes 736 city bus lines and 5487 bus site; in order to reflect the geographical topology characteristics of the web site of the bus, we adopt L space method, namely, each bus station is defined as a node, if two nodes in the same bus line when they are connected by an edge. As shown in Figure 4, we show the cumulative degree distribution of urban traffic scheduling. In Figure 4, the vertical axis represents the probability of accumulative degree, and the horizontal axis represents a certain number of degrees.

The smart city traffic scheduling network model of Internet of Things communication is used for digital simulation, and the results of digital simulation are compared with the results of empirical analysis. Table 3 shows the simulated calculation values of network characteristic statistical parameters. The simulated data are basically consistent with the empirical analysis data. This network model can better explain the characteristics of smart city traffic scheduling in the Internet of Things communication.

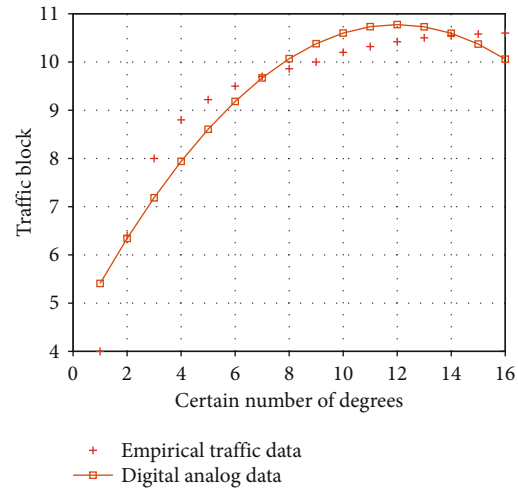


Figure 5: Comparison between simulation data and empirical data of urban transport network cumulative degree distribution.

The circular data curves in Figure 5, respectively, represent the simulation graph of the accumulation degree distribution with the urban transport network. It can be seen that the simulation graph is consistent with the results of the empirical analysis, showing a segmented degree distribution law.

According to the simulation results of network nodes supplemented by random attacks (Figure 6), the network evolution law is basically similar. As stated in the simulation modeling method, random attacks cover the phase transition points of system collapse after 50 configuration averaging, so it is impossible to make quantitative comparison. However, according to the evolution law of the average path length, there is an obvious rising process in the attack process, because there are fewer routes but more stations in each route.

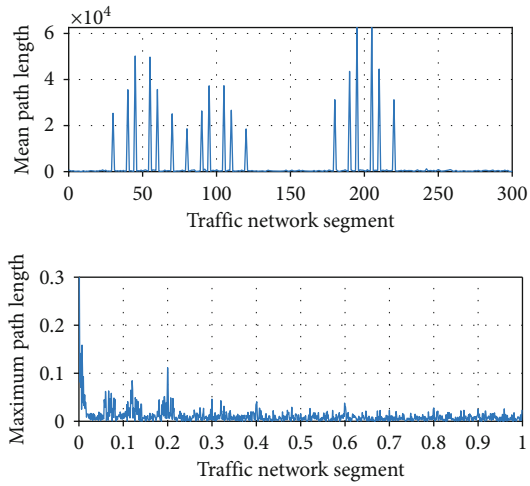


Figure 6: Simulation results of the robustness of random attacks complementing each other.

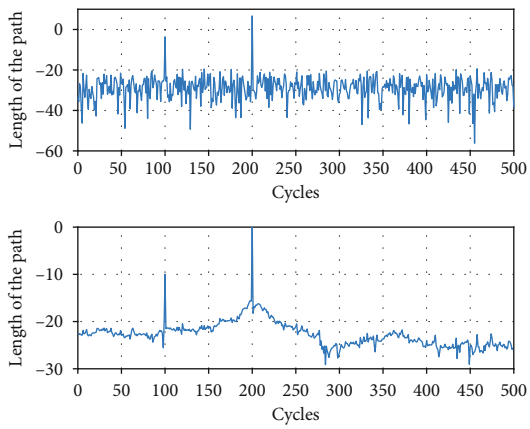


Figure 7: Search process of improved scheduling algorithm.

Parameter settings of smart city traffic scheduling algorithm are based on Internet of Things communication. Figure 7 shows the search process of scheduling algorithm and shows the optimal path length found by scheduling in each round of search.

6. Conclusion

To sum up, on the basis of demand analysis and feasibility analysis, combined with the principles and objectives of system construction, this paper establishes the urban traffic model based on the Internet of Things and puts forward the architecture and application architecture of the intelligent transportation system focusing on system integration and application. It is necessary to apply the intelligent transportation system of Internet of Things technology to improve the low efficiency of traffic information collection and the traditional intelligent system of imperfect traffic data, so as to promote the efficient development of urban traffic system. The results show that complementary networks can significantly improve the robustness of networks. These conclusions provide a theoretical basis for the construction of smart city public transport network traffic scheduling and provide model support for the operation and management of smart city public transport.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] X. Fei and G. Z. Tian, "Optimization of communication network fault identification based on NB-IoT," *Microprocessors and Microsystems*, vol. 80, pp. 103531–103545, 2021.

[2] L. Huang, X. Yuan, J. Zhang, N. Zhang, J. Li, and L. Wang, "Research on internet of things technology and its application in building smart communities," *Journal of Physics: Conference Series*, vol. 1550, no. 2, pp. 22029–22031, 2020.

[3] M. M. Rathore, A. Ahmad, A. Paul, and S. Rho, "Urban planning and building smart cities based on the Internet of Things using big data analytics," *Computer Networks*, vol. 101, pp. 63–80, 2016.

[4] X. Meng, "Urban planning and building smart cities based on the Internet of Things using big data analytics," *Computing Reviews*, vol. 58, no. 10, pp. 608–618, 2017.

[5] W. Li, W. Xue, and S. Hou, "Analysis on bilingual public signs in the view of audience theory—a case study of city traffic public signs in Xi'an," *Open Journal of Modern Linguistics*, vol. 5, no. 2, pp. 181–186, 2015.

[6] Y. Zhao, J. Ma, L. Shen, and Y. Qian, "Optimizing the junction-tree-based reinforcement learning algorithm for network-wide signal coordination," *Journal of Advanced Transportation*, vol. 2020, Article ID 6489027, 11 pages, 2020.

[7] F. Tang, "Intelligent bionic optimization algorithm based on the growth characteristics of tree branches," *International Journal of Cognitive Informatics and Natural Intelligence*, vol. 15, no. 2, pp. 34–46, 2021.

[8] C. X. Mavromoustakis, G. Mastorakis, and J. M. Batalla, "Internet of things (IoT) in 5G mobile technologies signal processing techniques for energy efficiency, security, and reliability in the IoT domain," in *Modeling and Optimization in Science and Technologies*, pp. 419–447, Springer-Verlag, 2016.

[9] W. H. Wei, Y. N. Zhu, T. M. Xue, S. Y. Xie, C. R. Yuan, and T. Y. Wang, "Optimization of the frame and performance analysis of solar modules based on Ansys workbench and Solidworks," *Key Engineering Materials*, vol. 693, pp. 169–173, 2016.

[10] G. Li, S. Fang, J. Ma, and J. Cheng, "Modeling merging acceleration and deceleration behavior based on gradient-boosting decision tree," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 7, pp. 2723–2737, 2020.

[11] Y. Y. Pan, "Lagrangian relaxation for the multiple constrained robust shortest path problem," *Mathematical Problems in Engineering*, vol. 2019, Article ID 3987278, 13 pages, 2019.

[12] X. Du, Z. Zhu, J. Chen, C. Qi, and X. Guo, "Route configuration method for highway passenger hubs from the perspective of transportation integration: a case study of Nanjing, China," *Symmetry*, vol. 12, no. 7, pp. 1194–1205, 2020.

[13] M. Jarrah, M. Jaradat, Y. Jararweh, M. Al-Ayyoub, and A. Bousselham, "A hierarchical optimization model for energy data flow in smart grid power systems," *Information Systems*, vol. 53, no. 12, pp. 190–200, 2015.

[14] C. Lin, Y. Bi, H. Zhao, Z. Liu, S. Jia, and J. Zhu, "DTE-SDN: a dynamic traffic engineering engine for delay-sensitive transfer," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 5240–5253, 2018.

[15] J. Jiao, Y. Sun, S. Wu, Y. Wang, and Q. Zhang, "Network utility maximization resource allocation for NOMA in satellite-based Internet of Things," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3230–3242, 2020.

[16] S. Xu, X. Wang, G. Yang, J. Ren, and S. Wang, "Routing optimization for cloud services in SDN-based Internet of Things with TCAM capacity constraint," *Journal of Communications and Networks*, vol. 22, no. 2, pp. 145–158, 2020.

[17] M. A. Al Mamun, M. A. Hannan, and A. Hussain, "Theoretical model and implementation of a real time intelligent bin status monitoring system using rule based decision algorithms," *Expert Systems with Applications*, vol. 48, pp. 76–88, 2016.

[18] X. Tian, Y. Cheng, D. M. Shila, and A. Wolisz, "Guest editorial special issue on enabling a smart city: internet of things meets AI," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7469–7472, 2019.

[19] H. J. Feng, L. G. Chen, Z. H. Xie, and F. R. Sun, ""Volume-point" heat conduction constructal optimization based on minimization of maximum thermal resistance with triangular element at micro and nanoscales," *Journal of the Energy Institute*, vol. 89, no. 2, pp. 302–312, 2016.

[20] M. Yuan, Y. Li, L. Zhang, and F. Pei, "Research on intelligent workshop resource scheduling method based on improved NSGA-II algorithm," *Robotics and Computer-Integrated Manufacturing*, vol. 71, no. 12, pp. 102141–102149, 2021.

[21] T. Guo and Z. Deng, "An improved EMD method based on the multi-objective optimization and its application to fault feature extraction of rolling bearing," *Applied Acoustics*, vol. 127, pp. 46–62, 2017.

[22] J. Yang, J. Zhou, Z. Lv, W. Wei, and H. Song, "A real-time monitoring system of industry carbon monoxide based on wireless sensor networks," *Sensors*, vol. 15, no. 11, pp. 29535–29546, 2015.

[23] H. Hu, B. Tang, X. Gong, W. Wei, and H. Wang, "Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 2106–2116, 2017.

[24] W. Wei, Q. Xu, L. Wang et al., "GI/Geom/1 queue based on communication model for mesh networks," *International Journal of Communication Systems*, vol. 27, no. 11, pp. 3013–3029, 2014.

[25] F. Orujov, R. Maskeliūnas, R. Damaševičius, W. Wei, and Y. Li, "Smartphone based intelligent indoor positioning using fuzzy logic," *Future Generation Computer Systems*, vol. 89, pp. 335–348, 2018.

[26] W. Wei, S. Liu, W. Li, and D. Du, "Fractal intelligent privacy protection in online social network using attribute-based encryption schemes," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 3, pp. 736–747, 2018.

[27] L. Jiang, L. Chen, W. Wang, W. Wei, Z. Lv, and H. Wang, "Advanced network representation learning for container shipping network analysis," *IEEE Network*, vol. 35, no. 2, pp. 182–187, 2021.

[28] N. A. Khan, N. Z. Jhanjhi, S. N. Brohi, R. S. A. Usmani, and A. Nayyar, "Smart traffic monitoring system using unmanned aerial vehicles (UAVs)," *Computer Communications*, vol. 157, pp. 434–443, 2020.

[29] F. Outay, H. A. Mengash, and M. Adnan, "Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: recent advances and challenges," *Transportation Research Part A Policy and Practice*, vol. 141, pp. 116–129, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9985972, 15 pages
https://doi.org/10.1155/2021/9985972

WILEY | Hindawi

Research Article

Reinforcement Learning for Joint Channel/Subframe Selection of LTE in the Unlicensed Spectrum

Yuki Kishimoto, Xiaoyan Wang, and Masahiro Umehira

Graduate School of Science and Engineering, Ibaraki University, 4-12-1 Nakanarusawa, Hitachi, Ibaraki 316-8511, Japan

Correspondence should be addressed to Xiaoyan Wang; xiaoyan.wang.shawn@vc.ibaraki.ac.jp

Received 25 March 2021; Accepted 18 May 2021; Published 2 June 2021

Academic Editor: Keping Yu

In recent years, to cope with the rapid growth in mobile data traffic, increasing the capacity of cellular networks is receiving more and more attention. To this end, offloading the current LTE-advanced or 5G system's data traffic from licensed spectrum to the unlicensed spectrum that is used by WiFi systems, i.e., LTE-Licensed-Assisted-Access (LTE-LAA), has been extensively investigated. In the current LTE-LAA system, a Listen-Before-Talk (LBT) approach is implemented, which requires the LTE user also perform carrier sense before the transmission. However, fair LTE-WiFi coexistence is still hard to guarantee due to their unbalanced frame sizes and traffic loads. In the LTE-LAA system, the optimal channel selection and subframe number adjustment are the keys to realize efficient spectrum utilization and fair system coexistence. To this end, in this paper, we propose a reinforcement learning-based joint channel/subframe selection scheme for LTE-LAA. The proposed approach is implemented at the LTE access points with zero knowledge of the WiFi systems. The results of extensive simulations verify that the proposed approach can significantly improve the fairness and packet loss rate compared with baseline schemes.

1. Introduction

In the past few years, we have witnessed a phenomenal growth in mobile data traffic. This growth is accelerated by the increasing number of mobile and Internet of Things (IoT) devices and the popularity of spectrum-hungry wireless applications such as online games and high definition videos. Since most of the mobile data traffic is carried by cellular networks, both the academia and industry have made many attempts to increase the capacity of LTE networks to accommodate this surging growth and exploit the possible enhancement on the future 5G networks [1–3]. Originally, the mobile communication system LTE is capable of providing 150 (Mbps) data rate with a maximum bandwidth of 20 (MHz). As the demand for high-speed communication is further increasing, currently used LTE-advanced system utilizes Carrier Aggregation (CA) technology to speed up the communication by bundling multiple 20 (MHz) LTE carriers.

The highest licensed spectrum used for downlink LTE communication in Japan is 3.5 (GHz), and the unlicensed spectrum is 5 (GHz). Since the communication capacity is proportional to the frequency bandwidth, aggregating multiple noncontiguous channels in the unlicensed 5 (GHz) band enables higher capacity of LTE networks. However, the unlicensed band is already used by other wireless systems such as WiFi networks. Based on the fact that LTE is a schedule-based technology, which would severely degrade the performance of WiFi by letting its transmission backoffs continuously, it is necessary to modify LTE to enable the fair coexistence between different wireless systems. To this end, Licensed-Assisted-Access (LAA) technology [4] has been proposed in 2013, which uses Listen-Before-Talk (LBT) approach to let LTE system assess the channel state before transmitting. Additionally, in March 2020, 3GPP (Third-Generation Partnership Project) committed to 5G NR (New Radio) in unlicensed spectrum in Release 16, which extends the LAA from LTE to 5G [5, 6].

However, even with LBT, the fair coexistence issue between LTE and WiFi systems on unlicensed spectrum is still nontrivial, due to their unbalanced frame size and traffic volume. For instance, the transmission duration for LTE varies from 2 to 10 milliseconds, but a typical WiFi transmission only lasts for a few hundreds of microseconds [7]. Considering a dense scenario that multiple LTE and WiFi systems

colocate to share multiple channels, the optimal channel and subframe number (In this article, we use the term "subframe number" to indicate "the number of subframes"). Selections are the keys for efficient and fair spectrum utilization. Furthermore, in a dynamic network, it is very important and challenging to dynamically adjust the optimal channel and subframe number according to the varying environment.

The channel selection and subframe number adjustment problems have been widely investigated in recent years. The most common channel selection method is to select a channel with minimum received power by using channel assessment. This method, unfortunately, suffers from low channel utilization efficiency and fails to well support the dynamic network environment. A learning-based channel selection mechanism for LTE operation in unlicensed bands was proposed in [8]. However, it needs global information of the coexistence system and does not take fairness into consideration. In [9], Challita et al. proposed a proactive channel selection scheme for LTE-U system by exploiting deep learning. However, it requires a WiFi traffic load distribution dataset as input and thus is hard to be applied in a dynamic environment. In [10], an online learning distributed channel selection scheme for 5G NR-U has been proposed, which focuses on optimal channel selection for uplink traffic offloading by formulating it to a noncooperative game. Regarding the work related on subframe number adjustment, the idea of blank LTE subframe was first proposed by Almeida et al. [11]. The goal is to achieve fair medium access by giving more transmission opportunities to WiFi systems. Based on [11] and recent advances in learning techniques [12–14], a Q-learning-based muting period selection scheme was proposed for fair LTE-WiFi coexistence in [15]. However, it focuses on maximizing the LTE throughput and can only work in single channel scenario.

Recently, joint channel and subframe number selection problem has been investigated. A joint user association and resource allocation approach for LTE-WiFi coexistence was proposed in [16], which aims at maximizing the number of users supported by LTE. However, they do not consider the traffic balance problem between LTE and WiFi systems. In [17], a double Q-learning-based scheme was proposed to achieve efficient LTE-WiFi coexistence by jointly considering channel selection, discontinuous transmission, and transmit power control. However, the goal of this research is efficiently utilizing the idle time as much as possible, instead of achieving fair medium access. In [18], the authors proposed a duty cycle optimization scheme by considering both the fairness and the throughput. However, the proposed scheme assumed that the throughput information needs to be exchanged between LTE and WiFi systems to perform Q-learning, which is hard to realize in reality. In [19], a deep reinforcement learning-based dynamic resource allocation algorithm to reduce the latency of devices has been proposed, which focuses on reducing the latency of mission critical devices in accessing uplink resources of the small cell network. In [20], a novel framework that uses flying UAV-enabled networks to provide service for VR users in an LTE-U system has been proposed, which does not take into account the fairness between LTE and WiFi systems. In [21], the coexistence of LTE and ZigBee networks at the unlicensed frequency band of 2.4 GHz is studied, which focuses on performance evaluations. To summarize, the aforementioned works cannot be applied on an LTE-LAA system in a distributed and dynamic way, with the purpose of achieving fairness and efficiency simultaneously.

To this end, in this paper, we propose a joint channel/-subframe number selection scheme for LTE-LAA system by exploiting reinforcement learning technique. The proposed scheme is distributedly implemented at LTE Access Points (APs) and requires zero knowledge from the WiFi systems. In the proposed scheme, the LTE AP monitors the traffic volume on the currently used channel and dynamically learns the optimal channel and subframe number selections to efficiently utilize the spectrum resource and maintain each system's throughput to its target value as close as possible. To minimize the frequent channel switching created by traffic variation, we further propose an enhanced scheme with channel switch penalty. The effectiveness of the proposed scheme is verified by extensive simulation results. We compare the proposed scheme with baseline schemes to show its superiority in terms of fairness and packet loss rate. Although the proposed joint channel/subframe number selection scheme focuses on LTE-LAA system, it could be easily extended to 5G NR-U system.

The rest of the paper is organized as follows. Section 2 briefly introduces the preliminaries of LTE-LAA system. Section 3 presents the system model and the proposed reinforcement learning-based scheme. Section 4 provides the simulation scenario and evaluation results. Section 5 presents the issue on undesired channel switches in a complicated dense scenario and presents an enhanced scheme with evaluation results. Finally, Section 6 concludes this paper.

2. Preliminaries of LTE-LAA System

2.1. Channel Access Policy. In LTE-LAA systems, efficient and fair channel access methods are required. LBT mechanism [4] has been proposed, with the purpose that LTE communications in the unlicensed spectrum do not significantly deteriorate the performance of nearby WiFi systems. With LBT, LTE system also performs carrier sense before the transmission, and the transmission is initiated only when the channel is idle. Figure 1 shows the concept of LBT.

2.2. Channel Selection Scheme. Appropriate channel selection is important to fully utilize the spectrum resource, especially when there are multiple wireless systems colocate nearby. In a common channel selection scheme [22], APs periodically monitor all the channels and calculate the average received power at a constant interval. The channel with the minimum received power in the current interval will be selected to access at the next interval. Sensing-based channel selection scheme has the following drawbacks. Firstly, since channel sensing and communications cannot be performed simultaneously, the channel utilization efficiency is low. Secondly, since the sensing period is randomly assigned, the sensing result may not accurately reflect the real channel utilization

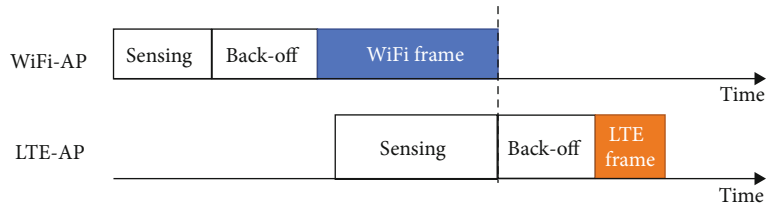


Figure 1: Conceptual diagram of LBT.

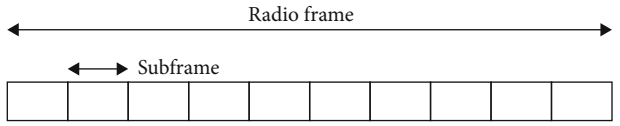


Figure 2: Radio frame structure for TDD downlink.

states. Last but not least, it works poor in a dynamic network environment due to the fixed long sensing cycle.

2.3. Radio Frame Structure for LTE. Similar as most of the previous researches [8, 9, 11, 15, 16], only downlink communication is considered in this work. Additionally, we assume that the LTE communications use Time Division Duplex (TDD), in which the radio frame structure is illustrated in Figure 2. The maximum number of subframes in one LTE frame is 10, and the length of one subframes is 1 (ms). Some of the subframes could be muted to give more channel access opportunities to potentially colocated systems such as WiFi [11, 23]. By dynamically varying subframe number, colocated LTE and WiFi systems could share the medium in a fair way.

2.4. LTE-U and LAA. An attempt to provide LTE communication in unlicensed frequency bands, i.e., LTE-Unlicensed (LTE-U) [24], has been originally proposed in 3GPP release 10. LTE-U mainly performs channel selection and duty cycle dynamic adjustment based on power measurements of the surrounding environment. However, the introduction of LTE-U may significantly degrade the performance of WiFi systems. To solve this problem, LTE-LAA was standardized in 3GPP release 13, which uses LBT to perform carrier sense before transmission. In this paper, the proposed scheme is based on LTE-LAA, in which the LBT is used.

3. Proposed Reinforcement Learning-Based Joint Channel/Subframe Selection Scheme

3.1. Reinforcement Learning. The proposed scheme is based on a typical reinforcement learning algorithm, i.e., Q-learning [25, 26]. Reinforcement learning is one of the three basic machine learning paradigms, by which the agent learns the optimal behavior through repeated interactions with the environment in discrete time steps. The learning is performed in an online fashion, and it does not need large amount of labeled data with correct input/output pairs. Figure 3 shows a conceptual diagram of Q-learning. In time step t, the agent observes the environmental state s_t and receives a reward r_t. The agent chooses an action a_t, and the environment evolves to the next state s_{t+1} and feeds reward r_{t+1} back to the agent. The action could be either

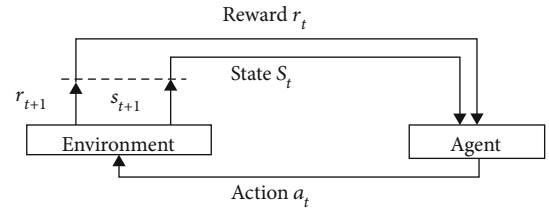


Figure 3: Conceptual diagram of the Q-learning.

Table 1: Definition of the proposed reinforcement learning-based scheme.

Parameter	Definition
State s	Low/high traffic volume
Action a	(channel, subframes number) selection pair
Reward R	$\begin{cases} (X_{\text{Own}}/L_{\text{frame}}) & \text{if (State : Low traffic volume)} \\ (X_{\text{Own}}/X_{\text{Own}} + X_{\text{Other}}) - \beta \cdot \rho_{\text{fair}} & \text{else} \end{cases}$

exploring the action space or exploiting the optimal action that gives the most cumulative reward as a result of a series of continuous actions. The optimal action is chosen based on a Q-table, in which the cell's value $Q(s_k, a_k)$, i.e., Q value, represents the value of the (state, action) pair. At the end of time slot t, the Q value is updated by

$$Q(s_t, a_t) \longleftarrow Q(s_t, a_t) + \alpha\{r_{t+1} + \gamma \max Q(s_{t+1}, a_t) - Q(s_t, a_t)\}, \quad (1)$$

where α is a learning rate and γ is a discount factor.

3.2. Proposed Scheme

3.2.1. Definition of the Proposed Scheme. In this paper, we propose a reinforcement learning-based joint channel/subframe number selection scheme, which is performed at individual LTE-APs. The state, action, and reward in the proposed scheme are defined as in Table 1. Specifically, we consider two states depending on the traffic volume of agent AP's currently used channel, i.e., low traffic volume state and high traffic volume state. The agent is in low traffic volume state if $X_{\text{Own}} + X_{\text{Other}} \leq L_{\text{frame}}$, i.e., the channel used by the AP, is unsaturated, or in high traffic volume state otherwise. Here, X_{Own} represents the agent AP's average subframe number in one learning cycle, and X_{Other} represents the average subframe number of all the other APs who share the same channel and are within the interference range with the agent

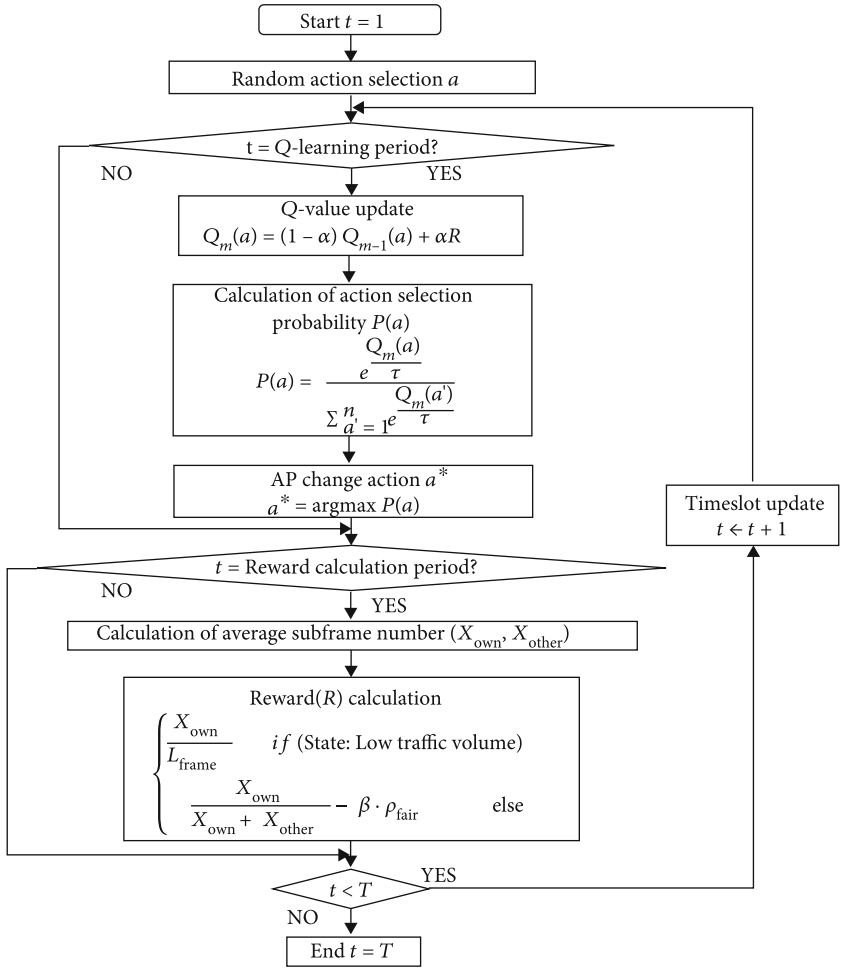


Figure 4: Flowchart of the proposed scheme.

AP. And L_{frame} is the subframe number in one frame. The actions are the (channel, subframe number) selection pairs. Notice that similar as the previous work [15, 17, 18], we focus on the selection of subframe number in this paper and leave the issue of selecting which subframe to use as the future work (In the simulations, the subframes are selected continuously in ascending order. More practical implementations that based on 3GPP standardization [4] will be considered in our future work). The reward functions are separately defined based on agent AP's current states as shown in Table 1. In low traffic state, the reward only focuses on its achieved throughput. But in high traffic state, both the throughput and fairness are taken into consideration. Here, β is a weight factor, ρ_{fair} denotes the fairness penalty which is calculated by Equation (2).

$$\rho_{\text{fair}} = \left| \frac{N_{\text{other}}}{1 + N_{\text{other}}} - \frac{X_{\text{Other}}}{X_{\text{Own}} + X_{\text{Other}}} \right|. \quad (2)$$

Here, $N_{\text{other}}/(1 + N_{\text{other}})$ represents the target fairness factor, where N_{other} is the number of APs that share the same channel and are within the interference range of the agent AP. Additionally, $X_{\text{Other}}/(X_{\text{Own}} + X_{\text{Other}})$ is the achieved fairness factor. The difference between them is defined as the fairness penalty, the lower it is, the fairer coexistence among different systems. The basic idea of the designed reward function is that in the low traffic volume state, all the APs could ideally send all the packets, but in the high traffic volume state, a fairness penalty is introduced to assure that all the APs using this channel could have equal opportunity to send the packets. Note that X_{Other} could be easily obtained by the agent AP thru channel sensing when it is not sending packets. N_{other} could be obtained by decoding the header the of the WiFi frames, which could be performed periodically when the scenario of the network is not highly dynamic. The interaction between LTE system and WiFi system is not required.

3.2.2. Flowchart and Action Selection Probability. Figure 4 shows the flowchart of the proposed scheme which is performed distributedly in each LTE-AP. Initially, the agent AP selects a random channel and subframe number. After that, when the time slot t equals the reward calculation period, the average subframe number $X_{\text{Own}}, X_{\text{Other}}$ in one cycle is calculated, and the corresponding reward R is obtained. When the time slot t equals the Q-learning period, the Q value ($Q(a)$) of action a is updated by Equation (3).

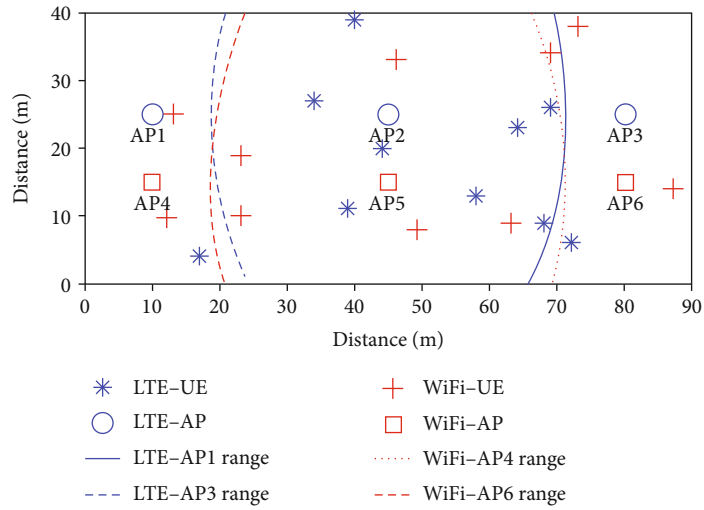


Figure 5: Simple LOS indoor scenario.

Table 2: Major parameters for simulations [15, 23].

Parameter	Value	Unit
Transmission power (P_{tx})	15	dBm
Antenna gain (G_{tx})	5	dB
Frequency band (f_c)	5	GHz
Bandwidth	20	MHz
Timeslot duration	9	μs
Timeslot number	40,000,000	
DIFS (T_{DIFS})	34	μs
SIFS (T_{SIFS})	16	μs
Backoff	[0,31]	Timeslot
Buffer size	10	Packets
Throughput calculation interval (T_{thr})	50,000	Timeslot
Radio frame length (L_{radio})	10	ms
Subframe length	1	ms

$$Q_m(a) = (1-\alpha)Q_{m-1}(a) + \alpha R, \tag{3}$$

where m is the number of times of Q-learning and α is the learning rate. It is considered that the larger the value of α, the easier it is to adapt to changing network conditions. In this paper, α evolves by Equation (4).

$$\alpha = 1 - \delta m, \tag{4}$$

where δ is a parameter that adjusts the changing speed of α. After the Q value is updated, the selection probability $P(a)$ for each action a is calculated by Equation (5).

$$P(a) = \frac{e^{(Q_m(a))/\tau}}{\sum_{a'=1}^{n} e^{(Q_m(a'))/\tau}}, \tag{5}$$

where n is the number of possible actions and τ is a design parameter representing the range of possible values of $P(a)$. The agent AP switches to the action $a^* = \text{argmax}\, P(a)$ to select the channel and subframe number.

To balance the exploration and exploitation, the parameter τ in Equation (5) gradually decreases by Equation (6)

$$\tau = \frac{\tau_0}{\log_2(1 + m/Z)}, \tag{6}$$

where τ_0 is an initial value of τ and Z is a parameter indicating the changing speed of τ. It is obvious that τ gradually decreases as m increases.

When τ is large, all actions are selected with almost the same probability regardless of the Q value (exploring for Q value). But when τ is small, the action with the largest Q value will be selected more easily (exploitation of Q value).

4. Simulation Results

4.1. Simulation Environment. We firstly consider a simple indoor LOS (Line-Of-Sight) environment to validate the performance of the proposed scheme. Similar to the previous work [8, 9, 11, 16, 17, 27], we only consider downlink communication in this work. Figure 5 shows the considered simple indoor LOS scenario with size $40(\text{m}) \times 90(\text{m})$, where three LTE-APs (APs 1-3) and three WiFi-APs (APs 4-6) are colocated to share two channels. The proposed joint channel/subframe number selection scheme is implemented at AP1 and AP3. The available channels are CH1 and CH2, and the selectable subframe numbers are [2, 4, 6, 8, 10]. To validate if the proposed scheme can adapt to the network's traffic load variations, we consider a dynamic network environment, in which the traffic volume at LTE-AP2 increases in the middle of simulations. To be specific, AP2 increases its subframe number from 1 to 10 at 180(s) and uses CH1 fixedly. Three WiFi APs use static channels, i.e., AP4, AP5, and AP6 assigned to CH1, CH1, and CH2, respectively. There are 10 LTE-UEs and 10 WiFi-UEs, each randomly placed in the network. Table 2 shows the major parameters

for simulations. WiFi system uses 802.11n standard protocol and is performed with frame aggregation.

Next, based on [28], we calculate the ranges of APs, beyond which the transmissions are unable to detect. The Threshold Level (TL) (/1 MHz) at which the carrier sensing of LBT is possible is calculated by $-73 + (23 - P_H)$, where P_H represents the transmission power P_{tx} when the antenna gain $G_{tx} = 0$ (dBi). According to the settings given in Table 2, P_H (dBm) of AP is calculated as $P_H = P_{tx} + (G_{tx} - 0) = 15 + 5 = 20$. Therefore, TL (/1 MHz) is calculated as $\{-73 + (23 - 20)\} = -70$. Since a 20 (MHz) bandwidth is assumed in the simulation, TL (dBm) is obtained as $10 \log_{10}(20/10^7)$. According to the settings given in Table 2, the received power P from the AP is calculated by Equation (7) according to [29].

$$P = -16.9 \log_{10} d - 12.8 - 20 \log_{10} f_c, \quad (7)$$

where d represents the distance (m) between the transmitting station and the receiving station and f_c represents the frequency band (GHz) used for communication. By considering the condition that P equals TL, we have the ranges as approximately 61.3 (m), beyond which the transmissions are unable to detect. The ranges of AP1, AP3, AP4, and AP6 are shown in Figure 5. Therefore, APs 3 and 6 are unable to detect the transmissions of APs 1 and 4 and vice versa.

The WiFi parameters used in this paper are shown in Table 3 based on [15, 30, 31]. Accordingly, we have the WiFi data packet length T_{data} and the ACK length T_{ACK} as 704 (μs) and 28 (μs), respectively. Table 4 shows the parameters for the proposed scheme and common channel selection method. Here, the sensing period and sensing time for the sensing-based scheme are set based on [22], which corresponds to a 1% sensing time. The learning period and reward calculation period used in the proposed scheme are set to 50000 timeslots, which indicates that the AP may change its action every 450 (ms). The other parameters in Table 4 are Q learning's parameters which are carefully adjusted to balance the performance, convergence, and adaptive capacity. The LTE and WiFi packet arrival intervals follow an exponential distribution $\lambda e^{-\lambda x}$ where x is the packet arrival interval and $\lambda = 1/\mu$. The average packet arrival interval (ms) of LTE-AP, μ_{LTE} is set as

$$\mu_{\text{LTE}} = T_{\text{DIFS}} + B_M + L_{\text{radio}} + T_{\text{ACK}} + T_{\text{SIFS}} = 10.2195, \quad (8)$$

where B_M is the average backoff. And the average packet arrival interval of WiFi-AP, μ_{WiFi}, for WiFi APs 4, 5, and 6 are set to 1.42 and 5.68 (ms), respectively. These average packet arrival intervals correspond to LTE's subframe numbers 1 and 4, respectively.

4.2. Evaluation Metrics. We compare the proposed scheme with two baseline channel selection schemes, i.e., sensing-based scheme and max-throughput scheme. Specifically, sensing-based scheme selects the channel with the minimum received power at a fixed interval. The sensing time is 1% of the whole cycle which is randomly assigned. Additionally, the max-throughput scheme is an ideal method, which chooses the channel that obtains the maximum system throughput.

Table 3: Major parameters for WiFi.

Parameter	Value	Unit
PLCP preamble + headers duration (T_{plcp})	20	μs
PLCP service field (L_S)	16	Bits
MAC header ($L_{\text{MAC\_h}}$)	224	Bits
Tail bits (L_t)	6	Bits
ACK length (L_{ack})	112	Bits
Payload (D)	12000	Bits
OFDM symbol duration (T_{sym})	4	μs
Number of bits per OFDM symbol (n_{sym})	72	Bits

Table 4: Major parameters for the proposed scheme and sensing-based scheme [22].

Parameter	Value	Unit
Learning period	50,000	Timeslot
Initial Q value	0.5	
τ_0	0.4	
Z	35	
β	2	
δ	0.00025	
Reward calculation period	50,000	Timeslot
Sensing period	4,000,000	Timeslot
Sensing time	40,000	Timeslot/period

To evaluate the effectiveness of the proposed scheme, we consider three evaluation metrics, i.e., throughput, fairness, and packet loss rate.

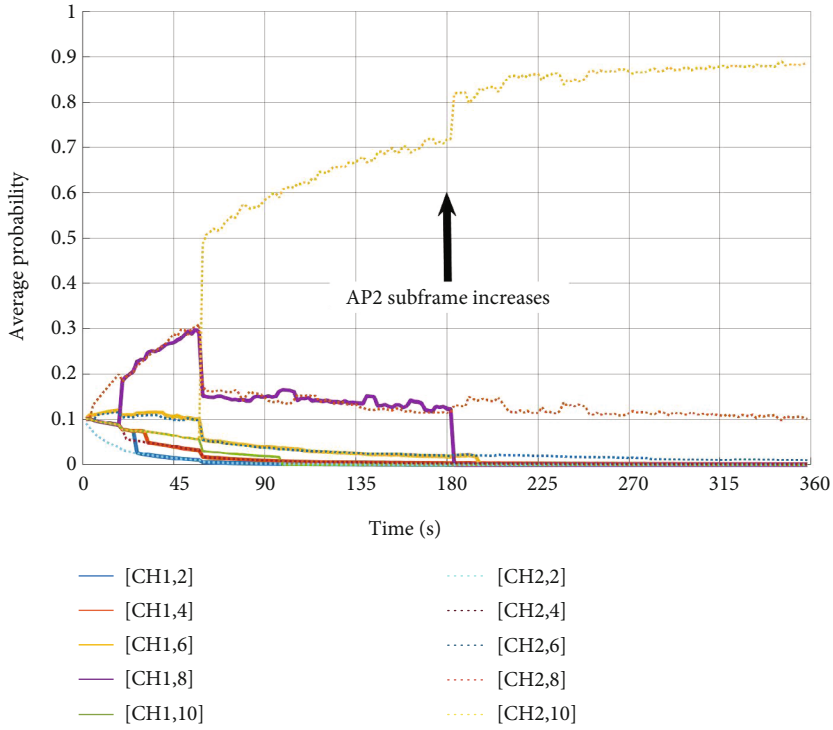
The throughput (Γ) is calculated by Equation (9).

$$\Gamma = r \times \frac{T_{\text{trans}}}{T_{\text{thr}}}, \quad (9)$$

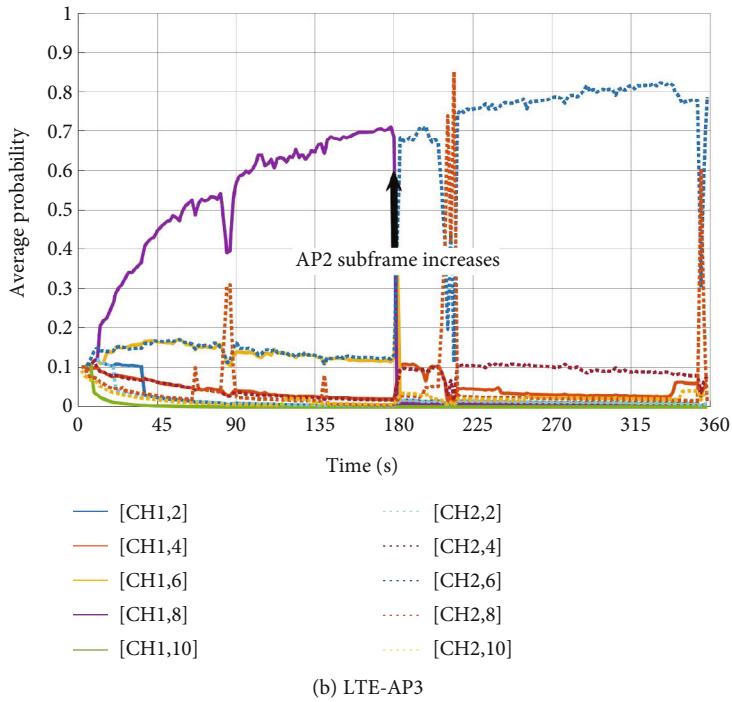
where r is the data rate and T_{trans} and T_{thr} indicate the AP's transmission time and the throughput calculation interval, respectively. QPSK (Quadrature Phase Shift Keying) modulation is adopted, and thus, the data rates for LTE and WiFi communications are 15.6 and 18 (Mbps), respectively [30]. In this work, the fairness (Ψ) needs to take into consideration the different traffic volume at different APs, which is defined by Equation (10).

$$\Psi = \frac{\Gamma}{\Gamma_{\text{ideal}}}, \quad (10)$$

where Γ_{ideal} is the throughput achieved when the AP operates in an isolated manner, i.e., access the channel without sharing it with any other systems. A low fairness value indicates that the AP's throughput is significantly affected by other systems that share the same channel.



(a) LTE-AP1



(b) LTE-AP3

Figure 6: Action selection probability for the proposed scheme.

The packet loss occurs when the queuing packets exceed the buffer size, and the lost packets plus the packets in the buffer are counted when we calculate the packet loss rate.

4.3. Simulation Results. First, we confirm the adaptive capacity of the proposed scheme in a dynamic network environment. Figure 6 shows the variance of action selection probability over time of LTE-AP1 and AP3. To make it clear, the illustrated probability is averaged by 4 times. The legend denotes (channel, subframe number). The arrow in the middle indicates the timing at which LTE-AP2 changes its subframe number proactively. We can observe that the action selection probability converges as time goes by. This indicates that the process of exploring for the Q value is gradually changing to the process of exploiting. From Figure 6(a), we can observe that AP1's highest action selection probability

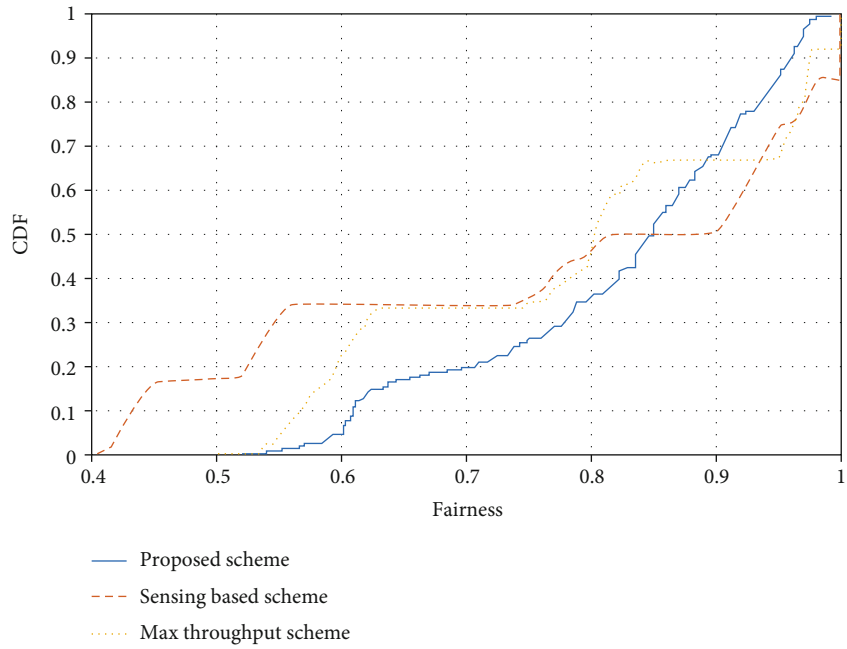


Figure 7: CDF of fairness for different schemes.

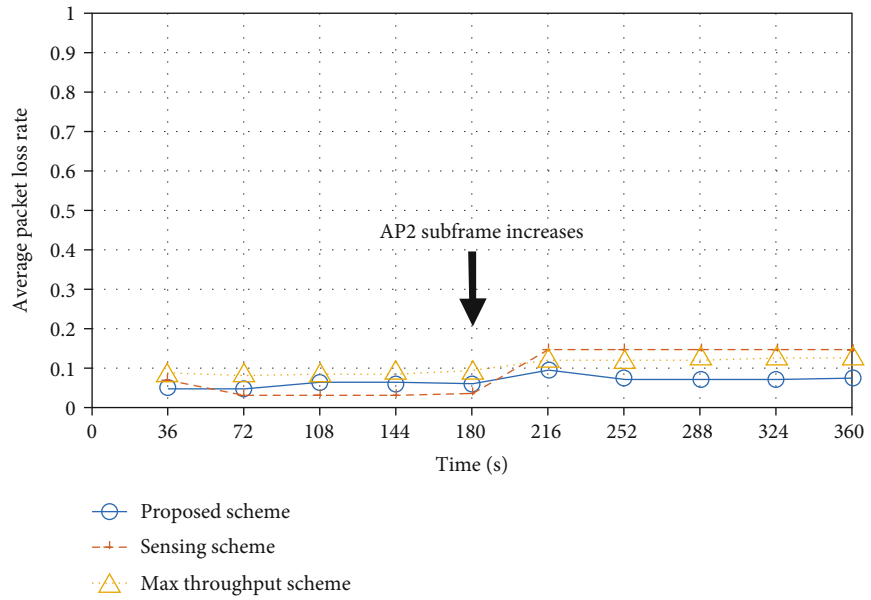


Figure 8: Average packet loss rate of all APs for different schemes.

is [CH2,10], which is the optimal result as expected. This is because there is no other AP in the range of AP1 that shares CH2, and thus, $X_{\text{Other}} = 0$. Therefore, large X_{Own} value can be obtained in the range of $X_{\text{Own}} + X_{\text{Other}} \leq L_{\text{frame}}$, which maximizes the reward. From Figure 6(b), for AP3, the probability of action [CH1,8] is the highest in most of the first half time. This is because subframe number of AP2 is 1, and thus, the value of X_{Other} is smaller in CH1 than that in CH2. Therefore, large reward could be obtained in the range of $X_{\text{Own}} + X_{\text{Other}} \leq L_{\text{frame}}$. However, when the subframe number of AP2 increases from 1 to 10 in the middle of the simulation, AP3's state changes from low traffic volume state into high traffic volume state. Accordingly, AP3's action with the highest probability changes to [CH2,6] immediately, which is as expected. This is because that using CH2 by sharing with AP6 could achieve higher reward. In addition, we can observe that the probability of [CH2,8] becomes high temporarily at around 210 (s). The reason is that the traffic volume of other APs within the transmission range is temporarily reduced at that time, due to the exponential distribution; thus, the reward of [CH2,8] becomes higher. Notice that the temporary changing of the action does not lead to

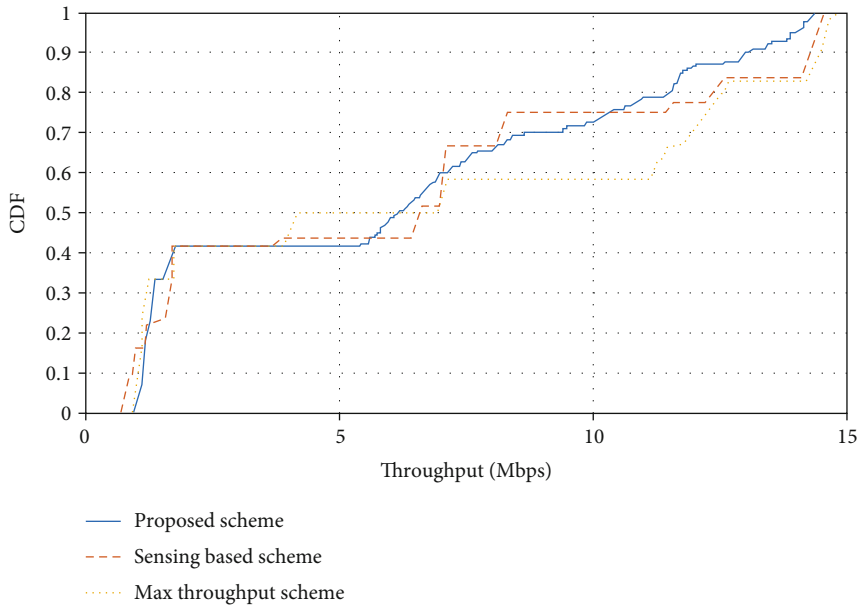


Figure 9: CDF of throughput for different schemes.

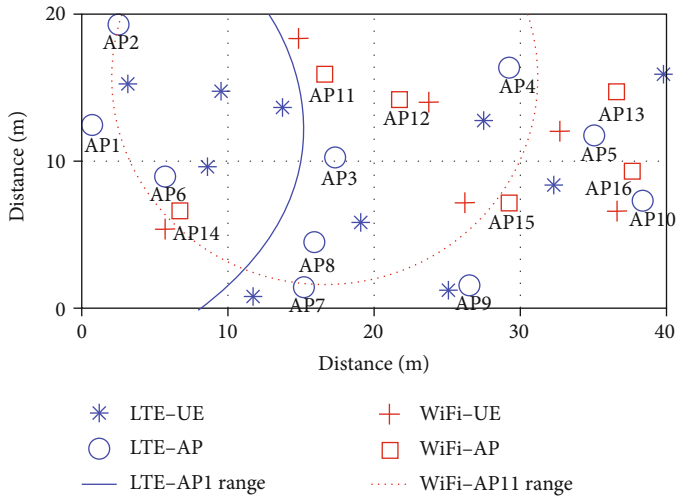


Figure 10: Dense indoor NLOS scenario.

channel switching, only the subframe number varies. Based on the previous analysis, we can conclude that the proposed scheme can adapt to the dynamic network environment automatically by selecting the ideal channel and subframe number in both low and high traffic volume states.

Next, in Figure 7, we show the Cumulative Distribution Function (CDF) of fairness for the proposed scheme, sensing-based scheme, and max-throughput scheme. All the results are averaged by three runs. We can observe that the proposed scheme significantly improves the fairness with low values, i.e., equal or lower than 0.84. This means that the proposed scheme can benefit the APs with poor performance. Additionally, the whole system's average fairness of the proposed scheme is 0.82, which is better than 0.77 for the sensing-based scheme and 0.79 for the max-throughput scheme.

Next, Figure 8 shows the comparison of the average packet loss rate of the proposed scheme, sensing-based scheme, and max-throughput scheme. We can observe that average packet loss rate of the proposed scheme is lower than that of two baseline schemes, when the system traffic volume is high.

Finally, Figure 9 shows the CDF of throughput of the proposed scheme, sensing-based scheme, and max-throughput scheme. As expected, the average throughput for all APs of the max-throughput scheme is the highest, i.e., approximately 6.8 (Mbps), compared with that of the proposed scheme and sensing-based scheme, i.e., 6.1 (Mbps) and 6.2 (Mbps), respectively. By the observations on the specific throughput results on different APs which have not been shown in this article due to limited space, we found that the low- and high-throughput values shown in Figure 9

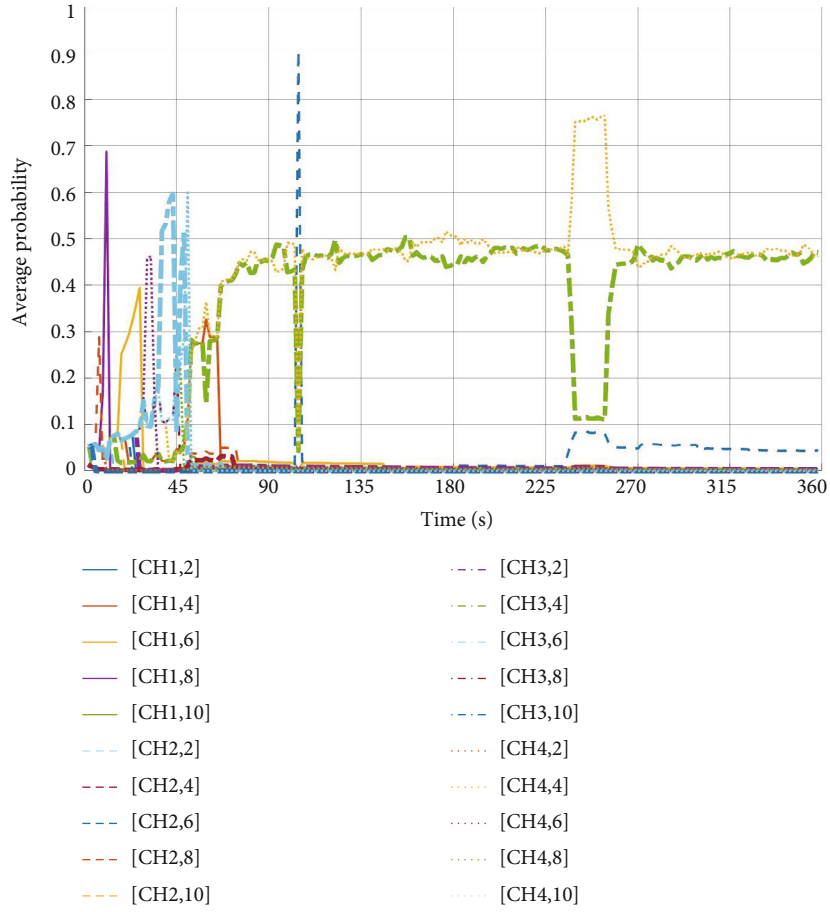


Figure 11: Action selection probability of AP5 for the original proposed scheme ($Z = 1$, $\delta = 0.001$).

correspond to WiFi and LTE systems' throughput, respectively. From Figure 9, we can observe that the throughput improvement of the max-throughput scheme mainly comes from the LTE-AP whose throughput is higher than 7 (Mbps). The reason is that in the max-throughput scheme, the LTE-AP uses the channel that leads to its maximal throughput with subframe number 10, without considering the negative impacts on WiFi systems. Notice that the max-throughput scheme is an ideal scheme which could only be realized in simulations.

5. Enhanced Joint Channel/Subframe Number Selection Scheme

5.1. Existing Problems and the Enhanced Proposed Scheme. To validate if the proposed scheme can work in a complicated scenario, we consider an extremely dense NLOS (Non-Line-Of-Sight) indoor environment in which the traffic is saturated in all channels [32]. As illustrated by Figure 10, there are 8 rooms (each of size 10 (m) × 10 (m)), and 10 LTE-APs (APs 1-10) and 6 WiFi-APs (APs 11-16) colocate to share 4 channels. There are 10 LTE-UEs and 6 WiFi-UEs in this scenario, which are associated to the closest APs. In this scenario, multiple APs are located within each other's communication range; thus, the traffic variance of one AP will affect the performance of the whole system significantly. The proposed scheme is implemented at all LTE-APs (APs 1-10). The available channels are CH1 to CH4, and the selectable subframe numbers are [2, 4, 6, 8, 10]. We still consider a dynamic network environment, in which the assigned channel at one WiFi-AP changes in the middle of the simulations. To be specific, AP14 proactively switches from CH3 to CH4 at 180 (s).

As an example, the ranges of LTE-AP1 and WiFi-AP11 are shown in Figure 10, beyond which the transmissions of AP1 and AP11 cannot be detected. These ranges at NLOS scenario are calculated as follows. TL (/1 MHz) at which the carrier sensing of LBT is possible is given by $-73 + (23 - P_H)$, so the receiving power is calculated by Equation (11) according to [29].

$$P = -43.3 \log_{10} d + 8.5 - 20 \log_{10} f_c. \quad (11)$$

By considering the condition that P equals TL, we have the ranges for NLOS scenario as approximately 14.4 (m), beyond which the transmissions cannot be detected.

μ_{WiFi} for WiFi AP 11, 13, 14, 16, 12, and 15 are set to 2.39 and 1.43 (ms), respectively. These average packet arrival intervals correspond to LTE subframe numbers 3 and 5, respectively.

In this extremely dense NLOS indoor scenario, we found that some of the APs' action selection probabilities do not

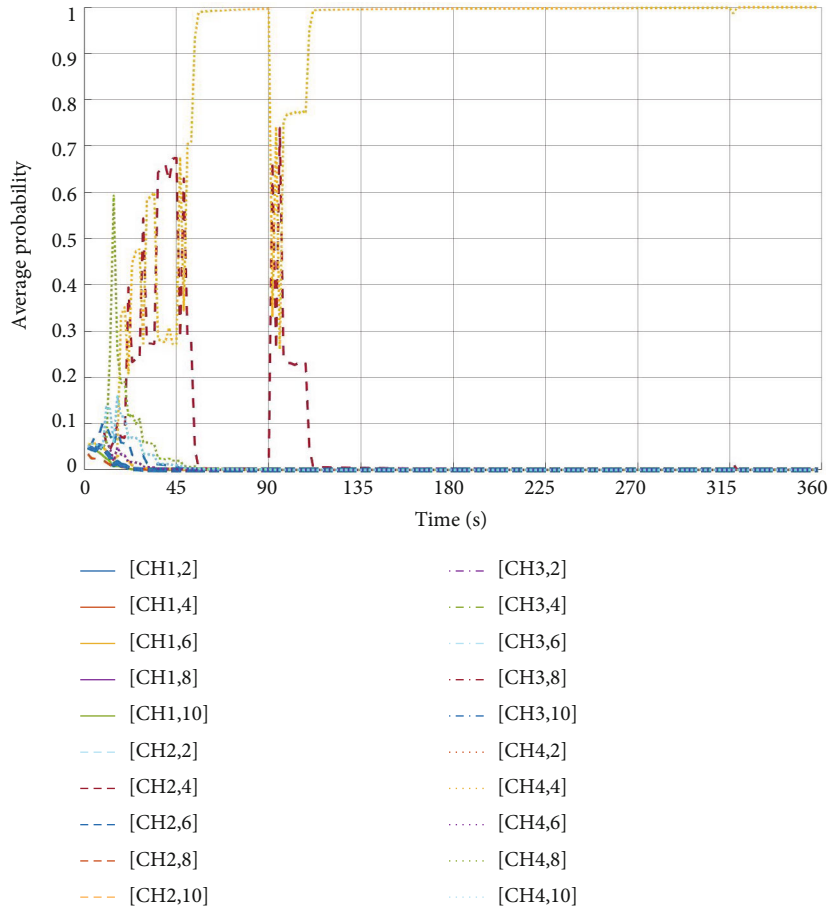


Figure 12: Action selection probability of AP5 for the enhanced proposed scheme.

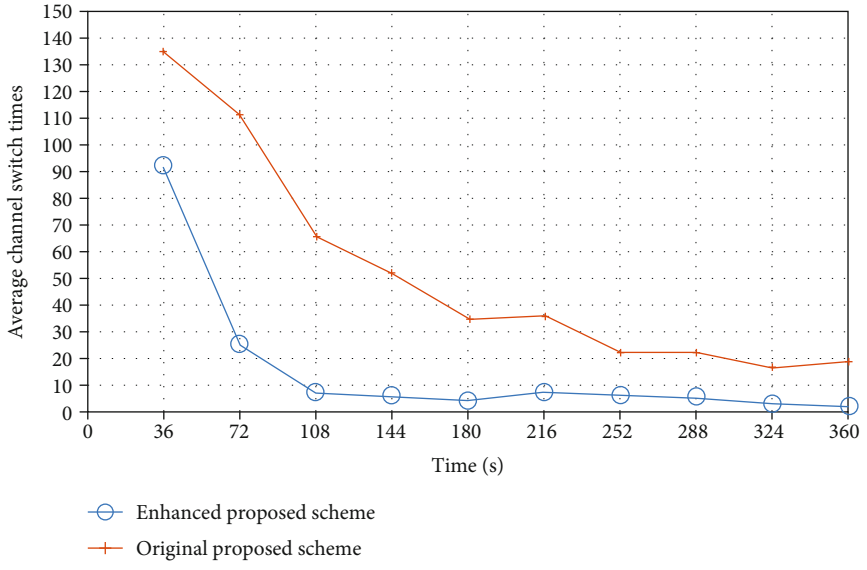


Figure 13: Average number of channel switches of all APs for different schemes.

converge by using the original proposed scheme. For instance, Figure 11 shows the variance of action selection probability over time for LTE-AP5. We can observe that CH3 and CH4 have the similar probabilities, and thus, the proposed scheme does not converge to one optimal action. The reason is that the traffic volume of CH1 and CH2 is almost saturated, and there is approximately the same amount of idle time in CH3 and CH4. This result means that using either CH3 or CH4 would lead to the same performance for AP5. However, in practice, it is not desirable to

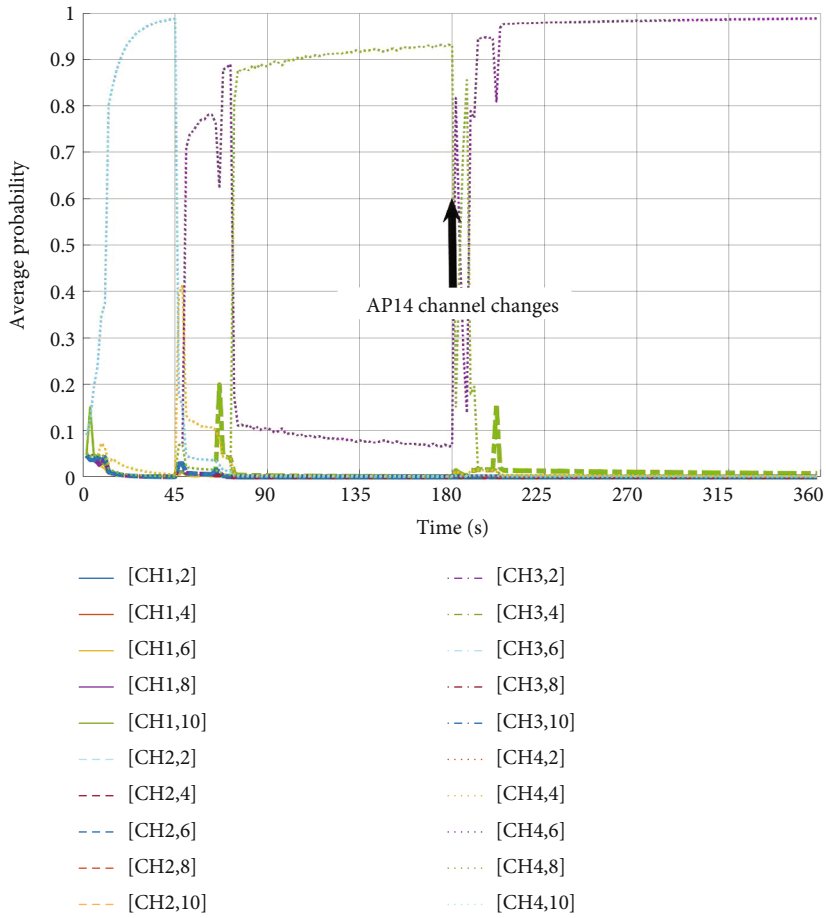


Figure 14: Action selection probability of AP1 for the enhanced proposed scheme.

switch channels frequently during communications. Noted that the sensing of channels also requires periodical channel switches. However, the channel switching during communications leads to not only the AP changes its channel but also all the connected end users.

To solve this problem, we further propose an enhanced joint channel/subframe number selection scheme to deal with this undesired frequent channel switches. In the enhanced proposed scheme, after the action selection probability is derived based on Equation (5), if the channel switch happens, the Q value is further updated by Equation (12).

$$Q_m(a^{\sim}) = Q_m(a^{\sim}) + \rho, \tag{12}$$

where $a^{\sim}$ represents the actions of channels other than the current channel. ρ represents a channel switch penalty which is set to -0.1 in the simulations. Since the channel switch penalty lowers the Q value for channels other than the channel currently used by the AP, the action of switching to another channel is suppressed. Additionally, in the enhanced proposed scheme, Z and δ are set to 10 and 0.001, respectively.

5.2. Simulation Results. Firstly, we verify that the enhanced proposed scheme could reduce the undesired frequent channel switches. In Figure 12, we show the variances of action selection probabilities over time of LTE-AP5. Compared with the result shown in Figure 11, we can observe that the action selection probability converges on [CH4,4] and approaches to 1 gradually. Thanks to the introduction of channel switch penalty, the frequent channel switch problem is solved. At around 90 (s), there is a temporary variance for action selection probability from [CH4,4] to [CH2,4]. The reason is that the traffic volume for all the APs is dynamically changing based on the exponential distribution. This temporary probability changing is undesired, since it leads to channel switches. The balance between stability and adaptability of the proposed scheme is our future work.

Next, Figure 13 shows the comparison of average number of channel switches between the enhanced proposed scheme and original proposed scheme. All the results are averaged by three runs. We can observe that average channel switch times of the enhanced proposed scheme are significantly lower than that of original proposed scheme. It could suppress the channel switch times to at most 10 times per 36 (s) after the scheme converges from 108 (s).

Next, we confirm that if the introduction of channel switch penalty would bring undesired negative effects on the adaptive capacity in a dynamic network environment. Figure 14 shows the variances of action selection probabilities over time of LTE-AP1 as an example. The arrow in the

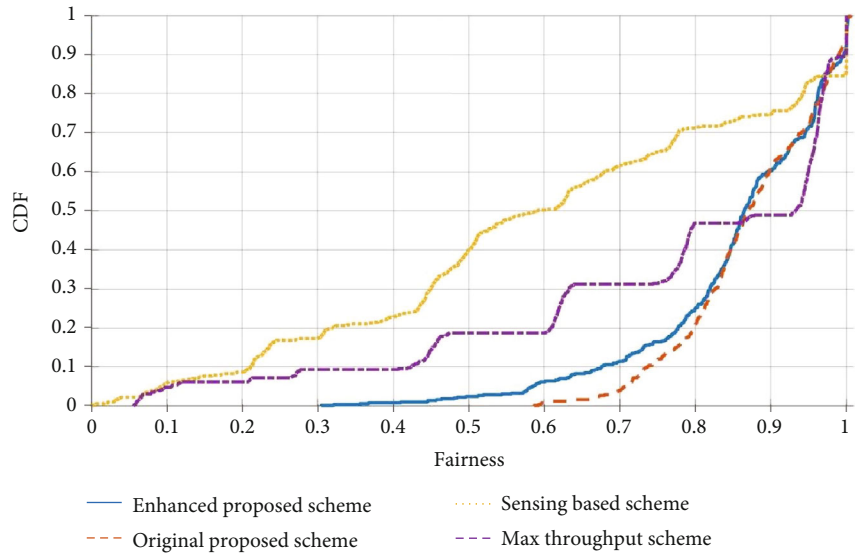


Figure 15: CDF of the fairness for different schemes.

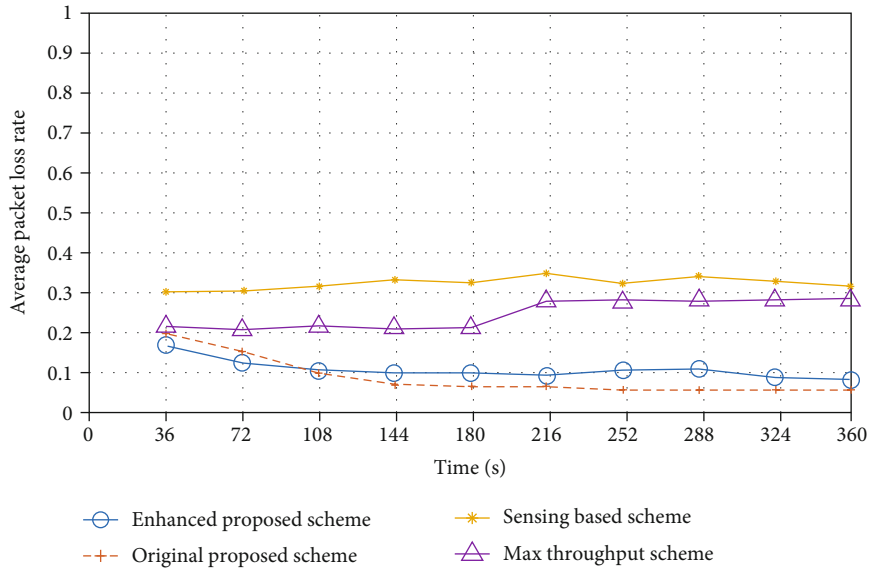


Figure 16: Average packet loss rate of all APs for different schemes.

middle indicates the timing that WiFi-AP14 proactively switches its channel. From Figure 14, we can observe that the probability of action [CH4,8] is the highest after the convergence in the first half time. This is because there is no other AP in the range of AP1 that shares CH4, and thus, $X_{\text{Other}} = 0$. Therefore, large X_{Own} value can be obtained in the range of $X_{\text{Own}} + X_{\text{Other}} \leq L_{\text{frame}}$, which maximizes the reward. However, when AP14 switches from CH3 to CH4 in the middle of the simulation, AP1's action with the highest probability changes to [CH4,6] accordingly, which is as expected. This is because that using CH4 by sharing with AP14 could achieve higher reward. Additionally, as for the selection results of other APs in this dynamic scenario, we confirm that they all make appropriate action adjustments according to this environment variance.

Next, we compare the performance of the enhanced proposed scheme, original proposed scheme, sensing-based scheme and max-throughput scheme in this dense NLOS scenario. In Figure 15, we show the CDF of fairness. We can observe that both the enhanced proposed scheme and original proposed scheme can significantly improve the fairness, especially for low fairness values, i.e., equal or lower than 0.96. Additionally, the whole system's average fairness of the enhanced proposed scheme is 0.85, which is better than 0.60 for the sensing-based scheme, 0.77 for the max-throughput scheme, and almost equal to 0.87 for the original proposed scheme.

Next, Figure 16 shows the comparison of the average packet loss rate. We can observe that the average packet loss rate of the enhanced proposed scheme is extremely lower

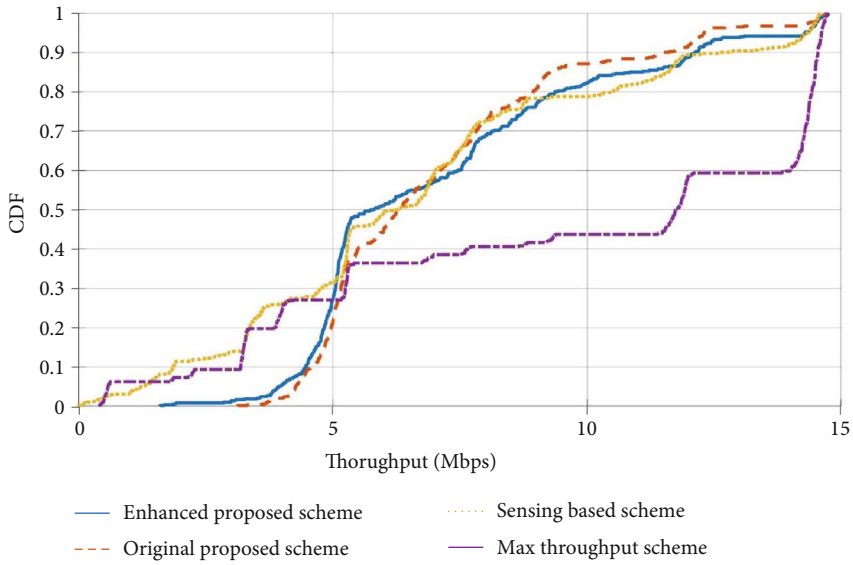


Figure 17: CDF of the throughput for different schemes.

than that of the sensing-based scheme and max-throughput scheme and slightly higher than that of the original proposed scheme. Moreover, for the max-throughput scheme, we can observe that the average packet loss rate increases by about 10% in the second half time. This indicates that the max-throughput scheme is unable to handle the channel switch of AP14. The average packet loss rates of the enhanced proposed scheme and the original proposed scheme are 0.11 and 0.090, respectively, which significantly outperform the sensing-based scheme and the max-throughput scheme, i.e., 0.33 and 0.25, respectively.

Finally, Figure 17 shows the CDF of average throughput. We can observe that both the enhanced proposed scheme and original proposed scheme could improve the throughput with low values, i.e., less than 5.0 (Mbps). Additionally, we can confirm that the average throughput of the max-throughput scheme is the highest, i.e., 9.5 (Mbps). And the average throughput of the enhanced proposed scheme is 7.2 (Mbps), which is better than 6.7 (Mbps) for the sensing-based scheme and 7.0 (Mbps) for the original proposed scheme.

To summarize, the enhanced proposed scheme can achieve almost the same fairness, packet loss rate, and throughput as the original proposed scheme, with significantly reduced channel switch times. Most importantly, the enhanced proposed scheme can realize system stability and adaptability simultaneously.

6. Conclusions

In this paper, we proposed a joint channel/subframe number selection scheme for the LTE-LAA system. It is able to achieve efficient channel utilization and fair system coexistence by exploiting reinforcement learning technique. We evaluated the effectiveness of the proposed scheme by computer simulations in two dynamic indoor environments and compared it with two baseline schemes. By using the proposed scheme, the optimal channel/subframe number can be selected even when the network conditions dynamically change. Compared with baseline schemes, both the fairness and packet loss rate are significantly improved. Especially in the extremely dense NLOS indoor scenario, by introducing the channel switch penalty, the system stability and adaptability are realized at the same time. In future work, we will consider using real WiFi traffic dataset in the simulations and optimize the learning parameters.

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.

Acknowledgments

This research was supported by Grant-in-Aid for Scientific Research (C) (20K11764), the Telecommunications Advancement Foundation, and ROIS NII Open Collaborative Research (21FA01).

References

[1] W. S. H. M. W. Ahmad, N. A. M. Radzi, F. S. Samidi et al., "5G technology: towards dynamic spectrum sharing using cognitive radio networks," *IEEE Access*, vol. 8, pp. 14460–14488, 2020.

[2] K.-L. A. Yau, J. Qadir, C. Wu, M. A. Imran, and M. H. Ling, "Cognition-inspired 5G cellular networks: a review and the road ahead," *IEEE Access*, vol. 6, pp. 35072–35090, 2018.

[3] X. Wang, M. Umehira, B. Han, H. Zhou, P. Li, and C. Wu, "An efficient privacy preserving spectrum sharing framework for internet of things," *IEEE Access*, vol. 8, pp. 34675–34685, 2020.

[4] *Study on Licensed-Assisted Access Using LTE*, 3GPP Study Item RP-141397, Edinburgh, Scotland, 2014.

[5] J. Oh, Y. Kim, Y. Li, J. Bang, and J. Lee, "Expanding 5G new radio technology to unlicensed spectrum," in *2019 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, Waikoloa, HI, USA, December 2019.

[6] K. Yu, L. Lin, M. Alazab, L. Tan, and B. Gu, "Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020.

[7] "IEEE Standard for information technology—telecommunications and information exchange between systems local and metropolitan area networks—specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," in *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)*, pp. 1–3534, IEEE, 2016.

[8] O. Sallent, J. Perez-Romero, R. Ferrus, and R. Agusti, "Learning-based coexistence for LTE operation in unlicensed bands," in *2015 IEEE International Conference on Communication Workshop (ICCW),*, London, UK, June 2015.

[9] U. Challita, L. Dong, and W. Saad, "Proactive resource management for LTE in unlicensed spectrum: a deep learning perspective," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4674–4689, 2018.

[10] Y. Shi, Q. Cui, W. Ni, and Z. Fei, "Proactive dynamic channel selection based on multi-armed bandit learning for 5G NR-U," *IEEE Access*, vol. 8, pp. 196363–196374, 2020.

[11] E. Almeida, A. M. Cavalcante, R. C. D. Paiva et al., "Enabling LTE/WiFi coexistence by LTE blank subframe allocation," in *2013 IEEE International Conference on Communications (ICC),*, Budapest, Hungary, June 2013.

[12] Y. Li, J. Zhou, J. Tian, X. Zheng, and Y. Y. Tang, "Weighted error entropy based information theoretic learning for robust subspace representation," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2021.

[13] Y. M. Li, J. T. Zhou, X. W. Zheng, J. Y. Tian, and Y. Y. Tang, "Robust subspace clustering with independent and piecewise identically distributed noise modeling," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2019.

[14] J. Zhang, K. Yu, Z. Wen, X. Qi, and A. K. Paul, "3D reconstruction for motion blurred images using deep learning-based intelligent systems," *Computers, Materials & Continua*, vol. 66, no. 2, pp. 2087–2104, 2021.

[15] V. Maglogiannis, D. Naudts, A. Shahid, and I. Moerman, "A Q-learning scheme for fair coexistence between LTE and Wi-Fi in unlicensed spectrum," *IEEE Access*, vol. 6, pp. 27278–27293, 2018.

[16] J. Tan, S. Xiao, S. Han, Y. Liang, and V. C. M. Leung, "QoS-aware user association and resource allocation in LAA-LTE/-WiFi coexistence systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2415–2430, 2019.

[17] A. Glanopoulos, F. Foukalas, and T. A. Tsiftsis, "Efficient coexistence of LTE with WiFi in the licensed and unlicensed spectrum aggregation," *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 2, pp. 129–140, 2016.

[18] Y. Su, X. Du, L. Huang, Z. Gao, and M. Guizani, "LTE-U and Wi-Fi coexistence algorithm based on Q-learning in multi-channel," *IEEE Access*, vol. 6, pp. 13644–13652, 2018.

[19] M. Elsayed and M. Erol-Kantarci, "Deep reinforcement learning for reducing latency in mission-critical services," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, December 2018.

[20] M. Chen, W. Saad, and C. Yin, "Echo state learning for wireless virtual reality resource allocation in UAV-enabled LTE-U networks," in *2018 IEEE International Conference on Communications (ICC)*, Kansas City, MO, USA, May 2018.

[21] I. Parvez, N. Islam, N. Rupasinghe, A. I. Sarwat, and I. Guvenc, "LAA-based LTE and Zig Bee coexistence for unlicensed-band smart grid communications," in *SoutheastCon 2016*, pp. 1–6, Norfolk, VA, USA, March-April 2016.

[22] S. Sengottuvelan, J. Ansari, P. Mähönen, T. G. Venkatesh, and M. Petrova, "Channel selection algorithm for cognitive radio networks with heavy-tailed idle times," *IEEE Transactions on Mobile Computing*, vol. 16, no. 5, pp. 1258–1271, 2017.

[23] A2A Research Inc, "LTE/LTE-A basics," 2020-7-2, http://www.a2a.jp/resources/lte_lte-A.pdf.

[24] S. Zinno, G. Di Stasi, S. Avallone, and G. Ventre, "On a fair coexistence of LTE and Wi-Fi in the unlicensed spectrum: a survey," *Computer Communications*, vol. 115, pp. 35–50, 2018.

[25] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[26] H. Zhou, X. Wang, M. Umehira, X. Chen, C. Wu, and Y. Ji, "Wireless access control in edge-aided disaster response: a deep reinforcement learning-based approach," *IEEE Access*, vol. 9, pp. 46600–46611, 2021.

[27] Y. Li, R. Liang, W. Wei, W. Wang, J. Zhou, and X. Li, "Temporal pyramid network with spatial-temporal attention for pedestrian trajectory prediction," *IEEE Transactions on Network Science and Engineering*, 2021.

[28] *ETSI, 2014a, EN 301 893, V1.7.2., Broadband Radio Access Networks (BRAN); 5 GHz High Performance RLAN; Harmonized EN Covering the Essential Requirements of Article 3.2 of the R & TTE Directive*, Sophia Antipolis: ETSI, 2014.

[29] *3GPP TR 36.814, V9.0.0, Further Advancements for E-UTRA Physical Layer Aspects*, 3GPP TR 36.814 V9.0.0 Release 9, 2010.

[30] M. Mehrnoush, V. Sathya, S. Roy, and M. Ghosh, "Analytical modeling of Wi-Fi and LTE-LAA coexistence: throughput and impact of energy detection threshold," *IEEE/ACM Transactions on Networking*, vol. 26, no. 4, pp. 1990–2003, 2018.

[31] S. Kubota and M. Morikura, *802.11 High-Speed Wireless LAN Textbook*, Impress, 2005.

[32] A. M. Voicu, L. Simić, J. P. de Vries, M. Petrova, and P. Mähönen, "Risk-informed interference assessment for shared spectrum bands: a Wi-Fi/LTE coexistence case study," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 3, pp. 505–519, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9403874, 11 pages
https://doi.org/10.1155/2021/9403874

Research Article

News Information Platform Optimization Based on the Internet of Things

Hongyun Tan and Yiping Li

School of Journalism & Communication, Jinan University, Guangzhou 510632, China

Correspondence should be addressed to Hongyun Tan; thayer99@stu2018.jnu.edu.cn

Received 15 April 2021; Revised 11 May 2021; Accepted 19 May 2021; Published 29 May 2021

Academic Editor: Wei Wang

The Internet of Things device online recommendation system has been applied in some Internet of Things operating companies and has achieved good results. In the process of designing and implementing the Internet of Things equipment online promotion system, this article uses the news information protocol transmission structure language to explain the use case analysis and activity diagram analysis of the Internet of Things equipment online promotion system and uses the Spring Hibernate (SH) integration framework in the field of the news information topology layer under the Internet of Things. Then, design and implement the technical architecture and main functional modules of the Internet of Things device online recommendation system, effectively improving the development efficiency and operating quality of the Internet of Things device online recommendation system. The system design concept and implementation ideas can be used as a reference for related industries when developing enterprise applications. This article uses the news information protocol transmission structure language to explain the demand analysis of the Internet of Things equipment online promotion system and mainly discusses the analysis of the use case of the Internet of Things equipment online promotion system and the analysis of core business activity diagrams. We completed the design of the Internet of Things device online promotion system based on the news information topology layer platform under the Internet of Things, mainly using the integration framework in the field of the news information topology layer under the Internet of Things to design the technical architecture of the Internet of Things device online promotion system, and design the Internet of Things of the function module structure and data table structure of the equipment online recommendation system. At the same time, they complete the realization of the main functions of the Internet of Things device online recommendation system; elaborate on the realization process of core modules such as Internet of Things device category management, Internet of Things device information management, announcement management, and device recommender management; and discuss the system testing process and application effects.

1. Introduction

With the development of the news information industry, the volume of news information business continues to grow rapidly. Therefore, the national news communication application market and the development of the Internet of Things are also very rapid. There are now more than 300 million news information users in the world, and there are also millions of Internet users. Experts predict that in recent years, the number of users applying network technology and Internet of Things user technology will continue and rapidly increase. It can be said that these two technologies will greatly change our lifestyle in the next century. News users in the destination country are a very large potential IoT user market, and the development of national news data services has a good foundation. Traditional news and information dissemination channels (such as newspapers, television, and radio) can no longer satisfy people. The demand for information updates has changed with the popularization of the Internet and the emergence of news information management platforms.

As a global strategic emerging industry, the Internet of Things is highly valued by the country and society. The Internet of Things technology uses the advantages of a new

generation of information technology and communication technology to deeply integrate with traditional industries and promote the revolutionary transformation of traditional industries. The design meets the requirements of national industrial development. Information processing measures will promote the prosperity of the information service industry and achieve the intelligence of the information service industry [1]. With the declining investment cost and the increasing maturity of the two basic conditions of communication technology advancement, the advantage of the Internet of Things to replace people with things can be extended to those scenarios that are limited by time and space. In industries such as logistics and express delivery, transportation, commerce, and retail, the Internet of Things gradually has a certain scale of layout [2]. Even though news communication can keep in touch with the outside world anytime and anywhere, users usually use voice services, while news data services have not been widely used. The commonly used short message service (SMS) only allows to send about 160 bytes of simple information, such a narrow bandwidth is not conducive to Internet access [3]. In order to connect news equipment and the Internet to realize the Internet of Things user data service, people have proposed many solutions, one of which is the Internet of Things user application protocol (news information) [4]. At present, the national news communication is in the development stage of 4G, and then it will be the 5G era [5]. Graphics-based information can be provided in the 3G environment. For example, news information, mobile games, and other services are developing rapidly, mobile TV services are starting to take off, news industry applications for all walks of life in the society are beginning to rise, and the Internet of Things news value-added information services are becoming a fast stage of development; multimedia-based information can be provided in the 4G environment, including video phones, video mailboxes, IP phones, video conferences, video streaming media, and online games, among which video services will be the leading application of 4G networks [6]. The wide application of Internet of Things user network technology has laid a solid foundation for the Internet of Things news value-added information telecommunications business [7].

This article introduces the business types of IoT news value-added information services. In order to realize these services, the protocol transmission structure of the news information platform for IoT users is given on this basis. It further introduces the three components of the logical structure of the IoT user news information platform: access interface layer, function scheduling layer, and business interface layer, and analyzes the role of these three parts in the platform. Finally, the network topology diagram of the news information platform for Internet of Things users is given. As an IoT device operating company, it should vigorously promote IoT devices and use advanced software development technology to build an online recommendation system for IoT devices. Through this recommendation system, increase the promotion of the IoT product market and deliver IoT products and technologies. The news information topology layer under the Internet of Things is a set of mature technical architecture, which provides an excellent mechanism for the establishment of scalable, flexible, and maintainable enterprise applications, which can simplify and optimize the development and deployment of application systems. The platform integrates a variety of advanced IT technologies such as database, Internet, and telecommunications; discusses WAP technology and XML technology; and uses the MVC architecture model to implement the WAP website through the WML language. Among them, the Internet of Things news information is based on the multiple frameworks of the news information topology layer platform under the Internet of Things, including the technical framework of the database persistence layer, which is conducive to the development of application systems based on the news information topology layer specification under the Internet of Things; news information protocol transmission structure is incorporated with new ideas, new methods, and new technologies in the field of software engineering; to achieve this effect, designing and implementing a news information management platform have very important economic and social benefits; it is suitable for describing any type of system with object-oriented technology and is especially suitable for requirement specification description. Therefore, this topic uses the news information topology layer and news information protocol transmission structure technology under the Internet of Things to study the construction process of the Internet of Things equipment online promotion system.

2. Related Work

Domestic research on IoT news big data started late, and the research content is relatively small, and the coverage is relatively narrow. Especially in the field of IoT news big data news, domestic research focuses on overcoming technical difficulties, and there is even less research on the resulting information sharing and privacy issues. Aazam et al. [8] pointed out in their paper that in the era of big data news in the Internet of Things, the news and health industry will generate a large amount of data every year, and the combination of news and information technology will become closer and closer. Zhang and Wen [9] pointed out that the development of IoT news big data technology in the field of news will become deeper and more extensive. However, problems such as inconsistent data standards, difficulty in data sharing, patient privacy protection, and information security will also arise. The research of Wang and Cai [10] showed that a huge challenge brought by the big data era of the Internet of Things news is the issue of personal privacy protection. People will leave data footprints when they use the Internet. These data footprints are related to each other, leaving a person behind. When the data footprint is large enough, criminals can discover personal private information by integrating and analyzing these data. Talari et al. [11] said that big data of Internet of Things news provides strong data support for clinical research, biotechnology development, and personalized treatment of patients, but medical ethics and data privacy must be considered. They believe that in order to balance the benefits and risks brought by technological innovation, relevant departments must follow the basic principles of the use of IoT news big data, use more advanced

network security technologies, and establish and improve laws and regulations. Sisinni et al. [12] pointed out in their research that it is necessary to learn from the US and EU's protection models for the privacy information of patients with news Internet of Things news big data, combined with actual situation, and carry out hierarchical and classified protection on the basis of a clear hierarchical structure. Swan [13] pointed out that the national news information news Internet of Things news big data resources need to be further explored and utilized. Current research focuses on theory and experiment and needs to be transformed into practice and application. The theories and methods used in this article have been fully used in other fields. Wang et al. [14] said that based on signal transmission theory and communication privacy management theory, it is concluded that perceived privacy risk has a significant impact on consumers' willingness to provide personal information, and the structural equation model is used to analyze consumers' willingness to participate in online group buying. There are also researchers based on the theory of planned behavior and the theory of technology acceptance, using structural equations to analyze the influencing factors of personal Internet of Things news sharing willingness [15]. Some scholars have used the random news platform database to regress and analyze the factors affecting the information dissemination of the Internet of Things [16].

The news IoT application focuses on providing extensive interoperability, allowing IoT devices to be shared across domains, and reusing knowledge available on the web. In order to solve the problem of interoperability of shared IoT devices across domains, it is proposed to implement semantic interoperability between heterogeneous devices in news network technology and intelligent body technology [17]. It achieves a "Global Understanding Environment" (GUE), and you can achieve interoperability between different devices. The Internet of Things needs to be pushed to a more open, interoperable, and collaborative news Internet of Things. From this, it can be clearly seen that the Internet of Things will gradually form the trend of the news Internet of Things. The prototype news matcher of the news Internet of Things supports news Web technology and implements standard and nonstandard reasoning tasks and a knowledge base of moderate expression. The reason why the news IoT prototype mobile matcher is proposed is because of the architecture and performance of the news IoT [18]. It is impractical to use available inference engines to process news-based information in popular computing scenarios. Scholars put forward the concept of Wisdom Web of Things (W2T), which is an extension of the Wisdom Web in the era of the Internet of Things. The meaning of "wisdom" is that every object in the Internet of Things can know for itself, and other objects can provide the right services to the right objects at the right time and under the right conditions [19]. W2T emphasizes the data cycle, that is, from data to data, from data to information, from information to knowledge, from knowledge to wisdom, from wisdom to service, from service to human beings, and then back to things. The establishment of a smart Internet of Things environment with ubiquitous computing that connects things to things has given enlightenment to the proposal of the news Internet of Things. Some scholars have proposed the news Internet of Things search engine (WOTS2E), which is a SWoT search engine, based on web crawlers, that can discover associated data endpoints, and use them to enable WoT-enabled devices and services [20–23]. The demand for data generated by equipment and services in the real world has led to the emergence of the Internet of News Things. News Internet of Things is an emerging network that combines the news network and the Internet of Things (IoT/WoT). Its purpose is to associate news-rich and easily accessible information with objects in the real world. The news Internet of Things brings together the News Network and the Internet of Things and links the information marked in news with the physical devices, services, and data that support the Web to achieve seamless data integration. To better understand real-world information, a standardized, scalable, and flexible way is needed to enable real-time Web-connected embedded devices to be discovered globally [24–26].

3. System Architecture Based on the News Information Platform of the Internet of Things

3.1. Hierarchical Distribution of News Information under the Internet of Things. The Internet of Things system can be divided into three levels: sensing layer, network layer, and application layer. The sensing layer is used to solve data acquisition problems. It consists of various sensors (temperature and humidity sensors, RFID tags, light intensity sensors, etc.) and sensor gateway constitutes; its core technology mainly includes sensor technology, radio frequency identification technology, Internet of Things user network networking technology (Zigbee network), and fieldbus control (Fieldbus Control System) technology; network layer, also known as the transmission layer, solves the problem of remote transmission of data obtained by the perception layer and integrates sensor networks, news communication technology, and Internet of Things technology [27–29]; the application layer is mainly responsible for information processing and friendly man-machine interface in the Internet of Things system and processing the bottom layer. The incoming data meets their own business needs, and the data resource database is updated in real time to provide unified shared resource support for various businesses and ultimately meet the applications of various industries and fields of the Internet of Things. The wireless news information platform is connected with the short message dispatch center through the standard SMPP protocol or the standard CMPP protocol (Monternet protocol) and can be connected with the short message dispatch center supporting the SMPP/CMPP protocol. According to the concept and characteristics of the Internet of Things, the Internet of Things can be divided into a sensor layer, a network layer, and an application layer according to a bottom-up process.

An information fitting regression classification model is usually applied to binary classification tasks; that is, for each independent variable X in the model, the value of the

dependent variable y corresponding to it has two cases, namely, 0 and 1. The information fitting regression model can solve the binary classification task well and has high accuracy and precision. In order to characterize the timeliness of information, this paper introduces the concept of mathematical expectation of timeliness of information (abbreviated as expectation of timeliness), which represents the lifetime of information, that is, the time period during which information truly reflects the state of the source. Given a set of possible messages $A = (1, 2, \cdots, n)$ and its probability distribution is P, let f denote the probability that message i keeps the information valid in the time interval of 0-t function, referred to as the aging probability function. Then, the time-sensitive amount C of message i can be expressed as

$$C = (x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n). \tag{1}$$

Formula (1) is the expression of expectation of ageing when the ageing probability function is continuous. The time-effect probability function is the time-effect expectation expression of the discrete case, which can be expressed as follows: it is given the set of possible messages A-a, and its probability distribution is P. The time point sequence T of the first change of message i-a is given in minutes. Let the timeliness probability distribution of the first change of message i-a be corresponding to T-i:

$$p(y \mid x_1, \cdots, x_n) = \frac{1}{1 + e^{-a_1 \times x_n}}. \tag{2}$$

Therefore, this paper adopts the information fitting regression model as a model to analyze the public's willingness to share news data. Information fitting regression belongs to supervised learning, so it is necessary to use labeled data as the training set to train the function. The training set is a set of training data composed of the features and labels of the data:

$$p(i) = \frac{e^{-(a+b_1 \times x_1 + \cdots + b_i \times x_i)}}{e^{a+bi+xi}}. \tag{3}$$

Among them, x is an m-dimensional vector, Y takes a value in {0,l}, and the classification task is simplified to how to find a decision function Y, so that it can have good enough performance on an unknown data set and make a more accurate prediction. In this paper, the information fitting regression model is a regression model of binary variables. The willingness to use dependent variables is a binary variable, and the influencing factors of independent variables are classified data. The dependent variable sharing willingness Y takes the values 0 and 1, and the independent variable influencing factors are set to $x_1, x_2, \cdots, x_n$, where the value of Y indicates a strong willingness to share and a value of 0 indicates a weak willingness to share. Then, the information fitting function is

$$1 - p(n) = \frac{1}{1 + e^{a+b_1 \times x_1 + \cdots + b_n \times x_n}}. \tag{4}$$

The conditional probability that the f sample has a strong willingness to share is denoted as P, and then it can be transformed from equation (4) into

$$\frac{p(i)}{1 - p(i)} = e^{a+b_i \times x_i}. \tag{5}$$

From equation (5), we can see that p-i is a nonlinear function composed of independent variables, which can be transformed into a linear function through a certain transformation. Define events with weak willingness to share; that is, when y is set to 0, there are

$$\ln\left(p(i)\right) = a + b_1 \times x_1 + \cdots + b_i \times x_i. \tag{6}$$

Calculate the ratio of the probability of a strong willingness to share and the probability of a weak willingness to share:

$$w(x, y) = \{F(x, y), G(x, y)\}. \tag{7}$$

The ratio of the probability of an event occurring to the probability of not occurring is called the odds. When odds $>$ l, it indicates that the situation with strong willingness to use is more likely to occur, and when O $<$ odds $<$ l, it indicates that the situation with weak willingness to use is more likely to occur. Take the natural logarithm of odds to convert it into a linear function:

$$\begin{aligned} F(i) &= \{f_1, f_2, \cdots, f_n\}, \\ G(i) &= \{g_1, g_2, \cdots, g_n\}. \end{aligned} \tag{8}$$

Therefore, equation (8) is also called the functional form or logit form of the information fitting model. For modeling and analysis of dichotomous dependent variables, information fitting regression models have become a common modeling method. Through a large amount of literature reading and analysis, it is found that the information fitting regression model can perform well in the modeling and analysis of individual characteristic factors and study the public's willingness to use the news IoT news big data platform.

3.2. Linear Topology Module Based on IoT Platform. The linear framework based on the Internet of Things platform is a framework that is different from the previous application system research and development. It includes a series of components that can simplify and optimize the research and development, installation, and operation of enterprise applications, thereby improving portability, security, and reuse value. The news information topology layer under the Internet of Things adopts a multilayer distributed application model. At the client layer, the components on the client computer that access the enterprise application, the interface logic between the user and the enterprise application, use the http protocol to access and browse the application server. The presentation layer, a component running on the application server of the news information topology layer under the Internet of Things, interacts with the business logic layer to

output the business data required by the user in a suitable method. The business logic layer, like the presentation layer, is also a component running on the application server of the news information topology layer under the Internet of Things. As an important part of the wireless news information platform, the news information platform based on WAP technology focuses on the design and implementation of the WAP news website, the data interface module with mobile operators, and the supporting data release and data collection function modules.

Under the news information topology layer specification under the Internet of Things, programmers can develop enterprise-level application systems according to the technical requirements of the news information topology layer under the Internet of Things and a variety of news information topology layer providers under the Internet of Things jointly abide by the specifications of the news information topology layer before and after the version under the Internet of Things to ensure the unity between the news information topology layer platform and enterprise applications under the various Internet of Things. In other words, enterprise applications based on the news information topology layer architecture under the Internet of Things can basically be installed and run on different application servers, without or only a small amount of code editing, which can greatly improve the portability of enterprise applications. The topology of the IoT news platform is shown in Figure 1. For programmers, they only need to focus on the business logic and architecture design of various enterprise applications.

Internet of Things news information is an open source project provided by the Apache software organization. It provides a model-view-controller (MVC) framework for Java Web enterprise applications, which is especially suitable for the development of large-scale and easy-to-expandable Web enterprise applications. The Internet of Things news information provides a common framework for Web enterprise applications, allowing programmers to focus on how to deal with actual business problems. The submission and delivery of short message information are realized through the bidirectional communication interface with the short message center provided by the standard protocol. With the help of the short message bearing function provided by the short message center system, the transmission of information between the information processing server and the mobile platform is realized, and various information services are provided for users. In addition, the Internet of Things news information framework provides many places for expansion and customization. Enterprise applications can easily expand the framework to better adapt to the actual needs of various users. In order to reduce the coupling of code and improve the efficiency of enterprise application research and development, here is a brief description of a news information topology layer enterprise application research and development strategy under the Internet of Things that integrates the Internet of Things news information framework and the Hibernate framework.

The news information platform based on the Internet of Things is a reusable MVC 2 design composed of a group of cooperating classes, Servlet, and JSP tags. This definition describes that IoT news information is a framework, not a library, but the IoT news information framework also includes a rich tag library and utility classes independent of the framework. The Internet of Things news information is a concrete realization of the JSPModel2 design specification. The logic algorithm of IoT news information is shown in Figure 2. This part of the properties is related to what is a one-to-one correspondence between the user input items in the page form. In the Internet of Things news information framework, the ActionForm class is used to submit the form information entered by the user to the controller; the JSP page is the main view component in a typical MVC application, including the functions of information display and the controller's solution result display. In addition to the above, the IoT news information framework also provides a powerful IoT news information tag library to assist users in processing view logic and uses ActionForm components to transfer information to the control layer (IoT news information core controller).

3.3. Structure Optimization Based on News Information Protocol Transmission. The basic application steps of the protocol transmission structure are based on the news information platform. It worked by using the Unified Modeling Language for application visualization modeling, that is, to describe and document the project that uses object-oriented methods to develop enterprise applications. The news information protocol transmission structure belongs to the third generation of modeling and specification language, which is a method for describing, visualizing, and establishing object-oriented enterprise applications in the enterprise application research and development stage. As a type of modeling language, the news information protocol transmission structure enables software project developers to concentrate on creating the software model and technical structure instead of considering which language and algorithm to choose to complete. After the software model is created, the model can be automatically converted into the selected programming language code by the integrated development tool of the news information protocol transmission structure.

As a subset of the main platform, the news information logical substructure is mainly used to mark and explain the Internet information and user interface sent and received by the news information and news terminal, enabling developers to define the user interface of the news information application in a device-independent manner. Moreover, in most cases, news information uses scripting languages such as WMLScript to directly process warnings and other messages on the news terminal to avoid data interaction between the news terminal and the remote server, thereby reducing the consumption of bandwidth resources. Figure 3 shows the hierarchical structure of the news information platform under the Internet of Things. The news information gateway is the news information agency. The user layer means that the system uses users and system administrators to manipulate the underlying data through a graphical interface. This form not only allows users to operate and edit but also ensures that users do not directly manipulate data packaging data through such an interface isolation method. The news

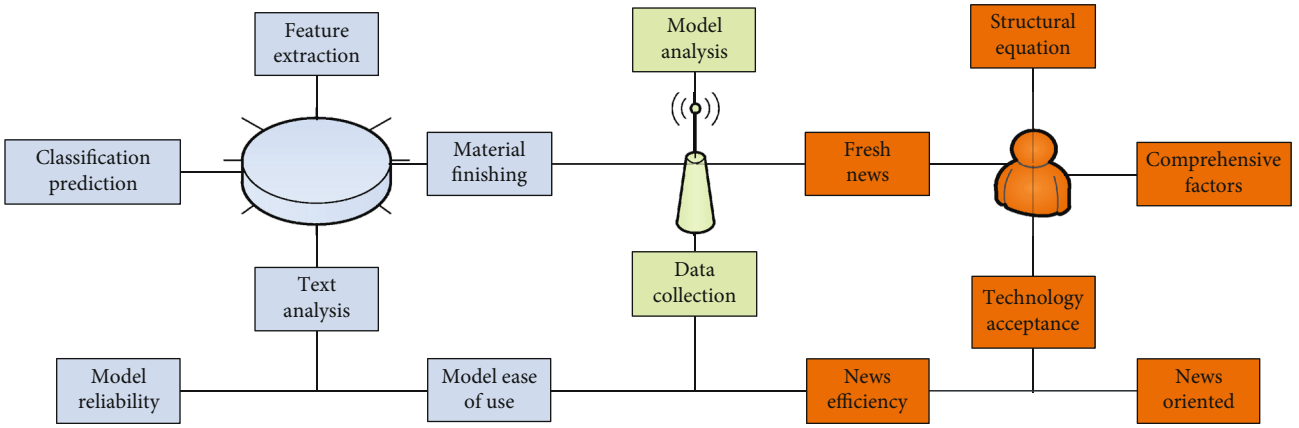


Figure 1: Topology structure of IoT news platform.

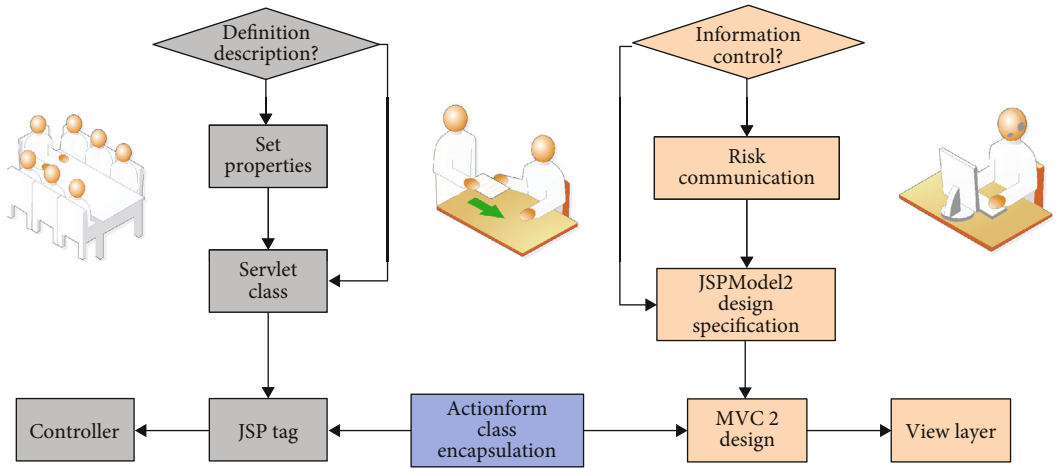


Figure 2: The logic algorithm of IoT news information.

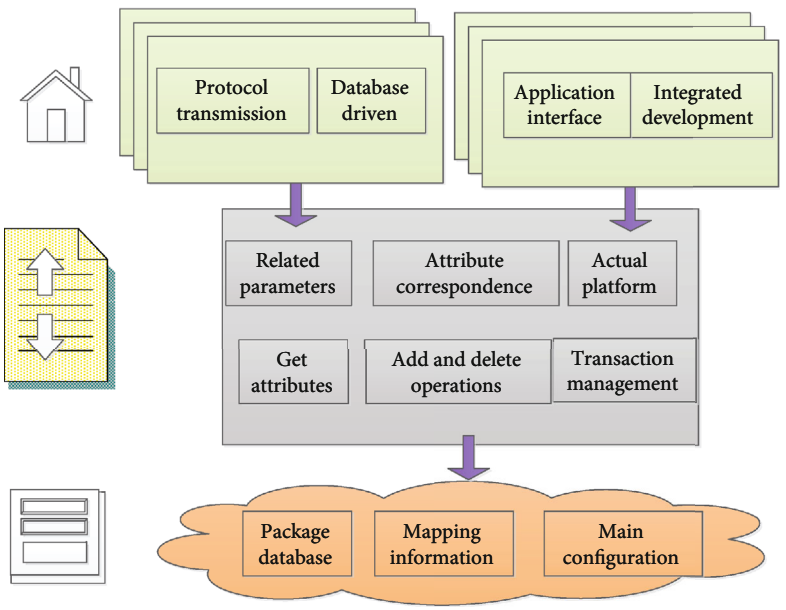


Figure 3: Hierarchical structure of news information platform under the Internet of Things.

information gateway is used as the gateway between the Internet and the news network to realize the conversion between the Internet protocol and the news information protocol. The news information agency also provides an information content code, which can compress and encode news information data, thereby reducing network data traffic and maximizing the use of the slow data transmission rate of current Internet of Things users. In addition, the news information agent also uses error correction technology to ensure that news information communications will not be seriously affected by changes in the quality of the Internet of Things users' telecommunication lines during network browsing and data transmission.

4. Design and Implementation of News Information Platform Based on the Internet of Things

4.1. System Function Module Simulation. System realization is the process of specifically writing information system function modules according to the predesign, which mainly includes the coding of the function modules, the installation and debugging of the Internet of Things equipment online recommendation system, and the testing of the system. The Internet of Things device online recommendation system is an information system developed in a specific operating environment. The Internet of Things device online promotion system is built on the Struts-Hibernate integration framework of the news information topology layer under the Internet of Things, the development tool is Eclipse, and the plug-in myeclipse is installed. It can be installed, debugged, configured, and run on Windows, Unix, Linux, and other operating systems, using this part of the tools and operating environment to develop and deploy the Internet of Things device online promotion system, which can be used in local area networks, metropolitan area networks, wide area networks, and other Internet of things In the network environment where the equipment online promotion agency is located. The mapping between the sale information tables is able to draw and describe the correspondence between the entity objects of the Internet of Things device online recommendation system and the data table. Table 1 shows the data table of the IoT news entity object.

Table 1: Data table of IoT news entity object.

Sample attributes	Scale range	Distribution number	Rate
Sample 1	A1	99	0.31
Sample 2	A2	122	0.21
Sample 3	A3	82	0.08
Sample 4	A4	45	0.4

From the perspective of the overall test effect of the model, the model setting is relatively reasonable and the overall fitting effect is good. The prediction accuracy rate of the model reached 79.52%. From further analysis of the regression results, we can see that the status of news information dissemination, the time of exposure to the Internet of Things, the experience of using the Internet to transmit news information, and the habit of recording body data, the signs of the coefficients of these variables are consistent with the actual situation, and they are all consistent with the actual situation. Then, they passed the empirical analysis test based on individual characteristic factors at a significance level of 10%. The regression coefficient of the Internet of Things is 0.5806, and the P value is 0.073. At a significance level of 0.1, the Internet of Things has a positive effect on the willingness to share news information. That is, IoT users are more willing to share their news data than non-IoT users. The reason may be that IoT users have a higher risk appetite than non-IoT users and are more willing to try new things. The regression coefficients of bachelor's degree and master's degree and above in the educational level are 0.7450 and 1.0486, respectively, and the corresponding P values are 0.032 and 0, respectively. They are both less than the significance level of 0.05. The degree of education has a positive effect on the willingness to share news information; that is, respondents with a bachelor degree or above are more willing to share news information. In the status of body news information, the regression coefficients for scoring one's body into 3 and 5 points are 3.3122 and 3.6256, respectively, and the corresponding P values are 0.004 and 0, respectively. Figure 4 shows the sharing of news information based on regression coefficients. At a significance level of 0.01, it has a positive effect on the willingness to share news information; that is, respondents with better physical conditions are more willing to share news information than those with poorer physical conditions.

In the item of time of exposure to the Internet of Things, the regression coefficient of exposure to the Internet of Things for more than ten years is 1.3135, and the P value is 0.016. At a significance level of 0.05, it has a positive effect on the willingness to share news information. The news section correspondence table is used to record each news and its corresponding section node. The same news can correspond to multiple section nodes at the same level, but in general, a news item corresponds to only one section node. The main information includes news number and section number. That is, respondents who have been exposed to the Internet for a longer time are more willing to share news information on the Internet of Things. In the experience of using the network to transmit news information, the regression coefficient is 0.6282, and the P value is 0. 053; at a significance level of 0.1, it has a positive effect on the willingness to share news information; that is, respondents who have experience of delivering news information online for medical treatment are more willing to share news information on the Internet of Things. In the item of whether there is a habit of using news equipment to record body information, the regression coefficient is 0.8576, and the P value is 0. 011. Figure 5 shows the distribution of the significance level test for different sample groups. At the significance level of 0.5, it has a positive effect on the willingness to share news information; that is, respondents who have the habit of recording physical data are more willing to share news information.

The path coefficient from perceived ease of use to perceived usefulness is 0.80, indicating that perceived ease of

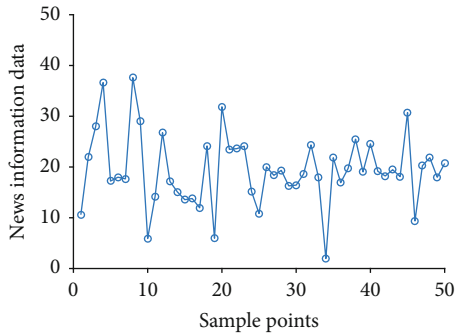


Figure 4: News information sharing based on regression coefficients.

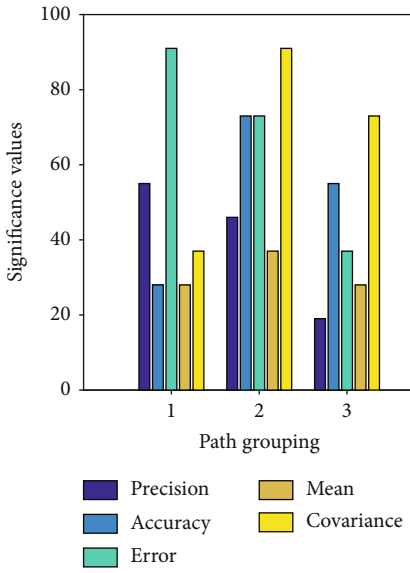


Figure 5: The distribution of the significance level of different sample groups.

use has a significant positive effect on the public's perceived usefulness of news information sharing, which is consistent with the technology acceptance model proposed by Davis. Respondents believe that news information sharing is a useful technology; the premise is that news information sharing is easy to realize and use. The path coefficient from perceived ease of use to willingness to share is 0.05, indicating that the perceived ease of use has a positive effect on the public's willingness to share news information using the news IoT news big data platform, but it is not significant. The path coefficient from perceived usefulness to sharing willingness is 0. 04. It shows that the perceived usefulness has a positive effect on the willingness to use the news big data platform to share news information, but it is not significant. The path coefficient from perceived risk of willingness to share is 0.14, indicating that the perceived risk has a negative impact on the public's willingness to share news information. The working method of the news information management platform developed in this paper adopts the dynamic web page management method, which is developed after solving the various shortcomings of the static state. In general, the results of structural equation analysis are consistent with the assumptions.

4.2. Entity News Platform Database Design. The physical news platform database is a single classifier, which has a bottleneck in performance improvement. News information integrated learning, that is, the random news platform database algorithm, is to gather a single entity news platform database and determine the category of the sample to be classified by collecting statistics on the classification results of each entity news platform database. It is equivalent to that each classifier will have a classification result. There are two important randomizations in the Random-News platform database, as shown in Figure 6, which is the training efficiency of the news information dataset algorithm. Since the process of randomly extracting training sample sets and feature subsets when constructing each entity news platform database is independent, the overall is the same, and each entity news platform database is an independent and identically distributed sequence of random variables. Therefore, the training of the random news platform database can be realized through parallel processing, which ensures the efficiency and scalability of the random news platform database algorithm.

First, use the train test split in the library to randomly divide all data into training subsets and test subsets. The train test split function is a commonly used function in cross-validation. The function is to randomly select train data and test data from the sample proportionally. The test size (the proportion of samples) selected in this paper is 20%, and the number of random seeds is 2. In the access type field, a value of "1" indicates that the user's access type is viewing and browsing; a value of "2" indicates that the user's access type is sharing. There are other new access types that can be added at any time in the future. If n estimators are too small, it is easy to underfit, and if n estimators are too large, it is easy to overfit. Generally, choose a moderate value for the random news platform database, increasing the "number of submodels" (n estimators) which can significantly reduce the variance of the overall model. The accuracy of the model will increase as the "number of submodels" increases. Since the decrease is the second term of the overall model variance formula, there is an upper limit for the improvement of accuracy. The random state is the seed used by the random number generator.

4.3. System Test and Application Effect. First, average the scores of the four questions in the sharing willingness part of the questionnaire, and classify those with an average score greater than or equal to 4 into one category, which indicates strong willingness to share, and classify those with an average score less than or equal to 2 into one category, indicating the willingness is weak. Putting the average score between 2 and 4 points into the first category indicates that the willingness to share is average. After processing, the respondents were divided into three categories, with the number of persons in each category as shown in Figure 7, which is the statistical data of sample correlation. It can be seen from the figure that

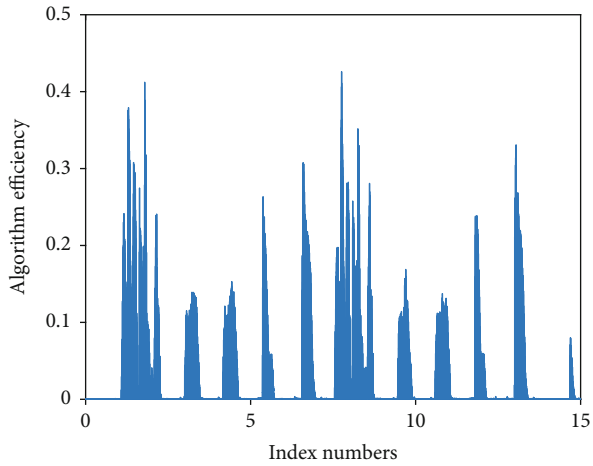


Figure 6: Algorithm training efficiency of news information data set.

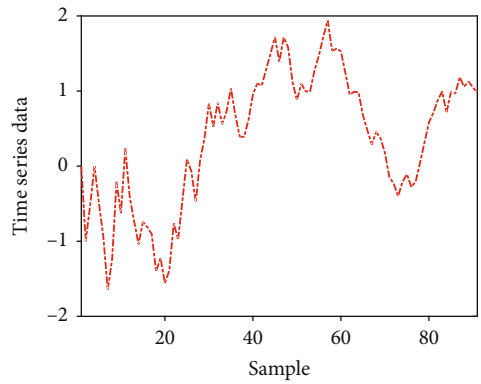


Figure 7: Statistical data of sample correlation.

there are 53 respondents with low willingness to share, 122 respondents with general willingness to share, and 154 respondents with strong willingness to share. The number of the three types of respondents is quite different, and the distribution is not very uniform, which poses a certain challenge to the subsequent classification. It can be seen from the figure that the distribution of the remaining six variables is not uniform, except for the number of Internet of Things user surveyed, non-IoT user surveyed, and the survey's habits of recording body data. By drawing a heat map of the correlation between the eight variables, it can be seen that the correlation between these eight variables is mostly in the normal range. The following result indicates that the correlation is not obvious, which also poses a challenge to the subsequent classification.

CMIN/DF is called relative chi-square or normalized chi-square. It is the minimum sample difference divided by the degree of freedom. Generally, this value is allowed to reach 5 as an appropriate fit. The relative chi-square value in this model is 3.786, which is less than 5, which meets the requirements, so this model is acceptable. RMSEA is the root mean square of the approximate error, and its value is less than or equal to 0.05 which means that the model fits very well, and its value is less than 0.08, which means that the model fits

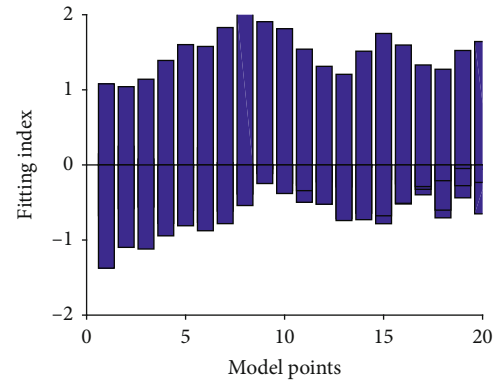


Figure 8: Histogram of IoT model fitting index.

properly. The root mean square of the approximate error in this model is 0.092, which is greater than 0.08, which does not meet the requirements, so the model should be rejected. GFI is the goodness of fit index, and the range is between 0 and 1. To accept the model, GFI should be greater than or equal to 0.90. The goodness of fit index of this model is 0.911, which is greater than 0.90, which meets the requirements, so this model is acceptable. IFI is a value-added fitting index. The range of its value is not guaranteed to be between 0 and 1. Figure 8 shows the histogram of the fit index of the IoT model. IFI close to 1 indicates a good fit, and it is greater than 0.90 which indicates an acceptable fit. The value-added fitting index of this model is 0.933, which is greater than 0.9. It meets the requirements, so this model is acceptable. The crawler module is a core module for the system to search for news, so it is important to ensure that this module can run correctly. The picture shows the functional test description of the crawler module and the specific display of the use cases. CFI is a comparative fitting index. A value close to 1 indicates a good fit, and a value greater than 0.90 indicates an acceptable fit.

The comparative fitting index of this model is 0.936, which is greater than 0.90. It meets the requirements, so this model is acceptable. TLI is the Tucker-Lewis coefficient, and the range of its value is not guaranteed to be between 0 and 1. TLI close to 1 indicates a good fit, and it is greater than 0.90 which means that the fit is acceptable. The TLI of this model is 0.919, which is greater than 0.90. The empirical analysis is based on individual subjective factors in the requirements, so this model is acceptable. AGFI is an adjusted goodness-of-fit index, and the range of variation is between 0 and 1, which is greater than 0.50 which means that the fit is acceptable. The adjusted goodness of fit index of this model is 0.655, it is greater than 0.5 which means it meets the requirements, so this model is acceptable. It can be seen from the above indicators that although the value of RMSEA is not in the ideal range and is slightly larger than the maximum value, the rest of the fitting indices are in the ideal range and meet the requirements.

The accuracy of classification is calculated by the predicted classification of the test data set and the actual classification of 300,000 square meters of data. It is based on the empirical analysis of individual characteristic factors. For example, when a certain data belongs to the first category,

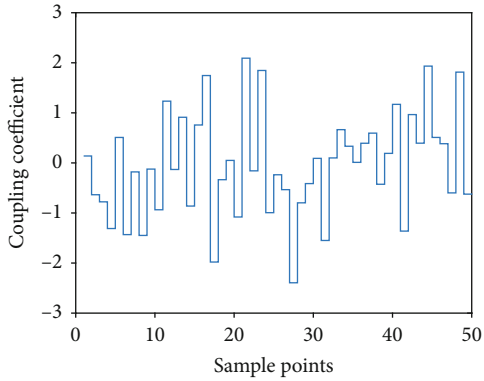


Figure 9: Distribution of news information data coupling factor.

the program divides it into the second or third category which is the classification error, and the correct rate of the predicted classification is obtained by calculating the proportion of the correct number of classifications to the total number. Through the calculation of the program, on this data set, the accuracy rate of the random news platform database model classification is 0.455, which exceeds 0.333, indicating that the classification has a certain effect, but it is still not very prominent. Eventually, 3 of the 12 samples of the first category were classified into the first category, 5 were incorrectly classified into the second category, and 4 were incorrectly classified into the third category; out of the 20 samples of the second category, 16 of them were finally classified in the second category, and 4 samples were mistakenly classified into the third category; finally, 11 of the 34 samples in the third category were classified into the third category, 1 was incorrectly classified into the first category, and 22 were mistakenly classified into the second category. Figure 9 shows the distribution of news information data coupling factors. Therefore, the second type is easier to distinguish, with higher accuracy, and the first and third types have lower accuracy and higher error rate. It is more intelligent, and the timeliness has also been greatly improved, which greatly improves the work efficient, and can provide a more intelligent site search function, a better user experience.

5. Conclusion

The main design goal of the Internet of Things user news information platform is to achieve a platform that makes full use of the information resources of traditional media and realizes multiple Internet of Things news value-added information services including SMS, MMS, and news information. This paper will focus on the design ideas of a news information platform for Internet of Things users based on news information technology and take news information as an example to study how to implement a news information platform based on news information technology under the GPRS network built by operators. The news information platform is based on news information technology, as an important part of the news information platform for Internet of Things users, it focuses on the design of news information news websites and data interface modules with news operators, and it supports data release and data collection function modules. With the implemented solutions, the platform integrates a variety of advanced IT technologies such as database, Internet of Things, and telecommunications, discusses news information technology, and uses architecture pattern language to implement a news information website. The significance level supports the null hypothesis, and the two variables of age and occupation cannot pass the test. At the same time, the prediction accuracy of the model is 79.52%, indicating that the model has a better fitting effect. At the same time, the random news platform database classification method is used to analyze the relationship between individual characteristics and the willingness to share. In summary, the development of this article focuses on user experience improvement and adopts a dynamic web management method to achieve uninterrupted service to users when news information is released. The platform has good functional scalability and builds a system architecture with strong subsequent scalability. The design and implementation of this article have been directly applied in the project, and also provide a research foundation and certain practical experience for further expanding the value-added services of the Internet of Things news and significantly expand the functions of the value-added information service platform system of the Internet of Things news.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Jinan University.

References

[1] C. Tao and L. Xiang, "Municipal solid waste recycle management information platform based on internet of things technology," *Multimedia Information Networking and Security*, vol. 3, pp. 729–732, 2019.

[2] H. C. Y. Chan, "Internet of things business models," *Journal of Service Science and Management*, vol. 8, no. 4, p. 552, 2019.

[3] J. Ye, B. Chen, Q. Liu, and F. Yu, "A precision agriculture management system based on Internet of Things and WebGIS," *IEEE Geoinformatics*, vol. 7, pp. 1–5, 2020.

[4] J. Lee, S. H. Kim, S. B. Lee, H. J. Choi, and J. J. Jung, "A study on the necessity and construction plan of the internet of things platform for smart agriculture," *Journal of Korea multimedia society*, vol. 17, no. 11, pp. 1313–1324, 2019.

[5] A. Bahga and V. K. Madisetti, "Blockchain platform for industrial internet of things," *Journal of Software Engineering and Applications*, vol. 9, no. 10, pp. 533–546, 2018.

[6] J. Mineraud, O. Mazhelis, and X. T. Tarkoma, "A gap analysis of Internet-ofings platforms," *Computer Communications*, vol. 89, pp. 5–16, 2018.

[7] Z. D. R. Gnimpieba, A. Nait-Sidi-Moh, D. Durand, and J. Fortin, "Using Internet of Things technologies for a collaborative supply chain: application to tracking of pallets and containers," *Procedia Computer Science*, vol. 56, pp. 550–557, 2020.

[8] M. Aazam, I. Khan, A. A. Alsaffar, and E. N. Huh, "Cloud of things: integrating Internet of Things and cloud computing and the issues involved," *Applied Sciences & Technology*, vol. 5, pp. 414–419, 2018.

[9] Y. Zhang and J. Wen, "The IoT electric business model: using blockchain technology for the internet of things," *Peer-to-Peer Networking and Applications*, vol. 10, no. 4, pp. 983–994, 2017.

[10] K. Wang and K. Cai, "Design of field information monitoring platform based on the internet of things," *Internet of Things*, vol. 312, pp. 597–602, 2019.

[11] S. Talari, M. Shafie-khah, P. Siano, V. Loia, A. Tommasetti, and J. Catalão, "A review of smart cities based on the internet of things concept," *Energies*, vol. 10, no. 4, p. 421, 2017.

[12] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial internet of things: challenges, opportunities, and directions," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4724–4734, 2018.

[13] M. Swan, "Sensor mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0," *Journal of Sensor and Actuator Networks*, vol. 1, no. 3, pp. 217–253, 2019.

[14] P. Wang, S. Chaudhry, and L. Li, "A visualization platform for internet of things in manufacturing applications," *Internet Research*, vol. 3, pp. 18–23, 2018.

[15] I. Yaqoob, E. Ahmed, I. A. T. Hashem et al., "Internet of things architecture: recent advances, taxonomy, requirements, and open challenges," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 10–16, 2017.

[16] D. Bandyopadhyay and J. Sen, "Internet of things: applications and challenges in technology and standardization," *Wireless Personal Communications*, vol. 58, no. 1, pp. 49–69, 2019.

[17] A. J. C. Trappey, C. V. Trappey, U. Hareesh Govindarajan, A. C. Chuang, and J. J. Sun, "A review of essential standards and patent landscapes for the Internet of Things: A key enabler for industry 4.0," *Advanced Engineering Informatics*, vol. 33, pp. 208–229, 2017.

[18] J. A. Stankovic, "Research directions for the Internet of Things," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 3–9, 2019.

[19] S. M. R. Islam, D. Kwak, M. D. H. Kabir, M. Hossain, and K. S. Kwak, "The internet of things for health care: a comprehensive survey," *IEEE access*, vol. 3, pp. 678–708, 2018.

[20] S. Li, T. Tryfonas, L. Li et al., "The Internet of Things: a security point of view," *Internet Research*, vol. 26, no. 2, pp. 337–359, 2016.

[21] M. Cui, J. Lu, A. Zhao, and C. Wei, "Application research of information platform of coal enterprise based on internet of things," *Coal Technology*, vol. 1, pp. 32–39, 2021.

[22] Y. Lu, Y. Qi, S. Qi et al., "Secure deduplication-based storage systems with resistance to side-channel attacks via fog computing," *IEEE Sensors Journal*, vol. 1, no. 1, p. 1, 2021.

[23] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the Internet of Things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

[24] W. Wei, Q. Ke, J. Nowak, M. Korytkowski, R. Scherer, and M. Woźniak, "Accurate and fast URL phishing detector: a convolutional neural network approach," *Computer Networks*, vol. 178, article 107275, 2020.

[25] A. Zielonka, A. Sikora, M. Wozniak, W. Wei, Q. Ke, and Z. Bai, "Intelligent Internet of Things system for smart home optimal convection," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4308–4317, 2021.

[26] C. S. Yang, "Adoption of software testing in Internet of Things: a systematic literature mapping," *Computing Reviews*, vol. 61, no. 2, pp. 72-73, 2020.

[27] A. al-Qerem, M. Alauthman, A. Almomani, and B. B. Gupta, "IoT transaction processing through cooperative concurrency control on fog–cloud computing environment," *Soft Computing*, vol. 24, no. 8, pp. 5695–5711, 2020.

[28] M. Shafiq, Z. Tian, Y. Sun, X. du, and M. Guizani, "Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city," *Future Generation Computer Systems*, vol. 107, pp. 433–442, 2020.

[29] D. Li, L. Deng, B. Bhooshan Gupta, H. Wang, and C. Choi, "A novel CNN based security guaranteed image watermarking generation scenario for smart city applications," *Information Sciences*, vol. 479, pp. 432–447, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 8492759, 13 pages
https://doi.org/10.1155/2021/8492759

Research Article

The Value Cocreation Influence Mechanism of Network Freight Transport Platform in IoT-Based Environments: Under the Service-Dominant Logic

Pengxia Bai,[1,2] **Qunqi Wu**,[2,3] **Qian Li**,[3] **Lei Zhang**,[4] **Yahong Jiang**,[1,2] **and Bo Chen**[1,2]

[1]*School of Transportation Engineering, Chang'an University, Xi'an 710064, China*
[2]*Integrated Freight Transportation Economics and Management Research Center, Chang'an University, Xi'an 710064, China*
[3]*School of Economics and Management, Chang'an University, Xi'an 710064, China*
[4]*Massey Business School, Massey University, Palmerston North 4442, New Zealand*

Correspondence should be addressed to Qian Li; q.li@chd.edu.cn

Received 12 April 2021; Revised 3 May 2021; Accepted 6 May 2021; Published 24 May 2021

Academic Editor: Wei Wang

The Internet of Things (IoT) has brought many benefits to the development of industries, from manufacturing to services. In the context of IoT, the network freight transport platform is equivalent to an integrator, integrating the scattered resources such as carriers and shippers in the supply chain, providing a solution, improving the use efficiency of vehicles, effectively reducing logistics cost, and creating value for the freight transport platform and bilateral customers. The purpose of this research is to explain the influence mechanism of transport demand subjects' participation in value cocreation (VCC) under the service-dominant logic (SDL). We proposed a conceptual framework and 10 hypotheses and used SEM to measure the direct and mediating relationship among SDL, transport demand subject (TDS) participation, relational embeddedness (RE), platform VCC, and TDS value in the online business environment. The findings show that (1) SDL, TDS participation, and RE do not directly affect the TDS value, but indirectly through platform VCC; (2) SDL affects platform VCC not only directly but also indirectly through TDS participation and RE; (3) TDS participation affects platform VCC not only directly but also indirectly through RE; (4) RE indirectly affects TDS value by affecting platform VCC; and (5) platform VCC plays a mediating role in SDL, TDS participation, RE, and TDS value. Hence, transport enterprises should pay attention to the innovation concept, guide TDS participation, jointly improve platform VCC, give play to the network effect of the freight transport platform, and achieve TDS value through the deepening of TDS relationship embedding.

1. Introduction

The rapid growth of the Internet of Things (IoT) causes a sharp growth of data [1]. Enormous amounts of networking sensors are continuously collecting and transmitting data to be stored and processed in the cloud [2]. Such data can be environmental data, geographical data, astronomical data, logistic data, and so on. Mobile devices, transportation facilities, public facilities, and home appliances are the primary data acquisition equipment in IoT. It is a trend to use information technology and the Internet to transform the traditional transport industry. Network freight transport is developed on the basis of the vehicle and cargo matching platform developed by using the Internet technology [3]. Through the Internet platform and digital technology reference, with the intervention of big data, the whole network freight transport is interconnected and interoperated and the industry is upgraded with the help of the smart logistics model. In addition, the Ministry of Transport proposed to use the "Internet" to develop efficient logistics. Therefore, through the development of network freight, freight operators can effectively solve the tax burden, freight settlement, and insurance concerns in the field of distribution and tally. Network freight transport can not only help enterprises to achieve efficient allocation of transport capacity resources but also help logistics enterprises to truly achieve cost reduction and efficiency increase [4].

Although the emergence of a network freight transport platform has solved many problems for China's transport industry, there are still many problems. For example, high investment cost, poor connection of logistics links, and low efficiency lead to low performance of the whole logistics service supply chain. Therefore, this paper explores the way to improve the performance of the logistics service supply chain from the perspective of the customers. It must be emphasized that the customers in this paper refer to manufacturing enterprises. Under the demand-oriented logic, transport services become more and more important. In order to increase customer value, customer engagement is very important. Driven by the Internet, the emergence of a freight transport platform has brought great benefits to the output of the customer value. Therefore, how to increase the value output of customers is the main problem of this paper. From the perspective of customer demand, this paper explains the influence mechanism among service-dominant logic, customer participation, relational embeddedness, value cocreation, and customer value.

The rest of the article is organized as follows. Section 2 is about the theoretical basis and research hypothesis. Section 3 is about the research method. The research adopted a questionnaire survey approach, including questionnaire design, data collection, reliability, and validity test. In Section 4, results analysis, hypothesis verification, and model interpretation are presented. The last section contains the research conclusions and management implications.

2. Theoretical Basis and Research Hypothesis

2.1. Theoretical Basis

2.1.1. Service-Dominant Logic (SDL). Service-dominant logic (SDL) replaces the traditional product-dominant logic. Based on SDL, reexamine the relationship between services and products and then think about the core issues in economic exchange and value creation [5]. SDL is a service-centered alternative to the traditional goods-centered paradigm. This service-centered view suggests that market exchange is a process in which the parties utilize their expertise for the benefit of each other, i.e., mutual provision of services [6]. Since Vargo and Lusch proposed it in 2004, SDL has been revised three times in 2006, 2008, and 2016 and has developed from the original 8 basic propositions to 11.The main content includes the following: the application of operational resources (knowledge and skills) is the foundation of all exchanges. Thus, all economies are service economies. SDL establishes the priority of operational resources over objective resources. Operational resources refer to the resources that generate revenue through the operation of other resources, such as knowledge, skills, and capabilities. Objectivity resources refer to the resources that must be used to generate benefits, such as natural resources, commodities, and money. That is, operational resources such as knowledge and skills are potential sources of value. In addition, SDL believes that the resources for creating value are not limited to the company; customers, suppliers, and other stakeholders also constitute operational resources and contribute to the creation of value [7]. Value is always created with the customer and others, so the business cannot create and deliver value alone. They can only make value propositions and provide services as inputs to realize the value. Ultimately, value is determined contextually by the service beneficiary (i.e., the customer). Vargo and Lusch further improved the theory of SDL and believed that value cocreation is coordinated through institutions and institutional arrangements created by participants.

2.1.2. Customer Participation (CP). Customer participation (CP) refers to the input provided by the customer, including effort, time, knowledge, or other resources [8–11]. However, some scholars believe that CP is "the act of customers investing various resources to create value with suppliers " [4, 12]. In essence, the basic exchange unit is service rather than product. In this way, suppliers can cocreate value by exchanging services with customers [13]. They actively create value with suppliers to better meet their individual needs and improve their satisfaction [14]. With the change of customer demand and the progress of technology, the role of customers and enterprises in value creation has changed [15]. On the one hand, compared with the past, customers are more willing and able to participate in the development of the company's products and services [13, 16]. For customers, being actively involved in value cocreation makes them more satisfied than those who are not, encourages them to be more innovative, and increases their motivation and willingness to take risks [17]. On the other hand, enterprises have gradually realized that only by guiding the active participation of customers can they timely understand the needs of customers and better realize the "customization" of customer needs [3, 18]. Due to the influence of customer satisfaction and loyalty, customer participation in value cocreation is crucial to help enterprises gain competitive advantages.

2.1.3. Relational Embeddedness (RE). Relational embeddedness (RE) is a core concept in the research of new economic sociology and an important tool for the study of a corporate relationship network [9]. It has a certain impact on enterprises' resource acquisition, knowledge sharing, technological innovation, and other behaviors and further affects the improvement of corporate performance and the shaping of competitive advantages [18, 19]. This paper proposes that RE refers to the degree to which transport demand subjects participate in value cocreation. In the context of freight logistics service, with the changes in the competitive environment and differentiated needs of customers, small- and medium-sized freight logistics firms continue to carry out alliance innovation. Due to the lack of relevant theoretical research, it is difficult to support the management and sustainable development of freight logistics firms. The purpose is to promote mutual learning between value cocreation subjects to create value and reduce transaction costs at the same time, so as to make the network platform produce the best profit [20]. This study focuses on Manbang company, which is a professional organization whose primary value creation activities include the accumulation, creation, and dissemination of knowledge to provide

customized services or solutions that meet the need of transport demand subjects.

2.1.4. Value Cocreation (VCC). Value is cocreated in the customer experience at different points of interaction, where products, distribution channels, technology, and people are seen as gateways to the experience [9, 21]. Therefore, customers are active partners in the process of shared value creation, rather than simply passive acceptance of the value created by others. Value cocreation (VCC), as a general concept describing the collaboration between multiple stakeholders, has attracted the attention of scholars and practitioners [1, 22]. On the one hand, enterprises should open research and development, production, marketing, sales, and other value links to meet the needs of customers and carry out more adequate communication and interaction with customers [23]. On the other hand, customers (demanders) can apply their own knowledge, skills, and experience to value creation activities, such as putting forward demands, purchasing, and using products or services, so as to achieve rapid demand satisfaction and better experience [2]. Suppliers and demanders are involved in each other's value creation, so as to realize the integration and collaboration of both sides in resources, capabilities, and knowledge, so as to better meet customer needs and improve enterprise performance [16, 19]. The main activity involved in platform value cocreation is resource integration [24, 25]. According to the view of SDL, both suppliers and demanders can act as resource integrators [26]. In addition, other social and economic actors can also act as resource integrators. Customer participation is the microbasis of value cocreation, resource integration and participation platforms play a central role in value cocreation, and service ecology and business models are the macroexpression of value cocreation [27]. Therefore, this paper proposes that value cocreation refers to the multiparties such as customers, suppliers, and platform providers to create network effects through mutual cooperation and resource sharing through the network platform and jointly realize value creation.

2.2. Research Hypothesis and Theoretical Framework

2.2.1. SDL and Transport Demand Subject Value (TDS Value). VCC is an important issue in the study of SDL. Many propositions in SDL are descriptions of VCC. For example, service (the use of operational resources such as knowledge and skills) is the fundamental basis of exchange [28]. Value is created by the cocreation of multiple participants [12]. One of the most prominent views of cocreation is the SDL, which emphasizes the concept of coproducing value and the need to cocreate value. In essence, the basic unit of exchange is a service rather than a product, so that vendors can cocreate value by exchanging services with customers [16]. SDL also states that value is created through institutions created by resource integrators (all economic and social actors) and cocreated under institutional arrangements [17, 23]. Under the guidance of the SDL concept, platform enterprises more clearly realize that it is difficult to achieve value creation only by relying on their own ability and it requires the joint participation of customers, channel partners, employees, suppliers, and other stakeholders [3]. Through distributors or staff for the feedback of the customer requirements and delivery, the enterprise through the network platform integrates the resources of the upstream suppliers and develops and provides products and services, through the downstream distributors, to meet the needs of the customer; at the same time, the enterprise and stakeholders to achieve the win-win situation do common realization of value creation [29]. Therefore, the following hypothesis is proposed: H1—SDL has a significant positive effect on TDS value.

2.2.2. SDL and Transport Demand Subject Participation (TDS Participation). According to the SDL, value is by nature collaborative and cocreated rather than created by a single actor. As a result, customers are seen as valuable creators, as opposed to passive responders or receivers. The customers are active participants in the relationship communication, and they work with the organization to cocreate value through the entire service value chain [18]. In this role, the customers may go beyond the activities and responsibilities that traditionally belong to the sales organization, such as self-service or generating ideas for service improvement [30]. In this case, they are always involved in the VCC and play a key role in defining the relevant interests of each product or service solution [31]. TDS participation can play a variety of roles to help themselves, others, and the supplier [20]. Customers use their own skills to integrate resources and create value [32]. SDL shows that suppliers and demanders always work together to create value [33]. Therefore, the following hypothesis is proposed: H2—SDL has a significant positive effect on TDS participant.

2.2.3. SDL and RE. SDL holds that all economies are service economies and all participants of economic activities and social activities are resource integrators [34]. Since service economy involves many subjects and organizations, each subject integrates resources in economic activities in order to achieve its own goals [35]. It has become one of the indispensable resources for enterprises to participate in the process of market competition that they have good social relations or occupy favorable social network position (such as network centrality) [5]. Because of the strong resource allocation effect of the social network, enterprises can obtain more scarce resources at a lower cost and faster speed [21]. Therefore, the social network has become an important way for enterprises to obtain resources [36]. According to SDL, "the view of service center must be customer-oriented and relational." "Customer orientation" emphasizes the core position of customers in the service economy, and "relationship" emphasizes the relationship among customers, enterprises, and VCC subjects. The relationship between the enterprise and network members is the cornerstone of analyzing and exploring the influence of network embeddedness on enterprise behavior [22]. Therefore, the following hypothesis is proposed: H3—SDL has a significant positive effect on RE.

2.2.4. SDL and VCC. SDL defines service as "the ability to apply for the benefit of others," which is regarded as the

theoretical basis for the development of service science and the research of the service system [8]. The main body of VCC uses the service system to provide resource sharing, including information sharing, work sharing, risk sharing, and commodity sharing [37], and realizes the connection and interaction between the parties through the ability of the service system (such as knowledge, skills, and technology) [38]. Under the common value proposition, VCC is realized through resource complementarity and ability integration [39]. In the new service system, the purpose and driving force of interaction and exchange is value creation [40]. The service ecosystem is an ecosystem approach based on SDL, which focuses on service exchange and resource integration [41]. The service ecosystem is considered to be the research direction of SDL in the next 10 years. VCC in the service ecosystem means that a wide range of social and economic participants (including customers, suppliers, and distributors) creates value for themselves and other systems through service exchange and resource integration in the ecosystem [42]. First of all, SDL holds that an institution is an important factor to realize VCC [43]. Institutions promote cooperation and coordination of activities in ecosystems by providing a shared code of conduct [44]. The system in this paper refers to the win-win business model and value chain integration mechanism of the service ecosystem. Second, SDL focuses on dynamic resources (such as knowledge and skills), which are important resources to enhance the core competence of an enterprise. Through the network platform, the enterprise aggregates the knowledge and skills accumulated by various subjects and forms the core competence, which has the characteristics of value, uniqueness, and irreplaceability [2]. Finally, VCC under the service ecosystem needs the network platform as the intermediary and the VCC subject realizes value creation through the network platform. VCC in this paper contains three dimensions, namely, information sharing (IS), responsible behavior (RB), and the interaction between the supplier and demander (ISD). The network platform needs to use new technologies, improve the platform capacity, absorb and integrate multiple heterogeneous resources, meet the needs of multiple subjects of resource integration and service exchange, and finally achieve VCC. Therefore, the following hypotheses are proposed: H4—SDL has a significant positive impact on VCC. H4a—SDL has a significant positive impact on IS. H4b—SDL has a significant positive impact on RB. H4c—SDL has a significant positive impact on ISD.

2.2.5. TDS Participation and VCC. The VCC process assumes that all participants are willing and ready to share complementary expertise and skills in order to create a win-win situation for each of the actors involved [45]. In this context, TDS is seen as a key resource in the value creation process. TDS participation needs to jointly develop the value proposition, which is the result of the negotiation between the parties. Suppliers' ability to engage in positive dialogue and interaction with demanders increases their potential to support value creation, and suppliers also use their expertise to provide products and services to demanders. In addition to paying money, demanders must sacrifice time and effort to obtain products and services; demanders also contribute resources such as knowledge in the use of products and services to create the best use value [10]. Therefore, TDS needs to achieve the best balance between the benefits of use value and the cost of effort. There are three forms of customer participation: one is customer communication (that is, customer-to-customer communication), which refers to word of mouth (positive and negative), and helping others [8]. The second is customer complaint, which refers to the interactive behavior of customers giving feedback, information communication, and seeking solutions to the company in the process of using the product or service [4]. The third is customer collaboration, such as customer participation in product development and innovation [11]. Research shows that VCC benefits both suppliers and demanders. Therefore, the following hypothesis is proposed: H5—TDS participation has a significant positive effect on platform VCC.

2.2.6. TDS Participation and RE. TDS participation refers to the behavior that TDS invests all kinds of resources to create value with the suppliers [19]. The resources invested by TDS include their efforts and investment in time, spirit, intelligence, entity, and emotion. Whether TDS participate in product development or service delivery or the communication and interaction between customers, they all need to solve problems together through information sharing on the basis of mutual trust. TDS participation can increase the level of network embeddedness [15]. RE reflects the relationship characteristics between suppliers and demanders, such as relationship quality [46]. Granovetter measures RE using interaction frequency, intimacy, relationship duration, and mutual service content. RE can increase the mutual recognition and trust between suppliers and demanders, as well as between demanders, and enhance the interaction and sharing between them [17]. Therefore, the following hypothesis is proposed: H6—TDS participation has a significant positive effect on RE.

2.2.7. TDS Participation and VCC. Companies can benefit from customer innovation by creating dedicated platforms to promote customer creativity. The network platform has the characteristics of interactivity, continuity, rapidity, and flexibility, which can be used by enterprises to attract customers to actively participate in VCC [47]. On the one hand, in order to realize customer participation in enterprise VCC, enterprises need to develop a special platform to support the exchange and interaction between suppliers and demanders and demanders and demanders. On the other hand, in order to meet the needs of customer, suppliers not only need to mobilize internal forces but also need to integrate the forces and resources of distributors and demanders. This requires the platform to have a fast, flexible, value chain integration mechanism, with the core ability to quickly respond to customer needs; at the same time, the platform also has strong technological innovation ability. Therefore, TDS participation can enhance platform capabilities, including value chain integration capabilities, organizational core capabilities, and technology research and development capabilities. Accordingly, the following hypotheses are proposed: H7—TDS

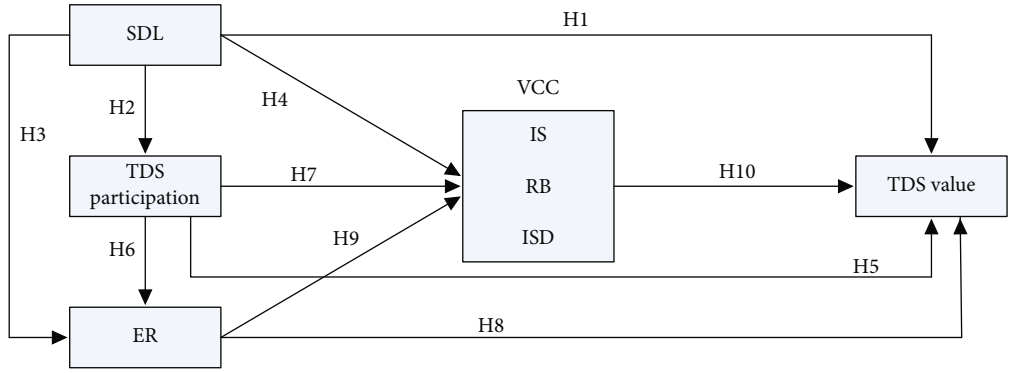


Figure 1: The theoretical model.

participation has a significant positive impact on VCC. H7a—TDS participation has a significant positive impact on IS. H7b—TDS participation has a significant positive impact on RB. H7c—TDS participation has a significant positive impact on ISD.

2.2.8. RE and TDS Value. This paper proposes that TDS participation has influence on VCC from the aspects of cognitive embedding, structural embedding, and RE. First of all, cognitive embedding is beneficial for the VCC subjects to reach a consensus and contribute to the generation of VCC behaviors of each subject. Secondly, structure embedding is conducive to platform enterprises to obtain more heterogeneous information and resources more efficiently and promote platform enterprises to realize value creation through the integration and utilization of internal and external resources. Structural embedding is to embed economic relations into the structure of the social relations network. The larger the network scale, the more partners the platform enterprises connect, and the richer the network resources. Third, RE is beneficial to the establishment of cooperation and trust among VCC subjects and to the acquisition of knowledge, technology, and other resources necessary for VCC from the relationship network [48]. In short, RE can absorb more heterogeneous resources such as knowledge, information, and technology and promote VCC. Therefore, the following hypothesis is proposed: H8—RE has a significant positive effect on TDS value.

2.2.9. RE and VCC. Different structural characteristics of embedded networks have different effects on the acquisition, integration, and reconstruction of resources [49]. Some scholars found that enterprises can improve the dynamic capability of the organization through RE, so as to grow in the turbulent environment. First of all, RE is conducive to the establishment of cooperation between enterprises, helps to build a business model with platform providers as the core, and promotes the improvement of value chain integration ability. Secondly, RE is beneficial to the improvement of the core competence of the organization. Through the learning and integration of external knowledge, RE can improve the organizational flexibility and core dynamic capabilities of enterprises, such as timely understanding of external environment changes and timely response to customer needs. Third, RE can promote the improvement of enterprises' technological innovation ability. RE helps platform enterprises to develop and provide new products and services by integrating multiple resources. In this process, knowledge should be transformed into capabilities through technology and productivity should be formed to realize value creation. Therefore, the following hypotheses are proposed: H9—RE has a significant positive impact on VCC. H9a—RE has a significant positive impact on IS. H9b—RE has a significant positive impact on RB. H9c—RE has a significant positive impact on ISD.

2.2.10. VCC and TDS Value. Platform plays an intermediary role in VCC, and VCC has a significant impact on TDS value [50]. First, IS has a significant impact on TDS value. IS means to share a win-win benefit distribution mechanism among VCC subjects, integrate the resources of upstream and downstream partners of the value chain, and form a synergistic and complementary ability with platform providers. Secondly, RB has a significant impact on TDS value. RB includes data sharing such as product development, market operations, and customer service, as well as human, financial, and technical management data. RB is formed by enterprises for a long time [51]. It has the characteristics of heterogeneity, value, difference, and irreplaceability and can bring unique value to customers. Therefore, RB can promote the creation of demander value. Finally, ISD has a significant impact on TDS value. ISD is an important link of the platform. On the one hand, to realize the interaction between demanders and suppliers, it is necessary to provide a platform and tools for VCC to realize value creation among all parties. On the other hand, platforms need to be capable of creating new products and services through continuous technological iteration. Therefore, this paper puts forward the following hypotheses: H10—VCC has a significant positive impact on TDS value. H10a—IS has a significant positive impact on TDS value. H10b—RB has a significant positive influence on TDS value. H10c—IDS has a significant positive influence on TDS value.

According to theoretical analysis, a theoretical framework is constructed to explain the influence mechanism among SDL, TDS participation, RE, VCC, and TDS value (see Figure 1).

Table 1: Reliability and validity.

Latent variable	Dimension	Average value	Standard deviation	Factor loading	Cumulative variance contribution rate (%)	KMO	Cronbach's α
SDL		6.76	0.58	0.73	74.58	0.86	0.92
TDS participation	TDSP1	6.74	0.51	0.81	78.32	0.87	0.91
	TDSP2	6.78	0.52	0.84	69.99	0.84	0.91
	TDSP3	6.71	0.51	0.78	83.12	0.82	0.93
	TDSP4	6.73	0.53	0.79	81.39	0.84	0.94
	TDSP5	6.77	0.53	0.76	79.85	0.83	0.89
RE	RE1	6.78	0.54	0.81	81.45	0.87	0.88
	RE2	6.72	0.55	0.83	84.42	0.91	0.91
	RE3	6.71	0.57	0.82	78.65	0.89	0.93
VCC	IS	6.74	0.54	0.79	82.67	0.74	0.94
	RB	6.69	0.56	0.71	81.34	0.79	0.91
	ISD	6.75	0.51	0.84	76.45	0.82	0.95
TDS value		6.74	0.56	0.81	73.28	0.76	0.89

3. Research Design

3.1. Sample Selection and Data Collection. This study takes the Manbang Group as a freight transport platform provider to conduct an investigation. In this study, the Manbang Group as a freight transport platform provider was taken as an example to conduct an investigation. The respondents were manufacturing enterprises, and the samples were from 21 provinces (municipalities) in Mainland China. The survey lasted from January 2021 to March 2021, and a total of 857 questionnaires were collected, and 786 valid questionnaires were collected after the unqualified questionnaires were eliminated. As a whole, the survey subjects are widely distributed and all of them have experience in the use of a Manbang freight transport platform. The validity of the data is guaranteed and the basic requirements of the research are met.

3.2. Questionnaire Design and Measurement Scale. On the basis of referring to the previous research scale, the authors designed the scale items combined with the actual. As most of these variables are difficult to be measured objectively and quantitatively, this study adopts the Likert seven-grade scale scoring method to measure them. A numerical scale of 1 to 7 indicates disagreement to agreement or a transformation from low to high. 4 represents the intermediate state [52]. Through soliciting opinions from sampling transport demand enterprises, consulting experts in the transport industry, and small-scale pretesting, the initial questionnaire was improved and the final survey questionnaire was formed on this basis.

3.3. Homologous Deviation and Integration Test. The questionnaire survey is conducted online, so homology deviation test should be carried out on the latent variable data, and four common factors with characteristic roots greater than 1 should be extracted. The explanatory variance of the first common factor is 32.9%, and the cumulative explanatory variance is 71.37%, which does not exceed the standard of 40%, so it can be used for data analysis. All the variables involved are the investigation of the phenomenon at the overall level of the organization. According to the suggestions of Shen et al. [27], the James formula is adopted to calculate the internal consistency coefficient RWG for testing and RWG (0.71~0.78) is greater than 0.70, indicating that it is feasible to integrate individual measurement data into data at the organizational level [52].

4. Research Results

4.1. Reliability and Validity. After the research team received the completed questionnaires, reliability analysis and validity analysis are generally two initial analyses that must be conducted to verify. The software SPSS 25.0 is used to analyze reliability and validity. For reliability, it is crucial to produce stable and consistent results. The higher the test reliability is, the more reliable the result is. There are many ways to measure the reliability. Cronbach's coefficient alpha is widely used in almost all questionnaire reliability analyses. Generally speaking, the alpha coefficient should be above 0.8 and 0.7~0.8 is an acceptable range. Nevertheless, the alpha coefficient of the subscale should be above 0.7 and 0.6~0.7 is acceptable. If the reliability coefficient is lower than 0.6, the scale should be modified [52]. In this study, the Cronbach's coefficients are all greater than 0.7 and are within the acceptable range (see Table 1). Therefore, it indicates that the reliability of the questionnaire is qualified [52].

Validity is the degree to which the measured results reflect the content to be examined. The validity of the questionnaire could be checked by the constructs of criterion validity and coherency. Bartlett test of sphericity and KMO (Kaiser Meyer–Olykin measure of sampling adequacy) are used to analyze the structure validity in this paper. Results of KMO are all greater than 0.71 (see Table 1), and the Bartlett statistic is different from 0 [52]. Therefore, the validity of the questionnaire is excellent.

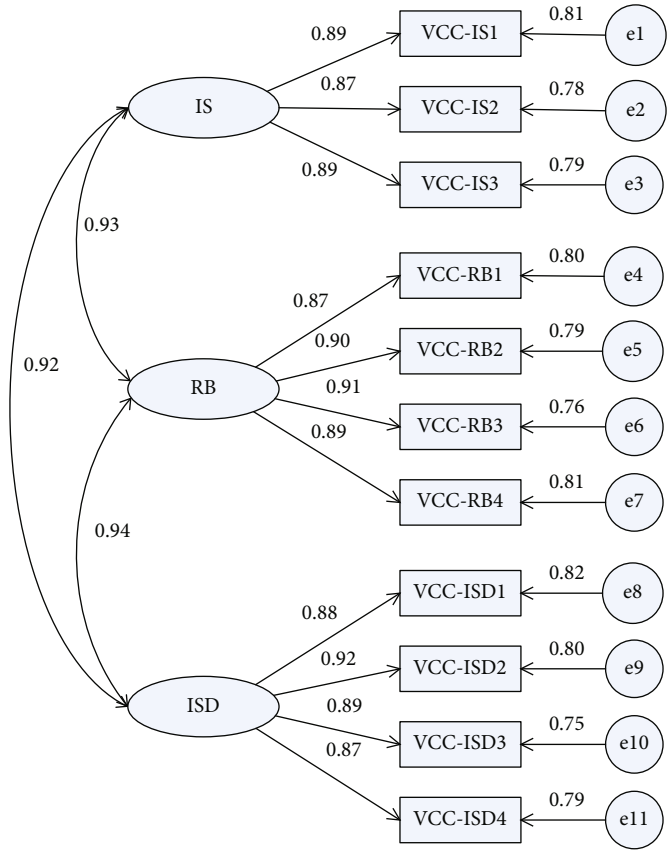


Figure 2: VCC first-order factor model.

4.2. CFA Test. By using software AMOS22 to conduct confirmatory factor analysis, it is found that the factor loading values of IS, RB, and ISD in value cocreation are within the range of 0.91~0.94 (see Figure 2), indicating that there is a correlation among the three potential variables and the correlation is relatively high (greater than 0.5). Second-order confirmatory factor analysis can be continued.

The second-order CFA test was performed on IS, RB, and ISD. Figure 3 shows that the three strongly correlated first-order factors (IS, RB, and ISD) are generally affected by a high-order factor (VCC) and the factor load coefficient is in the range of 0.95~0.98 (greater than 0.7). It is shown that the second-order latent variable VCC can represent three first-order latent variables, namely, IS, RB, and ISD.

4.3. SEM Test

4.3.1. Structural Equation Analysis of the Integrated Model. AMOS22 was used to construct the structural equation model, modify the MI index of the model, and analyze the model. The fitting indexes and their values of the integrated model were shown in Table 2, and the fitting indexes all met the requirements.

The path coefficient and hypothesis test in the structural equation of the integrated model are shown in Table 3.

The following conclusions can be drawn from Table 3:

(1) SDL, TDS participation, and RE do not directly affect TDS value but indirectly through platform VCC

(2) SDL indirectly influences VCC through RE and then realizes TDS value through VCC. Under the guidance of the SDL concept, transport enterprises, manufacturing enterprises, and other value-creating entities realize resource integration with the support of VCC through RE and finally realize TDS value

(3) SDL indirectly significantly affects TDS value through TDS participation and through VCC

(4) TDS participation significantly affected RE. The path coefficients all reach the significant level, indicating that the demand of TDS can directly affect the transportation service providers and promote the transportation enterprises to update the transportation service, instead of the traditional model from product to customer

(5) RE significantly affected VCC, and RE significantly affected TDS value through VCC

Therefore, it is assumed that H2, H3, H4, H6, H7, H9, and H10 are all supported. Moreover, from the perspective of TDS value, value cocreation plays an important intermediary role in TDS participation in service innovation of a freight transport platform.

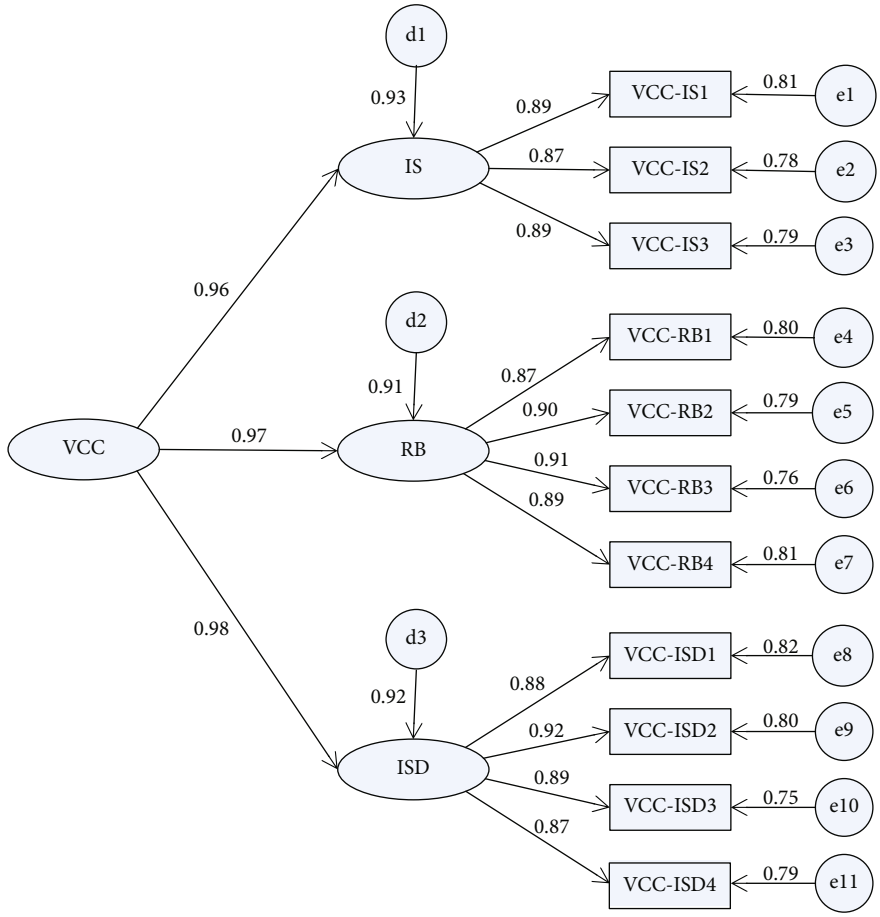


Figure 3: VCC second-order factor model.

Table 2: Fitting results of the integrated model.

Items	χ^2/df	RMSEA	IFI	CFI	TLI	PGFI	PNFI	PCFI
Index value	4.684	0.065	0.932	0.927	0.918	0.784	0.854	0.863
Measuring standard	<5 (large sample)	<0.08	>0.9	>0.9	>0.9	>0.5	>0.5	>0.5

Table 3: Path coefficient and hypothesis test of the integrated model.

Path	Standard coefficient	T value	P value	Results
H1: SDL → TDS value	0.132	0.615	>0.05	Nonsupport
H2: SDL → TDS participation	0.684	5.968	<0.001	Support
H3: SDL → RE	0.236	4.712	<0.001	Support
H4: SDL → VCC	0.271	4.891	<0.001	Support
H5: TDS participation → TDS value	0.035	0.698	>0.05	Nonsupport
H6: TDS participation → RE	0.765	6.537	<0.001	Support
H7: TDS participation → VCC	0.198	3.526	<0.001	Support
H8: RE → TDS value	0.023	1.684	>0.05	Nonsupport
H9: RE → VCC	0.798	14.358	<0.001	Support
H10: VCC → TDS value	0.912	32.145	<0.001	Support

Table 4: Fitting results of the subdimensional model.

Items	χ^2/df	RMSEA	IFI	CFI	TLI	PGFI	PNFI	PCFI
Index value	4.798	0.066	0.928	0.919	0.916	0.769	0.851	0.859
Measuring standard	<5 (large sample)	<0.08	>0.9	>0.9	>0.9	>0.5	>0.5	>0.5

Table 5: Path coefficient and hypothesis test of the subdimensional model.

Path	Standard coefficient	T value	P value	Results
H4a: SDL → VCC-IS	0.668	5.685	<0.001	Support
H4b: SDL → VCC-RB	0.784	5.923	<0.001	Support
H4c: SDL → VCC-ISD	0.694	5.742	<0.001	Support
H7a: TDS participation → VCC-IS	0.587	5.821	<0.001	Support
H7b: TDS participation → VCC-RB	0.854	14.251	<0.001	Support
H7c: TDS participation → VCC-ISD	0.765	6.537	<0.001	Support
H9a: RE → VCC-IS	0.918	13.526	<0.001	Support
H9b: RE → VCC-RB	0.823	15.684	<0.001	Support
H9c: RE → VCC-ISD	0.798	14.354	<0.001	Support
H10a: VCC-IS → TDS value	0.453	4.876	<0.001	Support
H10b: VCC-RB → TDS value	0.378	4.325	<0.001	Support
H10c: VCC-ISD → TDS value	0.402	5.134	<0.001	Support

4.3.2. Structural Equation Analysis of the Subdimensional Model. AMOS22 was used to test the role of the VCC subdimension among SDL, TDS participation, RE, and TDS value. The corresponding fitting indexes and their values are shown in Table 4.

The fitting indexes all met the requirements. The path coefficients and hypothesis test in the structural equation of the subdimensional model are shown in Table 5.

Table 5 shows that SDL, TDS participation, and RE have significant effects on IS, RB, and ISD. At the same time, IS, RB, and ISD have significant influences on TDS value. Therefore, hypotheses H4a~H4c, hypotheses H7a~H7c, hypotheses H9a~H9c, and hypotheses H10a~H10c are valid.

In conclusion, SDL, TDS participation, and RE do not directly affect TDS value but indirectly through VCC and its three subdimensions.

4.3.3. Test of Mediating Effect. In this part, the method proposed by Baron, Zhonglin Wen, and other scholars is adopted. And Mplus8.0 is used to test the mediating effect, in which SDL is X1, TDS participation is X2, RE is X3, VCC is M, IS is M1, RB is M2, ISD is M3, and TDS value is Y. The results are shown in Table 6.

As can be seen in Table 6, the model has a good fitting degree. According to the parameter estimates, the second-order latent variable VCC and its three first-order latent variables, namely, IS, RB, and ISD, all have significant mediating effects on SDL, TDS participation, RE, and TDS value. Therefore, the above hypotheses are valid and VCC has a mediating effect on TDS participation in service innovation of the freight transport platform.

5. Conclusions and Management Implications

5.1. Conclusions and Discussions. This research puts forward a theoretical framework about the relationships among SDL, TDS participation, RE, platform VCC, and TDS value. Taking platform VCC as the mediating variable, 10 hypotheses are proposed to explain the influence mechanism of value cocreation of the network freight transport platform from the perspective of TDS participation. Through empirical analysis of the questionnaire survey, the research conclusions are as follows:

(1) SDL does not directly affect TDS value, but indirectly through VCC. Structural equation model analysis shows that H4 and H10 pass the significance test, but H1 does not pass the significance test. On the one hand, SDL provides the conceptual guidance for TDS participation in VCC. On the other hand, VCC provides strong support for the innovation development of transport enterprises in the IoT-based environment

(2) SDL affects not only VCC directly but also indirectly through TDS participation and RE. First, the concept of SDL unifies the value cocreation subject's understanding of the importance of VCC. The main body of the value network, such as customers, suppliers, and distributors, should take the platform provider as the core and jointly build and improve the platform ecological system, so as to achieve better value cocreation. Second, TDS participation promotes value creation. SDL emphasizes that customers are the participants of value cocreation and customer

Table 6: Results of the mediating effect test.

Path	Estimated coefficient	Standard deviation	95% confidence interval		P	Test of goodness for fit
			Lower limit	Upper limit		
X1 → M → Y	0.987	0.061	0.654	0.821	≤0.001	CMIN/DF = 1.342, CFI = 0.989, TLI = 0.987, RMSEA = 0.021
X2 → M → Y	0.921	0.094	0.821	1.084	≤0.001	CMIN/DF = 1.627, CFI = 0.977, TLI = 0.968, RMSEA = 0.026
X3 → M → Y	0.961	0.201	0.801	1.062	≤0.001	CMIN/DF = 1.378, CFI = 0.984, TLI = 0.988, RMSEA = 0.022
X1 → M1 → Y	0.358	0.115	0.203	0.649	0.002	CMIN/DF = 2.849, CFI = 0.948, TLI = 0.931, RMSEA = 0.047
X1 → M2 → Y	0.534	0.172	0.234	0.937	0.002	
X1 → M3 → Y	0.512	0.124	0.365	0.798	≤0.001	
X2 → M1 → Y	0.301	0.101	0.162	0.491	0.002	CMIN/DF = 1.957, CFI = 0.952, TLI = 0.953, RMSEA = 0.033
X2 → M2 → Y	0.412	0.126	0.198	0.612	≤0.001	
X2 → M3 → Y	0.453	0.135	0.261	0.698	≤0.001	
X3 → M1 → Y	0.247	0.171	0.047	0.568	0.016	CMIN/DF = 1.491, CFI = 0.978, TLI = 0.981, RMSEA = 0.024
X3 → M2 → Y	0.384	0.164	0.098	0.931	0.022	
X3 → M3 → Y	0.462	0.142	0.453	0.729	≤0.001	

participation accelerates the information flow between customers, platform enterprises, and suppliers and greatly improves the response speed and flexibility of the platform. Third, RE promotes VCC.Participants interact through the network platform and create value together

(3) TDS participation does not directly affect TDS value, but indirectly through VCC. On the one hand, TDS participation can continuously improve the function of the platform, realize IS and ISD, and improve the response speed and innovation ability of the platform, so that the platform can better support the main body of VCC to realize value creation. On the other hand, TDS participation is conducive to information communication and knowledge sharing among demanders, between demanders and suppliers and between suppliers on the platform, so as to realize multiparty resource integration and complementarity and promote the performance of the logistics service supply chain

(4) TDS participation affects not only VCC directly but also indirectly through ER. TDS participation has a direct impact on the platform capability, including the platform's value chain integration, organizational core, and technological innovation capabilities. At the same time, TDS participation also directly affects RE and indirectly affects VCC through RE. On the one hand, demanders provide their demands and suggestions to suppliers through the platform, which helps suppliers integrate the resources of partners on the value chain, innovate, and develop products and services to meet the needs of TDS. On the other hand, TDS participation indirectly affects VCC through RE. TDS participation improves the RE between demanders and suppliers, and RE indirectly promotes the realization of resource complementarity among various subjects and promotes the improvement of value chain integration ability

(5) RE does not directly affect TDS value, but indirectly through VCC. First, RE promotes the cooperation between VCC subjects. Second, RE improves the core competence of the organization. Each subject realizes the complementarity and integration of knowledge, technology, and resources in the network embedding, and the injection of heterogeneous resources can promote the improvement of enterprise core competence. Third, RE promotes technological innovation and enhances technological innovation ability. Therefore, RE influences TDS value by promoting VCC

(6) Mediating role of VCC. The empirical analysis results show that the overall VCC plays a mediating role in the relationships between SDL and TDS value, TDS participation and TDS value, RE, and TDS value. At the same time, IS, RB, and ISD also play a mediating role in the relationship between SDL, TDS participation, RE, and TDS value

5.2. Management Implications. The results show that the key point of the transportation enterprise reform is the VCC of TDS participation in the network freight transport platform. In order to improve TDS satisfaction and loyalty, VCC needs to play an intermediary role and the deepening of VCC needs to be jointly promoted through the innovation of the SDL concept, the guidance of TDS participation, and the strengthening of the network relationship.

(1) Establish symbiotic organizations and build the platform ecological system. First, the diversified and personalized TDS needs make demanders become the initiator of demand and the participant of value creation. The C2B model is on the rise, while the traditional B2C model is facing severe challenges. Second, as a provider of products and services, a freight transport enterprise cannot win in the competition only by its internal resources and strength and it is even more difficult to achieve new growth. Therefore, it is necessary for the freight transport enterprise to build a value network based on TDS participation in value creation, regard the members of the value network as a community of destiny, and build a symbiotic organization. Thirdly, in order to realize VCC, symbiotic organizations need the support of the platform, and the VCC platform needs to be jointly created by symbiotic organizations and entities constantly interact

(2) It is pointed out that TDS value is the core of enterprise management and the focus of the common concern of all stakeholders. Peter Drucker said that "the purpose of business is to create customers." Therefore, VCC activities start from the needs and preferences of customers. Secondly, choose partners or suppliers with corresponding resources or capabilities. Third, the use of the corresponding technology and resources to achieve product development and service provision. Finally, the product or service is delivered to the TDS through the distributor close to the TDS or the TDS himself

(3) Guide TDS to participate and improve the platform capability. Enterprises should build a sustainable competitive advantage, establish a demander-oriented culture and machine, and provide support for TDS participation. One is to establish the "TDS-centered" service concept and create a cultural atmosphere of TDS participation. Second is to establish the incentive mechanism and operation process of TDS participation. The third is to provide tools and platforms for TDS participation

(4) Optimize network structure and enhance platform capability. First, build a culture of the shared vision, values, and identity to achieve relationship embedding. Secondly, optimize the network to expand the scale of the network and encourage more participants to participate, to achieve more heterogeneous resource integration. The frequency of contact and interaction among partners such as TDS, platform providers, and suppliers can improve the relationship quality among VCC subjects and promote the improvement of platform capability through good relationship embedding

5.3. Research Limitations and Prospects. Although this research has analyzed and studied the influence mechanism of VCC of the network freight transport platform from the aspects of SDL, TDS participation, RE, and VCC, there are still some shortcomings. It is mainly reflected in the following aspects: first, only the overall analysis of TDS participation is carried out and the specific dimensions such as the willingness, motivation, and behavior of TDS to participate in VCC of freight transport platform are not analyzed and displayed. Future research can strengthen this part of the content. Second, there is no in-depth analysis and research on the influence of various dimensions of RE on VCC and TDS value. Future research can increase the influence analysis of trust, communication, and commitment on VCC and TDS value. Third, the scope of the investigation is small. Only one industry is taken as an example. Future research can expand the scope of investigation to multiple industries and enterprises and test the theoretical model with multicase analysis. Fourth, the environmental factor is not considered in this study and subsequent studies should focus on the moderating effect of environmental dynamics.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Conceptualization was done by P.B. and Q.L. Data curation was done by P.B. and Q.L. Funding acquisition was done by Q.L. and B.C. Investigation was one P.B. and Y.J. Methodology was done by P.B. and B.C. Project administration was done by Q.L. and B.C. Resources were acquired by Q.L. and L.Z. Supervision was done by Q.W. Validation was done by Q.W. Visualization was done by L.Z. and Y.J. Writing and original draft preparation were done by P.B. and Q.L. Writing, review, and editing were done by Q.L. and L.Z.

Acknowledgments

This research was supported by the National Social Science Foundation of China (no. 19CGL004), Shaanxi Education Department Foundation (no. 19JZ011), Shaanxi Social Science Foundation (nos. 2019D013 and 2020D028), research program funded by Shaanxi Provincial Philosophy and Social Sciences Major Theoretical and Practical Issues (no. 2021ND0444), and Central Universities Fundamental Research Foundation for Chang'an University (nos. 300102231603, 300102231615, and 300102341677).

References

[1] Y. W. Lee, H.-C. Moon, and W. Yin, "Innovation process in the business ecosystem: the four cooperations practices in the media platform," *Business Process Management Journal*, vol. 26, no. 4, pp. 943–971, 2020.

[2] Y. K. Lew, J. Kim, and Z. Khan, "Technological adaptation to a platform and dependence: value co-creation through

partnerships," *Asian Journal of Technology Innovation*, vol. 27, no. 1, pp. 71–89, 2019.

[3] B. Peng, Y. Wang, S. Zahid, G. Wei, and E. Elahi, "Platform ecological circle for cold chain logistics enterprises: the value co-creation analysis," *Industrial Management & Data Systems*, vol. 120, no. 4, pp. 675–691, 2020.

[4] K. D. Bahn†, K. L. Granzin, and M. Tokman, "End-User contribution to logistics value co-creation: a series of exploratory studies," *Journal of Marketing Channels*, vol. 22, no. 1, pp. 3–26, 2015.

[5] K. A. Vakeel, E. C. Malthouse, and A. Yang, "Impact of network effects on service provider performance in digital business platforms," *Journal of Service Management*, vol. ahead-of-print, no. ahead-of-print, 2020.

[6] L. V. Ngo and A. O'Cass, "Innovation and business success: the mediating role of customer participation," *Journal of Business Research*, vol. 66, no. 8, pp. 1134–1142, 2013.

[7] Y. Fernando and C. Chukai, "Value co-creation, goods and service tax (GST) impacts on sustainable logistic performance," *Research in Transportation Business & Management*, vol. 28, pp. 92–102, 2018.

[8] S.-T. D. Yuan, S. Y. Chou, W. C. Yang, C. A. Wu, and C. T. Huang, "Customer engagement within multiple new media and broader business ecosystem – a holistic perspective," *Kybernetes*, vol. 46, no. 6, pp. 1000–1020, 2017.

[9] M. L. Santos-Vijande, J. Á. López-Sánchez, and P. Pascual-Fernández, "Co-creation with clients of hotel services: the moderating role of top management support," *Current Issues in Tourism*, vol. 21, no. 3, pp. 301–327, 2018.

[10] V. Marino and L. Lo Presti, "Engagement, satisfaction and customer behavior-based CRM performance," *Journal of Service Theory and Practice*, vol. 28, no. 5, pp. 682–707, 2018.

[11] J. Carlson, M. Rahman, R. Voola, and N. de Vries, "Customer engagement behaviours in social media: capturing innovation opportunities," *Journal of Services Marketing*, vol. 32, no. 1, pp. 83–94, 2018.

[12] C.-C. Chu, Y. F. Cheng, F. S. Tsai, S. B. Tsai, and K. H. Lu, "Open innovation in crowdfunding context: diversity, knowledge, and networks," *Sustainability*, vol. 11, no. 1, p. 180, 2019.

[13] N. Hasan and A. A. Rahman, "Ranking the factors that impact customers online participation in value co-creation in service sector using analytic hierarchy process," *International Journal of Information Systems in the Service Sector*, vol. 9, no. 1, pp. 37–53, 2017.

[14] C. Drummond, H. McGrath, and T. O'Toole, "The impact of social media on resource mobilisation in entrepreneurial firms," *Industrial Marketing Management*, vol. 70, pp. 68–89, 2018.

[15] J. Frempong, J. Chai, E. M. Ampaw, D. O. Amofah, and K. W. Ansong, "The relationship among customer operant resources, online value co-creation and electronic-word-of-mouth in solid waste management marketing," *Journal of Cleaner Production*, vol. 248, p. 119228, 2020.

[16] T. Hurni, T. L. Huber, J. Dibbern, and O. Krancher, "Complementor dedication in platform ecosystems: rule adequacy and the moderating role of flexible and benevolent practices," *European Journal of Information Systems*, pp. 1–24, 2020.

[17] M. J. J. Lin and C. H. Huang, "The impact of customer participation on NPD performance: the mediating role of inter-organisation relationship," *Journal of Business & Industrial Marketing*, vol. 28, no. 1, pp. 3–15, 2013.

[18] V. Ramaswamy and K. Ozcan, "What is co-creation? An interactional creation framework and its implications for value creation," *Journal of Business Research*, vol. 84, pp. 196–205, 2018.

[19] J. Frempong, J. Chai, and E. Ampaw, "Effects of waste management customer online value co-creation on sanitation attitude and advocacy: a customer-enterprise dyadic perspective," *Sustainability*, vol. 10, no. 7, p. 2557, 2018.

[20] M. Sartas, P. van Asten, M. Schut et al., "Factors influencing participation dynamics in research for development interventions with multi-stakeholder platforms: a metric approach to studying stakeholder participation," *PLoS One*, vol. 14, no. 11, article e0223044, 2019.

[21] H. Widjojo, A. Fontana, G. Gayatri, and A. W. Soehadi, "Value co-creation for marketing innovation: comparative study in the Sme community," *International Journal of Innovation Management*, vol. 24, no. 3, p. 2050030, 2020.

[22] J. Yu, Y. Wen, J. Jin, and Y. Zhang, "Towards a service-dominant platform for public value co-creation in a smart city: evidence from two metropolitan cities in China," *Technological Forecasting and Social Change*, vol. 142, pp. 168–182, 2019.

[23] H. Pan, "Online community value co-creation," *Online Information Review*, vol. 44, no. 3, pp. 645–669, 2020.

[24] A. Abdelgawad and K. Yelamarthi, "Internet of Things (IoT) platform for structure health monitoring," *Wireless Communications and Mobile Computing*, vol. 2017, 10 pages, 2017.

[25] C. Wang and C.-H. Chen, "An E-commerce economic dynamic data evaluation model based on multiuser demand constraints," *Wireless Communications and Mobile Computing*, vol. 2021, 9 pages, 2021.

[26] S. Wu, H. Kou, C. Lv, W. Huang, L. Qi, and H. Wang, "Service recommendation with high accuracy and diversity," *Wireless Communications and Mobile Computing*, vol. 2020, 10 pages, 2020.

[27] G. Shen, W. Wang, Q. Mu, Y. Pu, Y. Qin, and M. Yu, "Data-driven cybersecurity knowledge graph construction for industrial control system security," *Wireless Communications and Mobile Computing*, vol. 2020, 13 pages, 2020.

[28] J. A. Fehrer, H. Woratschek, and R. J. Brodie, "A systemic logic for platform business models," *Journal of Service Management*, vol. 29, no. 4, pp. 546–568, 2018.

[29] T. G. Pittz and T. Adler, "An exemplar of open strategy: decision-making within multi-sector collaborations," *Management Decision*, vol. 54, no. 7, pp. 1595–1614, 2016.

[30] N. Rubio, N. Villaseñor, and M. J. Yagüe, "Sustainable co-creation behavior in a virtual community: antecedents and moderating effect of participant's perception of own expertise," *Sustainability*, vol. 12, no. 19, p. 8151, 2020.

[31] B. Sarmah, S. Kamboj, and J. Kandampully, "Social media and co-creative service innovation: an empirical study," *Online Information Review*, vol. 42, no. 7, pp. 1146–1179, 2018.

[32] S. P. Singaraju, Q. A. Nguyen, O. Niininen, and G. Sullivan-Mort, "Social media and value co-creation in multi-stakeholder systems: a resource integration approach," *Industrial Marketing Management*, vol. 54, pp. 44–55, 2016.

[33] R. R. Sinkovics, O. Kuivalainen, and A. S. Roath, "Value co-creation in an outsourcing arrangement between manufacturers and third party logistics providers: resource commitment, innovation and collaboration," *Journal of Business & Industrial Marketing*, vol. 33, no. 4, pp. 563–573, 2018.

[34] L. T. Tuan, "Under entrepreneurial orientation, how does logistics performance activate customer value co-creation

behavior?," *The International Journal of Logistics Management*, vol. 28, no. 2, pp. 600–633, 2017.

[35] M. B. Tudose, G. Agafitei, and S. Avasilcai, "New research direction on performance and co-creation: a literature review," *IOP Conference Series: Materials Science and Engineering*, vol. 898, 2020.

[36] B. Xu, H. Zheng, Y. Xu, and T. Wang, "Configurational paths to sponsor satisfaction in crowdfunding," *Journal of Business Research*, vol. 69, no. 2, pp. 915–927, 2016.

[37] S. Ahmad, A. Badwelan, A. M. Ghaleb et al., "Analyzing critical failures in a production process: is industrial IoT the solution?," *Wireless Communications and Mobile Computing*, vol. 2018, 12 pages, 2018.

[38] A. Biørn-Hansen, T. M. Grønli, G. Ghinea, and S. Alouneh, "An empirical study of cross-platform mobile development in industry," *Wireless Communications and Mobile Computing*, vol. 2019, 12 pages, 2019.

[39] A. Busson and I. Lahsen-Cherif, "Impact of resource blocks allocation strategies on downlink interference and SIR distributions in LTE networks: a stochastic geometry approach," *Wireless Communications and Mobile Computing*, vol. 2018, 15 pages, 2018.

[40] P. Chamoso, A. González-Briones, S. Rodríguez, and J. M. Corchado, "Tendencies of technologies and platforms in smart cities: a state-of-the-art review," *Wireless Communications and Mobile Computing*, vol. 2018, 17 pages, 2018.

[41] S. Chen and H. Cheng, "Analysis of customization strategy for E-commerce operation based on big data," *Wireless Communications and Mobile Computing*, vol. 2021, 11 pages, 2021.

[42] S. Delle Monache, L. Comanducci, M. Buccoli et al., "A presence- and performance-driven framework to investigate interactive networked music learning scenarios," *Wireless Communications and Mobile Computing*, vol. 2019, 20 pages, 2019.

[43] D. Deng, C. Li, L. Fan, X. Liu, and F. Zhou, "Impact of antenna selection on physical-layer security of NOMA networks," *Wireless Communications and Mobile Computing*, vol. 2018, 11 pages, 2018.

[44] J. J. Escudero-Garzás and C. Bousoño-Calzón, "An analysis of the network selection problem for heterogeneous environments with user-operator joint satisfaction and multi-RAT transmission," *Wireless Communications and Mobile Computing*, vol. 2017, 13 pages, 2017.

[45] M. Pasha and S. M. W. Shah, "Framework for E-health systems in IoT-based environments," *Wireless Communications and Mobile Computing*, vol. 2018, 11 pages, 2018.

[46] T. Leclercq, I. Poncin, and W. Hammedi, "The engagement process during value co-creation: gamification in new product-development platforms," *International Journal of Electronic Commerce*, vol. 21, no. 4, pp. 454–488, 2017.

[47] M. Imran, H. A. Khattak, D. Millard, T. Tiropanis, T. Bashir, and G. Ahmed, "Calculating trust using multiple heterogeneous social networks," *Wireless Communications and Mobile Computing*, vol. 2020, 14 pages, 2020.

[48] H. Kou, F. Wang, C. Lv et al., "Trust-based missing link prediction in signed social networks with privacy preservation," *Wireless Communications and Mobile Computing*, vol. 2020, 10 pages, 2020.

[49] A. Lakas, M. E. A. Fekair, A. Korichi, and N. Lagraa, "A multi-constrained QoS-compliant routing scheme for highway-based vehicular networks," *Wireless Communications and Mobile Computing*, vol. 2019, 18 pages, 2019.

[50] W. Li, L. Yang, Z. Wen, J. Chen, and X. Wu, "On the optimization strategy of EV charging station localization and charging piles density," *Wireless Communications and Mobile Computing*, vol. 2021, 13 pages, 2021.

[51] X. Li, X. Jin, Q. Wang, M. Cao, and X. Chen, "SCCAF: a secure and compliant continuous assessment framework in cloud-based IoT context," *Wireless Communications and Mobile Computing*, vol. 2018, 18 pages, 2018.

[52] P. Bai, Q. Wu, Q. Li, C. Xue, and L. Zhang, "Mediating effect of organizational learning capacity on the relationship between relational embeddedness and innovation performance in freight logistics service," *Complexity*, vol. 2021, 18 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9965781, 11 pages
https://doi.org/10.1155/2021/9965781

Research Article

Motion Direction Inconsistency-Based Fight Detection for Multiview Surveillance Videos

Chuang Yao,[1] **Xiaoyan Su**,[1] **Xuehua Wang**,[1] **Xinyi Kang**,[1] **Jun Zhang**,[2] **and Jiankang Ren**[3]

[1]*School of Economics and Management, Dalian University of Technology, Dalian 116023, China*
[2]*Graduate School of Education, Dalian University of Technology, Dalian 116023, China*
[3]*School of Computer Science and Technology, Dalian University of Technology, Dalian 116023, China*

Correspondence should be addressed to Xiaoyan Su; sxy@dlut.edu.cn

Received 6 March 2021; Revised 12 April 2021; Accepted 19 April 2021; Published 17 May 2021

Academic Editor: Keping Yu

Nowadays, with the increasing number of surveillance cameras, human behavior detection is of importance for public security. Detection of fight behavior using video surveillance is an essential and challenging research field. We propose a multiview fight detection method based on statistical characteristics of the optical flow and random forest. Cyberphysical systems for monitoring can obtain timely and accurate information from this method. Two novel descriptors named Motion Direction Inconsistency (MoDI) and Weighted Motion Direction Inconsistency (WMoDI) are defined to improve the performance of existing methods for videos with different shooting views and solve the misjudgment on nonfight, such as running and talking. First, YOLO V3 algorithm is applied to mark the motion areas, and then, the optical flow is computed to extract descriptors. Finally, Random Forest is used for classification based on statistical characteristics of descriptors. The evaluation results on CASIA dataset demonstrate that the proposed method can improve the accuracy and reduce the rate of missing alarm and false alarm for the detection, and it is very robust against videos with different shooting views.

1. Introduction

With the increasing public demand for security and the development of machine vision technology, research in the field of intelligent monitoring continues to deepen, and many scholars are committed to use the powerful computing power of computers to process surveillance video data to strengthen security control. Nowadays, large-scale monitoring cyberphysical systems surround us and the massive videos obtained from these systems should be analyzed to guarantee public demand for security. Nevertheless, security systems relying on human observers are inefficient, and it may cause missed alarms due to the limited human capability to monitor surveillance video continuously, resulting in an urgent demand to a research of automatic alarm technology for the abnormal behaviors like fighting. And the result of proposed method will be a foundation for the cyberphysical systems to monitor and control of entities in the physical world [1].

As a key step to achieve behavior recognition, behavior analysis can explain and describe the motion of the target by using machine vision. Generally, the behavior analysis methods are categorized in two classes: the methods based on high-level human body structure and the methods based on low-level image information. The information used in high-level methods includes body point model, 2D model of human body [2], and 3D model of human body [3–5]. Moreover, in order to obtain more information, multiview videos or a 3D camera are required for this method. However, it makes the model complicated and computationally intensive. And it is difficult to guarantee the stability of this kind of method due to the immaturity of the human posture estimation technology. On the other hand, in the behavior analysis methods based on low-level image information, motion trajectory [6], shape features [7], texture features [8], optical flow features, and other image information are used to perform the behavior analysis. This kind of method

describes the target's behavior from a macroperspective, and it has the characteristics of simplicity and low complexity. Compared with the methods based on high-level human body structure, it performs better in real-time processing.

Fight as a typical abnormal behavior not only threatens the safety of citizens, but also has a negative influence on public. Therefore, it is crucial to achieve the real-time and reliable fight behavior recognition to assist security agents and ensure safety. The optical flow information extracted from the changes in pixels of the video sequences has satisfactory spatial and temporal characteristics, which is widely used to describe the target's motion tendency in video processing [9]. In recent years, the fighting detection based on the optical flow has been widely concerned. Some researchers detect fight by setting threshold on optical flow features. After that, scholars focus on employing machine learning algorithms for fight detection.

Based on the extracted optical flow information, some detection methods based on descriptors and threshold are proposed. Lin et al. [10] combined texture features and optional flow features to improve recognition rate in crowed scenes. The amplitude-based weighted direction histogram can suppress the direction confusion caused by noise in a small area and the entropy of this histogram can characterize the degree of chaos in the motion. However, it does not consider the changes of the optical flow energy when fight occurs [11]. Xu et al. [12] combined the weighted reconstruction trajectory and histogram entropy to detect fight behavior. The changes in the optical flow energy of the target object can be used to identify abnormal behavior. Nevertheless, it does not make a good distinction between fighting and running [13]. Lin et al. [7] used the motion features and shape features to detect fight, which can reduce the misjudgment rate of anomaly recognition effectively. However, it still has the problem of misjudgment in the conversation behavior. Based on the statistics of the amplitude and direction of optical flow, Huang et al. [14, 15] adopted mean and variance to recognize fight, providing a novel idea for definition of the features. Though the detection approach based on threshold method is simple and performs well in real-time, it performs not very well in videos with different shooting views and can hardly avoid misjudgment on behaviors like running and talking.

Besides, some scholars investigated fighting detection by using machine learning models for purpose of distinguishing fight and other behaviors effectively. Wang et al. [16] proposed a new feature, called Acceleration Violent Flow (AViF), and adopted support vector machine (SVM) and random forest to detect fight behavior. However, this method has a low recognition rate of the fight with insignificant movements such as punching etc. Yang et al. [17] proposed a real-time method based on the optical flow histograms, which calculates a scale descriptor and a rotation invariant feature descriptor histogram of optical flow orientation (HOFO) and uses the SVM classifier for detection. However, this method also has the problem of misjudgment in conversation. Considering the change of optical flow, Gao et al. [18] proposed the oriented violent flows (OViF) method for violence detection, which adopts AdaBoost and SVM classifiers for detection. However, this method performed poorly in crowded scenes. Recently, Mahmoodi et al. [19] proposed a new feature descriptor named Histogram of Optical flow Magnitude and Orientation (HOMO) to improve existing violence detection. Xu et al. [20] proposed a localization guided framework which exploits optical flow maps to extract motion activation information for detecting fight actions in surveillance videos. Febin et al. [21] proposed a cascaded method of violence detection based on motion boundary SIFT (MoBSIFT) and movement filtering for action recognition.

Approaches based on machine learning models show great performance in many scenes, but most of them do not consider the robustness in multiview which is essential for fight detection in real scenes. The existing algorithms present some limitations in solving misjudgment problems on running, overtaking, etc. which are similar to fighting in terms of motion characteristics sometimes. In terms of robustness on shooting view of videos and misjudgment on nonfight behaviors, we propose two descriptors, named Motion Direction Inconsistency (MoDI) and Weighted Motion Direction Inconsistency (WMoDI), by analyzing the motion characteristics of fight. In addition, a deep learning framework is used for motion region marking in this paper. We combine three existing descriptors and two novel descriptors, and the final feature with six statistical characteristics of these descriptors are calculated to improve the accuracy and robustness in multiview.

Experimental results on CASIA Action Dataset and UT-Interaction Dataset show that the proposed descriptors are discriminative enough to distinguish fight and nonfight behaviors in videos with different shooting views. In addition, the proposed method has certain performance advantages in accuracy and can recognize fight in different datasets effectively.

In summary, the main contributions of this paper are in the following aspects: (1) the robustness of fight behavior recognition from multiple video perspectives is considered in our proposed method. (2) The misjudgment rate of fight behavior detection can be reduced effectively by using MoDI and WMoDI. Part of this paper has appeared in [22]. As the extension of [22], we improve the method in motion regions marking and descriptors defining and redesign several substantial experiments to evaluate the performance of our method in different scenes.

The remainder of this paper is structured as follows. Section 2 discusses the fighting detection and describes the proposed method. Section 3 introduces the data sets used and provides experimental results. Finally, in Section 4, we summarize our paper and outline some promising future research directions.

2. Fighting Detection

As studies mentioned in Section 1, optical flow as a low-level image information can express the pattern of motion, and based on it, scholars proposed a series of descriptors for fight detection. Compared with nonfight behaviors, e.g., talking and walking, fight behaviors usually have the characteristics of suddenness, violence, and messy movements, which

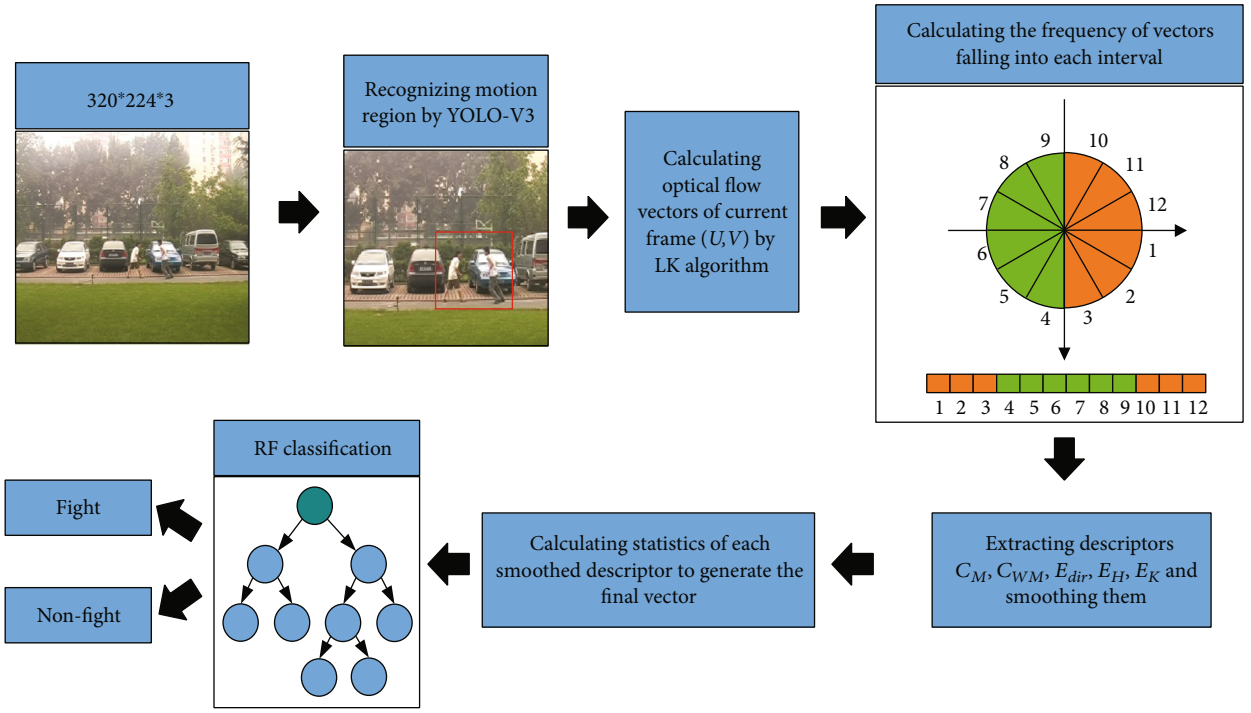


Figure 1: Overview of the proposed detection method.

means it is difficult to define the fight behavior with standard actions [12]. Actually, accelerated and sudden movement is often associated with states of high activation (happiness, anger) [23] which confirms that movement such as punches, kicks, or blows in a frame can be identified through suitable features. But the challenge of this work is to distinguish fight and other behaviors with high speed and sudden movement such as running.

To solve this challenge, we analyze the consistency and intensity of different motions in surveillance videos with different shooting views by calculating the displacement of each corner point. The result shows two typical distinctions. Compared with corner points in nonfighting (talking, walking, and running) videos whose inclination degree of displacement is consistent, corner points in fighting videos show chaotic direction changes. In addition, the velocity amplitude of corner points of nonfight tends to be stable while that of fight is larger and shows irregular changes. Therefore, velocity, acceleration and orientation are considered as the important factors for identifying fight behavior.

In our method, instead of recognizing the body structure and the movement of limbs, we adopt optical flow information to generate descriptors and Random Forest algorithm to identify the fight behavior. Figure 1 shows the overview of the proposed method. Firstly, the state-of-art deep learning model Yolo-V3 is adopted to mark the motion regions of passengers. Then, optical flow vectors are calculated, and a series of descriptors including two novel descriptors are extracted. Next, we smooth the descriptors, compute several statistical characteristics of them, and generate the final feature vector for training by concatenating the statistical characteristics. Finally, Random Forest algorithm is used to classify the frames.

2.1. Motion Regions Marking and Optical Flow Extraction. With the purpose of reducing the impact of noise points, the first step performed in the method is motion region marking. Xu et al. [20] used optical flow to extract motion activation boxes from consecutive frames. Recently, Zhang et al. [24] proposed an improved algorithm based on ViBe (visual background extraction) algorithm [25] to achieve moving target edge detection. In this paper, we employ You Only Look Once (YOLO) V3 [26], which is a deep learning framework based on Darknet-53 [27] and detects objects at three different scales, to mark the minimum circumrectangles of passengers. Figure 2 shows the architecture of it.

After obtaining the motion regions, the optical flow magnitude and orientation are estimated by Lucas-Kanade (LK) algorithm [28] which detects any rapid change of key points at the pixel level. By solving Equation (1), we can get the optical flow vectors of frame I.

$$\begin{bmatrix} I_{x1} & I_{y1} \\ I_{x2} & I_{y2} \\ \vdots & \vdots \\ I_{xn} & I_{yn} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_{t1} \\ I_{t2} \\ \vdots \\ I_{tn} \end{bmatrix}, \tag{1}$$

where I_{xi}, I_{yi}, and I_{ti} are partial differentials with respect to the position (x_i, y_i) and time t evaluated at key point i. The result (u, v) is shown in Equation (2).

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum I_{x_i}^2 & \sum I_{x_i} I_{y_i} \\ \sum I_{x_i} I_{y_i} & \sum I_{y_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum I_{x_i} I_{t_i} \\ -\sum I_{y_i} I_{t_i} \end{bmatrix}. \tag{2}$$

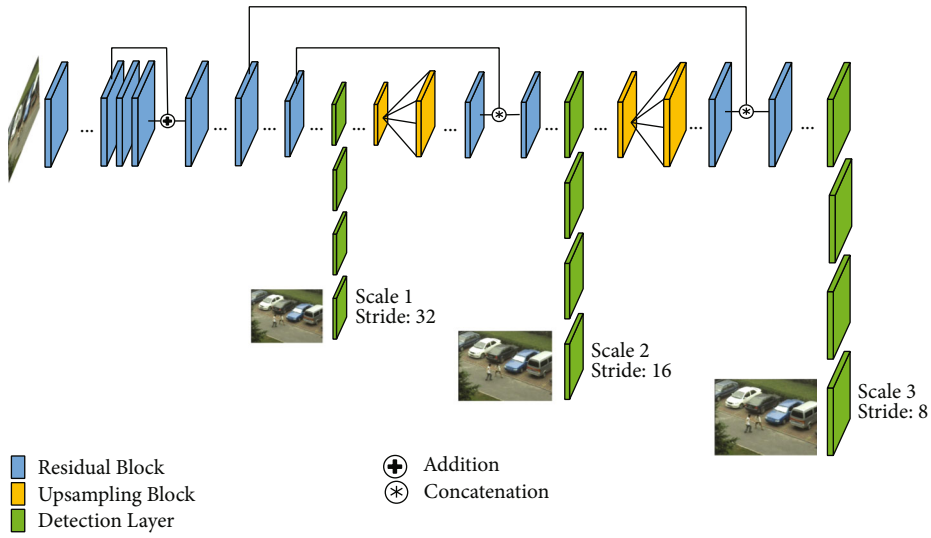


Figure 2: YOLO-V3 network architecture for motion regions detection.

Then, (u, v) is converted to the polar coordinate form as in Equations (3) and (4).

$$\rho_i = \sqrt{(u_i)^2 + (v_i)^2}, \tag{3}$$

$$\theta_i = \arctan \frac{v_i}{u_i}, \tag{4}$$

where (ρ_i, θ_i) represents the velocity magnitude and the orientation of vector i.

2.2. Feature Extraction. Considering input sequences captured from different views in real scenes, we evaluate the robustness of descriptors in multiview. According to Deniz et al. [29], when a sudden movement with high acceleration occurs, velocity magnitude of the key points in motion areas is usually larger than normal behaviors such as walking and talking. Based on that finding, we utilize optical flow to calculate two proposed and three existing descriptors [10–13], which are discriminative enough for fight behavior identification.

Based on the differences between fight and nonfight mentioned above, two novel descriptors are defined to emphasize the change in orientation and magnitude of velocity and ensure the multiview robustness. Next, we give the extraction process of two proposed descriptors.

Firstly, the polar coordinates are divided into 12 intervals as shown in Figure 1, where each interval's size is 30° [30]. The frequency and magnitude of the vectors falling into left intervals and right intervals can be calculated as follows.

$$\mathrm{N}_{\mathrm{left}} = \sum_{i=1}^{B} \delta(a(i) - j) j = 4, 5, 6, 7, 8, 9, \tag{5}$$

$$N_{\mathrm{right}} = \sum_{i=1}^{B} \delta(a(i) - j) j = 1, 2, 3, 10, 11, 12, \tag{6}$$

$$U_{\mathrm{left}} = \sum_{i=1}^{B} \rho_i \delta(a(i) - j) j = 4, \cdots, 9, \tag{7}$$

$$U_{\mathrm{right}} = \sum_{i=1}^{B} \rho_i \delta(a(i) - j) j = 1, 2, 3, 10, 11, 12, \tag{8}$$

where B represents the number of key points in motion region R, $a(i)$ is the angle interval in which the vector i falls, j represents the angle interval, and δ is the Kronecker delta function.

Then, the definitions of the proposed descriptors Motion Direction Inconsistency (MoDI) and Weighted Motion Direction Inconsistency (WMoDI) are given.

Definition 1. Given the frequency N_{left} and N_{right}, the MoDI of current frame, denoted by C_{M}, is defined as

$$C_{\mathrm{M}} = 1 - \frac{\left| N_{\mathrm{left}} - N_{\mathrm{right}} \right|}{N_{\mathrm{primary}}}, \tag{9}$$

where N_{primary} is the lager one in N_{left} and N_{right}.

Definition 2. Given the magnitude U_{left} and U_{right}, the WMoDI is defined as

$$C_{\mathrm{WM}} = 1 - \frac{\left| U_{\mathrm{left}} - U_{\mathrm{right}} \right|}{U_{\mathrm{primary}}}, \tag{10}$$

where U_{primary} is the lager one in U_{left} and U_{right}.

In addition to MoDI and WMoDI, three existing descriptors that are proved to be effective in detecting fight behavior are adopted in the method. Here, we give the definitions of them.

Definition 3. Given the normalized frequency N_{nor_j}, the Optical Flow Direction Entropy is defined as

$$E_{\text{dir}} = -\sum_{i=1}^{12} N_{\text{nor}_j} \log_2 N_{\text{nor}_j}. \tag{11}$$

Definition 4. Given the normalized magnitude H_{nor_j}, the Weighted Optical Flow Direction Entropy is defined as

$$E_H = -\sum_{i=1}^{12} H_{\text{nor}_j} \log_2 H_{\text{nor}_j}. \tag{12}$$

Definition 5. Given the velocity magnitude ρ_i and the number of corner points B, the Average Kinetic Energy is defined as

$$E_K = \frac{\sum_{i=1}^{B} \rho_i^2}{B}. \tag{13}$$

Normalized frequencies $N_{\text{nor}\,j}$ and $H_{\text{nor}\,j}$ are calculated as follows.

$$\begin{aligned} N_{\text{nor}_j} &= \frac{\sum_{i=1}^{B} \delta(a(i) - j)}{\sum_{j=1}^{12} \sum_{i=1}^{B} \delta(a(i) - j)}, \\ H_{\text{nor}_j} &= \frac{\sum_{i=1}^{B} \rho_i \delta(a(i) - j)}{\sum_{j=1}^{12} \sum_{i=1}^{B} \rho_i \delta(a(i) - j)}. \end{aligned} \tag{14}$$

E_{dir} is calculated based on the optical flow direction histogram, and the higher the value, the more chaotic the motion in current frame. Besides, E_H based on magnitude and orientation of optical flow vector considers the effect of amplitude energy. E_K considers the intensity of action in different behaviors.

Finally, the above descriptors are smoothed according to the continuity of video stream to enhance the regularity and reduce the impact of noise. The FPS of the input video is 25, and we accumulate the descriptors of 10 consecutive frames as the accumulated descriptors of the nth frame (as shown in Equation (15)).

$$\begin{cases} IC_{\text{M}}(n) = \sum_{n-9}^{n} C_{\text{M}}, \\ \text{IE}_{\text{H}}(\text{n}) = \sum_{\text{n}-4}^{\text{n}} \text{E}_{\text{H}}, \\ IE_{\text{dir}}(n) = \sum_{n-9}^{n} E_{\text{dir}}\, \text{IE}_{\text{dir}}(\text{n}) = \sum_{\text{n}-4}^{\text{n}} \text{E}_{\text{dir}}, \\ IE_H(n) = \sum_{n-9}^{n} E_H, \\ IE_K(n) = \sum_{n-9}^{n} E_K. \end{cases} \tag{15}$$

Table 1: Descriptors and statistical characteristics for final feature.

Descriptors	Statistical characteristics
MoDI (proposed in this paper)	Mean, variance, median, upper quartile, lower quartile, mean absolute deviation
WMoDI (proposed in this paper)	
Optical flow direction entropy	
Weighted optical flow direction entropy	
Average kinetic energy	

The last step of feature extraction is to calculate statistical characteristics of accumulated descriptors. Statistical characteristics are commonly used in data analysis. Inspired by the reference [15], we combine the statistical characteristics of accumulated descriptors to generate the final feature vector, and it is fed into the classifier for training. Table 1 gives the descriptors and statistical characteristics used in our method. And Algorithm 1 explains the feature extraction.

2.3. Classification Model. In our method, Random Forest (RF) algorithm is adopted as the classifier. The algorithm was first proposed by Breiman et al. in 2001 [31]. In this paper, bootstrap samples and Classification and Regression Tree (CART) algorithm are utilized to build a decision tree, which deploys the Gini Index, also named Gini impurity, to measure the quality of a split.

Next, we discuss the values of some hyperparameters of classifier in our method. Note that the classifier is trained in CASIA action database, which has been widely used in the research of behavior analysis [10–12, 32, 33]. This data set contains 432 videos, including 15 types of behavior, captured by 3 still cameras in angle, horizontal, and top-down view. And 105 of them are interactive videos, and the remaining 327 are single-person videos.

According to existing work [34], the accuracy of the classifier is mainly affected by the number of trees and the maximum depth of RF. Therefore, the hyperparameters are adjusted from both two aspects. And the accuracy of the classifier is evaluated by using 10-fold cross validation.

2.3.1. The Number of Trees. It is the number of decision trees contained in RF, and it has a great influence on the accuracy. By setting other hyperparameters of the model as fixed, we evaluate the accuracy of the model with different number of trees. Figure 3 shows the accuracy with respect to the number of trees. The x-axis represents the number of trees starting from 10 to 200, whereas the y-axis represents accuracy result for that specific parameter. The result shows that the accuracy improves with the increase of the number of trees, and then, it is gradually stable when the value of x-axis is more than 80. Moreover, the accuracy value reaches the maximum value when the value of x-axis is 110. To this end, the number of trees is set to 110 in the experiments of Section 3.

2.3.2. The Maximum Depth of RF. This parameter is adjusted under the condition that the number of trees is 110. Figure 4 shows the accuracy with respect to the maximum depth of RF, where the x-axis represents the maximum depth of the

Inputs: the video $S(x, y)$
Outputs: the final feature vectors
for each frame of $S(x, y)$ **do**
  Compute optical flow vectors (u, v) by Equation (2)
  Converted (u, v) to the polar coordinate form by Equations (3) and (4)
  Compute N_{left}, N_{right}, U_{left}, U_{right} by Equations (5), (6), (7), and (8)
  if $N_{\text{left}} > N_{\text{right}}$ **then**
    $N_{\text{primary}} = N_{\text{left}}$
  else
    $N_{\text{primary}} = N_{\text{right}}$
  if $U_{\text{left}} > U_{\text{right}}$ **then**
    $U_{\text{primary}} = U_{\text{left}}$
  else
    $U_{\text{primary}} = U_{\text{right}}$
  Compute MoDI C_{M} by Equation (9)
  Compute WMoDI C_{WM} by Equation (10)
  Compute descriptors E_{dir}, E_H, and E_K by Equations (11), (12), and (13)
  Obtain smoothed descriptors by (15)
end for
Calculate the statistical characteristics of smoothed descriptors
Concatenate the statistical characteristics to generate the final vectors

Algorithm 1: Feature Extraction algorithm.

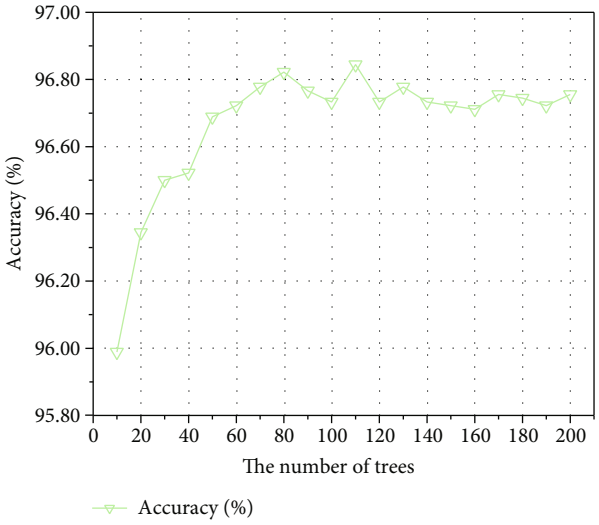


Figure 3: Accuracy with respect to the number of trees.

decision tree starting from 2 to 30, and the y-axis represents accuracy result for that specific parameter. The result shows that the accuracy increases with the increase of the maximum depth when the maximum depth is under 16. After the parameter reaches 16, the speed of accuracy converging slows down. And the accuracy reaches the maximum value when the parameter is 26. Therefore, 26 is selected as the maximum depth of RF in experiments of Section 3. Table 2 shows the hyperparameters of the RF.

3. Experiments

In this section, the data sets used in our experiments are introduced firstly. Then, in the experiments, we evaluate the performance of RF classifier. And the robustness of the proposed descriptors is verified by data visualization and ROC curves. Finally, the advantage of the proposed method is shown by comparing with some existing methods.

3.1. Datasets. CASIA Action Dataset and UT-Interaction Dataset [35] are used to evaluate the effectiveness and robustness of our method. From CASIA Action Dataset, all videos of fighting and 15 other videos in 5 categories are selected, and the details are shown in Table 3. With these videos, we verify the performance of the proposed method in different shooting views. Sample frames of fighting in three views of the dataset are shown in Figure 5.

In addition, UT-Interaction Dataset is selected as another dataset, which is taken by an outdoor still camera and contains 6 types of interactive videos. The videos of punching and kicking are used as samples. Figure 6 gives some sample frames.

3.2. Experimental Settings. We evaluate the performance of our classifier in terms of accuracy, missing alarm (MA), false alarm (FA), and F1-score and compare the RF classifier with other representative classifiers including support vector machine (SVM), adaptive enhancement (AdaBoost) and bagging guided aggregation (Bagging). The evaluation parameters are computed based four measures, named True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). Their definitions are given in Table 4.

The experimental results are shown in Table 5 and Figure 7. Note that the confusion matrices for each classifier are obtained from 300 frames sampled randomly from CASIA dataset. We can find that the RF outperforms other classifiers in terms of accuracy. In addition, the MA and FA of RF are suitable for fight detection in real scenes. Therefore, we adopt RF as classifier in our method.

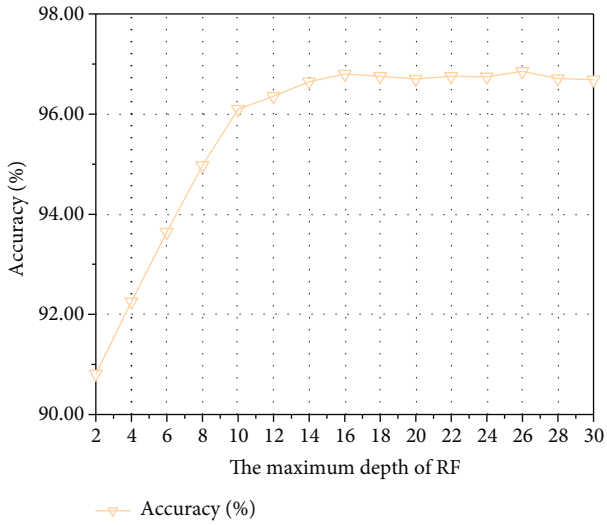


Figure 4: Accuracy with respect to the maximum depth of RF.

Table 2: Hyperparameters of RF.

Parameters	Number of trees	Maximum depth of RF
Value	110	26

Table 3: Detail information of CASIA Action Dataset.

Behavior categories	Number of videos	Number of frames
Fighting	$3 \times 4 = 12$	4580
Running	$3 \times 1 = 3$	160
Follow together	$3 \times 1 = 3$	757
Meet apart	$3 \times 1 = 3$	738
Meet together	$3 \times 1 = 3$	992
Overtaking	$3 \times 1 = 3$	694

Figure 5: Sample frames in three view of CASIA Action Dataset.

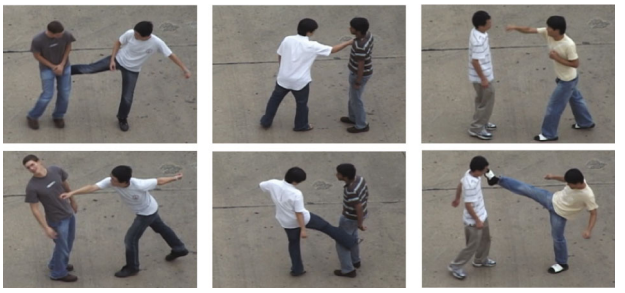


Figure 6: Sample "punching" and "kicking" frames of UT-Interaction Dataset.

Table 4: Evaluation parameters.

Parameters	Definition
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN} \times 100\%$
Missing alarm (MA)	$\frac{FN}{TP + FN} \times 100\%$
False alarm (FA)	$\frac{FP}{TN + FP} \times 100\%$
F1-score	$2 \times (P \times R/P + R)$, where $P = TP/TP + FP$, $R = TP/TP + FN$

Table 5: Comparison of classifier performance.

Classifiers	Accuracy	F1-score	MA	FA
SVM	92.00%	92.36%	11.58%	3.67%
AdaBoost	97.00%	96.95%	4.67%	1.33%
Bagging	92.67%	92.14%	1.40%	0.60%
RF	97.66%	97.66%	2.67%	2.00%

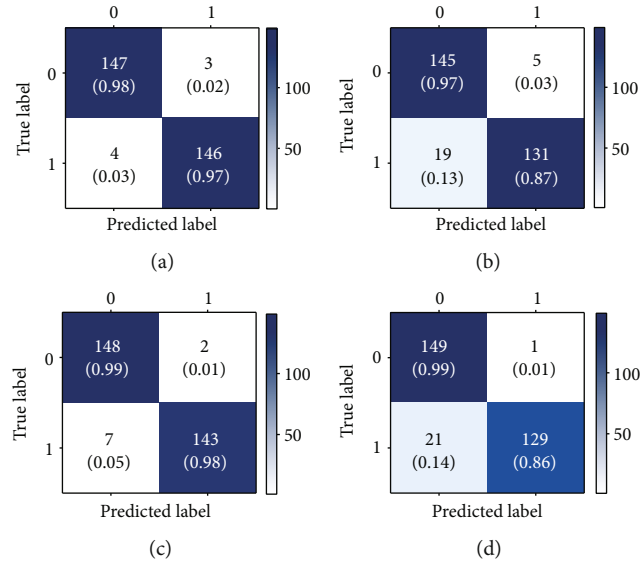


Figure 7: Confusion matrices for (a) RF, (b) SVM, (c) AdaBoost, and (d) bagging classifier (Label 1: fight (positive), label 0: nonfight (negative)).

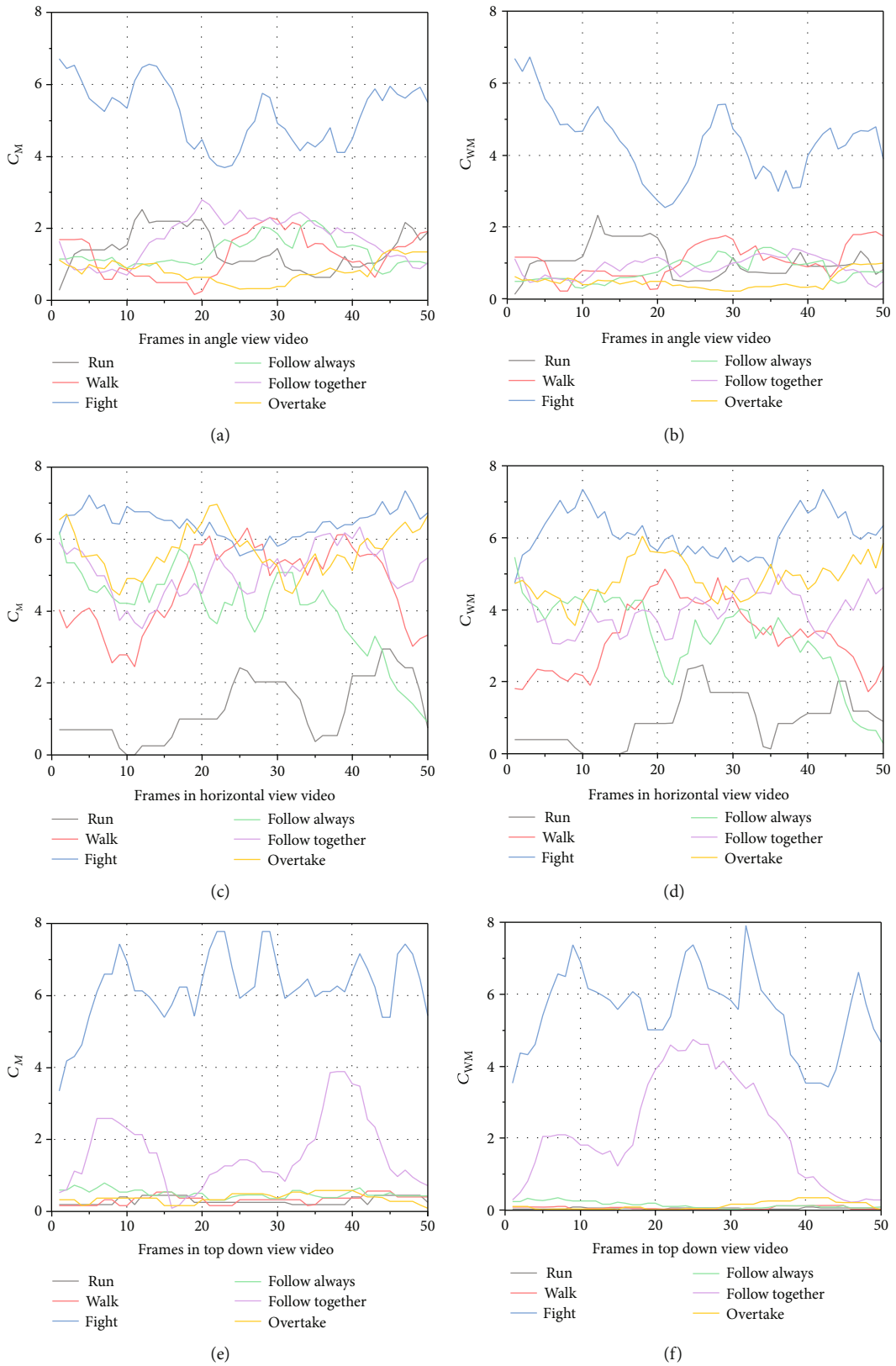


FIGURE 8: Proposed descriptors of behaviors in three views. (a) C_M of different behaviors in angle view. (b) C_{WM} of different behaviors in angle view. (c) C_M of different behaviors in horizontal view. (d) C_{WM} of different behaviors in horizontal view. (e) C_M of different behaviors in top down view. (f) C_{WM} of different behaviors in top-down view.

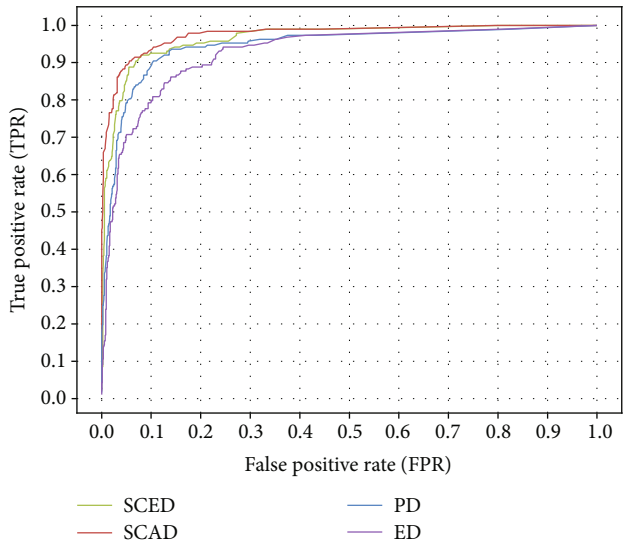


Figure 9: ROC curves of different descriptors. SCED: statistical characteristics of existing descriptors; SCAD: statistical characteristics of all descriptors; PD: proposed descriptors; ED: existing descriptors.

Table 6: Accuracy of proposed model and existing methods.

Method	Accuracy
Xia et al. [37]	86.4%
Wang et al. [38]	87.9%
MI_ULBP [39]	92.7%
Xu et al. [12]	92.79%
Our method	96.77%

Table 7: Detection results with UT-Interaction Dataset.

Action	Punch	Kick
Number of videos	10	10
Number of videos alarmed	9	9

Furthermore, 18 videos of six behaviors shot from 3 different views in CASIA data set are selected to analyze the robustness of the proposed descriptors. Figure 8 gives the quantitative results of descriptors of six behaviors over time. The results show that the descriptors of fight are certainly higher than that of others in most frames. To guarantee that the classifier can recognize fight effectively when the descriptors are not highest and improve the accuracy in frame level, we compute the statistical characteristics of descriptors, which represent the global changes of descriptors.

In order to evaluate the effectiveness of our proposed descriptors and the statistical characteristics used in our method, the Receiver Operating Characteristic (ROC) curves for each descriptor are illustrated in Figure 9. For the model with statistical characteristics of proposed and existing descriptors [10–13, 36], its Area Under Curve (AUC) is 0.976 that outperforms with other descriptors. The results show that statistical characteristics of descriptors are more discriminative than descriptors and can facilitate classifier distinguish fight and nonfight.

We evaluate the performance of our proposed model against the following state-of-art approaches:

(i) The method proposed by Xia et al. [37]: it is a multiview interactive based on self-similarity descriptors and graph shared multitask learning

(ii) The method proposed by Wang and Gao [38]: it is a deep learning based abnormal behavior recognition approach by utilizing interframe information and two-stream convolution neural network

(iii) MI_ULBP [39]: it combines moment invariants and uniform LBP feature and trains a binary support vector machine pattern classifier for the purpose of recognizing behaviors

(iv) The method proposed by Xu et al. [12]: it is a human abnormal behavior detection algorithm by combining the weighted reconstruction trajectory and histogram entropy

The accuracy of proposed method and existing methods is shown in Table 6. Our method gives the best result for the CASIA Action Dataset.

In addition, UT-Interaction Dataset is used to verify the applicability. Totally 20 videos of human-human interactions are selected, and half of them contain kicking occurring, and the rest contain punching occurring. Note that each video frames range from 80 to 120, and we suppose that there is fight occurring in the video when the frames predicted as positive are more than thirty percent of total frames. Table 7 shows the results, and it indicates the method proposed in this paper is also applicable to UT-Interaction Dataset and can effectively identify the fight behavior in videos.

In summary, the experimental results above show that the proposed method can effectively recognize fight behavior in videos with different shooting views. The reason mainly includes three aspects. Firstly, the feature vectors proposed in this paper have a high degree of distinction and multiview universality. Secondly, the smoothing process of extracted descriptors reduces the impact of noise and improves the robustness. Thirdly, as an efficient machine learning classifier, Random Forest shows the superiority for this task.

4. Conclusions

In this paper, in order to make the cyberphysical systems perceive the required environment information more effectively, we proposed an effective method for fight behavior detection which can reduce the misjudgment rate on nonfight behaviors and improve the robustness of detection with videos in different shooting views. Based on the analysis of motion characteristic, two novel descriptors MoDI and WMoDI are presented which are extracted from optical flow information. Instead of feeding the descriptors into classifier directly, 5 statistical characteristics are selected to generate the final vector. Experimental results demonstrate that the proposed

detection method performs well in videos with different shooting views, and it has the highest accuracy in CASIA dataset. As part of our future works, we will pay close attention to improve the early-warning capability and extend our method to poor light and even dark scenes.

Data Availability

You can get CASIA Action Dataset in website: http://www.cbsr.ia.ac.cn/english/action%20databases%20en.asp. And you can get UT-Interaction Dataset in website: https://cvrc.ece.utexas.edu/SDHA2010/Human_Interaction.html

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 (No. 2018YFC0807500), National Natural Science Foundation of China (No. 62072067), National Natural Science Foundation of China (No. 71904022), Social Science Foundation of Liaoning Province (No. L17CTQ002), and Fundamental Research Funds for the Central Universities (No. DUT21JC27).

References

[1] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: towards survivable cyber-physical systems," in *2008 The 28th International Conference on Distributed Computing Systems Workshops*, pp. 495–500, Beijing, China, June 2008.

[2] S. Ghazal, U. S. Khan, M. M. Saleem, N. Rashid, and J. Iqbal, "Human activity recognition using 2D skeleton data and supervised machine learning," *IET Image Processing*, vol. 13, no. 13, pp. 2572–2578, 2019.

[3] W. Ding, K. Liu, E. Belyaev, and F. Cheng, "Tensor-based linear dynamical systems for action recognition from 3D skeletons," *Pattern Recognition*, vol. 77, pp. 75–86, 2018.

[4] J. Dong, W. Jiang, Q. Huang, H. Bao, and X. Zhou, "Fast and robust multi-person 3D pose estimation from multiple views," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 921–929, Long Beach, CA, USA, June 2019.

[5] X. Wang, L. T. Yang, L. Song, H. Wang, L. Ren, and M. J. Deen, "A tensor-based multi-attributes visual feature recognition method for industrial intelligence," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2231–2241, 2021.

[6] M. L. T. L. Sandifort, J. Liu, S. Nishimura, and W. Hürst, "An entropy model for loiterer retrieval across multiple surveillance cameras," in *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*, pp. 309–317, Yokohama, Japan, June 2018.

[7] G. Lin, Y. Bai, and W. Zhang, "Recognition of abnormal interactions based on coupled hidden Markov models," *Journal of Southeast University (Natural Science Edition)*, vol. 43, no. 6, pp. 1217–1221, 2013.

[8] A. F. Villán, "Facial attributes recognition using computer vision to detect drowsiness and distraction in drivers," *Electronic Letters on Computer Vision and Image Analysis*, vol. 16, no. 2, p. 25, 2018.

[9] H. Chen, S. Ye, O. V. Nedzvedz, and S. V. Ablameyko, "Application of integral optical flow for determining crowd movement from video images obtained using video surveillance systems," *Journal of Applied Spectroscopy*, vol. 85, no. 1, pp. 126–133, 2018.

[10] Q. Lin and L. Zhang, "Crowed abnormal detection based on GLCM and optical flow," *Computer & Modernization*, vol. 24, no. 3, pp. 114–118, 2014.

[11] J. Du and L. Xu, "Abnormal behavior detection based on regional optical flow," *Journal of Zhe Jiang University (Engineering Science)*, vol. 45, no. 7, pp. 1161–1166, 2011.

[12] Z. Xu, Y. Luo, and P. Liu, "Abnormal behavior detection of joint weighted reconstruction trajectory and histogram entropy," *CAAI Transactions on Intelligent Systems*, vol. 13, no. 6, pp. 1015–1026, 2018.

[13] C. Luo, X. Li, and H. Li, "Research on human abnormal behavior detection based on optical flow energy," *Computer and Network*, vol. 40, no. 21, pp. 71–73, 2014.

[14] J. Ren, C. Lin, Q. Liu, M. S. Obaidat, G. Wu, and G. Tan, "Broadcast tree construction framework in tactile internet via dynamic algorithm," *Journal of Systems & Software*, vol. 136, pp. 59–73, 2018.

[15] J. Huang and S. Chen, "Detection of violent crowd behavior based on statistical characteristics of the optical flow," in *2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 565–569, Xiamen, China, August 2014.

[16] C. Wang and S. Fei, "Violence detection in crowded scene," *Industrial Control Computer*, vol. 33, no. 2, pp. 105-106, 2020.

[17] Z. Yang, T. Zhang, J. Yang, Q. Wu, and L. Yao, "Violence detection based on histogram of optical flow orientation," in *International Conference on Machine Vision*, p. 906718, London, UK, 2013.

[18] Y. Gao, H. Liu, X. Sun, C. Wang, and Y. Liu, "Violence detection using oriented violent flows," *Image & Vision Computing*, vol. 48-49, pp. 37–41, 2016.

[19] J. Mahmoodi and A. Salajeghe, "A classification method based on optical flow for violence detection," *Expert Systems with Applications*, vol. 127, pp. 121–127, 2019.

[20] Q. Xu, J. See, and W. Lin, "Localization guided fight action detection in surveillance videos," in *2019 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 568–573, Shanghai, China, July 2019.

[21] I. P. Febin, K. Jayasree, and P. T. Joy, "Violence detection in videos for an intelligent surveillance system using MoBSIFT and movement filtering algorithm," *Pattern Analysis and Applications*, vol. 23, no. 2, pp. 611–623, 2019.

[22] X. Wang, C. Yao, X. Su, J. Dong, and Y. Li, "Random forest based multi-view fighting detection with direction consistency feature extraction," in *2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (Green Com) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (Smart Data) and IEEE Congress on Cybermatics (Cybermatics)*, pp. 558–563, Rhodes, Greece, November 2020.

[23] F. A. Pujol, H. Mora, and M. L. Pertegal, "A soft computing approach to violence detection in social media for smart cities," *Soft Computing*, vol. 24, no. 15, pp. 11007–11017, 2019.

[24] H. Zhang, Y. Qian, Y. Wang, R. Chen, and C. Tian, "A ViBe based moving targets edge detection algorithm and its parallel implementation," *International Journal of Parallel Programming*, vol. 48, no. 5, pp. 890–908, 2019.

[25] O. Barnich and M. van Droogenbroeck, "ViBe: a universal background subtraction algorithm for video sequences," *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1709–1724, 2011.

[26] J. Redmon and A. Farhadi, "YOLOv3: an incremental improvement," 2018, http://arxiv.org/abs/1804.02767.

[27] J. Redmon, *Darknet: Open Source Neural Networks in C*, 2013, http://pjreddie.com/darknet/.

[28] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, Vancouver, BC, USA, August 1981.

[29] O. Déniz, I. Serrano, G. Bueno, and T. Kim, "Fast violence detection in video," in *International Conference on Computer Vision Theory & Applications*, Lisbon, Portugal, January 2015.

[30] W. T. Freeman, "Orientation histograms for hand gesture recognition," in *International Workshop on Automatic Face and Gesture Recognition*, pp. 296–301, Zurich, Switzerland, 1995.

[31] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[32] L. Ren, Z. Meng, X. Wang, L. Zhang, and L. T. Yang, "A data-driven approach of product quality prediction for complex production systems," *IEEE Transactions on Industrial Informatics*, p. 1, 2020.

[33] H. Zhu, T. Xie, Y. Guan, C. Deng, and X. Wang, "Hypergraph matching with an entropy barrier function," *IEEE Access*, vol. 7, pp. 16638–16647, 2019.

[34] C. Hu, Y. Chen, L. Hu, and X. Peng, "A novel random forests based class incremental learning method for activity recognition," *Pattern Recognition*, vol. 78, pp. 277–290, 2018.

[35] M. S. Ryoo and J. K. Aggarwal, *UT-Interaction Dataset*, 2010, http://cvrc.ece.utexas.edu/SDHA2010/Human_Interaction.html.

[36] J. Ren, G. Wu, X. Su, G. Cui, F. Xia, and M. S. Obaidat, "Learning automata-based data aggregation tree construction framework for cyber-physical systems," *IEEE Systems Journal*, vol. 12, no. 2, pp. 1467–1479, 2018.

[37] L. Xia, W. Guo, and H. Wang, "Interaction behavior recognition from multiple views," *Journal of Central South University*, vol. 27, no. 1, pp. 101–113, 2020.

[38] Z. Wang and B. Gao, "Spatio-temporal fusion convolutional neural network for abnormal behavior recognition," *Computer Engineering and Design*, vol. 41, no. 7, pp. 2052–2056, 2020.

[39] S. Nigam and A. Khare, "Integration of moment invariants and uniform local binary patterns for human activity recognition in video sequences," *Multimedia Tools and Applications*, vol. 75, no. 24, pp. 17303–17332, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5589505, 10 pages
https://doi.org/10.1155/2021/5589505

Research Article

Optimization of the Personalized Service System of University Library Based on Internet of Things Technology

Yi Zhuang

Library, Shandong Women's University, Jinan 250300, China

Correspondence should be addressed to Yi Zhuang; 22012@sdwu.edu.cn

Received 1 March 2021; Revised 16 April 2021; Accepted 26 April 2021; Published 12 May 2021

Academic Editor: Wei Wang

In library applications, radio frequency indentification (RFID) technology, sensors, and wireless transmission networks have been applied to various services such as self-service checkout and return systems, electronic reader cards, intelligent bookshelves, intelligent monitoring of library premises, augmented reality (AR) interactive picture books, physical corridors, and seat reservations; in regional library alliances, real crossregional and cross-system alliance cooperation through IoT technology is also becoming increasingly important. Continuous information resource sharing is an important means to maximize the effectiveness of library information resources and meet the information needs of various users. The development of IoT technology opens new ideas and methods for information resource sharing in regional library alliances, effectively expanding the scope of information resource sharing and improving the efficiency of information resource sharing. This paper briefly presents the relationship, architecture, and key technologies of IoT technology and the definition, characteristics, and types of regional library consortium and content. Analysis of the characteristics and principles of regional library consortium information resource sharing is in the context of IoT and the corresponding studies on information sharing between regional library consortia at home and abroad. We also propose strategies to establish a specialized agency for information resource sharing, establish a sound investment mechanism for information resource sharing, ensure the security of information resource sharing of the regional library consortium, and increase the publicity and training capacity of information resource sharing of the regional library consortium.

1. Introduction

With the rapid development of the Internet, the earth has been transformed into a global village. Information technology has penetrated into all walks of life and all fields of the world [1]. People cannot live and work without the Internet. However, under Moore's law, i.e., the information society will produce a revolution every fifteen years. The development of the Internet has hit a bottleneck and is suffering from various challenges while facing many opportunities. The Internet of Things (IoT) is an "internet of things" based on the traditional Internet, using technologies such as RFID and wireless sensor networks to build an "internet of things" that connects things and senses each other [2]. The birth of the Internet of Things is the inevitable product and wisdom crystallization of the development of the Internet to a certain stage of *W* and economic and social development and is a new technological revolution in the field of information technology. With the progress of the times, the number of books is growing exponentially. At the same time, with the improvement of people's ideological awareness, the borrowing volume of books is also increasing [3]. As an important carrier of books, the difficulty of book management in libraries has increased, and the traditional manual management method is no longer suitable for the needs of modern book management. How to improve the accuracy, security, and efficiency of book management at a lower cost is the most urgent task of many university libraries. Books are important in library book management because they are valuable, easily damaged, not lost, and not renewable [4].

After IBM put forward the concept of "Smart Earth" in 2008, smart transportation, smart city, smart medical care, smart education, and other related topics have emerged and received wide attention. In recent years, the concepts of

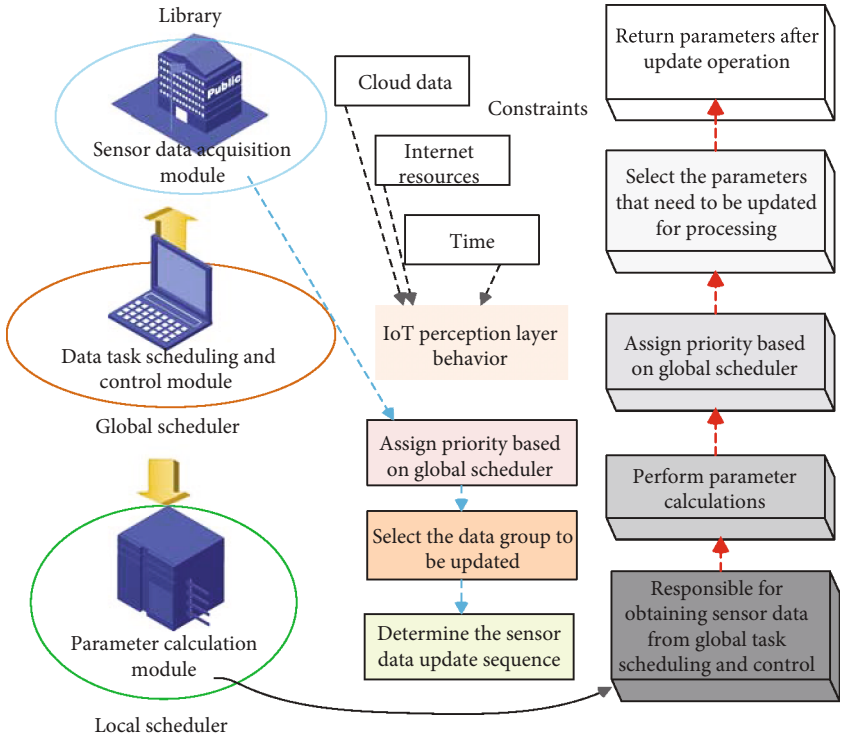


Figure 1: Intelligent library management.

"smart campus" and "smart learning" in domestic universities have also been proposed, and the research of smart library, as an integral part of smart campus, has become more and more the focus of modern library research, as shown in Figure 1. Under the era of wisdom, the essential characteristics of public welfare services cannot be changed, and the position of its main functions and related functions cannot be reversed [5]. However, university libraries must highlight "wisdom" in their future construction goals and services. With the deepening of Internet, cloud computing, and big data technology, libraries are developing in the direction of wisdom libraries. The construction and innovation of university wisdom libraries will inject new vitality, provide new ideas and expand new space for the new period of university libraries, and provide practical experience and models that can be referred to for the role innovation and function transformation of university libraries [6]. In the background of the intelligent era, the proposal of "intelligent library" is promoting the transformation of university libraries from traditional lending to intelligent service, from paper-based to digital resources-based, from centralized to distributed, from passive to active, from common to individual, and from stereotyped manual to flexible and intelligent, and the service of university libraries has gradually realized the transition from traditional type to digital and mobile libraries. The services of university libraries have gradually realized the leap from traditional libraries to digital and mobile libraries [7]. Knowledge diffusion has put forward higher-level requirements on the way of serving information resources in university libraries, and the needs of readers and users have shifted from simply searching for information resources to more efficient access to information resources that meet their diverse new needs [8–11]. In the age of the Internet, each information user is both a creator and a provider of information, and saving time and cost is also a service that patrons urgently need. Smart libraries aim to help every reader user find the library information resources they want quickly, efficiently, and accurately, so that users can take what they need, and libraries can take their responsibilities, optimize the work of the entire library system, effectively shorten the time of information processing, improve the diversity of information resources, and realize smart services [12].

The future development goal and need of university libraries are the integration of digital resources and integration of user services, providing readers and users with a convenient, fast, and prudent way of service delivery. From the traditional physical library to the digital library, the digital library to the complex library, and then to a higher-level intelligent library, no matter how technology evolves and changes, the library principles of "people-oriented," "reading-centered," and "service-oriented" still apply. Regardless of how technology advances and changes, the basic concepts of "people-oriented," "reader-centered," and "service" libraries remain unchanged. Therefore, it has been the responsibility of libraries to change from traditional "passive stereotypical service" to active "active intelligent service." This paper introduces the concepts and key supporting technologies of information construction of the recommendation system of the wisdom library, including wisdom library, recommendation system, and mainstream algorithm of the recommendation system, briefly compares and analyzes the advantages and disadvantages of common recommendation algorithms of the recommendation system, and finally elaborates the wisdom library service model based on the wisdom

recommendation system, which provides the theoretical basis for the construction of the wisdom library service model. Finally, we provide the theoretical basis for the construction of the service model of smart library.

2. Related Work

Scholars have mainly studied the definition, characteristics, service principles, service roles, service means, service targets, or service evaluation of smart libraries. At the conference on "Human-Computer Interaction Mobile Devices" in 2003, Aittola of Oulu University Library in Finland proposed that "Smart Library" refers to a mobile library that can break the limitations of location and space and be perceived by people and actively find the users' needs for them. Aittola, from the University of Oulu Library in Finland, suggested that a "Smart Library" is a mobile library that can break the limitations of location and space and be perceived by people, actively finding the library information resources for the users. In February 2010, the Smart Libraries Newsletter, a publication of the American Library Association (ALA), was launched with the title "Smart Libraries Through Technology." In February 2010, the Smart Libraries Newsletter, a publication of the American Library Association (ALA), featured a column entitled "Smart Libraries Through Technology", which was devoted to the latest information and progress in the development of smart library technology. The researchers examine the status and value of intelligent agent technology in the context of the goal of smart libraries. The researchers envision a model for smart libraries that can maximize the effectiveness of library services by supporting the integration and analysis of data and information resources. The researcher believes that in the context of the information age, the ability to adapt to the changing environment and thus provide the services required by the general public is the primary and core value for the existence and survival of libraries. Scholars have used questionnaires and other methods to study the digital resource use and learning behavior of library patrons, suggesting the strengthening of cooperative and informal library functions and summarizing and analyzing the smart library information service model from the perspective of patron users [13].

Researchers have proposed that smart libraries should have two service models, namely, the market model and the evaluation model. The marketplace model has obviousness selection characteristics and is based on effective communication between librarians and patrons; the assessment model has potential evaluation characteristics and is a description of the current state of patrons and prediction of their future choices [14]. The researchers explored various recommender systems in terms of characteristics and technical potential and described the application of various predictive tools in recommender technology as a guide for research and practice in the field of recommendation. Researchers proposed new algorithms for recommender systems based on time-decaying weights and matrix clustering that can better solve the data sparsity problem and better respond to the variable interests of users [15]. Researchers, on the other hand, have opened up new perspectives in the recommendation system development research, supporting the inclusion of user participation in the development and design process of recommendation systems to better tap into users' unique preferences. In recent years, foreign scholars have been studying library consortia since the 1970s, when they first mentioned the definition of "library consortia" and analyzed the development of university libraries in the United States, emphasizing the role of library consortia in library development. The previous results focus on the creation and maintenance of information resources by various institutions on the Internet of Things, and through interregional and intersectoral links, truly realize the common construction and distribution of resources, and provide readers with rapid and convenient access to digital resources.

Researchers mentioned that a library consortium is a collaboration of two and more libraries, a new consortium bound by a mutually agreed agreement for the purpose of sharing resources, reducing costs, and mutual benefits. Based on the adherence to the agreements among the alliance members formalizing the organization of the alliance at present, library association organizations have violated the tangible boundaries of libraries in the past, and the hierarchical distinction of libraries is not very different, and the members of alliance libraries enjoy rights and obligations equally. With each library's collection with its own characteristics and focus, each member library has a relatively deep understanding of common construction and sharing, for example, the United States has solved the logistics and distribution problems of interlibrary loan in the cooperation process of library union and truly realized the deep integration of resource sharing. Foreign research on information resource sharing has focused on library cooperation, library collections, library networks, and library consortia [16].

The Duchy Library of Weimar in Germany established an interlibrary loan relationship with the University Library of Jena in the late 18th century, which was one of the earliest models for libraries to share information resources. First proposed by Melvil Dewey in 1886, library cooperation, also known as interlibrary cooperation, has emerged as a library resource sharing activity and is often found in library-related literature and library conferences. In order to achieve resource sharing, member libraries should develop collections among themselves and develop various services based on the collections, and cooperative collections should be a way to reduce duplicate acquisitions and achieve economical acquisitions. Using interviews and document delivery tests, the researchers interviewed librarians in 15 university libraries to investigate their perceptions of inappropriate services. It was concluded that resource sharing would not operate in an equitable manner if the majority of users' requests were directed to only a small portion of the university libraries in Taiwan [17]. Inadequate library collections in any region or country will not allow resource sharing. The researchers pointed out the participation of the library of the Institute of Advertising and Marketing of Rio de Janeiro and the library of Petro Acutepolis, a member country of the CBIES/RJ Resource Sharing Association of the State of Rio de Janeiro, Brazil, in the consortium. The main results are scientific and technical knowledge sharing, collaborative

access to software and databases, development of an online collaborative catalog series, and information retrieval and access facilities. This experience has shown that library consortia are effective tools for information resource sharing and contribute to the quality of academic education.

3. Basic Theory of Library Information Resource Sharing in the Context of Internet of Things

3.1. Internet of Things Technologies. The huge increase in online information has led users into a sea of information explosion, and data overload has caused information scarcity, resulting in the inability of users to obtain information that is truly useful to them. The Internet of Things is a series of science and technology, using sensor devices to connect objects and computer terminals, through computer control to achieve automatic and intelligent objects, and to achieve the goods and goods information by the Internet to pass and express, is the Internet outward extension and expansion of the network, and is the user side of the exchange of information and communication between goods and good network. The architecture of IoT can be divided into three parts: sensing layer, network layer, and application layer, as shown in Figure 2.

The sensing layer is mainly sensed by various sensors and then transformed into signals to be transmitted, such as temperature and humidity sensors, position sensors, and acoustic sensors. The sensing layer is mainly used to collect the physical world and information, including physical quantities, signs, audio, and video. The role of the perception layer is equivalent to the role of human sensory organs to the human body system, by collecting information from the perceived external environment to identify objects and perceive physical-related information. The network layer is the communication module that transmits signals from the conversion of the sensory layer to each other through Internet technology, which mainly refers to the current communication technology and terminal technology that provides communication capability for terminals in various industries. The network layer is equivalent to the human nervous system, capable of transmitting the sensed information without barriers, with high reliability and security, and requires the mutual integration of sensor networks and mobile communication technologies and Internet technologies.

M2M is the transfer of data transmitted by sensors from one end machine to another, i.e., enabling machine-to-machine conversations in the Internet of Things. M2M, one of the core technologies of the Internet of Things, is now a prevalent form of IoT application that enables all machines and devices to be equipped with communication and networking capabilities, allowing for a truly seamless interface between machines and people and systems. Cloud computing technology is an Internet-based computing method through which some shared related hardware resources and network information technology are made available to other computers according to demand and is a product of the convergence of computer development needs. It is the core of the realization of the Internet of Things, using the cloud computing model, which provides a dynamic scalable virtualized resources of the computing model. It also has super storage capacity, which is also equivalent to the central nerve of the "brain" of IoT, with computing and storage capacity. IoT sensors and Internet operators interact with information through network lines and computer terminals to provide data to the cloud and receive technical services from the cloud.

3.2. Regional Library Service System Based on Internet of Things. Library consortia have been a hot topic of research in library science at home and abroad, reaching a peak in the 1990s and gradually stabilizing. The widely accepted definition is that a library consortium is a consortium of libraries in a certain range organized for the purpose of resource sharing and mutual benefit and governed by mutually agreed agreements and contracts. The regional library consortium is not only limited to the cooperation between libraries but also the cooperation with all sectors of the society within the geographical area that can be called a consortium, which is based on a mutually agreed agreement or contract, and is a cooperation subject to mutual constraints. This kind of cooperation can be interlibrary cooperation, digital resource sharing, library book loan and return, cooperation in purchasing databases, joint development of special databases, lecture alliance, training alliance, publicity activities alliance, etc., which are important initiatives for the common development of libraries and all sectors. The early library alliances mainly cooperated with libraries of the same system, and the content of the alliances was also based on basic business, with little communication with the outside world, little influence in the society, and low social status. Nowadays, the alliance tends to be more and more a big partnership across systems, industries, regions, and disciplines and is oriented to the society, with links and cooperation between groups such as government departments, research institutes, commercial companies, educational institutions, and new media industries, as shown in Figure 3. This alliance is the integration of libraries with talents, technologies, and social resources from all walks of life under IoT technology, which can effectively improve user satisfaction, social influence, and social status.

The early library alliances were mostly cooperative in terms of joint cataloging of literature, interlibrary loan, and coordination of book division in the collection. Modern library alliances under Internet and IoT technologies, on the basis of the original cooperation, also cooperate in areas such as librarian training, technical seminars, expert lectures, celebrity painting and calligraphy exhibitions, reference consultation, resource procurement, digital resource coconstruction and sharing, and smart library construction. Using more institutions, the form results focused much on the controlling of the information and crossconstruction. However, a hierarchical distinction of libraries has not yet been applied in the libraries. Based on this consideration, members should consider much on the digital resources compliance with the agreements among alliance members. Library consortium is an alliance between libraries and other groups based on

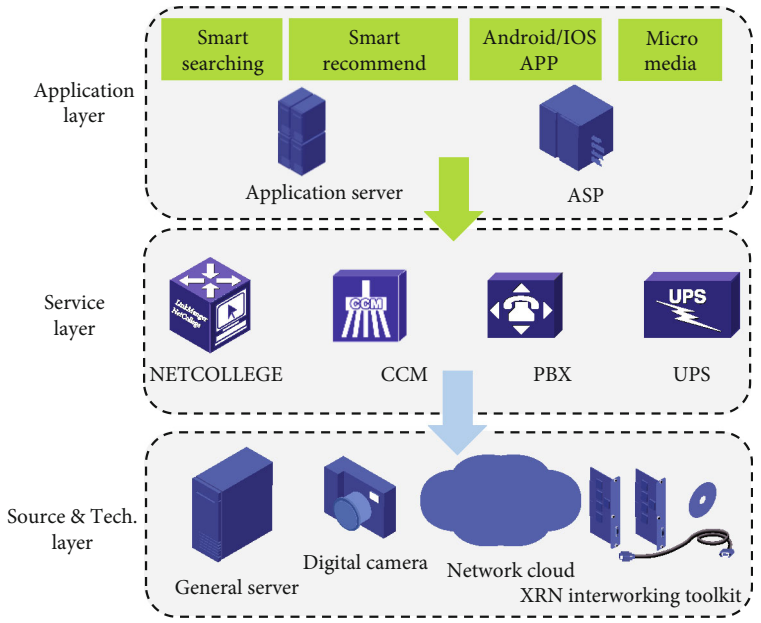


Figure 2: Internet of things architecture.

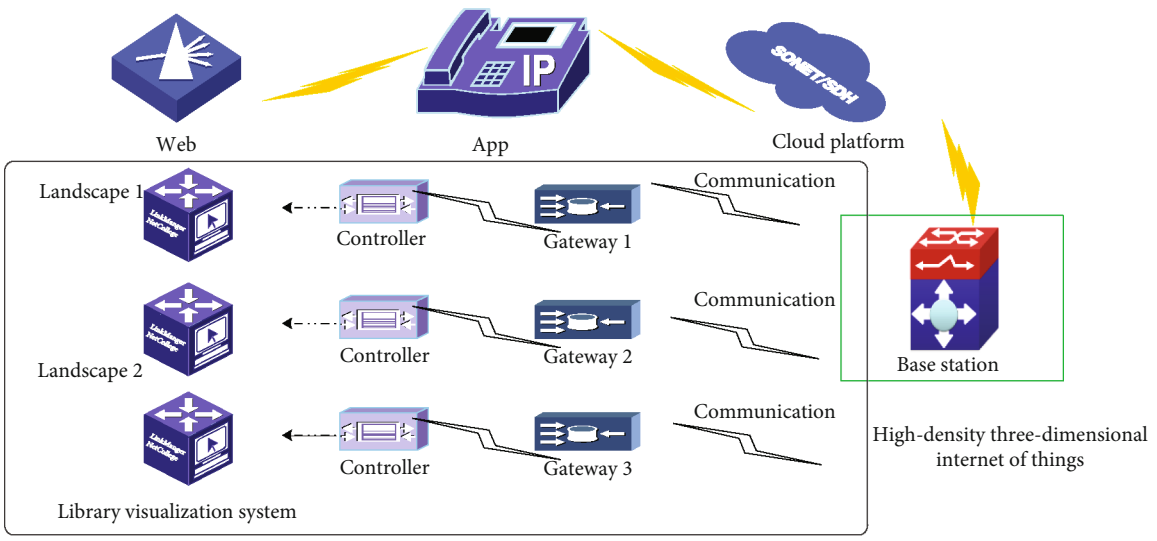


Figure 3: Regional library consortium service system.

cooperative relationship, and all member libraries join voluntarily and are established with the will of common building and sharing, mutual benefit, and mutual assistance. At the beginning of establishment, standardized constitution and system will be formulated through common discussion, and fixed full-time staff will be established for management, and if any unit fails to comply with or violates the rules and regulations, it will be punished or expelled. If flexible management is required for work in the operation process, member libraries should also actively cooperate in the management and communication work such as institutional integration, benefit and risk distribution, and business process reorganization. There is a wide variety of types of library alliances, and there is no more standardized and unified criteria to compare from some aspects that are to divide the same kinds of alliances into one type. Each type of alliance is an alliance formed by member libraries on the premise of three major elements: a foundation of trust, a consensus of culture, and an effective monitoring system for the purpose of sharing resources, reciprocal benefits, and promoting the overall development of libraries. The United States is the fastest growing country in terms of the number of library consortia and the number of types of development crisscrossing the country.

For each reader u belonging to the nearest candidate neighbor Mc, the set of books borrowed by the reader and the target reader in a certain statistical time period is found from the OPAC system library and is represented by the set U_n. Based on Mc and U_n, a user-item matrix is created, denoted by $A(m, n)$, which contains the set $V = (u_1, u_2, \cdots u_m)$ of m target readers and the set $I = (i_1, i_2, \cdots \cdots i_n)$ of n book information resources, in). The rating of reader u for

book i is represented by the matrix element R_{ui}. If reader u has no borrowing behavior for book i within a certain time statistic, $R_{ui} = 0$. If reader u has borrowing behavior within the statistic time period, the value of R_{ui} is determined based on the borrowing time because the length of borrowing time reflects the long-term preference of readers to a certain extent, which is calculated as follows: let the sequence of books borrowed by reader u within a certain statistic time be (item 1, item 2, ..., item i, ..., item n), the corresponding borrowing time sequence is (t_1, t_2, ... t_i, ..., t_n), and the starting time of the statistical period is time which is Tmin, the cut-off time is Tmax, and then the reader u for the book item $i$is the rating value calculation, such as the formula

$$Q_o = \frac{\sum_{i=0}^{n} t_i + T_{\min}}{T_{\max}}. \tag{1}$$

Based on the user-item matrix, the top-N readers who are the closest and most similar to the target readers are calculated as the set of nearest neighbors of the target readers by modifying the formula of cosine similarity. The calculation formula is as follows:

$$\max\left(Q_o, Qi\right) = \frac{\sum_{i=0}^{n}\sum_{i=0}^{n}\left(t_i + T_{\min}\right)_{ij}}{\sqrt{\left(t_i + T_{\max}\right)_{ij}}}, \tag{2}$$

where max (Q_o, Q_i) denotes the similarity between reader u_1 and reader u_2 tables the books jointly rated by reader u_1 and u_2,$R_{u1,c}$, denotes the rating of reader u_1 on book c, and u_2 denotes the average rating of reader u_1 and reader u_2 on the books.

Based on the rating data of book resources by the nearest neighbors of the target readers, the interest preference of the target readers for books $P_{u,i}$ can be predicted. Then, by setting the number of recommended books, or the threshold of readers' interest preference, book resource recommendations are generated for the target readers. Using this formula based on the weighted average of item means, the rating prediction of book resources is made within the nearest neighbor set of the target readers. Let the nearest neighbor set be U and be the interest preference of $P_{u,i}$ target reader u for book i. Then, $P_{u,i}$ can be calculated using the rating data of the target reader's nearest neighbor set for book i. The formula is as follows.

$$Q_{n,i} = \frac{R_q}{\sum_{j=0}^{m}\sum_{i=0}^{n}\mathrm{siru}\left(n_i, n_j\right)}. \tag{3}$$

3.3. Personalized Service Optimization in the Context of IoT. The intelligent library service model based on the recommendation system studied in this paper is centered on serving readers and users and focuses on relying on information technology to provide personalized and intelligent services for libraries. The overall block diagram of the intelligent library service model for universities is shown in Figure 4. Intelligent search is the premise and foundation of intelligent recommendation. Retrieval is the first step for users to obtain information, and intelligent retrieval in university libraries aims to enable readers to retrieve the required information more quickly, accurately, efficiently, and in an orderly manner. Compared with traditional retrieval, intelligent retrieval can record and analyze users' retrieval behavior, identify users' clear or potential demand preferences, and present users with the most relevant retrieval results; on the other hand, it can automatically correct retrieval strategies through users' feedback or evaluation of retrieval results, so that users can get the resource information closest to their needs and improve users' intelligent experience in retrieval. While traditional information retrieval is based on keywords or similarity-based retrieval, the basic concept of wisdom retrieval is to realize the established needs of established users. On the basis of traditional search, smart search intelligently filters out part of the information that may be invalid for users and helps readers locate the resource information they need more quickly and precisely. Not every user logs in to reader services every time they visit a library website, and most users simply try to quickly retrieve the books they need. Therefore, many users usually prefer the most time- and energy-efficient behavior and method to achieve their search purpose. Through intelligent search, it can make the search results more targeted and contain links directly to the final result page, thus not only can help users avoid the blindness of book search, shorten the time of book search and reduce the difficulty of search but also improve the efficiency of user search and enhance the accuracy and reliability of book search results.

Based on its own specificity and research needs, this paper synthesizes the survey results and classifies the recommendation services of university libraries into two categories, namely, personalized and nonpersonalized recommendation services. Personalized recommendation refers to the ability to provide users with personalized and intelligent recommendations by analyzing and mining user behavior information according to certain techniques and methods based on the characteristics and needs of different users. It can predict users' preferences based on their characteristics and needs. Such as guess your favorite recommendation services, the recommended content has special, targeted, predictive. Intuitively, from the perspective of recommendation results, different users get different recommendation results (which can be similar). Nonpersonalized recommendations can only meet the general needs of general users, but not for readers with different backgrounds, different purposes, and different time periods. Intuitively, from the viewpoint of recommendation results, different users get the same and nondifferentiated recommendation results.

Nonpersonalized recommendations are usually in the form of a simple page, presenting all the recommended information directly to all users, without targeting, such as new books to the library, lending ranking and other recommendations, and the recommended content is universal and general. Nonpersonalized recommendation is also subdivided into dynamic nonpersonalized recommendation and static nonpersonalized recommendation (hereinafter, referred to as

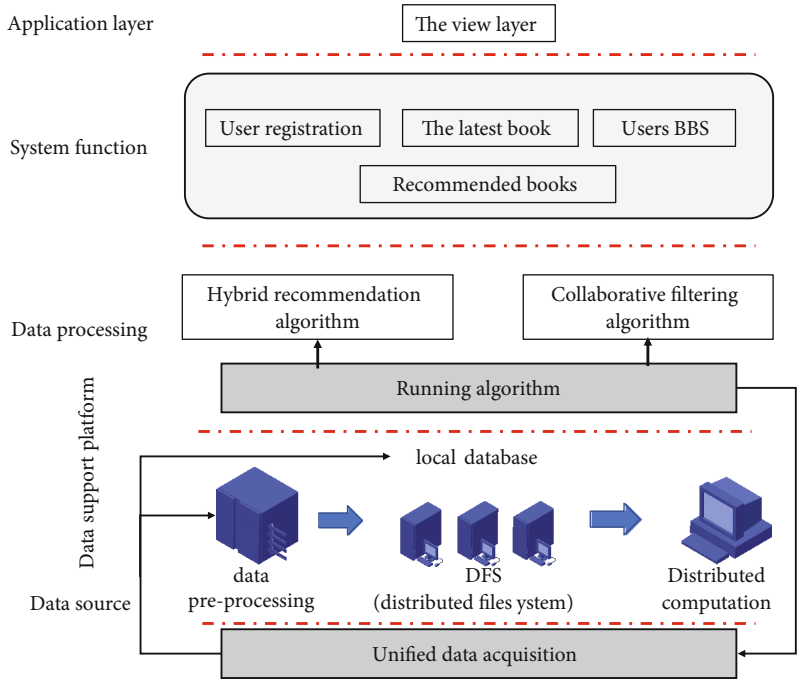


Figure 4: Overall block diagram of the personalized service model.

dynamic recommendation and static recommendation). Dynamic recommendation refers to the recommendation that different users get the same recommendation, but need to be realized by certain technology or algorithm, or a threshold as a scale, or a time period as a boundary, to dynamically display the recommended content, such as popular lending, lending ranking, and other recommendations. Static recommendation means that the librarians present the recommendation results to the users in a static page through the editing operation of the backend website. For example, the new book notice is presented to all readers in a simple static page, while the popular borrowing is based on certain dynamic data through statistical analysis and other techniques to come up with the recommendation results.

Of the 109 university libraries surveyed, all libraries offered some portion of the nonpersonalized service, as shown in Figure 5, representing 100% of the libraries surveyed. Sixty-nine libraries, or 63% of the libraries surveyed, provide a portion of the personalized recommendation service. Among the static recommendations, about 96% of the libraries provide new book announcement recommendation service, which shows that universities pay more attention to publicizing and promoting new books that come to their libraries. About 28% of the libraries in the surveyed universities provide the recommendation service for librarians, teachers, and students, about 19% of the libraries provide the recommendation service for course reference books, and about 10% of the libraries provide the recommendation service for classic books, which is the lowest percentage. In addition, 79% of the surveyed university libraries provide recommendation services to readers in the form of resource links, which, by providing a link related to searching for books, point to reading platforms such as Douban and Google. Among dynamic recommendations, about 62% of all surveyed colleges and universities provide popular search recommendation service, which is the highest percentage, about 58% of colleges and universities provide popular borrowing as a recommendation service method, about 57% of colleges and universities provide popular collection recommendation, about 55% of colleges and universities provide using popular evaluation service, about 46% of colleges and universities provide popular recommendation method, and about 34% of the colleges and universities provided the lowest percentage of recommendation methods of borrowing ranking. It is possible that there is a crossover between the various recommendation methods, and institutions have chosen only one or more of them.

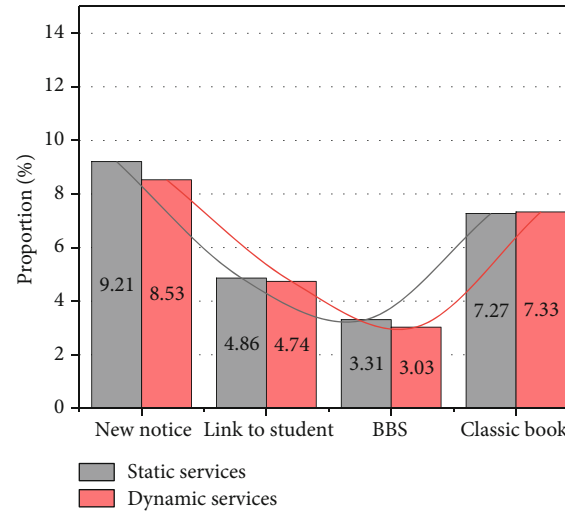


Figure 5: Static and dynamic personalization services.

4. System Optimization Test

There are many valuable or potentially valuable knowledge and rules in user behavior information. Through the analysis of users' retrieval behavior and implied association behavior, the intelligent retrieval service can use data mining, association rules, and other technologies to provide recommendation services to mass users who are not logged in at the corresponding position on the relevant detailed information page, recommending the same, similar, or related books they retrieve or browse, so as to help users find the books they need more quickly and accurately, and thus reducing the book search service is a kind of a nonsense search service. The essence of intelligent retrieval is a kind of impersonal recommendation, which is a practical application of data mining analysis technology in information resource retrieval processing and belongs to a kind of analysis and mining activity for information resources in the network.

At present, OPAC online public search catalog is the main network entrance for university libraries to search library resources, and it is the most important platform and window for readers to query and browse books with libraries. The design and implementation of its functions have the most direct impact on the quality of library services and material utilization. Therefore, the development of intelligent search technology relies on the OPAC system to a considerable extent. Users retrieve library resource information through OPAC search engine, and user query log information can be formed in the backend of the library website to describe the user's search behavior. The intelligent retrieval system can construct a user behavior model, and on the basis of querying and analyzing user behavior logs, mine potential information of user retrieval behavior then predict the retrieval results that users may visit or prefer and intelligently select and recommend information resources that are close to users' interests or behaviors. By correlating user retrieval with retrieval results and making implicit knowledge of behavior explicit for retrieval results, the intelligent retrieval service recommends hidden books to users, which can make users improve their accuracy in retrieval while feeling novel and diverse, as shown in Figure 6, so that the knowledge-filled correlation or similarity intelligent retrieval will be more tempting to users.

Smart retrieval is the first step to provide users with smart services, mainly serving general public users and belonging to a rough refinement. Smart recommendation is similar to the information retrieval system represented by search engine, but with more emphasis on personalized, diversified, and novel recommendation results. Search is to know exactly what you are looking for, but search is no longer a solution under information overload. The recommendation system is a "push" and "pull" interaction, that is, to recommend information resources to users, while providing and displaying information resources to users to help them choose information. Compared with the intelligent search engine which simply lists the search results on the basis of certain filtering, the intelligent recommendation system can study the behavioral preferences of readers, establish the reader user model, and discover the interest points of readers, so as to meet the

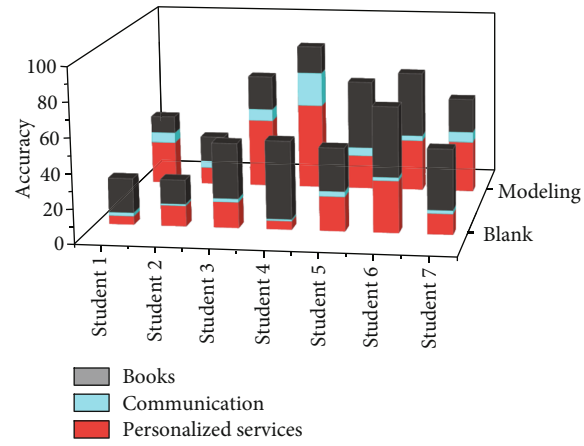


Figure 6: Personalized push accuracy.

diversified new demands of readers for information resources, improve the utilization rate of library resources, and enhance the intelligent processing ability of knowledge and information. With the core of integrating digital information resources to serve readers, the main task of the wisdom recommendation system is to link users and information from the passive of query to the active of recommendation, with humanized, personalized, and socialized features, helping users find valuable information and also allowing potentially valuable information to be presented to users, in order to achieve a win-win situation for knowledge producers and knowledge consumers. In addition, a high-quality wisdom recommendation system can generate wisdom recommendations to reader users on the one hand, and on the other hand, it can build a close connection with reader users, so that reader users can form a dependence on wisdom recommendations.

The information service system enables a more personalized recommendation function. Firstly, it enables a simple subject book push service based on readers' interest information. To meet the needs of readers, the information service system collects RSS feeds of various Chinese and foreign language journals and allows readers to customize the journal resources they are interested in, and after successful customization, the titles of relevant journals can be pushed to readers regularly to keep users abreast of the latest situation of journals. The existing OPAC system of university libraries can no longer meet the increasingly diversified information needs of patrons, let alone provide intelligent library information resource recommendation services. Although the OPAC system also provides services such as popular lending and new book recommendation, these basic recommendation services cannot differentiate between different readers and users, as shown in Figure 7, but provide the same book recommendation for all readers and users. The information service system, on the other hand, can use the information of each reader's personal professional background to provide differentiated recommendation services such as popular borrowing, compared with the simple recommendation attached to the OPAC system. Since university faculty and students have relatively stable disciplinary backgrounds, which are strongly related to their faculty backgrounds, the information service

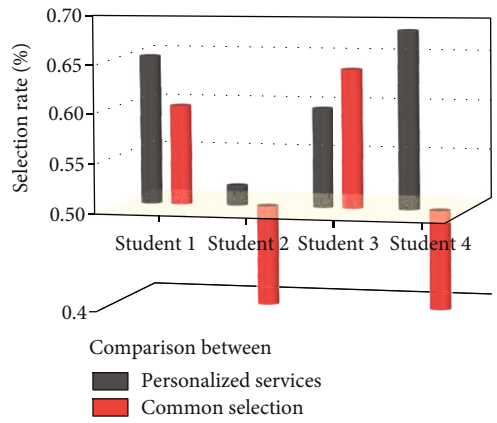


Figure 7: Degree of differentiation of personalized push.

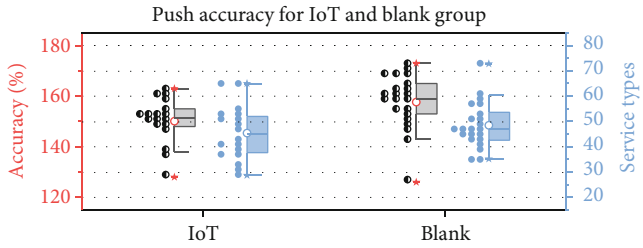


Figure 8: Comparison of push accuracy based on IoT technologies.

system was developed with a discipline-to-faculty correspondence table that gives users default disciplinary information based on their faculty backgrounds. If a reader does not set a subject preference when logging in to the information service system, the reader will receive the default subject library resources. If the reader has set his or her discipline preference, the system will push the book resources that are closer to the user's needs. The personalized recommendation service is made more valuable by setting your own special subject preferences based on your subject background, role information, etc.

The WeChat service provided by university libraries has many gaps from the real wisdom service, the content, means, ways, and planning of WeChat service are not perfect and mature, and it is still in the primary stage of exploration and experiment; university libraries should focus on relying on WeChat public platform to establish WeChat wisdom service. Under the micromedia environment, the wisdom WeChat service can expand the service scope, service content, and service mode of college libraries to better meet the different needs of different users. Traditional university libraries are often restricted by time or location, and for many readers who are not in the local service area or do not have computer terminals to access the Internet and during the nonworking hours of the library, the library does not provide basic services such as consultation and information sending. By using WeChat public service platform, university libraries can provide basic services to readers who have opened WeChat at anytime and anywhere and can also provide wisdom recommendation services for each different user to truly realize the wisdom service mode, as shown in Figure 8. Library wisdom WeChat service, firstly, should provide basic services for all users including library search and lending query, secondly, provide RSS customization service, etc., to carry out new book notification, regular lecture information push, etc., and again, provide wisdom service, including providing differentiated recommendation for users, and by collecting and analyzing various information of users, providing users that may be interested in or have demand for at the same time, it provides intelligent consulting services, automatic answering services, and sets up WeChat message services, etc. to enhance the interaction and communication between the library and the users and to meet the real-time and diversified service needs of the users in time. Finally, a user sharing module can be set up to increase library users and make recommendations for friends through users' own evaluation or sharing of books, using social networks and other media technologies.

5. Conclusion

This paper investigates and analyzes the basic status of recommendation service in university libraries, combines quantitative and qualitative analysis, makes screening, classification, and integration of data, summarizes the service status and shortcomings of recommendation service in university libraries, presents the specific findings of recommendation service provided by domestic university libraries with charts and graphs, and after the data investigation and results display, finds that the service of recommendation system is much better than that of the system before optimization. In addition, this paper also provides a summary of the current situation and shortcomings of the recommendation services provided by domestic university libraries. In addition, this paper also puts forward the idea of constructing the wisdom library service model based on wisdom recommendation. The four aspects of wisdom retrieval, wisdom recommendation, wisdom APP, and wisdom micro media are discussed, and the contents, ways, and structures of wisdom services that should be provided by the four parts of wisdom service mode are analyzed in detail. Then, through the optimized and improved recommendation algorithm, information such as readers' personal subject knowledge background is incorporated into the readers' model, and the recommendation system is envisioned to be improved and constructed, and the applicable recommendation for colleges and universities is pointed out. The application algorithm of the service is pointed out.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Q.-G. Chen, L. Chen, Q.-H. Zhong et al., "Optimization of urinary small extracellular vesicle isolation protocols: implications in early diagnosis, stratification, treatment and prognosis of diseases in the era of personalized medicine," *American Journal of Translational Research*, vol. 12, no. 10, pp. 6302–6313, 2020.

[2] J.-N. Kramer, F. Künzler, V. Mishra et al., "Which components of a smartphone walking app help users to reach personalized step goals? Results from an optimization trial," *Annals of Behavioral Medicine*, vol. 54, no. 7, pp. 518–528, 2020.

[3] A. Kapusta, Z. Erickson, H. M. Clever et al., "Personalized collaborative plans for robot-assisted dressing via optimization and simulation," *Autonomous Robots*, vol. 43, no. 8, pp. 2183–2207, 2019.

[4] R. Logesh, V. Subramaniyaswamy, D. Malathi et al., "Dynamic particle swarm optimization for personalized recommender system based on electroencephalography feedback," *Biomedical Research Tokyo*, vol. 28, no. 13, pp. 5646–5650, 2017.

[5] P. S. Dulai and C. A. Siegel, "Optimization of drug safety profile in inflammatory bowel disease through a personalized approach," *Current Drug Targets*, vol. 19, no. 7, pp. 740–747, 2018.

[6] X. Wang, J. Zhou, W. Yang et al., "Warpage optimization and influence factors analysis of 3d printing personalized Jjy tablets," *Drug Development and Industrial Pharmacy*, vol. 46, no. 3, pp. 388–394, 2020.

[7] F. Mofidi and H. Akbari, "Personalized energy costs and productivity optimization in offices," *Energy and Buildings*, vol. 143, pp. 173–190, 2017.

[8] R. Alkurd, I. Y. Abualhaol, and H. Yanikomeroglu, "Personalized resource allocation in wireless networks: an Ai-enabled and big data-driven multi-objective optimization," *IEEE Access*, vol. 8, pp. 144592–144609, 2020.

[9] G. W. Colopy, S. J. Roberts, and D. A. Clifton, "Bayesian optimization of personalized models for patient vital-sign monitoring," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 2, pp. 301–310, 2018.

[10] M. E. Fresard, R. Erices, M. L. Bravo et al., "Multi-objective optimization for personalized prediction of venous thromboembolism in ovarian cancer patients," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1500–1508, 2020.

[11] L. Xu, J. Zhang, B. Shi, and W. Meng, "Automating shift-scheduling calibration by using bionic optimization and personalized driver models," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4367–4376, 2019.

[12] 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.

[13] K. R. Devi and J. Bhavithra, "Personalized nutrition recommendation for diabetic patients using improved K-means and krill-herd optimization," *International Journal of Scientific & Technology Research*, vol. 9, no. 3, pp. 1076–1083, 2020.

[14] D.-K. Lee, V. Y. Chang, T. Kee, C.-M. Ho, and D. Ho, "Optimizing combination therapy for acute lymphoblastic leukemia using a phenotypic personalized medicine digital health platform: retrospective optimization individualizes patient regimens to maximize efficacy and safety," *Journal of Laboratory Automation*, vol. 22, no. 3, pp. 276–288, 2017.

[15] Y. Gao, Y. Feng, and J. Tan, "Exploratory study on cognitive information gain modeling and optimization of personalized recommendations for knowledge reuse," *Journal of Manufacturing Systems*, vol. 43, pp. 400–408, 2017.

[16] W. Zhao, S. Li, H. Yao et al., "Molecular optimization enables over 13% efficiency in organic solar cells," *Journal of the American Chemical Society*, vol. 139, no. 21, pp. 7148–7151, 2017.

[17] K. Cheng, Y. Liu, C. Yao, W. Zhao, and X. Xu, "A personalized mandibular implant with supporting and porous structures designed with topology optimization - a case study of canine," *Rapid Prototyping Journal*, vol. 25, no. 2, pp. 417–426, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5566647, 11 pages
https://doi.org/10.1155/2021/5566647

Research Article

Network Public Opinion Prediction and Control Based on Edge Computing and Artificial Intelligence New Paradigm

Ying Zhu

Hubei University of Technology, Wuhan Hubei 430060, China

Correspondence should be addressed to Ying Zhu; yingzh@hbut.edu.cn

Received 20 February 2021; Revised 25 March 2021; Accepted 7 April 2021; Published 19 April 2021

Academic Editor: Wei Wang

In this paper, an adaptive edge service placement mechanism based on online learning and a predictive edge service migration method based on factor graph model are proposed to solve the edge computing service placement problem from the edge computing dimension. First, the time series of the development of online chaotic public opinion is a platform for vectorized collection of keyword index trends using the theory of chaotic phase space reconstruction. Secondly, it is necessary to use the main index method to judge whether the time series has the chaotic characteristics of the network public opinion data. The simulation results show that network public opinion is the development characteristic of chaotic time series. Finally, the prediction model is improved by using complex network topology. Through the simulation experiment of network public opinion and chaotic time series, the results show that the improved model has the advantages of accuracy, rapidity, and self-adaptability and can be applied to other fields.

1. Introduction

In recent years, with the rapid development of social economy and science and technology in the world, many new technologies have emerged in the information and communication technology industry. Among them, there are two representative technologies which are widely considered to have a great impetus and far-reaching influence on human economy and society. First, the depth study of the representative as the field of artificial intelligence technology, benefit from the algorithm, calculate force and the progress of datasets, and so on, obtained the development which progresses by leaps and bounds in recent years, and in the unmanned, e-commerce, financial, and other fields, intelligent household and wisdom deeply changed people's way of life and improved the production efficiency [1, 2]. The other technology is edge computing technology evolved from traditional cloud computing technology [3]. Compared with cloud computing, edge computing sinks strong computing resources and efficient services to the edge of the network, thus having lower delay, lower bandwidth occupancy, higher energy efficiency, and better privacy protection.

Based on the dimension of AI-enabled edge computing, this paper first proposes an online service placement mechanism based on user adaptive management for the dynamic migration and placement of edge computing services. The mechanism can adapt to complex user behavior and changeable edge network environment through online learning artificial intelligence technology, so as to assist users to make efficient service migration decisions. It then shows how to use the factor graph model, an emerging artificial intelligence technique, to achieve user location prediction to improve the quality of dynamic migration decisions for edge services. The time series representing the development trend information of network public opinion is collected, and Fourier transform is carried out.

The phase space development time series of network public opinion was reconstructed by using CAO method and autocorrelation function method in phase space reconstruction theory, and the prediction model was established by using reconstructed time delay vector and reserve pool neural network. The complex network topology is used to improve and optimize the reserve pool structure of the reserve pool neural network, so as to adapt to the chaotic characteristics

and sensitivity of the development trend of network public opinion. Using intelligent algorithm to optimize the parameters of the new neural network, the problem of determining parameters is transformed into an optimization problem. In the aspect of swarm intelligence, the collaborative mechanism, evolution rule, and operator coupling optimization algorithm were designed, and the chaos initialization process was adopted to increase the population diversity, and the global search ability and local search ability were coordinated. The ability to implement algorithms jumps out of local optimality.

2. Related Work

In recent years, the problem of resource allocation in moving edge computing has attracted extensive research attention. Taking the power minimization of mobile devices as the optimization objective, the resource allocation problem of mobile edge computing is studied, and the backhaul capacity limit, interference, and tolerable waiting time are considered [4, 5]. Based on linear programming, an optimal allocation strategy for multiple resources (computing power and wireless bandwidth) is proposed, and performance improvements in system throughput and service delay are demonstrated [6]. The computational unloading formula is transformed into a constrained Markov decision process, which is aimed at minimizing the energy consumption of the user's equipment while meeting the maximum latency requirements of the application [7]. A low-complexity dynamic computing unloading algorithm based on Lyapunov optimization was proposed to optimize the task execution time [8, 9]. A heuristic algorithm is proposed to divide the computing tasks of users to minimize the average completion time of all users [10]. The computational unload decision is designed to optimize the time and energy consumption of the whole system by taking physical resource block allocation and computational resource allocation as optimization problems. However, none of these algorithms can well solve the problem of resource allocation in complex scenarios [11]. This paper studies the differences between network media and traditional media and clarifies its development advantages and approaches [12]. This paper starts from the concept of network public opinion to distinguish the boundaries of public opinion and discusses how the general will, that is, the will of all people, is generated on the basis of public will and private will. This paper makes a comparative analysis of the generalities related to public opinion and discusses the interaction and influence among public opinion, government policies, and media organizations [13]. By using a two-dimensional rectangular grid to abstract represent the real interpersonal network, the dynamic process of public opinion evolves within its grid space structure, and the evolution process of network public opinion is represented by simulating the interaction of opinions among adjacent vertices at a constant rate. This model is called the opinion exchange model [14]. A theoretical framework is constructed to discuss the conditions under which the common opinions among members are affected by the society over time, the formation of network contact methods defined by the environment, and the attributes of individual agents and time-related factors, as well as the influence of joint effect on the dissemination of public opinion [15]. Ignore the costs of resource is put forward to maximize the expected utility of fully rational decision makers the best decisions, puts forward a thermodynamic inspired the standardization of the limited rational decision, with limited rational decision problems can describe famous variational principle and reflect the thinking of the people in the spread of public opinion and decision, from another aspect to think [16] individual Internet users to the dissemination of public opinion or not. Starting from the two dimensions of political participation, interaction, and democratic supervision of the subjects of online public opinion, the influence of online public opinion on online democracy is theoretically analyzed and studied [17, 18]. From the perspective of mastering the initiative of online public opinion, this paper puts forward that if we want to grasp the law correctly, we need to collect and analyze the online public opinion scientifically to master the initiative of online public opinion, so as to guide the online public opinion correctly. It is believed that the main body of online public opinion is the Internet users, and the correct grasp of the main body characteristics of Internet users can better understand the transmission law, so the study of the main body characteristics of Internet public opinion is completed by analyzing the concept and connotation of Internet users [19, 20].

On the basis of the existing public opinion in China, by analyzing the influencing factors of online public opinion in emergencies, the concept of the number of people who know the information on the Internet is introduced to describe the influence of online public opinion, and a differential equation model of the evolution law of online public opinion in emergencies is established [21, 22]. To the spread of the network public opinion after the process as the research object, analyses of the main factors that influence the impact of network public opinion spread link, using system dynamics simulation method simulation evolution law of network public opinion spread, through qualitative and quantitative analyses, show that to improve the government credibility, the government social service efficiency can effectively reduce the degree of spread of public opinion [23]. In a word, researchers at home and abroad have conducted in-depth studies on the basic theory, supporting technology and evolution mechanism of online public opinion. Qualitative research methods are mainly used to analyze the evolution process of online public opinion from related fields such as sociology, psychology, and communication to elaborate its evolution law [24, 25]. The quantitative research methods mainly include cellular automata model, system dynamics method, social network analysis method, and statistical analysis method [26–28]. From different angles and disciplines, the network public opinion is discussed and studied. Optimization algorithm and neural network methods include the development trend of network public opinion prediction analysis method. Because the time series of the development trend of network public opinion is complex and nonlinear, the traditional statistical method for prediction has certain limitations, so it is necessary to adopt a new method to forecast the highly nonlinear time series of network public opinion.

3. Research Framework of Network Early Warning and Prediction Based on Edge Computing and Artificial Intelligence New Paradigm

3.1. Optimization Framework. Based on the idea of artificial intelligence-enabled edge computing, aiming at the dynamic migration and placement of edge computing services, a user adaptive management online service placement mechanism is introduced. This mechanism uses artificial intelligence to integrate the new artificial intelligence paradigm and adaptively learns the complex user behavior and the changeable edge network environment, so as to assist users to make efficient service migration decisions. As Figure 1 shows, EdGent's optimization logic is divided into three phases: the offline training phase, the online optimization phase, and the collaborative inference phase.

The above based on edge server and the depth of the new paradigm to study artificial intelligence model inference framework design train of thought is as follows: in the offline phase, training well meets the demand of task branch network, at the same time as the branch network layer neural network training is different in the regression model, to gauge neural network layer on the edge of the server and the terminal equipment run time delay. In the online optimization stage, the regression model will be used to find the exit points and model segmentation points that meet the task delay requirements. In the collaborative inference phase, the edge server and the end device will run the deep learning model according to the resulting scenario.

3.1.1. Offline Training Stage. In the offline training phase, EdGent needs to perform the following two initialization operations: to analyze the performance of edge servers and terminal devices and to generate a regression model-based delay estimation model for different types of deep learning model network layers (such as convolution layer and pooling layer). In estimating the network layer of the runtime latency, on each layer of the network layer of the Edgent modeling, modeling rather than on the depth of the learning model, the time delay of different network layer is by its own independent variables (such as the size of the input data and output data size); based on each layer of the independent variable, it can establish regression model and estimate the time delay of each layer of network layer. The branch network model with multiple exit points is trained to simplify the model. Here, the author adopts the Branchynet branch network structure. Under the Branchynet structure, a branch network with multiple exit points can be designed and trained to generate. It is important to note that performance analysis depends on the device (such as mobile phones, VR headsets, and smart watch when the different equipment to run the same deep learning model the performance of different), and deep learning model depends on the application (for example, different computer vision applications such as object recognition and classification of the depth of the corresponding learning model), so in a given deep learning applications and devices (that is, the finite edge server and terminal equipment), under the condition of the above two initialization in the offline phase needs to be done only once.

3.1.2. Online Optimization Stage. The main work of this stage is to find the exit points and model segmentation points in the branch network that meet the time delay requirements by using the regression model of off-line training. In order to maximize the accuracy of the scheme, in this stage, the author starts from the branch with the highest accuracy and iteratively finds out the exit points and segmentation points that meet the requirements. In this process, EdGent measures the network bandwidth of the link between the current mobile terminal and the edge server in real time to estimate the data transfer latency between the mobile terminal and the edge server. EdGent then traversed the different split points on each network branch from large to small in size and estimated the end-to-end delay and model accuracy for the selected branch network and the split points based on the current network bandwidth and different network layer computing times. After traversal of all branch networks and shard points, EdGent outputs the combination with maximum accuracy of all network branch and shard point combinations that meet the delay requirements.

3.1.3. Collaborative Inference Stage. In the collaborative inference stage, the edge server and the mobile terminal carry out collaborative inference on the deep learning model according to the optimal network branch and segmentation point combination output in the online optimization stage. Experiments have shown that EdGent performs well in improving the real-time performance of deep learning applications, enabling high-precision model reasoning under different computational latency requirements.

3.2. Adaptive Edge Computing Based on Online Learning and New Paradigm Algorithm of Artificial Intelligence. Based on the dimension of edge computing enabled by artificial intelligence, this paper proposes a network early warning and prediction mechanism for user adaptive management in order to solve the problem of network early warning and prediction of edge computing services. The mechanism can adapt to complex user behavior and changeable edge network environment through online learning artificial intelligence technology, so as to assist users to make efficient network early warning and prediction decisions, how to use factor graph model to realize user location prediction so as to improve the quality of edge network early warning and prediction decision.

The establishment of service placement policy is usually related to user behavior characteristics and network environment. Among them, the behavior characteristics mainly include the user's mobility model, the type of request service, and personal preference, and the network environment mainly includes the resources available in the edge network and the transmission bandwidth between nodes and other factors.

In order to measure the user service quality in the edge network effectively, we study the user-perceived delay and the cost of service migration. User perceptible delay is considered from two aspects: computing delay and communication

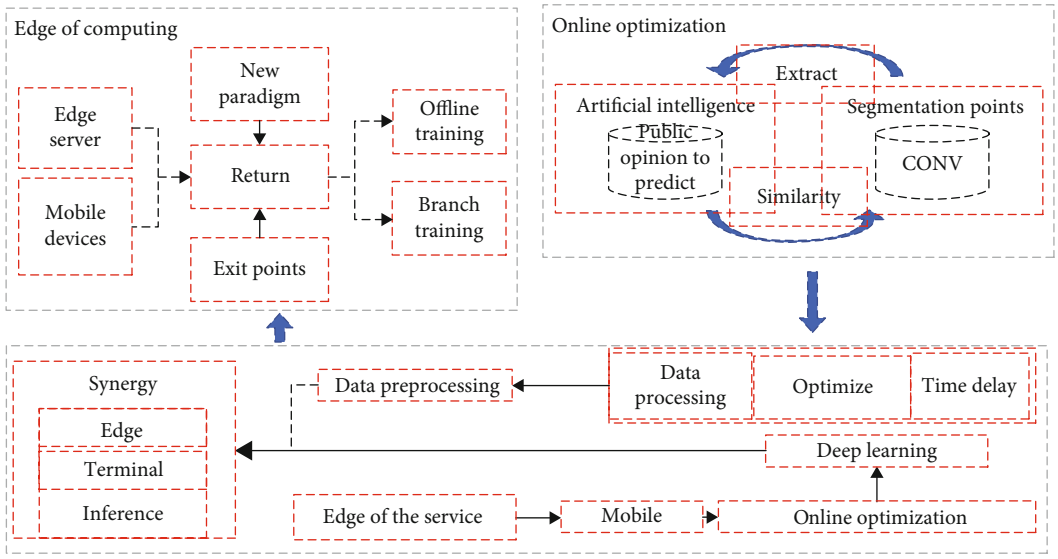


Figure 1: Running the inferential optimization framework EdGent based on the deep learning model of edge and terminal collaboration.

delay. The computing delay mainly depends on the amount of request service data to be processed and the available computing resources to place the node, such as the speed of the CPU. Communication delay is mainly composed of the user's current access delay and transmission delay. The access delay mainly depends on the location of the current user and the state of the edge router. In general, to reduce communication power consumption, users will choose to connect with the nearest edge router. The transmission delay is related to the network state (such as the bandwidth between edge nodes) and the network distance. The additional operating costs (such as bandwidth utilization) caused by service migration mainly depend on the node selection at the previous and current time. Thus, the user's service quality can be described.

$$\chi_1 C_1(\pi_1) + \chi_2 C_2(\pi_2 x_2) + \chi_2 S(\pi_3). \quad (1)$$

If the user behavior characteristics and network state of the future long-term time T can be accurately predicted, then the offline optimal strategy for long-term service can be obtained through dynamic programming method. However, in the actual environment, it is difficult to accurately predict the above users and network information. At the same time, for each decision moment, due to the lack of understanding of network environment parameters, users will consume additional communication costs to collect system information.

At the end of this time segment, the user gets a quality-of-service representation of the placement policy. At the same time, the adaptive management mechanism will use all the information within the time segment to update the user behavior characteristics and the potential relationship between placement strategy and service quality performance, that is, the network parameters of service placement strategy. The specific algorithm is as follows.

The validity of the proposed framework is verified by simulating the random driving of the network warning test in the edge network. Figure 2 records the service quality performance of the adaptive edge service placement mechanism based on online learning under different time segments and compares it with the theoretical optimal placement strategy under known long-term total information. It can be seen from the results that, with the increase of time, the proposed mechanism keeps approaching the optimal placement strategy, which indicates that the strategy formulation can be effectively optimized through online learning.

4. Research on Predictive Control Model of Network Public Opinion Based on New Paradigm of Edge Computing Artificial Intelligence

In the evolution process of network public opinion, the change of opinion and attitude and the diffusion of information complement each other, and they promote the evolution and development of network public opinion events together through the synergistic effect. In the evolution process of public opinion, the aggregation of opinions is closely related to the dissemination of public opinion information. Therefore, this paper conducts research on this basis, considers the two branches at the same time, studies their complex nonlinear evolutionary blending process, and analyzes their characteristic rules.

On the Internet, netizens should first be informed of public opinion information. When the public opinion information is obtained, it exchanges opinions with the neighborhood netizens. If the individual state is unknown, that is, he has not heard of the public opinion, then he has no attitude towards the public opinion, as a neutral attitude processing. When an individual never knows the state of public opinion into the state of public opinion, then the individual has his own attitude value towards public opinion events, and the attitude value is jointly affected by neighborhood individuals. In the process of constantly learning and paying attention to

Initialize compute nodes in the network, corresponding user characteristics, network estimation parameters, and their cumulative context.
For each time segment $t = 1, 2..., T$, do the following.
According to the network parameters of node I, combined with the current user characteristics $B(t)$, the corresponding cost is evaluated according to Thompson sampling.
The computing node with the least estimated cost is selected as the current service placement strategy, and the corresponding QoS performance is received at the end of the time segment.
Update and select the corresponding user characteristics of the compute node and its network estimation parameters.

Algorithm 1: New paradigm algorithm for adaptive artificial intelligence based on edge computing.

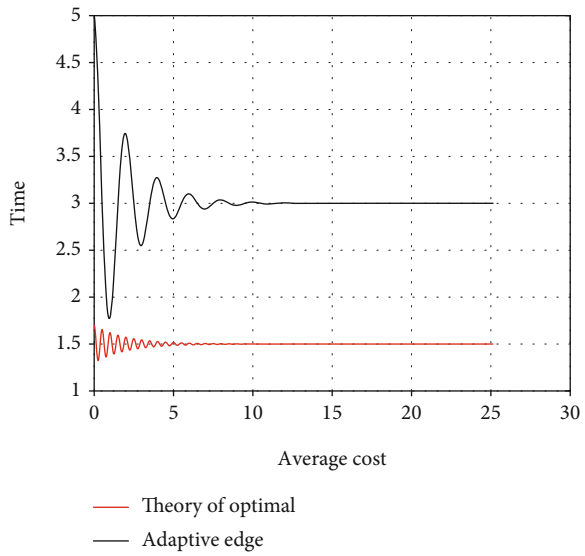


Figure 2: Comparison between the service placement strategy of the new paradigm of adaptive edge computing fusion artificial intelligence and the theoretical optimal placement strategy.

the evolution of the public opinion event, the attitude value and awareness of individual netizens are also changing and revising constantly. The individual netizen is taken as a node, and the neighborhood netizen group is composed of nodes that are bordered to the node.

Due to the complexity and redundancy of Internet information, the degree of understanding of individual nodes of Internet users to online public opinion information is set as the degree of knowledge, denoted by R. Individual netizens have different understanding processes of public opinion information, which will affect netizens' opinions and attitudes towards public opinion events. If the whole picture of the network public opinion events is well understood, the individual will have a higher coefficient of firmness towards the public opinion. In the process of continuous communication, the more we know about the event, the more likely we are to change our attitude according to the information we know. However, due to the different educational level and personality of each netizen, his firmness to public opinion information is also different, and the firmness coefficient is expressed as Q. At the same time, taking into account the influence of the neighborhood netizens on the individual node, in real life, people who are close to each other are more feasible to the information transmitted by each other, while people who are far from each other are not so strong to the information transmitted by each other, including their attitudes. The degree of influence of neighborhood individuals on the netizens of this node is set as neighborhood coefficient, which is represented by P. The neighborhood coefficient is determined by the distance of psychological distance and is expressed as follows:

$$P_{ij} = \frac{k}{p_{ij}}. \tag{2}$$

To sum up, the attitude state of the netizen node towards public opinion information at time t is expressed by S, so

$$S_I(i) = F(P, R, Q, S_{i-1}(i)). \tag{3}$$

The above formula indicates that the state of the node at the next moment is composed of the synergistic effect of the current state, knowledge, firmness coefficient, and neighborhood influence, and its specific relationship is

$$S_{I+1}(i) = R_I Q_I S(i) + (1 - q_i) \sum_j R_j P_{ij}. \tag{4}$$

With the occurrence of public opinion events, people's concern for public opinion events will decrease with the passage of time, that is to say, netizens' memory effect on events has a certain timeliness, and the memory effect coefficient is denoted as M. Therefore, considering the evolution law of online public opinion under the effect of memory effect, we can write

$$S_{I+1}(i) = R_I Q_I S(i) + (1 - q_i) \sum_j R_j P_{ij} \text{-} M_I(I) S_I(i), \tag{5}$$

$$\text{Accuracy} = \frac{\text{TP}}{\text{TP} + \text{FN} + \text{FP} + \text{FN}}, \tag{6}$$

$$\text{Precison} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{8}$$

$$F1 = \frac{2 * \text{TP}}{2 * \text{TP} + \text{FP} + \text{FN}}. \tag{9}$$

Table 1: NW network structure parameters.

Probability p	Mean path length	Clustering coefficient	Average degree
0.1	1.8094	0.1906	190.58
0.2	1.6395	0.3602	360.426
0.3	1.4891	0.5105	510.214
0.4	1.3576	0.6423	641.654

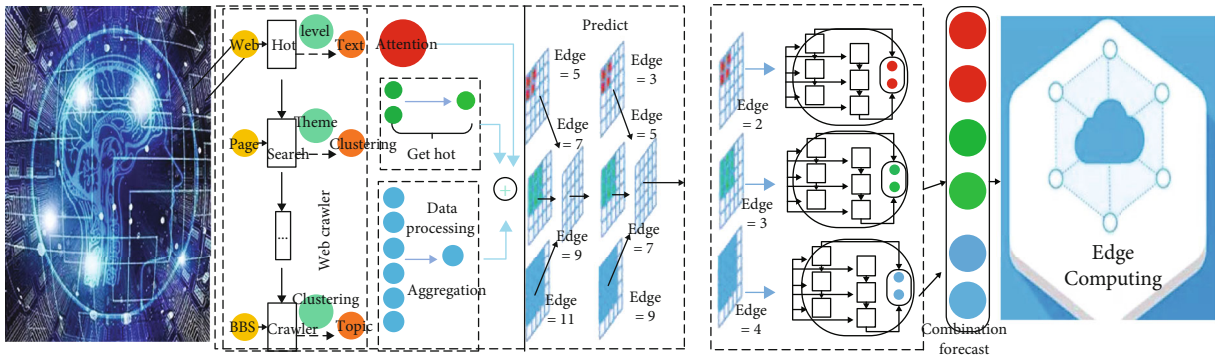


Figure 3: Architecture of network public opinion prediction system.

Firstly, according to the NW small-world model construction algorithm, the NW small-world network composed of N nodes is established, where N is 2000, and the reconnect probabilities are 0.1, 0.2, 0.3, and 0.4, respectively. The network structure parameters are shown in Table 1. Each node represents an individual netizen and is influenced by the nodes to which it has an edge. Then, dynamic evolution rules are used to simulate the changing trend of network public opinion.

The architecture of the network public opinion prediction system firstly captures the public opinion data from the Internet through the network crawler based on heuristic search. After the operation of data filtering, cleaning, and interference removal, text clustering is carried out, and the clustering method based on hierarchical topic tree is adopted. After clustering, the network public opinion data becomes cluster distribution, and then, the hot topic is acquired. Find hot topics from a variety of topics. The approach is based on topic attention techniques. Then, data aggregation is carried out to make the data become a discrete data series based on time series. The combined prediction method combining grey theory and weighted Markov is used to forecast the data. Use forecast data to support decision-making. The specific architecture is shown in Figure 3.

Step 1. Build a network crawler based on heuristic search to capture, remove impurities, and store network data.

Step 2. Divide the topic domain, then use the method of automatic threshold calculation to establish a hierarchical topic tree, conduct hierarchical clustering analysis, and evaluate the clustering quality with purity and F value.

Step 3. Calculate the media attention and public attention. Based on the given proportion balance factor X, calculate the comprehensive attention, and then determine and capture the hot spots. Then, the data aggregation software is used to generate discrete time series.

Step 4. Establish a GM (1,1) model for prediction and use the prediction results to establish Markov prediction. Finally, the predicted values are stored for decision support.

In order to make accurate location prediction, features of multiple dimensions are extracted from users' historical information, such as time, location, network status, and social frequency characteristics at different moments, and are effectively integrated into a unified framework by using factor graphs. Among them, the time feature for users from the timestamp time information (such as specific moments, weekday, or weekend), frequency of social characteristics is to the user in the period of all kinds of social behavior characteristics (such as release tweets, refresh forward tweet, tweet, preview multimedia video, and multimedia video watching) separately calculated frequency data, and network status characteristics is to point to the current status of network users. Deep learning model inferences optimization framework based on edge and terminal collaboration. As Figure 4 shows, EdGent's optimization logic is divided into three phases: the offline training phase, the online optimization phase, and the collaborative inference phase.

The convergence of the new paradigm algorithm of edge computing and artificial intelligence in optimizing average service delay and average energy consumption is evaluated, respectively. Among them, the number of edge servers is set as 5 and the number of mobile devices is set as 115. The computing power of edge servers has two settings under the two

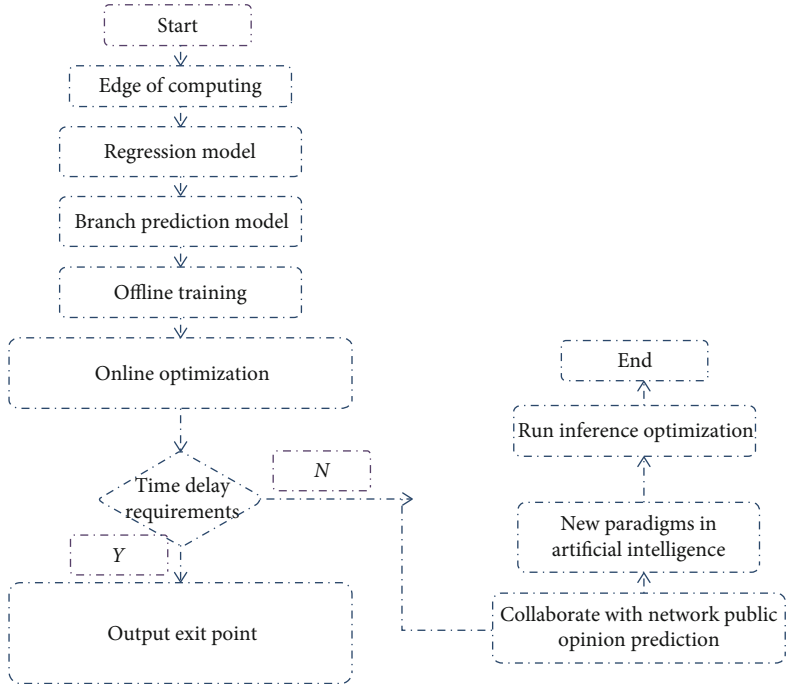


Figure 4: Network prediction process of the new paradigm of artificial intelligence.

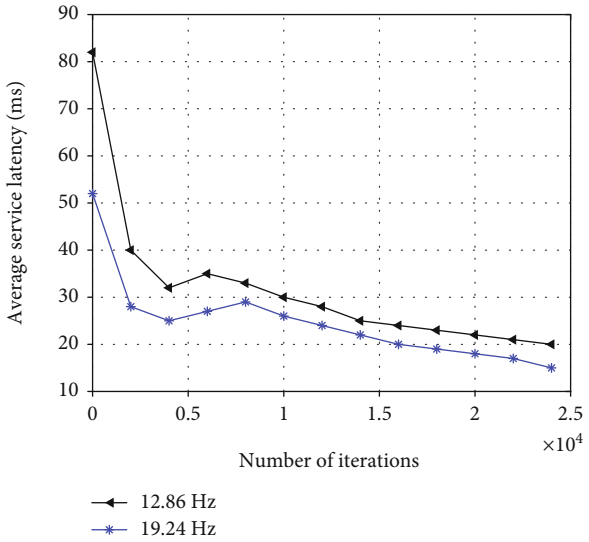


Figure 5: Convergence performance of average service delay optimization by edge computing fusion with new paradigm algorithm of artificial intelligence.

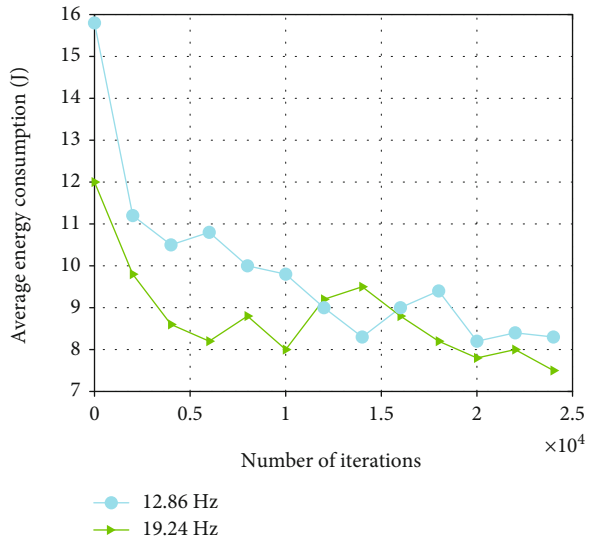


Figure 6: Convergence performance of edge computing combined with the new paradigm algorithm of artificial intelligence when optimizing average energy consumption.

optimization goals, respectively, 12.86 GHz and 19.24 GHz. As shown in Figure 5, the average service delay dropped sharply over the first 7000 iterations and then stabilized. As can be seen from Figure 6, when the number of iterations reaches 5000, the average energy consumption begins to decline at a slower rate. As the training iteration goes on, the new paradigm algorithm of edge computing fusion artificial intelligence will converge quickly. It can be seen from the two subgraphs that the computing power of the edge server has little effect on the convergence speed of the new paradigm algorithm of edge computing fusion artificial intelligence. The number of iterations was set to 4600 to balance performance and complexity.

4.1. Experimental Design. Although edge computing can solve the problem of limited user resources and excessive delay in cloud computing, the service coverage of edge nodes is small, and the movement of users will have a great impact on the service quality. As shown in Figure 7, when a user moves from one edge node servant region to another node

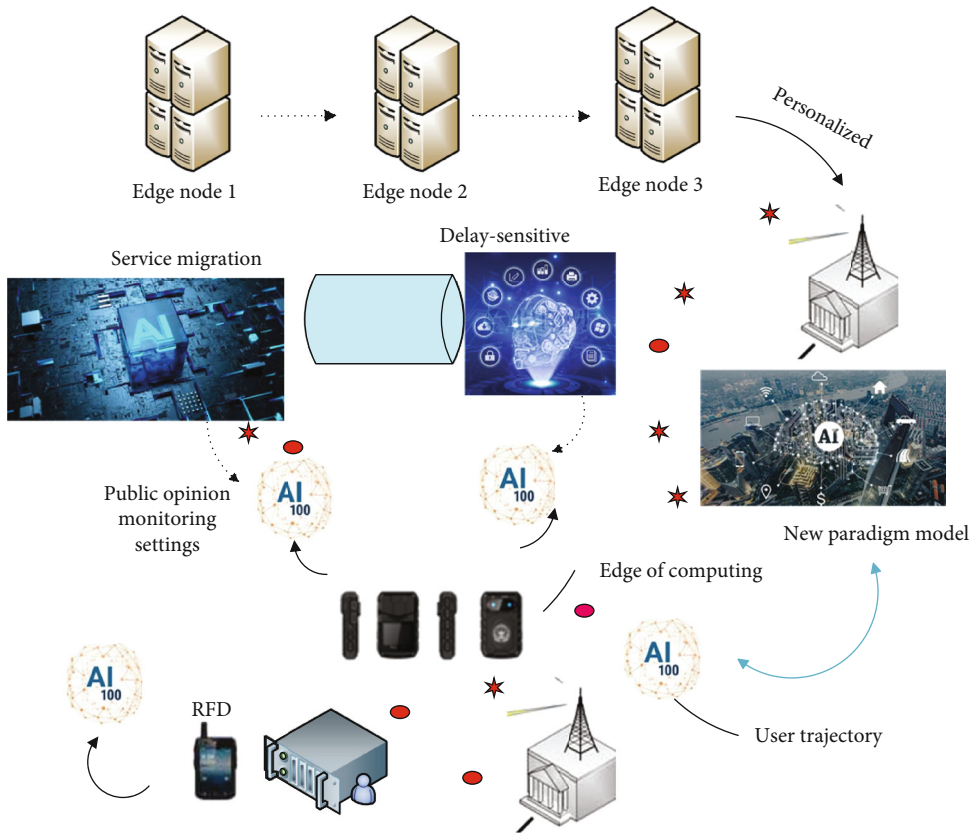


Figure 7: Experimental setting of network public opinion monitoring based on edge computing and artificial intelligence new paradigm.

servant region, it is necessary to consider whether to perform service migration to ensure satisfactory service quality.

On the one hand, users can choose to continue to allow the service to be processed in the original edge nodes, and the continuity of the service can be guaranteed through the data transmission between edge nodes. On the other hand, users can opt for service migration to reduce end-to-end latency. The former may cause a large transmission delay due to a long network distance, while the latter introduces additional overhead caused by service migration. Due to the personalized needs of users at the same time, the different service types and migration are increasing the difficulty of service place, such as the application of lightweight users tend to be more local processing, users tend to be more computationally intensive application cloud server processing, computationally intensive, and delay the application of the sensitive user more inclined to edge server processing. Therefore, an adaptive edge service placement mechanism based on online learning is proposed, which can effectively balance the trade-off between delay and migration cost.

Discrete time series of 30-day relevant data of a hot topic on the network after pretreatment is shown in Table 2.

Table 2: Observation data of an affair.

Time	Postteaching	Time	Postteaching	Time	Postteaching
1	5	5	203	9	345
2	67	6	223	10	456
3	146	7	256	11	567
4	175	8	278	12	768

5. Results and Analysis

Select "two sessions" as a key word statistics of a total of 124 days of search index as the development trend of online public opinion data for analysis. Time series data of 124 components were divided into training samples and prediction data samples. Among them, the first 94 data are used as training sample data, and the last 30 values are used as prediction data. The prediction experiment was conducted on the time series of network public opinion, and the experimental results are shown in Figure 7.

It can be seen from Figure 8 that the reserve pool neural network is effective in predicting the time series of network public opinion.

Assuming that the proportion of the initial informed people before the public opinion is not widely disseminated is 0.1, the proportion of the people with neutral attitude to the informed people is always 0. In L, the firm coefficient is 0.5, and the propagation probability is 0.6, and the linearly increasing dynamic memory effect coefficient M is adopted, and the evolution times is 20, simulation experiment on the evolution of public opinion under different initial states of informed people. The result is shown in Figure 9.

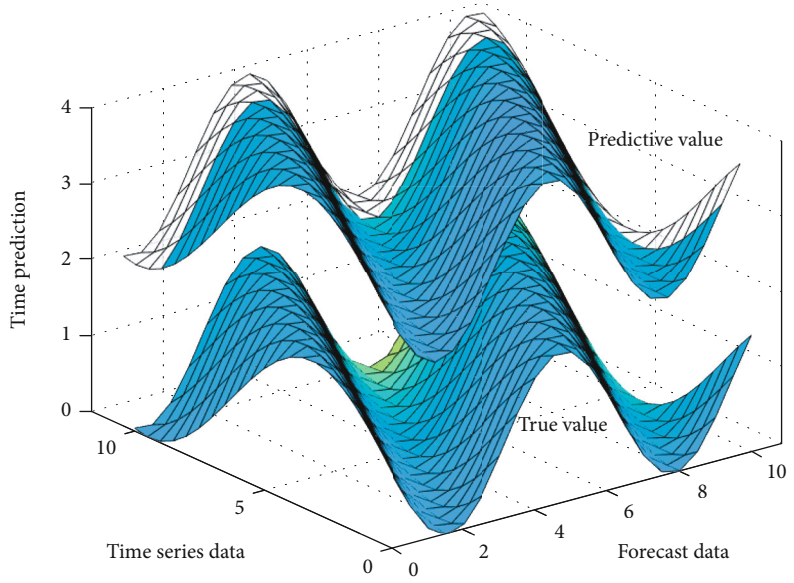


Figure 8: Experimental results of public opinion information.

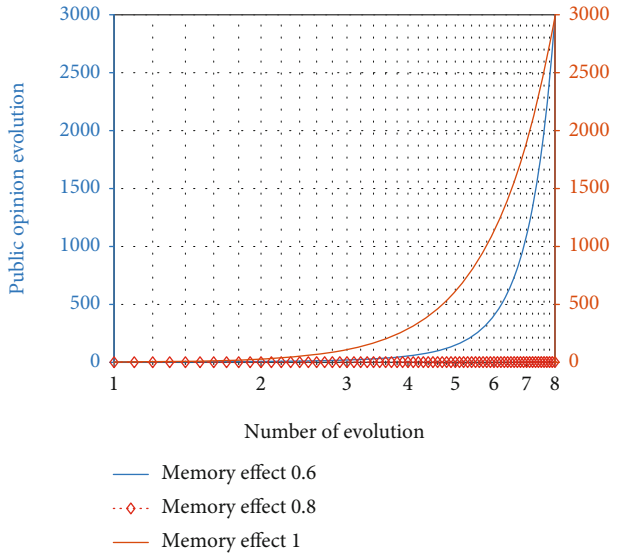


Figure 9: Evolutionary results of public opinion.

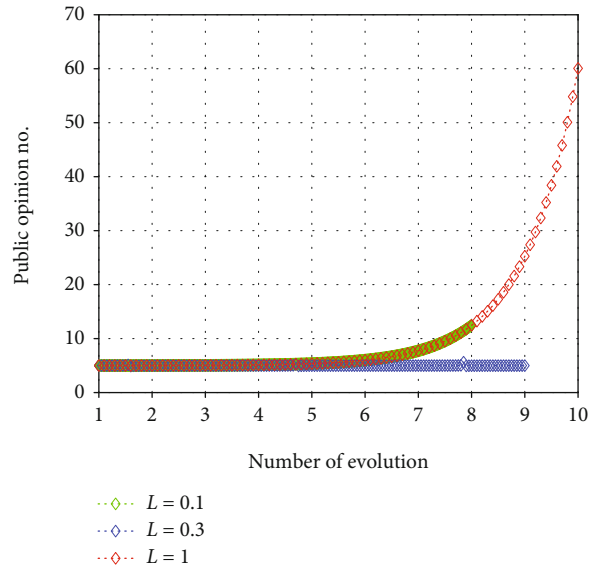
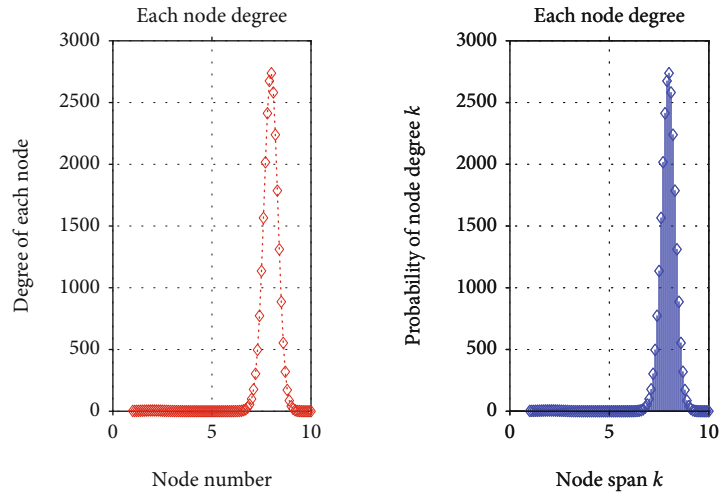


Figure 10: Evolvement result of static opinion leader coefficient public opinion.

It is assumed that static leadership coefficient $L = 0.1 : 0.3 : 1.0$ and linearly increasing dynamic leadership coefficient are adopted. The probability of small-world network connection is 0.1, and leaders show a positive attitude towards public opinion events, assuming their attitude value is 1. The number of evolutions is 50, and the other parameters are set in accordance with the above. Simulation experiments on the evolution of public opinion of informed people in different initial states are conducted. The result is shown in Figure 10.

The evolution times were set as 50, and the proportion of informed people was 0.1, among which the proportion of approving people to informed people was 0.3, the proportion of opposing people was 0.6, the certainty coefficient was 0.5, and the propagation probability was 0.5. Linear increasing dynamic memory effect coefficient and dynamic opinion leader coefficient were used to simulate it, and the experimental results are shown in Figures 11(a) and 11(b).

This paper proposes a more practical public opinion communication model. It is assumed that each user does not fully believe or adopt the views of his friends when communicating with his neighbors, but accepts the views of his friends with probability. Therefore, the probability of user interaction is introduced. Theoretical analysis and numerical simulation show that the proportion of people with different opinions in a grid network always maintains a stable level, and only for a certain probability can different opinions in a random network coexist; otherwise, all people will hold the same opinion. The study of population proportion model



(a) Node degree distribution diagram (b) Node degree probability distribution diagram

Figure 11: BA network characteristic diagram.

finds that the proportion of people holding various opinions in the grid network is related to the proportion of people in X and Y at the initial state, but has nothing to do with the communication probability. The existence of special parameters for a particular network can also make the system reach the equilibrium state.

6. Conclusion

From the perspective of edge computing and artificial intelligence-enabled edge computing, this paper introduces the adaptive edge service placement mechanism based on online learning and the edge service migration method based on location prediction, aiming at the dynamic service migration problem caused by the large-scale and high-density deployment of edge nodes. In the future, edge intelligence, as a new paradigm for edge computing and artificial intelligence to enable each other, will give rise to a large number of innovative research opportunities and has a wide range of application prospects in many fields such as intelligent Internet of Things, intelligent manufacturing, and smart city. First, introduce edge fusion calculation for reserve pool of artificial intelligence neural network topology structure improvement; secondly, borrow 2 norm regularization models to complete the network weight method, and then, use mixed collaborative evolutionary algorithm to optimize network model parameters, to establish the forecast model of improvement, and use it to network public opinion trend of time series prediction research. The improved prediction model is applied to chaotic time series prediction, and the superiority and effectiveness of the improved model are further verified and analyzed.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

We declare that there is no conflict of interest.

Acknowledgments

This study is supported by the Provincial Department of Education Project, Internet Public Opinion Prediction and Control in Big Data Environment (19Q065).

References

[1] Y. Miao, G. Wu, M. Li, A. Ghoneim, M. Al-Rakhami, and M. S. Hossain, "Intelligent task prediction and computation offloading based on mobile-edge cloud computing," *Future Generation Computer Systems*, vol. 102, pp. 925–931, 2020.

[2] M. S. Elbamby, C. Perfecto, and C. F. Liu, "Wireless edge computing with latency and reliability guarantees," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 145–167, 2019.

[3] M. Chen, Y. Miao, and H. Gharavi, "Intelligent traffic adaptive resource allocation for edge computing-based 5G networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 1, pp. 2890–2903, 2019.

[4] Y. Guang, L. Dongbo, and M. Chaofeng, "Design and implementation of roadside intelligent information interaction system based on edge computing," *Journal of Physics: Conference Series*, vol. 1486, no. 2, pp. 22022–22028, 2020.

[5] 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, pp. 101813–101819, 2020.

[6] L. Rutkowski, M. Korytkowski, and R. Scherer, "Lecture notes in computer science artificial intelligence and soft computing image descriptor based on edge detection and crawler algorithm," *Ingeniería y Universidad*, vol. 57, pp. 647–659, 2016.

[7] D. C. Corrales, J. C. Corrales, and A. F. Casas, "Towards detecting crop diseases and pest by supervised learning," *Ingeniería Y Universidad*, vol. 19, no. 1, pp. 207–228, 2015.

[8] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, "Edge intelligence: paving the last mile of artificial intelligence with edge computing," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1738–1762, 2019.

[9] 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.

[10] Y. Dai, S. M. Du Xu, G. Qiao, and Y. Zhang, "Artificial intelligence empowered edge computing and caching for internet of vehicles," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 12–18, 2019.

[11] O. Debauche, S. Mahmoudi, S. A. Mahmoudi et al., "Edge computing and artificial intelligence for real-time poultry monitoring," *Procedia Computer Science*, vol. 175, pp. 534–541, 2020.

[12] G. Han, M. Guizani, G. Jia, and J. Lloret, "Special section on emerging trends issues and challenges in edge artificial intelligence," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 4172–4177, 2019.

[13] X. Wang, Y. Han, V. C. M. Leung, D. Niyato, X. Yan, and X. Chen, "Convergence of edge computing and deep learning: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 869–904, 2020.

[14] M. Elmoulat, O. Debauche, S. Mahmoudi, S. A. Mahmoudi, P. Manneback, and F. Lebeau, "Edge computing and artificial intelligence for landslides monitoring," *Procedia Computer Science*, vol. 177, pp. 480–487, 2020.

[15] B. Hussain, Q. Du, A. Imran, and M. A. Imran, "Artificial intelligence-powered mobile edge computing-based anomaly detection in cellular networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 4986–4996, 2020.

[16] L. Yang, X. Chen, S. M. Perlaza, and J. Zhang, "Special issue on artificial-intelligence-powered edge computing for Internet of Things," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9224–9226, 2020.

[17] B. Hussain, Q. Du, and A. Imran, "Artificial intelligence-powered mobile edge computing-based anomaly detection in cellular networks," *IEEE Transactions on Industrial Informatics*, vol. 16, pp. 1551–3203, 2019.

[18] W. Sun, J. Liu, and Y. Yue, "AI-enhanced offloading in edge computing: when machine learning meets industrial IoT," *IEEE Network*, vol. 33, no. 5, pp. 68–74, 2019.

[19] C. X. Xue, T. W. Chang, and T. C. Chang, "Embedded 1-Mb ReRAM-based computing-in-memory macro with multibit input and weight for CNN-based AI edge processors," *Sensors*, vol. 9, no. 9, pp. 1–13, 2019.

[20] J. M. Cecilia, J.-C. Cano, J. Morales-García, A. Llanes, and B. Imbernón, "Evaluation of clustering algorithms on GPU-based edge computing platforms," *Sensors*, vol. 20, no. 21, pp. 6335–6346, 2020.

[21] M. S. Rahman, I. Khalil, M. Atiquzzaman, and X. Yi, "Towards privacy preserving AI based composition framework in edge networks using fully homomorphic encryption," *Engineering Applications of Artificial Intelligence*, vol. 94, pp. 103737–103756, 2020.

[22] W.-C. Chien, H.-Y. Weng, and C.-F. Lai, "Q-learning based collaborative cache allocation in mobile edge computing," *Future generation computer systems*, vol. 102, no. 1, pp. 603–610, 2020.

[23] S. S. Dash, P. C. B. Naidu, and R. Bayindir, "Advances in intelligent systems and computing artificial intelligence and evolutionary computations in engineering systems," *Engineering Applications of Artificial Intelligence*, vol. 46, pp. 473–483, 2018.

[24] M. Tamboli, S. Limkar, and M. Kalbande, "Energy efficient distributed topology control technique with edge pruning," in *Advances in Intelligent Systems and Computing*, pp. 527–535, Springer, Cham, 2015.

[25] A. Narayanan, A. S. De Sena, D. Gutierrez-Rojas et al., "Key advances in pervasive edge computing for industrial Internet of Things in 5G and beyond," *IEEE Access*, vol. 8, pp. 206734–206754, 2020.

[26] V. Gezer, J. Um, and M. Ruskowski, "An introduction to edge computing and a real-time capable server architecture," *International Journal of Intelligent Systems*, vol. 11, no. 1&2, pp. 105–124, 2018.

[27] S. H. K. Veni and L. P. Suresh, "An analysis of various edge detection techniques on illuminant variant images," *Advances in Intelligent Systems and Computing*, vol. 325, pp. 521–532, 2015.

[28] Z. Xiong, N. Xiao, and F. Xu, "An equivalent exchange based data forwarding incentive scheme for socially aware networks," *Journal of Signal Processing Systems*, vol. 93, no. 1, pp. 1–15, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5589711, 9 pages
https://doi.org/10.1155/2021/5589711

Research Article

ColorEmo: Color Hunt with Affective Words

Xiaohui Wang, Yi Song, Yue Chen, Chunduo Su, and Danqi Hu

School of Mechanical Engineering, University of Science and Technology Beijing, 100083, China

Correspondence should be addressed to Xiaohui Wang; wangxh14@ustb.edu.cn

Received 22 January 2021; Revised 26 March 2021; Accepted 5 April 2021; Published 16 April 2021

Academic Editor: Wei Wang

Color is a main channel to express emotions of an image. Selecting a good color scheme with color harmony and proper emotional expression is very important in many fields, such as design. There are many color palette tools to help users search, create, edit, save, and share color themes, but few tools are from the perspective of emotional expression and explicitly address the emotional effects of colors. In this paper, a color hunt system with affective words called ColorEmo is developed. Multiple types of input, including affective words, affective categories in image-scale space, and main colors, are allowed for users with accurate emotional description. Based on the dataset with 428,924 color themes, the system provides rich candidates. The system is designed for user research experiments. The user-friendly interactions are offered for easy color modifications. The affective matching and color harmony are evaluated in real time while changing colors. This system can be used in many scenarios of designs and applications.

1. Introduction

The use of colors is everywhere. With the proliferation of Cyberphysical Systems (CPSs) [1–4], finding a set of colors for matching different scenarios of practical applications can be tricky. Much of the feelings about an image are expressed by colors even if the context is also important. Colors can make us feel happy or sad or nervous or relaxed. Based on the study of human color perception, there are many color harmony rules to describe the matching relationship among colors in a color theme, and some color schemes are provided. For ordinary users, even designers, it is difficult to find the proper color theme when using color palettes. So, there are many color palette tools to help users find color themes from the perspective of different user demands. But few color palette tools are from the perspective of emotional expression and explicitly address the emotional effects of colors.

A color theme is a template of colors to describe the color composition of an image [5]. The most commonly used color themes are the 3-color theme with three colors and 5-color theme with five colors. In this paper, a 5-color theme with richer expression is adopted. To quantify emotions and establish the relationship between emotions and color themes, the image-scale space with two dimensions warm-cool and soft-hard is adopted [6], which has been applied to many applications [7–9].

In this paper, a color hunt system called "ColorEmo" is built to find proper color themes with affective words. First, how to represent the user's emotional expression as an accurate input is a tough question. The most intuitive input type is an affective word. Since it is sometimes difficult to find a word accurately, some emotional categories and affective words distributed in the image-scale space are provided to give instructions to users. Based on the affective input, abundant color themes are provided based on the dataset with 428,924 color themes and 180 affective words labeled by image-scale coordinates. The semantic similarity calculation is adopted to give the image-scale coordinates for any affective words. The interactions in ColorEmo are designed for easy operations, and the information is kept on one page as much as possible to avoid the cognitive load caused by page jumps. When users modify the colors, the affective matching and color harmony evaluation results based on the assessment algorithm are shown in real time. Furthermore, the recommendation and collection functions are provided in the system.

The main contributions include the following:

(i) A complete interactive system called "ColorEmo" is developed based on user research to find proper color themes based on affective words

(ii) An assessment algorithm is proposed to evaluate the affective matching and the color harmony of color themes, giving instructions when the users tune colors

The rest of the paper is organized as follows. Section 2 gives related work. Section 3 describes the user research for color hunt, and seven functions are obtained for system design. Section 4 shows the system design and dataset of ColorEmo. Section 5 presents the ColorEmo system and the assessment algorithm. Section 6 finally draws the conclusions.

2. Related Work

2.1. Color Palette Tools. There are many color palette tools to help users search, create, edit, save, and share color themes. Most of these tools provide popular colors and color palettes by the way of waterfall flow. The color palette contains five colors with HEX, HSV, or RGB representations, such as Color-hex [10], COLOURlovers [11], and Coolors [12]. Users can select suitable color themes to match with their application expectations, create their own palette based on the selected color theme, appreciate, share, and save other color themes.

Searching hundreds of thousands of color themes one by one is time-consuming. Some tools apply rules or provide semantics to narrow the search results. The color harmony rule is commonly used, which contains analogous, monochromatic, triad, complementary, compound, shades, and custom. One of the most authoritative tools Adobe Color [13] applies the color harmony rule to restrict the range of color adjustment. Color Hunt [14] provides color palettes arranged by new, hot, popular, and random. Color Palettes [15] divides color themes into four categories, warm, cool, pastel, and contrasting, in which each color theme is extracted from an image and contains a description. "I want hue" [16] provides the differentiation report and color vision deficiency by calculating the distances between all color pairs in the palette with some color difference metrics, such as Euclidean distance, Protanope distance, Deuteranope distance, and Tritanope distance.

Despite there being so many color themes related to different aspects, very few color themes explicitly address the emotional effects of colors. In this paper, a system is proposed to search color themes with affective words.

2.2. Color and Emotions. Colors can make us feel happy or sad or nervous or relaxed. How to quantitatively describe emotions is the key question in affective computing. The quantitative methods of emotions can be divided into two categories, the category method and the dimension method. In the category method, the emotions are divided into some categories, such as the six basic emotions proposed by Darwin et al. (happy, sad, surprise, fear, anger, and disgust) [17] and four positive emotions (amusement, awe, contentment, and excitement) and four negative emotions (anger, disgust, fear, and sad) proposed by Machajdik and Hanbury [18]. In the dimension method, the emotions are distributed in the emotional space, such as the PAD model with pleasure, arousal, and dominance and three orthogonal dimensions [19].

One of the most famous two-dimensional emotional systems widely used in art design is the image-scale space proposed by Kobayashi [6, 20]. It has two dimensions: warm-cool and hard-soft and contains some categories, such as romantic and clear. Based on the long-term psychophysics experiments, Kobayashi collected the 180 affective words and 490 color themes, which are distributed on the image-scale space, shown in Figure 1. The image-scale space is commonly used for emotional analysis in some applications [7–9].

2.3. Color Harmony. In art design, there are many color harmony rules and color harmony algorithms [21, 22]. Itten introduced a kind of color wheel to describe color harmony, especially on hue [5]. For example, complementary color schemes are that colors opposite each other on the color wheel, like red and green. Monochromatic color schemes are based on different tones of the same color. Analogous color schemes use colors that are next to each other on the color wheel. Triadic color schemes use colors that are evenly spaced around the color wheel. Rectangle or tetradic color schemes are four colors arranged into two complementary pairs. Matsuda summarized a set of 80 color schemes [23]. In this paper, we refer to these color schemes and quantify the color wheel to describe the unified degree in the assessment algorithm.

The hue channel mainly affects the color harmony [5]. Tokumaru et al. described how saturation and luminance channels affect color harmony [24]. In the assessment algorithm of this paper, the color harmony degree based on saturation and luminance channel is described as "contrast" and "saturation" by using rules as simple as possible.

The color harmony rules have been used in many color-related applications [25–27]. Cohen-Or et al. applied the harmonic models for image color harmonization [28]. Classic color harmony theory combined with a new deep learning method is used to evaluate the beauty, sentiment, and remembrance of art [29].

3. Method

As a designer, it is important to understand the psychological effects colors might have on an average person or your client's target audience. In the user research, we use the in-depth interview method to collect opinions and suggestions from designers for the color hunt. To collect information from the perspective of different design fields, designers in four areas (graphics, UI, original painting, and photography) are our target subjects of our in-depth interview.

The content of the in-depth interview includes two aspects: first, the interviewer directly asks the subjects questions; second, the subjects provide their commonly used

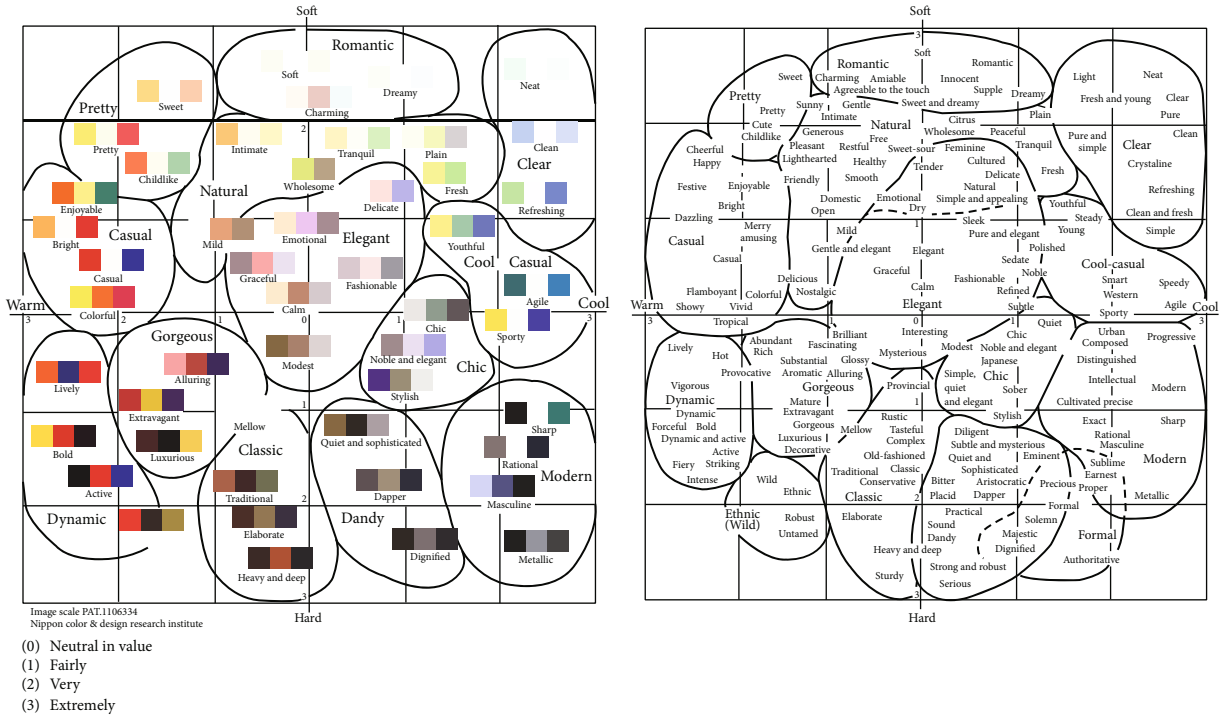


Figure 1: The image-scale space [6], a two-dimensional emotional system widely used in art design.

color palette tools and evaluate them. The questions are summarized into ten categories as follows.

(1) How do you select the color theme in normal work

(2) Do you think your color theme selection habits are good? What problems have you encountered when selecting color themes

(3) How do you filter the candidates when selecting a color theme? Describe the factors in detail

(4) Which color palette tools do you often use

(5) Do you know the color harmony rules? Give the rules in detail

(6) Do you consider the emotional factors when selecting color themes

(7) What functions would you expect if there is an emotion-based color hunt tool

(8) Do you usually collect some good color themes

(9) Do you collect some beautiful pictures to extract color themes from these pictures

(10) What is the interactive process when selecting color themes in the color palette tools

The in-depth interview results are summarized, shown in Figure 2.

The user demands and the expected interactions are extracted based on the results. Finally, the seven functions of the ColorEmo system are obtained, which are the basis of the system design. We introduce how to obtain each system function briefly as follows.

3.1. Input Types. The subjects mention that there are few tools to search color themes by emotions, but they want such a tool. Besides, they do not know how to give input. So, "emotion-based input" is obtained as the user demands. Multiple types of input are expected in the system interaction. Therefore, the input types should be well designed.

3.2. Color Theme Dataset. There are not enough candidate colors in some color palette tools. Some color themes lack science and authority. When designers select a favorite color theme, they do not know whether it meets the public's preference. An authoritative color theme dataset is very important, which is the data foundation of the system. In our system, there should be abundant and popular color themes with accurate labels.

3.3. Candidate Color Theme Display Form. When using the color palette tools, browsing and selecting color themes are the most time-consuming. And getting a satisfactory color theme requires fussy interactions. The subjects expect simple interactions. The key to reducing the number of operations is the user-friendly candidate color theme display form.

3.4. Modify Colors. There are some inflexible ways to modify color themes. In some tools based on certain color harmony rules, candidates cannot be modified to arbitrary colors to ensure that the harmony rules are not broken. The flexible modification of candidates is needed, and there are no restrictions on color modification. Based on the proposed assessment algorithm, when the emotion of the modification

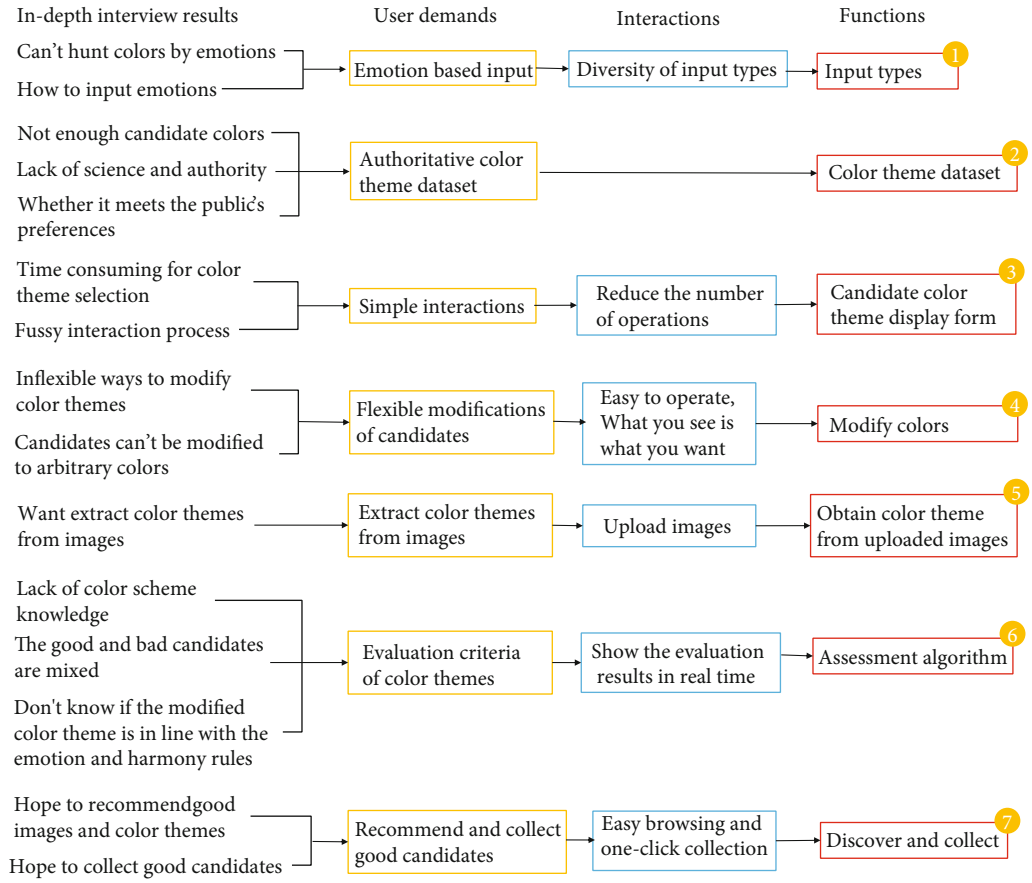


Figure 2: Summary of the user research.

is far from the input affective word, the assessment value will be very low to remind users. The need for interaction is that what you see is what you want. It is possible that an unintentional modification produces a very satisfactory result.

3.5. Obtain Color Theme from Uploaded Images. The subjects want to extract a color theme from their favorite images, which is also a function in many tools. In our system, we keep this function.

3.6. Assessment Algorithm. Most users, even designers, lack color scheme knowledge. They are hard to distinguish the good and bad candidate color themes, which are always mixed in most systems. Furthermore, after they modify the candidate color theme, they do not know whether it is in line with the emotion and color harmony rules. So, the evaluation criteria of color themes are needed. When users modify the colors, the evaluation results should be shown in real time. Therefore, the assessment algorithm to evaluate how well the color theme matches the emotion and color harmony rules is the key function of the system.

3.7. Discover and Collect. The subjects want the system to recommend good images and color themes and let them collect good candidates. To discover good images and color themes and provide a collection function is a must-have feature in a complete system. The interaction for discovering good candidates should be easy browsing. The interaction for collection is a one-click operation.

4. System Design

4.1. System Flow. According to the summarized functions based on the user research, we design the ColorEmo system, and the system flow is shown in Figure 3. The color theme generation flow is proposed to find the appropriate color theme with user demands.

The input of the system can be an affective word and a main color. The input affective word is selected from the image-scale space or directly entered. Based on our previous work [8], 428,924 color themes labeled by image-scale coordinates are collected in the dataset of ColorEmo. 180 affective words labeled by image-scale coordinates [6] are adopted. To give feedback for any user input, we use the semantic similarity calculation based on Hownet to obtain the coordinates of the input based on the 180 words. So, for any affective word, hundreds of color themes are filtered from the dataset. Besides, color themes can also be obtained from uploaded images like most color palette tools.

For simple interaction, hundreds of color themes are displayed by a slider instead of waterfalls to avoid the cognitive load caused by page jumps, which will reduce the number of operations. Furthermore, display color schemes by geometry for each color theme intuitively. To modify colors flexibly, we

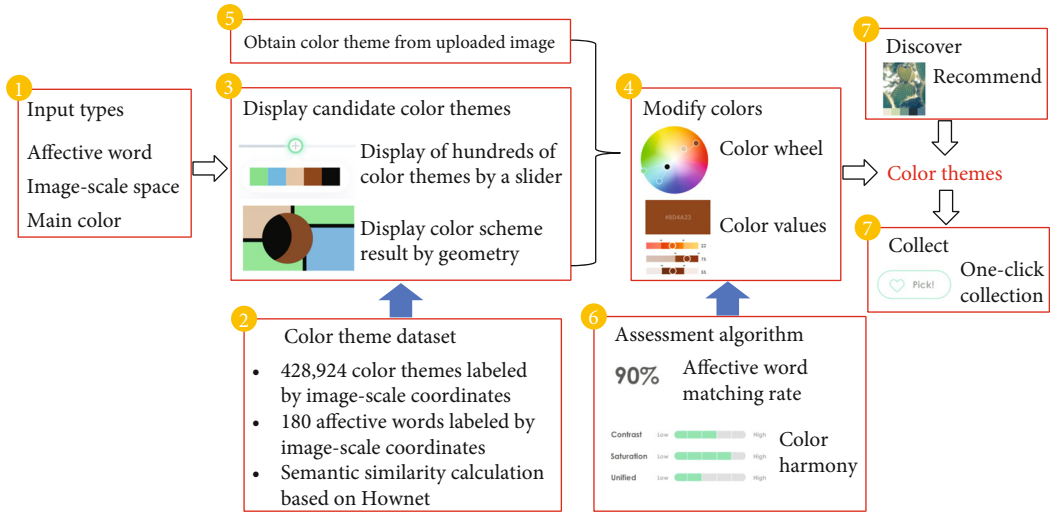


Figure 3: The system flow of ColorEmo.

use the color wheel to adjust colors roughly and to modify color values for fine adjustment while visualizing their harmonious relationship in the wheel. While changing colors, the evaluation results of the affective matching and the color harmony are shown in real time based on our proposed assessment algorithm. Therefore, the color values, the distribution of five colors in the wheel, the color scheme result rendered on the geometry, and the evaluation results are all obtained in real time, which realize that what you see is what you want.

When the satisfied color theme is obtained after modification, you can collect it. Besides, the system recommends lots of good images and their color themes.

4.2. Dataset. In our previous work [8], we collect more than 420,000 color themes from Adobe Kuler [30] (44,986 themes) and COLOURLovers [11] (383,938 themes) which are famous online communities to create and share color designs. The mapping between these color themes and image-scale space is built [8]. In this paper, we use mapping to find abundant color themes based on affective words. So, in the ColorEmo system, there are more than 420,000 color themes labeled by the two-dimensional coordinates in the image-scale space.

When users input an affective word, the two-dimensional coordinates of the word in the image-scale space are calculated first. Then, the color themes are obtained based on the image-scale coordinates from the above dataset. When users select an affective word from the image-scale space, the image-scale coordinates are obtained directly from Kobayashi's research [6]. If users input an arbitrary affective word which is not in the set of the 180 words collected by Kobayashi [6], the image-scale coordinates are calculated based on Hownet [31]. Hownet is a Chinese dictionary to reveal the relationship between concepts and the attributes of concepts, which is similar to WordNet in English [32]. We use Hownet to calculate the semantic similarity between the input word and the 180 words. The top n most similar words are selected, and the coordinates of the input word are the weighted average of the coordinates of the n words. After the experiments, n is set to 3.

5. ColorEmo System

The homepage of ColorEmo is shown in Figure 4. In the navigation bar at the top of the homepage, there are four buttons corresponding to four pages. The default page is the "generation" page to find and generate the appropriate color theme to meet the emotional expression of the input. The "discover" page shows the recommended color themes and images. The "assessment" page evaluates how well the color theme matches the emotion and color harmony rules. The "my pick" page shows all the user's collection while logged in.

In the "generation" page, there are two parts for user input: the affective word selection part and the main color selection part. In the affective word selection part, users can enter an affective word in the text box or select an affective category in the image-scale space. When describing emotions, it is sometimes difficult to find a word accurately. So, some affective categories [6] distributed in the image-scale space are provided to give users an intuitive concept to help them quickly describe the emotional expression they want. The color of the category is the main color distribution of this emotion, and the shape represents the location distribution. The main color can be selected on the right color table.

When a category is clicked, some affective words in this category are shown, taking the "Clear" category as an example shown in Figure 5. There are 180 words in total collected by Kobayashi [6].

After entering an affective word, press the "next" button in the bottom right corner of this page. The specific color theme generation page is shown in Figure 6. The layout of the color theme generation page focuses on the spatial and temporal relationship of the interaction [33].

For each function summarized from the user research, the interface of the system is carefully designed, highlighted by rectangles and numbers, which is consistent with the labels in Figure 2. Figures 4 and 5 are designed for system

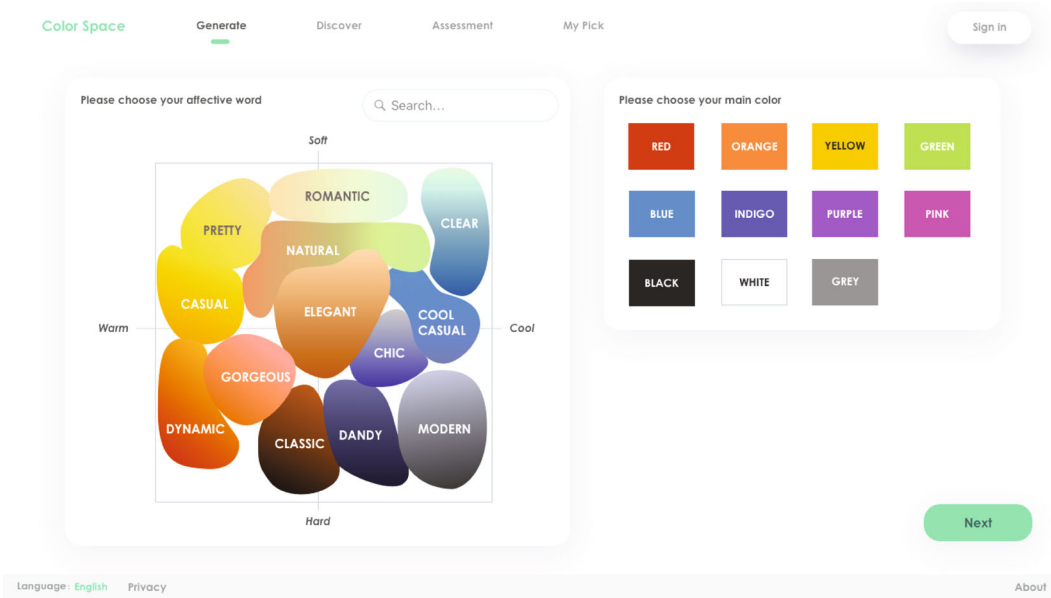


Figure 4: The homepage of ColorEmo.

Figure 5: The word distribution of "Clear" category.

function 1 "input types." The color theme dataset of function 2 is the data foundation of the system and is not directly displayed on the interface.

Instead of using a waterfall to show all the candidates, a slider is adopted to carry them shown in the box labeled by 3. Users slide the slider left and right to see all the candidates. If you like it, press the add ("+") button to collect it temporarily. All the collections are displayed in the box under the slider. Having all the candidates hidden behind the slider is to keep the information on one page as much as possible to avoid the cognitive load caused by page jumps. Besides, to show the color scheme effect of each color theme intuitively, we adopt an abstract geometry with random coloring by the color theme.

The main area of the interface is used for color modification. The color wheel is the basic tool for combining colors. In the ColorEmo system, the color wheel has two functions, one is to adjust colors roughly, and the other is to distribute the five colors in the wheel to visualize their harmonious relationship. Besides, the color values of each color under the color wheel are used for fine adjustment. The values in different color spaces are provided. Modification results are displayed in real time, which achieves that what you see is what you want.

Similar to most color palette tools, color themes can be extracted from images. In the ColorEmo system, the entrance of reference picture uploading is shown in the upper left of the interface.

Based on the assessment algorithm introduced in detail in subsection "assessment algorithm," the affective matching rate is shown in the assessment part in the right of the interface. Contrast, saturation, and unified evaluate how well it matches the color harmony rules. The color harmony value is quantified as five levels, namely, very low, low, medium, high, and very high. The evaluation results are shown in real time to give users instructions while changing colors.

If you like the color theme, you can click the "pick" button in the upper right of the interface. The color theme appears to the left of the button again for the user to confirm this color theme conveniently.

The "discover" page of the ColorEmo system is shown in Figure 7. There are lots of images, and their color themes are automatically recommended based on the user's input.

The "my pick" page shows all the user's collection while logged in. The interface is in the form of tiled favorite color themes and images, which is similar to the "discover" page.

5.1. Assessment Algorithm. The assessment algorithm includes two parts, the affective matching rate and the color harmony degree. Based on the color theme-image-scale coordinate relationship model in our previous work [8], the image-scale coordinates for any modified color themes can

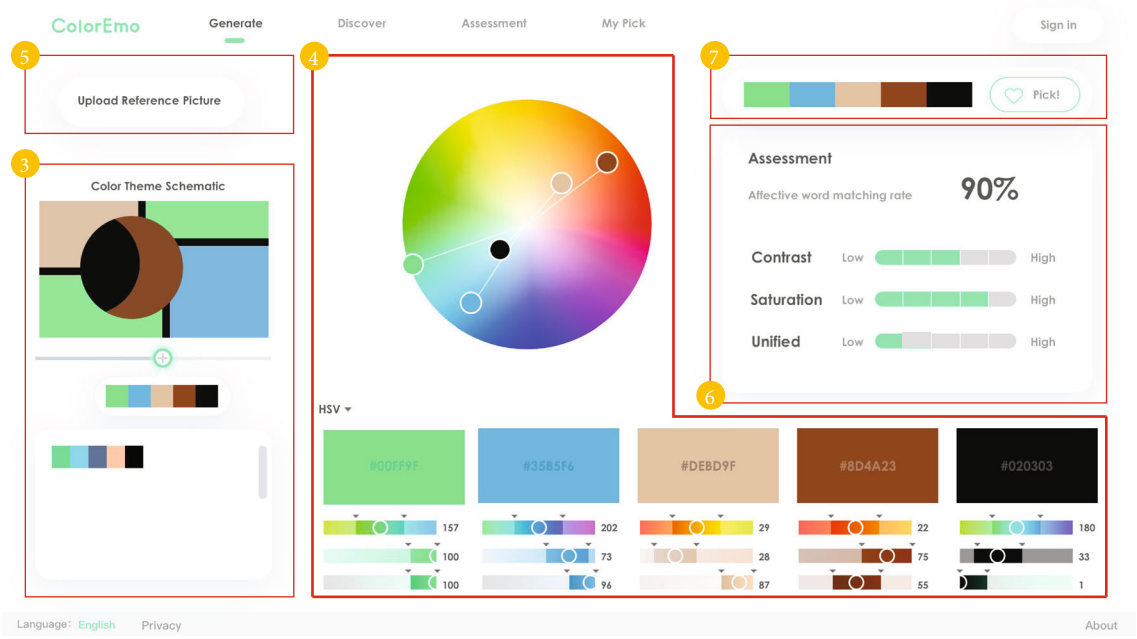


Figure 6: Color theme generation page.

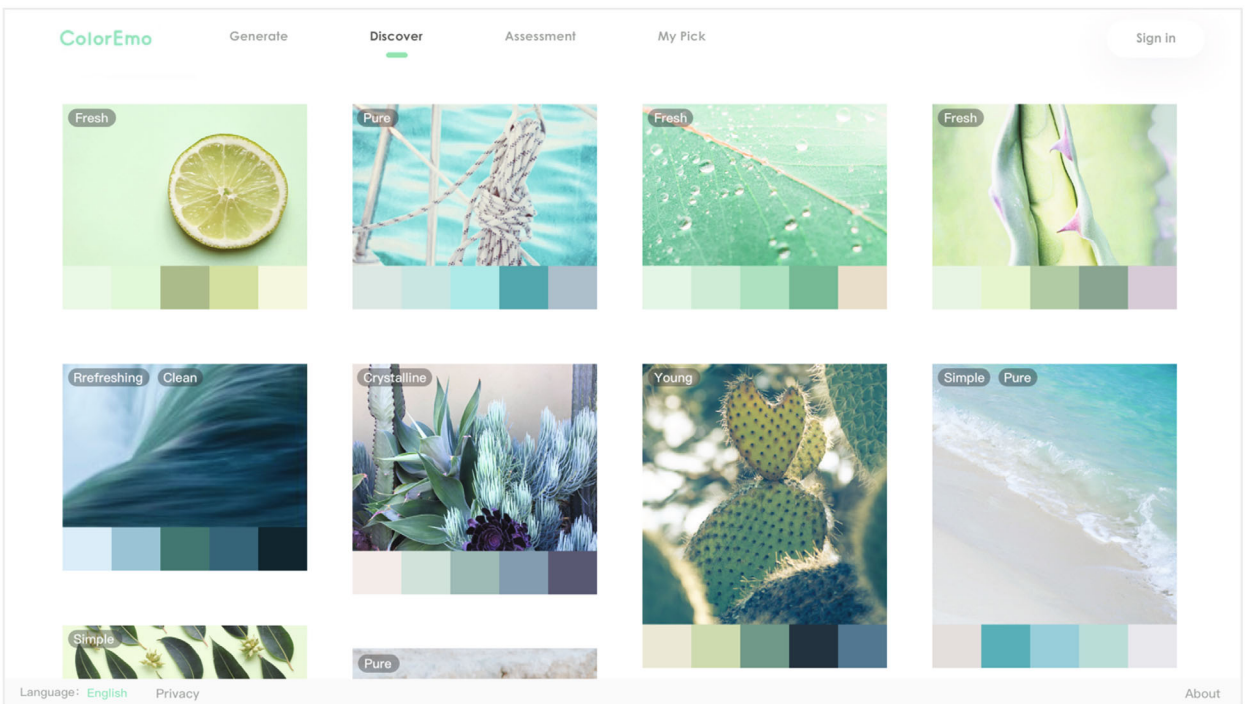


Figure 7: The discover page of ColorEmo.

be calculated. The affective matching rate is the Euclidean distance of coordinates between the input affective word and the modified color theme, which is normalized to be between *0* and *1*.

When applying color harmony theory to an application, a very specific theory is adopted. In this system, we adopt the color harmony theory as simple and versatile as possible to get the universal color themes, which can be subsequently applied to many applications. Contrast, saturation, and unified evaluate how well it matches the color harmony rules from luminance, saturation, and hue and three aspects in the HSL color space. The contrast degree describes the contrast of both luminance and saturation of the five colors in a color theme. The saturation degree reflects the average of the five saturations. The unified degree describes the distribution and contrast of the hue channels of the five colors. The color harmony value is quantified as five degrees, namely, very low, low, medium, high, and very high.

Based on Itten's contrasts in color [5], the color contrasts include the pure color (hue) contrast, the light-dark contrast, the cold-warm contrast, the complementary contrast, the simultaneous contrast, the contrast of quality (color saturation), and the contrast of quantity. In this paper, the contrast is a combination of the light-dark contrast and the contrast of quality, which is quantified by luminance and saturation. Without considering black, white, and gray colors, the maximum absolute value of the difference in saturation and luminance between two of the five colors in a color theme is set as v. If v is less than 45, the contrast degree is very low. If v is between 45 and 60, the contrast degree is low. If v is between 60 and 75, the contrast degree is medium. If v is between 75 and 90, the contrast degree is high. If v is more than 90, the contrast degree is very high.

The saturation degree evaluates the average saturation of five colors. The calculation method is to extract the saturation channel of each color and to take the average value of the five saturations. If the average is less than 20, the saturation degree is very low. If the average is between 20 and 40, the saturation degree is low. If the average is between 40 and 60, the saturation degree is medium. If the average is between 60 and 80, the saturation degree is high. If the average is more than 80, the saturation degree is very high.

In the RYB (or subtractive) color model, red, yellow, and blue are the primary colors. Green, orange, and purple, created by mixing two primary colors, are three secondary colors. Six tertiary colors are created by mixing primary and secondary colors. If the colors in a color theme include three different primary colors, the unified degree is very low. The unified degree describes the distribution and contrast of the hue channels of the five colors, which is less intuitive than luminance and saturation. The largest angles (180 degrees) in the color wheel are divided into five parts. The maximum absolute value of the difference in hue between two of the five colors in a color theme is set as h. If h is less than 36, the unified degree is very high. If h is between 36 and 72, the unified degree is high. If h is between 72 and 108, the unified degree is medium. If h is between 108 and 144, the unified degree is low. If h is more than 144, the unified degree is very low.

6. Conclusion

In this paper, a complete system called "ColorEmo" is proposed to hunt colors based on affective words. More than 400,000 five-color themes with image-scale coordinates and 180 affective words are data support of this system. Based on the user research experiments, seven system functions are summarized to find the accurate color theme with user demand. In the interactions of the system, slider-based color theme display, geometry-based color scheme display, and color wheel-based modification are designed for user-friendly color adjustment to reduce the cognitive load. While modifying colors, the affective matching rate and color harmony degrees are shown in real time based on the proposed assessment algorithm. The recommendation and one-click collection functions are also provided in the system. In the future, we will design more applications by incorporating other fields [34–37] based on the system.

Data Availability

Data is not open.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] X. Han, G. J. Shen, X. Yang, and X. J. Kong, "Congestion recognition for hybrid urban road systems via digraph convolutional network," *Transportation Research Part C*, vol. 121, article 102877, 2020.

[2] X. Kong, S. Tong, H. Gao et al., "Mobile edge cooperation optimization for wearable internet of things: a network representation-based framework," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 5050–5058, 2020.

[3] W. Wang, F. Xia, H. Nie et al., "Vehicle trajectory clustering based on dynamic representation learning of internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020.

[4] W. Wang, X. Zhao, Z. Gong, Z. Chen, N. Zhang, and W. Wei, "An attention-based deep learning framework for trip destination prediction of sharing bike," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020.

[5] J. Itten, *The Art of Color*, Van Nostrand Reinhold Company, New York, 1960.

[6] S. Kobayashi, *Color Image Scale*, Kosdansha International, 1991.

[7] Y. Ma, J. Jia, S. Zhou, J. Fu, Y. Liu, and Z. Tong, "Towards better understanding the clothing fashion styles: a multimodal deep learning approach," in *AAAI Conference on Artificial Intelligence*, pp. 38–44, San Francisco, California, USA, 2017.

[8] X. Wang, J. Jia, and L. Cai, "Affective image adjustment with a single word," *The Visual Computer*, vol. 29, no. 11, pp. 1121–1133, 2013.

[9] M. Solli and R. Lenz, "Color semantics for image indexing," in *European Conference on Colour in Graphics, Imaging, and Vision*, pp. 353–358, Joensuu, Finland, 2010.

[10] ColorHex, *Color-hex [DB/OL]*, 2020, http://www.color-hex.com/color-palettes/.

[11] COLOURlovers, *COLOURlovers*, 2020, http://www.colourlovers.com/palettes.

[12] F. Bianchi, *Coolors*, 2020, https://coolors.co/.

[13] AdobeColor, *Adobe Color*, 2020, http://color.adobe.com/.

[14] Colorhunt, *Color Hunt*, 2020, http://colorhunt.co/.

[15] Colorpalettes, *Color Palettes*, 2020, http://colorpalettes.net/.

[16] *I Want Hue*, 2020, http://tools.medialab.sciences-po.fr/iwanthue/.

[17] C. Darwin, P. Ekman, and P. Prodger, *The Expression of the Emotions in Man and Animals*, Oxford University Press, 2002.

[18] J. Machajdik and A. Hanbury, "Affective image classification using features inspired by psychology and art theory," in *International conference on Multimedia*, pp. 83–92, Firenze, Italy, 2010.

[19] A. Mehrabian, "Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in

temperament," *Current Psychology: Developmental, Learning, Personality, Social*, vol. 14, no. 4, pp. 261–292, 1996.

[20] S. Kobayashi, *Art of Color Combinations*, Kosdansha International, 1995.

[21] E. Jeon, Y. Han, and M. Nam, "How you see yourself influences your color preference: effects of self-construal on evaluations of color combinations," *Psychology & Marketing*, vol. 37, no. 7, pp. 980–994, 2020.

[22] M. Zaeimi and A. Ghoddosian, "Color harmony algorithm: an art-inspired metaheuristic for mathematical function optimization," *Soft Computing*, vol. 24, no. 16, pp. 12027–12066, 2020.

[23] Y. Matsuda, *Color Design*, Asakura Shoten, 1995.

[24] M. Tokumaru, N. Muranaka, and S. Imanishi, "Color design support system considering color harmony," in *IEEE International Conference on Fuzzy Systems*, pp. 378–383, Honolulu, HI, USA, 2002.

[25] C. Li and T. Chen, "Aesthetic visual quality assessment of paintings," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 236–252, 2009.

[26] L. Wang, J. Giesen, M. D. KT, P. Zolliker, and K. Mueller, "Color design for illustrative visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1739–1754, 2008.

[27] S. Odabasioglu and N. Olgunturk, "Effect of area on color harmony in simulated interiors," *Color Research and Application*, vol. 45, no. 4, pp. 710–727, 2020.

[28] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y. Q. Xu, "Color harmonization," *ACM Transactions on Graphics*, vol. 25, no. 3, p. 624, 2006.

[29] E. Cetinic, T. Lipic, and S. Grgic, "A deep learning perspective on beauty, sentiment, and remembrance of art," *IEEE Access*, vol. 7, pp. 73694–73710, 2019.

[30] AdobeKuler, *Adobe Kuler*, 2020, http://kuler.adobe.com/.

[31] Z. Dong and Q. Dong, *HowNet and the Computation of Meaning*, World Scientific Publishing Company, Singapore, 2006.

[32] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller, *Introduction to WordNet: An On-Line Lexical Database*, Oxford University Press, 1997.

[33] Y. M. Li, R. Q. Liang, W. Wei, W. Wang, J. T. Zhou, and X. Li, "Temporal pyramid network with spatial-temporal attention for pedestrian trajectory prediction," *IEEE Transactions on Network Science and Engineering (T-NSE)*, 2021.

[34] Z. Wang, X. Ye, C. Wang, J. Cui, and P. S. Yu, "Network embedding with completely-imbalanced labels," *IEEE Transactions on Knowledge and Data Engineering*, 2020.

[35] Z. Wang, X. Ye, C. Wang, Y. Wu, C. Wang, and K. Liang, "RSDNE: exploring relaxed similarity and dissimilarity from completely-imbalanced labels for network embedding," in *AAAI Conference on Artificial Intelligence*, vol. 32, New Orleans, Louisiana, USA, 2018.

[36] Y. M. Li, J. T. Zhou, J. Y. Tian, X. W. Zheng, and Y. Y. Tang, "Weighted error entropy based information theoretic learning for robust subspace representation," *IEEE Transactions on Neural Networks and Learning Systems (T-NNLS)*, 2021.

[37] Y. M. Li, J. T. Zhou, X. W. Zheng, J. Y. Tian, and Y. Y. Tang, "Robust subspace clustering with independent and piecewise identically distributed noise modeling," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, California, USA, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5586662, 11 pages
https://doi.org/10.1155/2021/5586662

Research Article

Design and Implementation of Intelligent English Electronic Dictionary System Based on Internet of Things

Wenming Yong

School of Foreign Languages, Xinxiang Medical University, Xinxiang, Henan 453003, China

Correspondence should be addressed to Wenming Yong; yongwenming@xxmu.edu.cn

Received 19 February 2021; Revised 22 March 2021; Accepted 31 March 2021; Published 12 April 2021

Academic Editor: Wei Wang

In this paper, the intelligent English electronic dictionary system is studied to design and implement the electronic dictionary system according to the advantages of the Internet of Things. The software architecture, the design, and implementation of the client and server-side and related technologies in the development process of the dictionary application are used as the research content to comprehensively discuss the development process of the electronic dictionary. The client and server-side is based on C/S technology architecture, and the server-side is a standard Maven Web project, which is managed by Maven and does not cause conflicts; the model-view-controller framework is built using Spring MVC to achieve the separation of user interface and application logic. Spring MVC is used to build a model-view-controller framework to separate user interface and application logic. Spring dependency injection is used to build a loosely coupled project, which helps to separate project components; Spring Data JPA is used to build a persistence layer to facilitate data access and maximize the developer's ability to automatically realize logical operations on data. After the overall performance test of the system, the performance is good under the platform, and the intelligence of trilingual word query is achieved, and the quickness and ease of use meet the requirements that can be applied.

1. Introduction

With the continuous development of modern human information society and the progress of high-tech innovation level, various new smart wearable devices have entered our daily life, and human society has completely entered the information age [1]. With the development of modern mobile Internet and network communication technology, the functions of smartphones are also becoming increasingly powerful. Internet of Things (IoT), a new and promising technology, deals with several devices that interact through the Internet, ensuring that smart terminals such as wearables, sensors, and cell phones can be connected to smart nodes that can communicate with each other [2]. The Internet of Things interconnects "objects" and enables machine-to-machine (M2M) communication, which means that heterogeneous devices can communicate with each other without human intervention [3]. Given the recent advances in ubiquitous computing technologies, there are many IoT applications for different environments that promise to enhance and improve the quality of users' daily lives, such as smart homes, smart cities, smart industries, and smart healthcare, which have different characteristics and they also have different latency and data rate requirements. For offline word search function depends on the local dictionary, and the update of local dictionary depends on the server-side update, this update process is usually slow and the update method is relatively simple, directly replacing the original dictionary. To achieve value-added services, many dictionary software provides article translation but charges for the high-quality translation provided, and there is also the phenomenon of checking the nondesired and poor quality of explanation. In addition to the core word search function, customer needs are positioned differently, too many advertisements or fees are pushed, and user experience is poor, so the software has limited vitality [4].

Users can create multiple dictionaries to enable multilingual search functions. By entering entries, definitions, and example sentences for different thesauri, multilingual word search is possible. Also, the user can view the collected

vocabulary off-site, i.e., after changing devices and logging in, which is convenient for the user [5]. Also, most of the current papers on electronic dictionaries are limited to discussing the implementation of the client-side functions and do not discuss the server-side implementation in detail. In this paper, in addition to introducing the client-side functions, a feasible solution is given for the server-side implementation. The software should be based on the principle of saving memory and blocking advertisements, thus increasing the vitality of the software. The specific measure is to use tab layout to achieve convergence when implementing client-side functions, with each tab corresponding to a functional module, thus reducing the amount of program code [6]. Different applications in the IoT environment can use small base stations to reduce the traffic load from macrocells, fill the coverage blind spots of macro base stations, significantly increase the system capacity, and improve the energy efficiency of the whole network. Also, small base stations are considered as an effective solution for handling massive data traffic in next-generation cellular networks due to their low transmission power and small size, which allow flexible site acquisition and thus are more cost-effective than traditional macro base stations in meeting the same traffic demand.

An IoT application traffic detection method based on deep learning. This method uses Wireshark for structured data storage of traffic data packets on IoT devices and cuts and filters Pcap traffic packets using the Split Cap tool to obtain different cuts of traffic packets Samples and, then, convert the traffic data with different fragment sizes into a binary graph file representation, and a more accurate application traffic identification can be obtained through the deep learning method. Therefore, a set of three-dimensional dense network planning schemes is needed to meet the coverage demand, capacity demand, and communication quality demand of users in a certain planning area, minimize the cost, and quickly and effectively deploy small base stations in dense network scenarios. In recent years, the network electronic dictionary based on modern multimedia network technology has been different from the traditional dictionary in form. Therefore, the main functions and technical advantages of the system include text fuzzy automatic query, automatic language recognition, fast automatic search, automatic database storage, and smartphone adaptation. It can provide good language communication language, and culture education is the most important component of its core. The current development of education in the history of higher education is in the beginning and development stage, higher education and facilities gradually improved, a higher level of postgraduate students at the higher education level, the proportion of postgraduate students in the total number of rising, and examples of the application of the dictionary, for the promotion of interethnic language and culture exchange and development has a positive role and significance.

2. Analysis of IoT Intelligent English Electronic Dictionary System

2.1. Related Work. As we all know, before the emergence of electronic dictionaries, people used a lot of paper-based word search tools. In terms of Chinese-English dictionaries, there are various versions of dictionaries, such as the Longman Dictionary [7]. Traditional paper-based tools are very inconvenient; in this case, people began to develop electronic dictionaries like calculators [8]. This hardware-based electronic dictionary has a display screen and an input keyboard and is relatively small and easy to carry [9]. The electronic dictionary of this period mainly provides simple words or phrases interpretation. With the development of electronic hardware technology, the computing speed and storage space of hardware electronic dictionaries have also been greatly improved, and there are electronic dictionaries with richer contents and more complete functions on the market. The users of electronic dictionaries are mainly young students [10]. It is launched by some traditional English learning websites, such as Jinshan Word Search. It provides online network word search and provides a simple offline word search function [11]. With China's full access to the Internet, computers became popular and more people started to use PCs to look up information from the Internet [12].

The current development technology of electronic dictionaries mainly uses Android [13]. Many scholars are concerned with the structure of the lexicon as well as the speed of word search, for example, studies based on the construction of word structures between three languages and studies on how to speed up the establishment of indexes carried out by word search [14]. Some foreign scholars on electronic dictionaries also focus on lightweight usability, which plays an auxiliary role in learning, and some applications oriented to special education, which tend to use machine learning methods in the development of technology [15]. Some companies that traditionally provide online translation provide APPs for free download to attract more users, such as Jinshan Word expert and Yadao Dictionary [16]. In terms of core functions, the current electronic dictionaries implement functions such as local word search, online word search, adding new words, and network word search [17]. Also, many electronic dictionaries have some special features. For example, word memorization function, article translation function, and human translation function.

The electronic dictionary discussed in this paper is developed based on the Android platform and is implemented using a combination of Java and MySQL, and the software is divided into a client-side as well as a server-side. To improve the viability of the software, the running efficiency of the client-side is increased based on the principle of memory saving.

2.2. IoT-Based System Design. Capacity and coverage are two key objectives for base station deployment, and the general base station deployment planning process in cellular networks starts with an analysis of the propagation model and traffic density in the target service area. The former is used to measure power coverage, where the goal is usually to ensure that the received power of users in a given area is above a threshold to correctly demodulate the expected information in a worst-case scenario, and the latter is used for capacity coverage, where the goal is that sufficient throughput can be provided for the users covered by the base

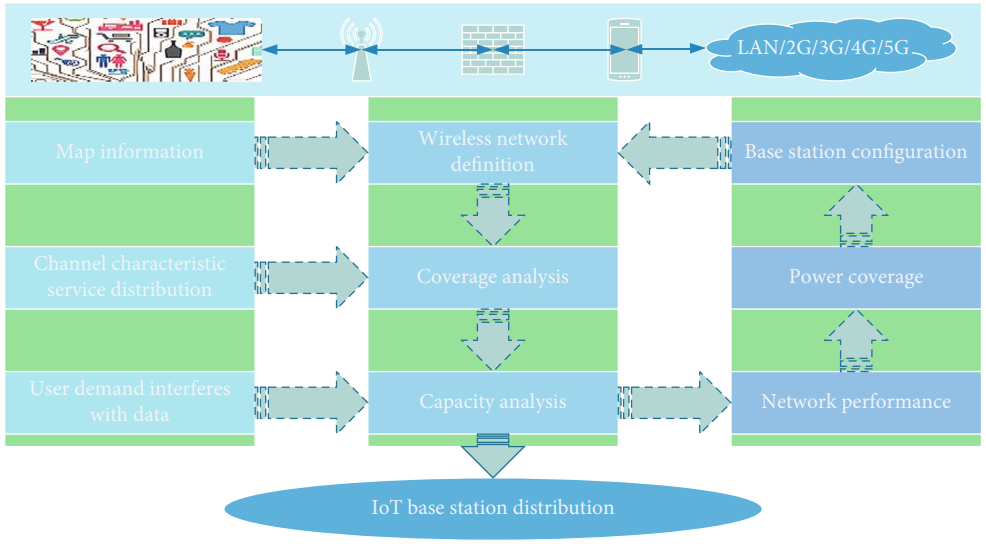


Figure 1: Flow chart of IoT base station deployment method.

station in a given area [18]. Both objectives should be achieved in the base station deployment to provide QOS-guaranteed service to the users, and from the overall system perspective, the deployment cost of the system will be minimized if both coverage and capacity can be optimally planned. However, compared to power coverage planning using proven propagation models and field measurements, capacity planning is more challenging because the traffic demand distribution is inherently dynamic and depends on user usage patterns and mobility.

The base station deployment process is shown in Figure 1. First, in the process of wireless network definition, suitable candidate sites are selected, and relevant base station parameters are configured according to the input map information; then, channel characteristics and service distribution are input, and suitable channel models are selected for coverage analysis, and then, the capacity analysis is performed according to user requirements, i.e., required data rate and interference. When the coverage and network performance cannot reach the standard, the iterative process is repeated to adjust the location and number of base stations until the coverage and network performance meet the demand, and then, the base station deployment is completed. The wireless network definition process in the base station deployment process should determine the candidate locations for deploying base stations based on terrain and population density. Coverage analysis is usually achieved by selecting enough sites from the candidate sites to deploy base stations and configuring their parameters to meet the minimum received power demand of users in a given planning area, and capacity analysis is achieved by designing the capacity allocation of base stations to meet all coverage. Capacity analysis is achieved by designing the capacity allocation of the base stations to meet the maximum rate demand of all users in the coverage area. Proper deployment of base stations not only improves the performance of the cellular network but also reduces capacity waste, reduces CAPEX and OPEX, and cost-effectively provides qualified QOS.

The exact method usually uses mathematical modelling or exhaustive search to select the global optimal solution of the base station deployment problem, such as using the simplex method or exhaustive method to find the optimal deployment location of the base station; although the result of this method is theoretically the global optimal solution, it is computationally inefficient for large-scale problems, and usually, an approximate optimal solution is sufficient without sacrificing computation time obtain a globally optimal solution. To balance the accuracy rate and efficiency, a heuristic algorithm can be used for base station planning to find a satisfactory deployment solution in a limited amount of time. Unlike heuristic algorithms, which are problem independent, i.e., do not exploit problem specificity, metaheuristics can be further divided into continuous metaheuristics, which are used to deal with continuous solution spaces, and discrete metaheuristics, which are used to deal with discrete solution spaces; although continuous metaheuristics are only applicable to continuous optimization problems, many researchers have improved them to further deal with discrete problems.

In a dense network in an indoor three-dimensional scenario, due to a large number of small base stations and the high number of floors and walls crossed between small base stations and terminals, the propagation model is chosen in this paper to describe the indoor propagation characteristics in a dense urban environment, with the path loss (in dB) expressed as

$$PL(d_m) = PL(d_0)^2 + 10m \log\left(\frac{d_{su}}{d_0}\right) + AF(D_{su})^2, \quad (1)$$

where d is the propagation distance between the mobile terminal and the small base station with which it communicates, and F is the near-ground reference distance (recommended to be 1 m) and, therefore, the first-meter loss, when the radio signal frequency is 2.4 GHz.

$$PL(d_0 = 1) = 60\text{dB}. \tag{2}$$

The expression in dB is

$$\text{AF}(d_m) = 10 \log (\mu d_m)^2. \tag{3}$$

U is determined by the environment per unit path length attenuation coefficient, generally taking the value of 0.2~0.7 (dB/m). Therefore, the numerical expression of the path loss can be derived by substituting formula (3.3) into formula (3) as follows:

$$PL(d_{sm}) = \text{K}_{su} d_{su}^{m+1}, \tag{4}$$

$$K_{su}^2 = \frac{u 10^{PL(d_0)/10}}{d_0^m}. \tag{5}$$

In the actual network planning scenario, the distribution density of users fluctuates over time and space due to the social attributes of users. In specific hotspot areas, user distribution density fluctuates more over time. For example, user density in residential areas is very high on weekends but smaller on weekdays, in contrast to work areas where user density is very high on weekdays, especially during the day. Later, we can turn on these small base stations during high traffic periods. In the low traffic period, we can also find the minimum number of small base stations that meet the current user demand and their optimal locations. Some of these small base stations have the same locations in the high and low traffic periods, and they have already been deployed in the high traffic period, so later, we only need to deploy some of the small base stations that are only turned on in the low traffic period, and the base stations that are duplicated in the high traffic period only need to be switched on.

The deep learning feature part uses the calculation of the average value of each word feature vector in the text to obtain the sentence feature vector of the same dimension. Since the vector value is in the range of real numbers, there are negative values. In the experiment, the depth feature is normalized to make its value. Both are between 0 and 1; the shallow features use the same combination of words, parts of speech, and dictionary features to obtain the corresponding feature set, and finally, the deep learning features and shallow learning features are weighted and merged, and the weight coefficient is 1 [19]. After the fusion, the features include not only the words, part of speech, and dictionary features of the shallow learning, and the calculation method remains unchanged, but also the deep learning feature vectors obtained by the tool.

Network sizing estimation is usually a comprehensive analysis from both coverage and capacity aspects to determine the required size, i.e., the number of sites, for network planning in a given area. On the one hand, coverage estimation combines link budget and propagation model to calculate the coverage radius of base stations to find the number of base stations required to cover the planning area; on the other hand, capacity estimation calculates the number of base stations required to meet all service demands by processing various actual services into some virtual equivalent services based on the resources provided by individual base stations. The larger number of required BTS in the coverage estimation and capacity estimation is taken as the number of sites for the size estimation.

$$R_{\text{SBS}} = \left\{ d_{\text{su}} \middle| PL(d_{\text{su}}) = \gamma_{\text{PL}}^2 \right\} + \text{DPL}. \tag{6}$$

The length of particles in the original particle swarm algorithm is fixed, which represents the dimension of the search space. In this paper, assuming that the number of deployed small base stations is N, each particle is a vector composed of the coordinates of N small base stations, and since the dense network planning in this paper is to find the optimal location and the minimum number of base stations at the same time, the number of base stations N is varied, so for the optimization problem in this paper, the particle swarm with varying length is used. The particle swarm is used to search for the optimal solution. The initial particle length is equal to three times the number of small base stations obtained from the network size estimation, and then, the optimal position is searched so that it satisfies all constraints, and then, the particle length is reduced.

2.3. Intelligent Electronic Dictionary System Design. One of the core ideas of software engineering is code reuse and decoupling. In this Android client design, we use a modular design; each module can be run independently from other modules, without affecting each other. Many functions common to the modules are written as tool classes, which are convenient for other modules to use, avoiding code redundancy and facilitating maintenance at the same time. For example, the login and registration operations need to access the remote database through the network, where the HTTP access can be written as a public class [20]. The SQLite database files are packaged and compiled together so that non-root users cannot modify the content, which guarantees the consistency and security of data. The client front-end UI design is simple, with word search input at the top of the screen and a menu switching module at the bottom of the screen; the user interface is intuitive, generous, clear, and easy to learn and use. Generating words is a difficult problem we often encounter in our daily life, and the Android client can maximize the actual work and learning needs of different users in each segment.

The Android-based electronic dictionary system is mainly developed in Java language, and the system architecture is divided into client-side and server-side, based on the MVC hierarchy. The view layer mainly realizes the application interface between client and server, and the control layer is used to extract, analyze, guide, and process the business logic of the interface layer; the logic layer is used to process and guide the underlying components and map the data with the underlying database. The data is encapsulated, delivered, and processed in JASON form. The data interaction between the client and server-side of the Android-based electronic dictionary system is mainly reflected in user login and registration. After logging in, users can look up words from other

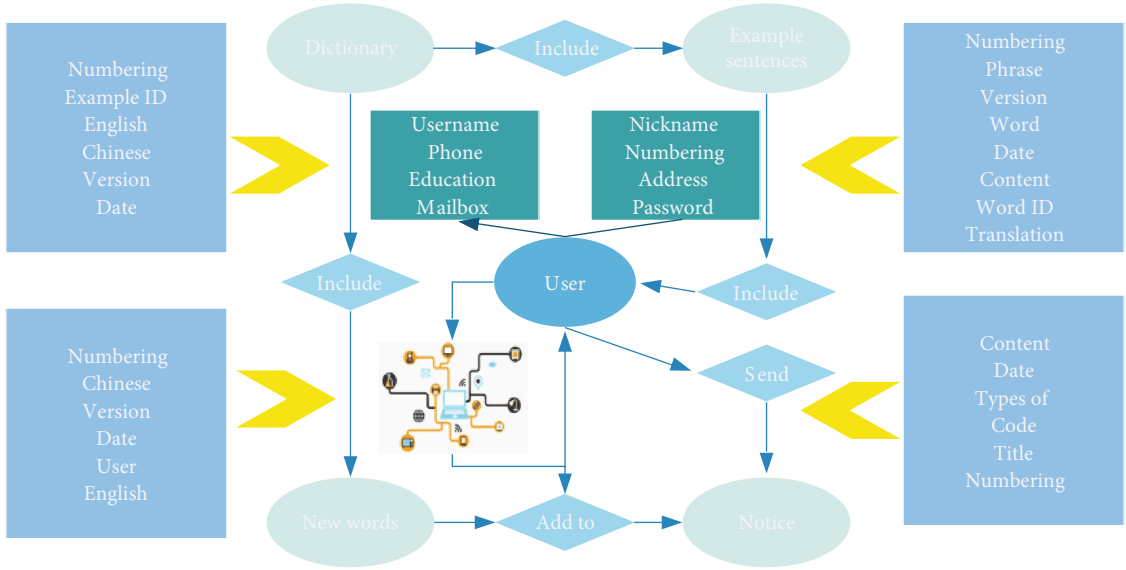


Figure 2: Server-side database ER diagram.

places, for example, after logging in on another phone, they can still view the saved word information.

There is a complete process for user login, which details the location, sequence, and conditions that must be met for each part of the user login process. When the user enters the login information, the user first performs a character check, such as whether the input information characters are correct, whether the input character length reaches the specified length value. If the check is passed, the network check is performed; otherwise, the user returns to the previous step; whether the network is connected is checked, if it is connected, the user name is checked in the next step, and then, the password matching stage is entered; if the password match is successful, the login success message is returned; otherwise, end the login and return the failure message.

The conceptual design is first carried out on the server-side to demonstrate the static data requirements through ER diagram modelling. Five entities are designed in the server-side database, which are User, Dictionary, Ec-sentence, New-word, and Notification. The server-side data entity relationships are shown in Figure 2.

A database was designed based on the ER diagram, as shown in Table 1.

Notification management is to let the administrator push some information to the client as needed, such as thesaurus upgrade information. The administrator edits the notification content in the notification management operation interface and saves it in the database. After selecting the edited notification content and clicking push, the message is transmitted to the client. The data is received and displayed in a pop-up window in the client APP.

An Android dictionary application contains several scenario screens, and the user can select different scenarios according to his actual needs. These functions are completely independent of all the scenarios, they do not know what scenarios are handled, and their job is only to control the program logic to exchange information and thus support the complete operation of the system.

Table 1: User table.

Column name ID	Type of data	Length	Values
Name	Int	3	5.5
Password	Varchar	5	2.8
Mail	Varchar	7	4.7
Address	Varchar	6	6.8
Nickname	Varchar	9	5.4
Content	Varchar	4	1.4
Education	Varchar	7	2.9

After the requirements analysis, the framework design of the system can be carried out based on the software requirements specification. Framework design is essentially an abstract design pattern, which does not consider the actual development operation, but only a general description and planning of the overall software development from the sketch. The framework design of a system will directly affect the robustness, scalability, usability, and other software characteristics of the software. In this paper, we design the English trilingual APP dictionary system, to consider the user habits and the ease of use of the software, while solving the technology of automatic language identification, trilingual word search, and fuzzy query, we also set up the function of automatic interface switching of different languages and automatic image text recognition interface, so that the overall design of the system and the framework of the system include four modules, among which the dictionary module is mainly the dictionary. The system can automatically recognize the language and query the corresponding result according to the user's input. The translation module mainly completes the translation function of long sentences. The image recognition module allows users to take pictures and recognize the

words or sentences in the images through the program; the setting module allows users to switch the language of the program interface and select different dictionary libraries; the description section provides instructions for the use of the software and some information about the development.

A collaborative product design project is a field of study that uses computer technology to support professional work, combined with advanced manufacturing technology to effectively manage and support the whole process of the product design project, which not only requires different theoretical knowledge and practical experience in the field of product design but also has a set of effective means and mechanisms to integrate and coordinate the fields of knowledge and experience to accomplish the integration of different product designs and tasks. In general, design experts believe that the basic elements of collaborative work are mutual collaboration, trust, communication, compromise, consistency, continuous improvement, and coordination. In general, a collaborative design-based product design project management approach is as follows: first, under the joint leadership of a complete product design organization and organization involved in completing a product design project, project information and design documents are placed on an information-sharing platform from the first day the project team is created and are viewed and managed by all members of the product design project team.

The system is divided into logical processing, data resources, system interface, button control, input system, local query, and network query according to the characteristics of the dictionary system that the system can work together. Data resources include thesaurus files, icons, and font files. The system interface is responsible for the display function of the system; button control and input system are the underlying interfaces of the system. And the query is the core function module of the system. Among them, logic processing is the core module, and the other six modules all need to be processed logically to play the role of application within the system.

The fuzzy query takes effect by detecting the information in the input information and registration information matching, and the fuzzy query function takes effect when the information is consistent or partially consistent. If the matching is completed, the matching result is judged to be successful or directly returned. Users directly choose to complete the query or continue the query according to the query result. The speed of fuzzy query is faster than the speed of exact query, so the efficiency of the fuzzy query is higher in massive data query, and the design of fuzzy query is one of the core functions of the system.

3. Analysis of Results

3.1. IoT Performance Analysis. For the deployment problem of edge gateways in IoT, this paper proposes a simulated annealing-based edge gateway deployment method. Firstly, we analyze the edge gateway coverage and the conditions of the service endpoint traffic generator, the factors affecting the computational task offloading delay, and the constraints to be satisfied by the edge gateway capacity allocation; then, an edge gateway deployment optimization model is established, and finally, an adaptive external penalty function is proposed, which is combined with the simulated annealing algorithm to solve the edge gateway deployment problem and obtain the optimal deployment scheme. The simulation results show that the algorithm proposed in this chapter can minimize the deployment cost, improve the resource utilization of the edge network, and achieve the edge gateway load balancing in the edge gateway deployment problem, which can provide some guidance for the later IoT edge gateway placement problem. In this section, we use simulation experiments to evaluate the performance of the algorithm proposed in this paper, we set a $1\,\text{km} \times 1\,\text{km}$ planning area, and the performance simulation results are shown in Figure 3.

We did two sets of experiments as shown in Figure 4, Figure 4(a) with 40 endpoint traffic generators randomly distributed in the planning area, and Figure 4(b) with 80 endpoint traffic generators randomly distributed in the planning area.

Based on the city's radio channel interference model, we set the channel gain. The original random perturbation method in the inner loop is to randomly select an edge gateway, cancel it if it is already deployed, and randomly deploy another edge gateway, or deploy it if it is not deployed or swap its state with one of the deployed edge gateways, but we found that this perturbation method is not effective in reducing the objective function during the simulation. Therefore, we change the random perturbation method in the inner loop to randomly select an edge gateway and directly cancel it.

To compare the performance of the adaptive external penalty-based simulated annealing algorithm with the generalized simulated annealing algorithm, we did 10 experiments on the two algorithms separately and took the average of the optimal number of deployed edge gateways for the 10 experiments, and the results are shown in Figure 5. The outer loop of both algorithms is from temperature to down, and the number of inner loops is 1000. The horizontal coordinates of Figure 5 represent the outer loop iteration steps, and the vertical coordinates represent the historical optimal number of deployed edge gateways at each iteration step. From Figure 5, we can see that the number of historical optimal deployment edge gateways decreases in both algorithms as the number of iteration steps increases, but since the algorithm incorporating the adaptive outlier penalty function may be able to reach some isolated feasible domains that are not reachable by the generic simulated annealing algorithm, also, when the algorithm proposed in this paper searches for solutions outside the feasible domain that are not far from the feasible domain, we perturb them to its neighbourhood to explore, and as the search proceeds, the solution moves in the direction of the optimal solution, which increases the probability of finding a better solution than the generic simulated annealing algorithm, so the iteration in which our proposed algorithm obtains a smaller number of optimal edge gateways than the original algorithm.

We increased the number of terminal traffic generators from 30 to 60 to verify whether the increase in the number

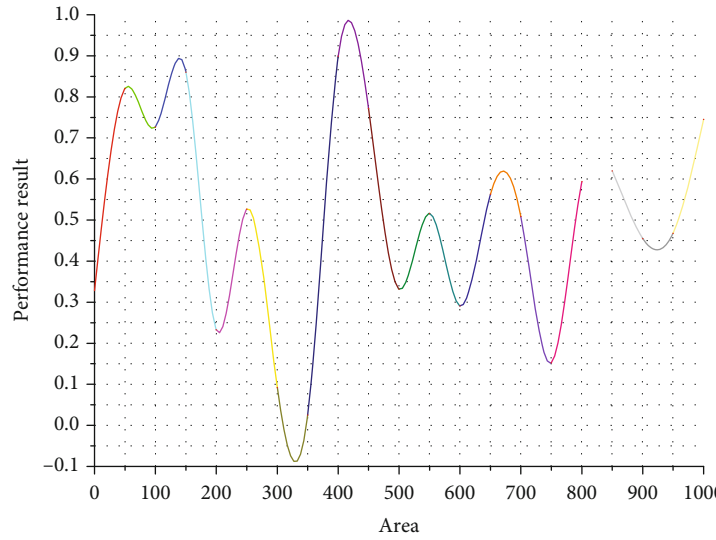


Figure 3: Simulation parameters.

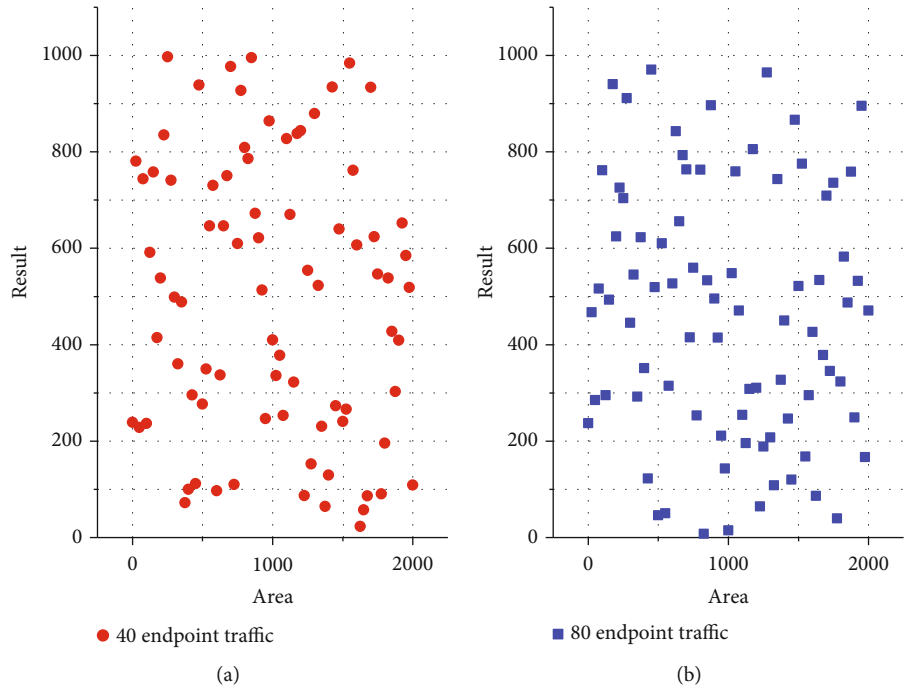


Figure 4: Distribution of terminal traffic generators.

of terminal traffic generators affects the performance of the two algorithms, and we did 10 experiments for each of the two algorithms and took the average value of the optimal number of deployed edge gateways for the 10 experiments. Therefore, the algorithm proposed in this paper will converge earlier than the general simulated annealing algorithm and find a solution with a smaller number of edge gateways. The algorithm proposed in this paper has a faster trend of decreasing target value in the early stage with 60 terminal traffic generators than with 30 terminal traffic generators, because the target value decreases faster in the early stage as the number of iterations increases due to more redundant coverage caused by the randomly generated initial solution with the increase of terminal traffic generators with the constant planning area size and edge gateway coverage.

3.2. Electronic Dictionary System Performance Results Analysis. The translation interface is like the dictionary query interface. Unlike dictionary query, the implementation of translation function is a bit more complicated, which involves many reasons such as sentence slicing, sentence sentiment analysis, and translation according to contextual meaning. This system puts the translation module into the server and reserves the interface to interface with it. As

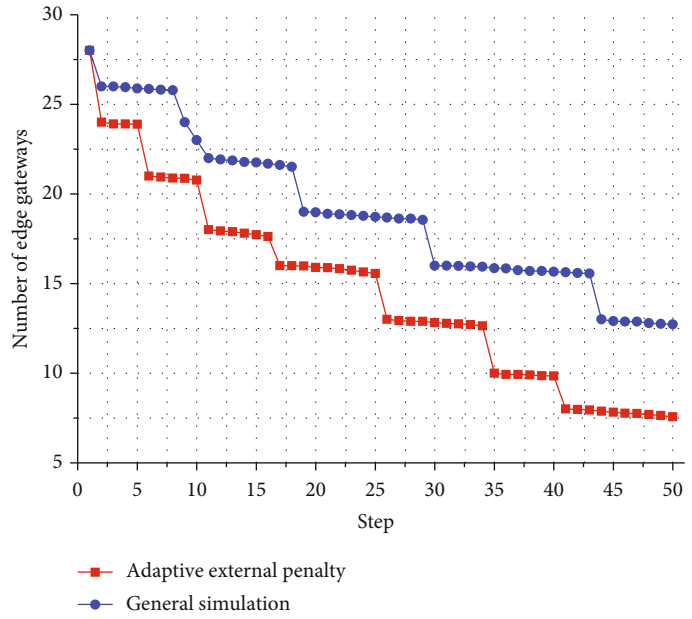


Figure 5: Comparison of two algorithmic solutions (30 endpoint traffic generators).

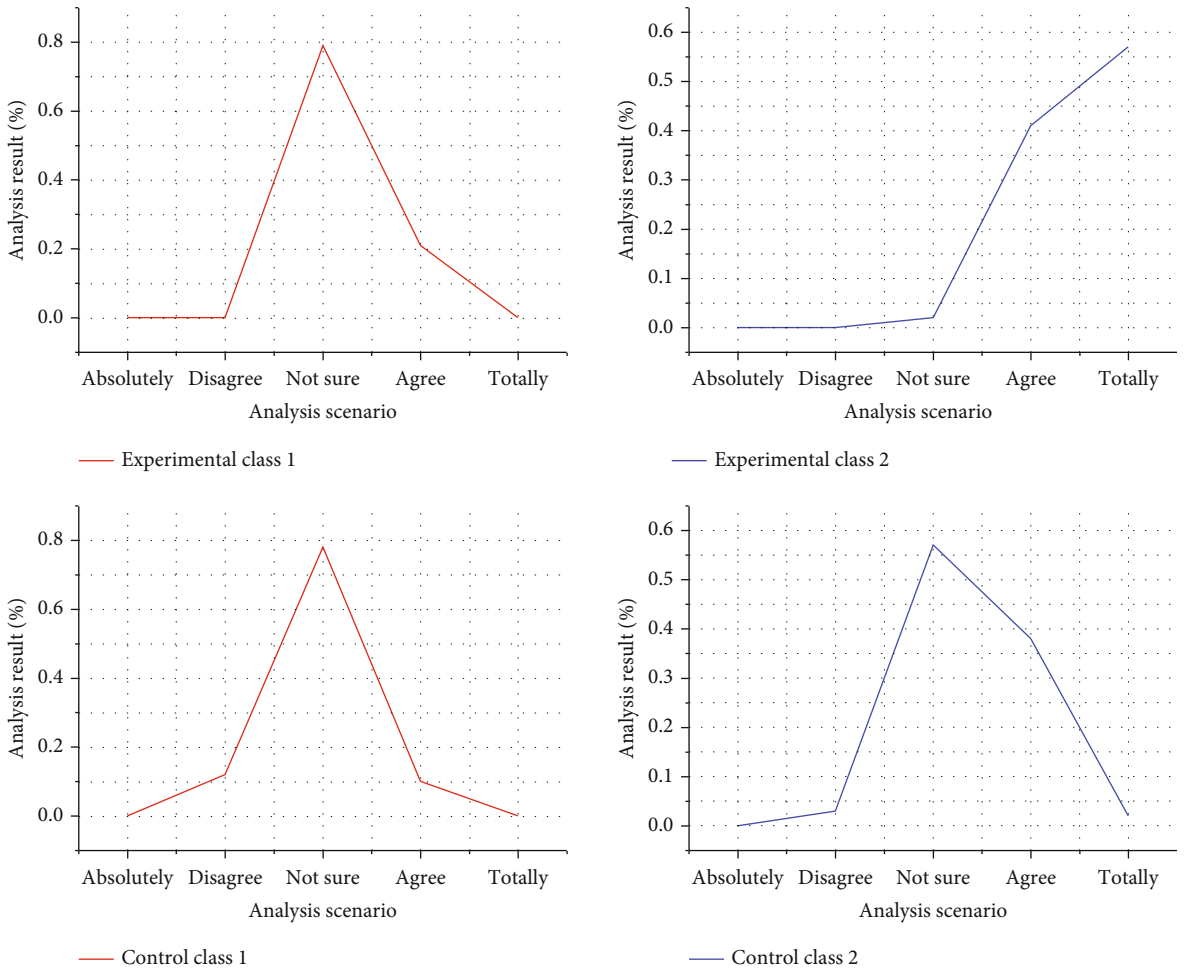


Figure 6: Comparison of postmeasurement data.

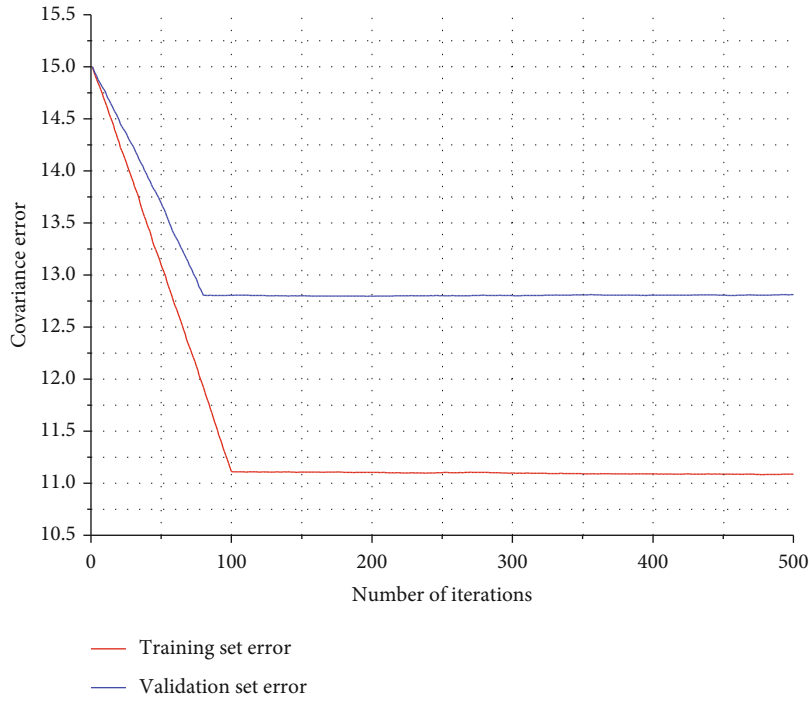


Figure 7: Mean square error variation curve.

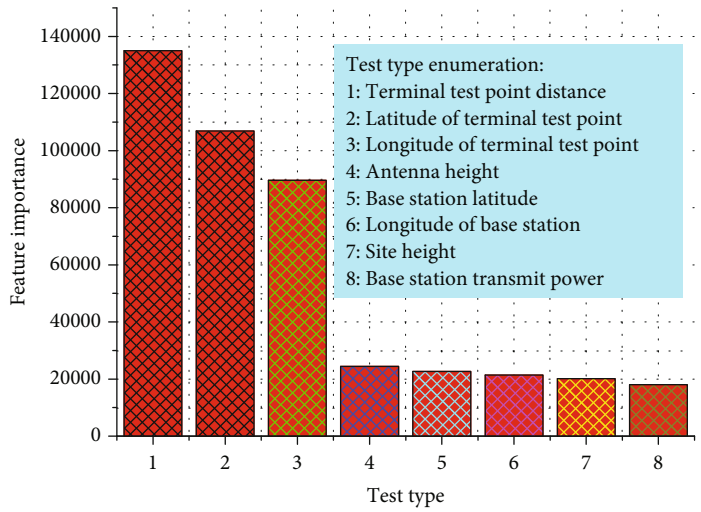


Figure 8: Feature importance results graph.

shown in Figure 6, 91.5% of the students in the experimental class said they wanted their teachers to set more interesting tasks and tasks close to life situations in the vocabulary class, because they had been trained in task-based vocabulary teaching method, while more than half of the students in the control class chose "not sure" about this statement in both the pretest and posttest, which also again, this reflects that students' attitudes and suggestions about vocabulary lessons have changed from their previous perceptions in the process of using task-based teaching methods.

To prevent overfitting during the training process, this paper adopts 5-fold cross-validation, i.e., the original data set is divided into 5 equal parts, 4 of which are selected as the training set and the remaining 1 as the validation set in each round of training. The mean square error of the training set and the validation set is shown in Figure 7. This is a normal situation because the model has seen all the samples in the training set, so its error is the lowest, but because the validation set is used as feedback to adjust the model parameters, the model is equivalent to refer to the data in the validation set, so the error is higher than the training set error.

Figure 8 shows the visualization results of the average feature importance of each feature in the 5 models. From the results in the figure, it can be seen that the feature importance in descending order is the distance between the terminal test point and the base station it is connected to, the latitude of the terminal test point, the longitude of the terminal test

point, the antenna height, the latitude of the base station, the longitude of the base station, the site height, and the base station transmit power; which is consistent with what we said in all factors affecting the received signal strength of mobile devices, the distance between the base station and the terminal is the most important due to the propagation characteristics of electromagnetic waves. The geographical location of the base station is less important than the geographical location of the terminal test point because the number of base stations in our collected data is 37 in total, which is much less than the number of terminal test points, and the variation of the value is relatively small, and the transmit power of the base station is the least important because, among the 37 base stations, the transmit power is only one is 31 dBm, and the others are 30 dBm, the data almost did not change, so there is almost no impact on the prediction results.

In addition to the unit tests, system integration tests were also performed. The focus was on testing the correctness of data interaction between the client and the server. The test results show that the client can normally login and register, get the thesaurus and complete local word search, and call the Yadao Dictionary interface to realize network word search and save raw words; the server-side can normally create thesaurus, add words, and push notifications. The server-side can create thesaurus, add words, and push notifications normally.

4. Conclusion

This article chooses the development of an electronic dictionary based on the Android system as the research content, combined with the existing Internet of Things technology, comprehensively applies the various integration tools of the Android development platform, and the back-end database is implemented by MySQL. By analyzing the mainstream word search tools, aiming at their shortcomings, a systematic realization goal is proposed. Focus on realizing the function of searching words and saving program memory. The client has functions such as local word search, network word search, adding new words, and logging in and registering. The server side mainly realizes the functions of thesaurus management, user management, and notification management. The entire Maven Web project is built around the software engineering ideas of "high cohesion" and "low coupling," and a flexible and loosely coupled Web application is established, which minimizes the coupling between components and facilitates the second project in the future. Expansion of development. This article mainly solves the problems of one-click query and fuzzy query of English vocabulary, the localized storage of SQLite database, the interface of translation and image recognition, and the automatic switch setting function of different language interfaces, which basically achieves intelligence, speed, and ease of use requirements. Through the research of this article, the Internet of Things technology and deep learning technology provide more intelligent, fast, and easy-to-use means in the design of intelligent English electronic dictionary system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I would like to thank Mr. Ma Junfeng for his help on artificial intelligence, without whom I would not have been able to complete this research.

References

[1] E. Rubio-Drosdov, D. Díaz-Sánchez, F. Almenárez, P. Arias-Cabarcos, and A. Marín, "Seamless human-device interaction in the internet of things," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 4, pp. 490–498, 2017.

[2] M. A. A. da Cruz, J. J. P. C. Rodrigues, J. Al-Muhtadi, V. V. Korotaev, and V. H. C. de Albuquerque, "A reference model for internet of things middleware," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 871–883, 2018.

[3] E. Siow, T. Tiropanis, and W. Hall, "Analytics for the internet of things," *ACM Computing Surveys*, vol. 51, no. 4, pp. 1–36, 2018.

[4] H. Habibzadeh, K. Dinesh, O. R. Shishvan, A. Boggio-Dandry, G. Sharma, and T. Soyata, "A survey of healthcare internet of things (HIoT): a clinical perspective," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 53–71, 2020.

[5] H. Zhang, F. Li, J. Wang, Z. Wang, C. Sanín, and E. Szczerbicki, "Experience-oriented intelligence for internet of things," *Cybernetics and Systems*, vol. 48, no. 3, pp. 162–181, 2017.

[6] R. Dantu, I. Dissanayake, and S. Nerur, "Exploratory analysis of internet of things (IoT) in healthcare: a topic modelling & co-citation approaches," *Information Systems Management*, vol. 38, no. 1, pp. 62–78, 2021.

[7] V. Subbarayalu, B. Surendiran, and P. Arun Raj Kumar, "Hybrid network intrusion detection system for smart environments based on internet of things," *The Computer Journal*, vol. 62, no. 12, pp. 1822–1839, 2019.

[8] H. Washizaki, S. Ogata, A. Hazeyama, T. Okubo, E. B. Fernandez, and N. Yoshioka, "Landscape of architecture and design patterns for iot systems," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10091–10101, 2020.

[9] A. Azmoodeh, A. Dehghantanha, and K. K. R. Choo, "Robust malware detection for internet of (battlefield) things devices using deep eigenspace learning," *IEEE Transactions on Sustainable Computing*, vol. 4, no. 1, pp. 88–95, 2018.

[10] C. K. Ng, C. H. Wu, K. L. Yung, W. H. Ip, and T. Cheung, "A semantic similarity analysis of internet of things," *Enterprise Information Systems*, vol. 12, no. 7, pp. 820–855, 2018.

[11] D. Pivoto, P. D. Waquil, E. Talamini, C. P. S. Finocchio, V. F. Dalla Corte, and G. de Vargas Mores, "Scientific development of smart farming technologies and their application in Brazil,"

Information processing in agriculture, vol. 5, no. 1, pp. 21–32, 2018.

[12] M. Lesani, M. Naderan, and S. E. Alavi, "Fuzzy ontology with ANFIS neural network for semantic sensor networks in smart homes based on internet of things," *International Journal of Web Research*, vol. 2, no. 1, pp. 26–38, 2019.

[13] M. Meddeb, A. Dhraief, A. Belghith, T. Monteil, K. Drira, and S. AlAhmadi, "Cache freshness in named data networking for the internet of things," *The Computer Journal*, vol. 61, no. 10, pp. 1496–1511, 2018.

[14] T. Tomiyama, E. Lutters, R. Stark, and M. Abramovici, "Development capabilities for smart products," *CIRP Annals*, vol. 68, no. 2, pp. 727–750, 2019.

[15] M. Harlamova, M. Kirikova, and K. Sandkuhl, "A survey on challenges of semantics application in the internet of things domain," *Applied Computer Systems*, vol. 21, no. 1, pp. 13–21, 2017.

[16] V. Borissova, "Digital transformation for digital competitiveness at a micro level," *Economic Studies Journal*, vol. 1, pp. 89–106, 2021.

[17] T. E. Brown, "Sensor-based entrepreneurship: a framework for developing new products and services," *Business Horizons*, vol. 60, no. 6, pp. 819–830, 2017.

[18] M. Ma, D. He, H. Wang, N. Kumar, and K. K. R. Choo, "An efficient and provably secure authenticated key agreement protocol for fog-based vehicular ad-hoc networks," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8065–8075, 2019.

[19] C. L. Lin, "Establishing environment sustentation strategies for urban and rural/town tourism based on a hybrid MCDM approach," *Current Issues in Tourism*, vol. 23, no. 19, pp. 2360–2395, 2020.

[20] A. Puška, I. Stojanović, and A. Maksimović, "Evaluation of sustainable rural tourism potential in Brcko district of Bosnia and Herzegovina using multi-criteria analysis," *Operational Research in Engineering Sciences: Theory and Applications*, vol. 2, no. 2, pp. 40–54, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5547944, 8 pages
https://doi.org/10.1155/2021/5547944

Research Article

Algorithm of Underground Personnel Positioning Based on Improved Monte Carlo

Bin Wu

Department of Information Engineering, Chengyi University College, Jimei University, Xiamen, China

Correspondence should be addressed to Bin Wu; binwu@stu.xmu.edu.cn

Received 2 February 2021; Revised 8 March 2021; Accepted 15 March 2021; Published 31 March 2021

Academic Editor: Wei Wang

In order to improve the positioning accuracy of underground targets, especially the positioning accuracy of moving targets, an improved weighted Monte Carlo positioning algorithm is proposed. In the sampling initialization stage, the beacon node gradually constructs the sampling area according to the RSSI size and combines the Monte Carlo method to further narrow the range and improve the sampling success rate. In the filtering stage, refer to the sampling area at time and further improve the sample quality at $t-1$ after two filterings. In the recollection stage, cooperate with invalid sample sets to reduce the number of recollections and weigh the final samples to improve the positioning accuracy of the nodes to be tested.

1. Introduction

The positioning of underground personnel in coal mines is essential to safe production. Therefore, the research of underground personnel positioning is also one of the research hotspots of the mine safety monitoring network [1–3].

There are many existing underground personnel positioning systems, which are roughly divided into two types: infrastructure-based positioning and wireless sensor-based positioning [4, 5]. Infrastructure-based positioning usually installs auxiliary positioning hardware facilities in environments such as underground roadways, car yards, and substations and realizes positioning through receiving equipment that personnel carry with them. Although this positioning technology has high accuracy, the application cost is too expensive to be applied in an unknown environment, and the transmitter is based on a wired network [6, 7].

Among the researches on personnel positioning based on WSNs, one type of research is based on track estimation algorithms [8]. This type of algorithm determines the position of the underground personnel based on parameters such as the number of steps, the length of the steps, and the course of the walking. However, these parameters are uncertain and random, leading to large errors in the estimated personnel position. In practice, in order to ensure the safety of underground operators, it is required to fully implement personnel positioning within the entire underground environment. Operators take a vehicle or cage to go down to the underground parking lot and then take a monkey car to reach the working surface. Therefore, the personnel do not walk down the well. The trajectory calculation algorithm cannot locate the personnel when they use the vehicle to move, resulting in the incomplete positioning system of the entire underground personnel.

The Monte Carlo [9] positioning method is currently a more popular positioning method for unknown node movement. When the Monte Carlo method solves the problem, it connects with a certain probability model or random process and calculates the statistical characteristics of the problem parameters through observation or sampling test and finally gives an approximate solution [10, 11].

The Monte Carlo Localization (MCL) method was originally applied to the mobile robot localization proposed by Frank et al. [12]. Robot MCL is a positioning method based on Bayesian filtering and combining the characteristics of robot movement and perception. It uses weighted sampling values to represent the robot's state distribution and determines the position of the robot through repeated predictions and updates.

The positioning of WSN mobile nodes is more complicated than robot positioning, but its positioning methods can be used for reference. Hu and Evans applied the MCL of robot positioning to the positioning of mobile nodes in WSNs, using nonranging methods, using a sample set with weights to estimate the posterior probability density distribution, and positioning the mobile nodes [13]. In UAV positioning, Wang et al. proposed an improved Monte Carlo method to improve positioning accuracy by eliminating position deviations [14]. Chen et al. proposed an adaptive Monte Carlo location method based on fusion posture estimation [15]. Some researchers have also studied the use of spectrum sensing technology in radio systems in mobile networks [16–18].

2. Basic Principles of MCL

The idea of MCL positioning is to budget the position at time t based on the position at time t-1. The MCL process is mainly divided into the prediction phase, the filtering phase, the resampling phase, and the importance sampling phase. The algorithm steps are as follows [19, 20].

2.1. Predict the Position at Time t. Assuming that the maximum moving speed of the node is $v_{\max}$, then at time t, the unknown node may be located in a circular area with the position at time $t-1$ as the center of l_{t-1} and the radius of $v_{\max}$, as shown in Figure 1.

N location samples are randomly collected in the area where the unknown node is located to form a sample set$L = \{l_t^1, l_t^2, \cdots, l_t^N\}$, where l_t^i represents the possible ith location of the node at time t.

Assuming that the moving speed of the unknown node satisfies a uniform distribution in the upper $[0, v_{\max}]$, the probability that the position of the unknown node at time t is estimated from the position at time $t-1$ also obeys the uniform distribution. As shown in the formula,

$$p(l_t \mid l_{t-1}) = \begin{cases} \dfrac{1}{\pi v_{\max}^2}, & d(l_t, l_{t-1}) < v_{\max}, \\ 0, & \text{else.} \end{cases} \quad (1)$$

Among them, $d(l_t, l_{t-1})$ represents the Euclidean distance between l_t and l_{t-1}, and $p(l_t \mid l_{t-1})$ represents the posterior probability. It can be seen from formula (1) that as the node $v_{\max}$ increases, the circular area becomes larger and larger, the position range at time t becomes larger and larger, and the posterior probability becomes smaller and smaller.

2.2. Filtering Stage. In the filtering stage, the unknown node judges and screens possible locations based on the observation values received at time t, filters out invalid sample points in the sampling area, and improves the accuracy of the sample. The method of sample filtering is based on whether the broadcast of the beacon node can be received within the communication range of the beacon node and filter out the sample locations that have not received the broadcast.

Suppose S is the set of beacon nodes with one-hop distance that unknown nodes can communicate, T is the set of beacon nodes with two-hop distances, and R_c is the communication radius of the node. Then, the filter condition of sampling sample l_t^i is

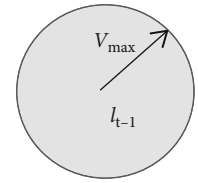


Figure 1: Location area at time t.

$$\text{filter}\left(l_t^i\right) = \left(\forall s \in S, d\left(l_t^i, s\right) \le R_c\right) \cap \left(\forall t \in T, R_c < d\left(l_t^i, t\right) \le 2R_c\right), \quad 1 \le i \le N, \quad (2)$$

where $d(l_t^i, s), d(l_t^i, t)$ represents the Euclidean distance between the sample l_t^i and the beacon node s and t. If $\exists s \in S$, $d(l_t^i, s) > R_c$ at time t, it means that at position l_t^i, the node cannot listen to the broadcast of a one-hop beacon node and then filter out such a position. If $\exists t \in T$, $d(l_t^i, t) > 2R_c$ or $d(l_t^i, t) \le R_c$, it means that at position l_t^i, the node cannot receive the broadcast of the two-hop beacon node, and l_t^i is filtered out from L.

The condition of sample filtering is to use one-hop and two-hop beacon nodes. The reason why three-hop or more beacon nodes are not used is to reduce the communication overhead and calculation amount of the network.

2.3. Sample Resampling. After filtering, the unqualified sample positions are filtered out; then, the number of samples required by the algorithm may not reach N, and samples need to be resampled to make up for the deficiency. In the sample resampling stage, repeat 2.1 and 2.2 until the number of samples N that meets the requirements is collected or the maximum number of sampling times is reached.

2.4. Importance Sampling Stage. In the final sample set, since it cannot be directly sampled from the posterior distribution, it is necessary to sample the importance of the sample set. Generally, the average value of all samples is taken as the position of the unknown node at time t.

Although the MCL method does not require distance measurement hardware, the calculation method is simple, and it is highly adaptable. It can even realize the positioning of mobile nodes when the beacon nodes are sparse. However, the success rate is low and the sampling process is repetitive, which leads to an increase in node calculations, causing greater energy waste. If the WSN nodes of mine tunnels work with high-frequency sampling, the energy consumption will be faster and the network life will be greatly shortened, which is unbearable. Therefore, the MCL method cannot be directly applied in practice [21, 22].

3. Mobile Positioning Based on Improved MCL

Many researchers have proposed improved algorithms for MCL, including Baggio and Langendoen who proposed the MCB algorithm based on the MCL method [10]. The algorithm limits the sample location area to the sample box by defining the beacon box and the sample box, which reduces the sampling range and improves the success rate of sampling samples. But when the probability of sampling samples distributed in beacon boxes is relatively small, the success rate of sampling is very low. Qiao improved on the basis of MCB and proposed the MCCB algorithm [23] based on the center of mass to further improve the positioning accuracy. Dil et al. proposed a high-precision MCL algorithm [24] by using the distance between measurement nodes in the filtering stage of the MCL method.

The literature shows that the existing improved methods based on MCL have achieved good positioning results, but these improved algorithms cannot be directly applied to the positioning of moving targets in coal mines, because the improvement method mostly assumes that the application scenario is an open area or evenly distributed beacon nodes or the moving range of moving targets is not limited and so on. In fact, the underground structure is complex, the roadway is narrow and long, and the range of personnel activities is limited. Most of the underground personnel are walking along the direction of the roadway, and the beacon nodes are deployed on both sides of the road wall. Therefore, a personnel positioning algorithm suitable for underground WSNs needs to be studied urgently.

3.1. IMCL Algorithm Description. By referring to the MCL algorithm, this article establishes the location sampling area of the mobile node according to the Received Signal Strength Indication (RSSI) received by the node. In order to facilitate the analysis, this article takes the horizontal roadway as the research object and abstracts the horizontal roadway into a two-dimensional plane; that is, the coordinate system is two-dimensional. U stands for underground personnel. Generally speaking, when U receives the RSSI of at least 3 beacon nodes, it can be positioned. The RSSI value between nodes is averaged after three measurements. The RSSI value of each node is different, which also determines that their influence on the location of the unknown node U is different. Suppose that at time T the unknown node U sends a broadcast to the surroundings and listens to feedback from multiple beacon nodes. U will sort the received RSSI values from large to small and select the beacon node A_1, A_2, A_3 corresponding to the largest three values as a positioning reference. The larger the RSSI, the closer the distance. Therefore, A_1, A_2, and A_3 are also the three beacon nodes closest to U, assuming that they are all within the range of one or two hops of U.

3.1.1. Initialization. Assume that $\mathrm{RSSI}(i, j)$ and $d(i, j)$, respectively, represent the RSSI that node i receives from node j and the distance between the two nodes. For the beacon node A_1, if $\mathrm{RSSI}(A_1, U) > \mathrm{RSSI}(A_1, A_2) > \mathrm{RSSI}(A_1, A_3)$, it means that U is the closest to A_1, A_2 is the second, and A_3 is the farthest. Since U is in the underground roadway, its possible location area should be the overlap between the circle with the center of circle A_1 and the radius of $d(A_1, A_2)$ and the roadway space. As shown in Figure 2, the gray part is the sampling area determined by A_1, denoted as R_1.

For the beacon node A_2, if $\mathrm{RSSI}(A_2, U) > \mathrm{RSSI}(A_2, A_1) > \mathrm{RSSI}(A_2, A_3)$, it means that U is the closest to A_2, A_1 is the second, and A_3 is the farthest. The unknown node U should be in the overlapping part of R_1 with the circle centered at A_2 and radius $d(A_2, A_1)$. As shown in Figure 3, the gray part is the sampling area jointly determined by A_1 and A_2, denoted as R_2.

For the beacon node A_3, if $\mathrm{RSSI}(A_3, A_2) > \mathrm{RSSI}(A_3, A_1) > \mathrm{RSSI}(A_3, U)$, it means that A_2 is the closest to A_3, A_1 is the second, and U is the farthest. The unknown node U should be located in the overlapping part of R_2 and the outer side of the circle with the center of A_3 and the radius of $d(A_3, A_1)$. As shown in Figure 4, the gray part is the sampling area jointly determined by A_1, A_2, and A_3, denoted as R_3.

Then, R_3 is expressed as

$$\begin{cases} \left(x - x_{A_1}\right)^2 + \left(y - y_{A_1}\right)^2 \le d(A_1, A_2)^2, \\ \left(x - x_{A_2}\right)^2 + \left(y - y_{A_2}\right)^2 \le d(A_1, A_2)^2, \\ \left(x - x_{A_3}\right)^2 + \left(y - y_{A_3}\right)^2 \ge d(A_1, A_3)^2, \\ y_{A_2} \le y \le y_{A_1}. \end{cases} \tag{3}$$

Suppose the coordinates of the four vertices of R_3 are A_1, A_2, A_3, and A_4, which can be solved according to formula (3).

According to the MCL method, the position of the unknown node U at time t is a circle with position l_{t-1} at time $t-1$ as the center and radius $v_{\max}$, denoted as R_4; then, $R_3 \cap R_4$ is taken as the sampling area of U at time t, as shown in Figure 5.

Since the gray part is an irregular figure, the bounding rectangle is taken as the initial sampling area. Let M and Q be the diagonal vertices of the rectangle, and the coordinates are $M(x_{\max}, y_{\max})$ and $Q(x_{\min}, y_{\min})$. M and Q coordinates satisfy

$$\begin{cases} x_{\min} = \max\left(\min\left(x_a, x_b, x_c, x_d\right), x_{t-1} - v_{\max}\right), \\ y_{\min} = \max\left(\min\left(y_a, y_b, y_c, y_d\right), y_{t-1} - v_{\max}\right), \\ x_{\max} = \min\left(\max\left(x_a, x_b, x_c, x_d\right), x_{t-1} + v_{\max}\right), \\ y_{\max} = \min\left(\max\left(y_a, y_b, y_c, y_d\right), y_{t-1} + v_{\max}\right). \end{cases} \tag{4}$$

Among them, (x_{t-1}, y_{t-1}) represents the position coordinates of the node at $t-1$. The initial sampling area is a rectangular area with M and Q as diagonal fixed points.

3.1.2. Prediction Stage. Randomly select N sample points in the initial sampling area, and the moving speed of U obeys a uniform distribution on $[0, v_{\max}]$. At time t, the probability of the position of U in the sampling area also obeys the uniform distribution:

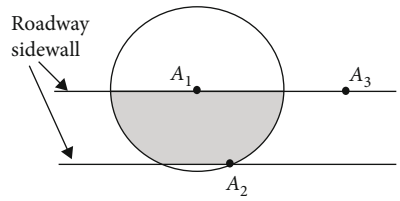


Figure 2: Sampling area determined by A_1.

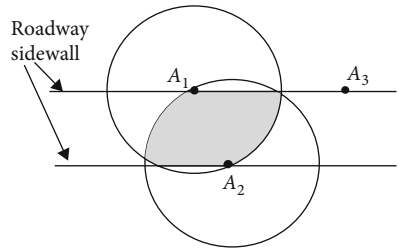


Figure 3: Sampling area determined by A_1, A_2.

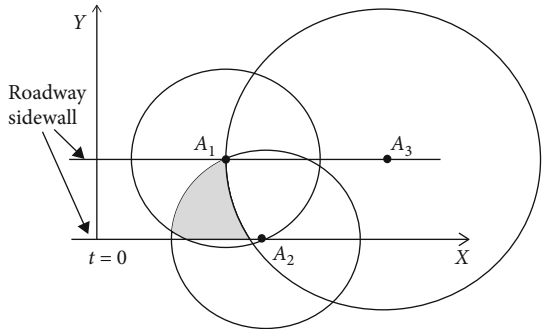


Figure 4: Sampling area determined by A_1, A_2, A_3.

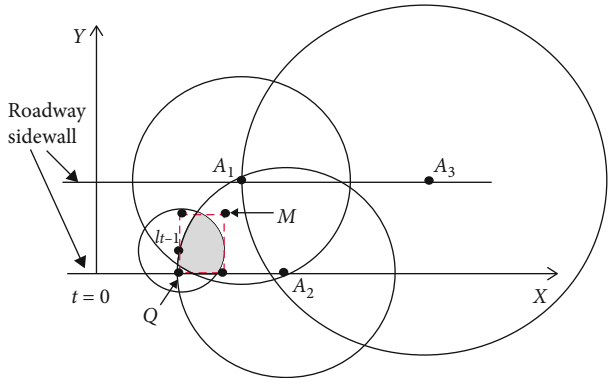


Figure 5: Initial sampling area.

$$p(l_t \mid l_{t-1}) = \begin{cases} \dfrac{1}{(x_{\max} - x_{\min})(y_{\max} - y_{\min})}, & x_{\min} \le X \le x_{\max} \cap y_{\min} \le Y \le y_{\max}, \\ 0, & \text{other.} \end{cases} \tag{5}$$

In order to reflect the contribution of the beacon node to the positioning and to avoid the influence of low-quality samples on the positioning accuracy, each sample is weighted. At $t = 0$, the weight of the sample is $w_0^i = 1/N$, $1 \le i \le N$.

3.1.3. Filtering Stage. Filter the N sample points according to formula (6) to remove invalid samples:

$$\text{filter}\left(l_t^i\right) = \forall s \in S_j, (j-1)R_c \le d\left(l_t^i, s\right) \le jR_c. \tag{6}$$

Among them, S_j represents the set of j-hop beacon nodes, and R_c represents the node communication radius. Assume that the filtered sample set is L_t and the invalid sample set is I_t. When the sample meets the filter condition L_t, it will be kept, and if it does not meet the filter condition, it will be removed to I_t.

In order to further improve the quality of the sample and improve the accuracy of positioning, another sample filtering is performed from L_t. Since the position of the unknown node at time t is determined based on the position at time $t - 1$, when L_t is filtered again, the sample set L_{t-1} at time $t - 1$ is used to determine whether the sample meets the condition, and when the sample that does not meet the condition, remove from L_t and save it to the invalid sample set.

Take any sample l_t^i from L_t, if there is a sample l_{t-1}^j in L_{t-1} such that $d(l_t^i, l_{t-1}^j) \le v_{\max}$, then keep l_t^i in L_t; otherwise, remove it from L_t.

The specific filter conditions are

$$\text{filter}'\left(l_t^i\right) = \exists l_{t-1}^i \in L_{t-1}, d\left(l_t^i, l_{t-1}^j\right) \le v_{\max}. \tag{7}$$

Combining formulas (6) and (7), the condition for filtering the originally collected N samples is

$$\begin{aligned} \text{filter}\left(l_t^i\right) = &\left(\forall s \in S_j, (j-1)R_c \le d\left(l_t^i, s\right) \le jR_c\right) \\ &\cap \left(\exists l_{t-1}^j \in L_{t-1}, d\left(l_t^i, l_{t-1}^j\right) \le v_{\max}\right). \end{aligned} \tag{8}$$

Because the quality of the filtered samples is different, the weights are differentiated. The beacon node closest to the unknown node has the greatest power to determine the positioning result, so the weight of the sample is set to the reciprocal of the distance to the nearest beacon node as the weight coefficient, $w_t^i = 1/d(A_i, l_t^i)$, where A_i represents the beacon closest to the unknown node. For nodes, the smaller $d(A_i, l_t^i)$, the greater the weight and the greater the influence of the sample position.

3.1.4. Retake Stage. After the filtering stage, the number of remaining samples may not reach N, and resampling is required. In order to improve the success rate of the resampling stage, during the resampling, the sampled sample is compared with the elements in I_t, and if it is in the set, it is discarded and resampled. After recollection, repeat prediction and filtering until the number of samples meets the requirements or reaches the maximum number of samples.

Algorithm based on improved Monte Carlo.
Step 1. The mobile node listens to the beacon node broadcast and selects the beacon node corresponding to the maximum 3 RSSI as the positioning node.
Step 2. The selected beacon node constructs the possible location area of the mobile node one by one according to the RSSI size.
Step 3. Take the intersection of the MCL sampling area and the location area in Step 2, which is the final sampling area of the algorithm.
Step 4. Select N samples at random and initialize the sample weights.
Step 5. Filter samples and update sample weights.
Step 6. Determine whether the number of remaining samples is sufficient or whether the maximum sampling times are reached. If yes, continue; otherwise, go to Step 4.
Step 7. Normalize the weight of the sample and calculate the coordinates of the mobile node.

Algorithm 1:

3.1.5. Location Estimate. Assuming that the final number of samples that meet the requirements is Num, the weights of the samples are normalized:

$$w_t^i = \frac{w_t^i}{\sum_{i=1}^{\text{Num}} w_t^i}. \tag{9}$$

Then, the position of the moving target at time t is shown in formula (10), where (x_t^i, y_t^i) is the coordinate of the ith sample at time t:

$$\begin{cases} x_t = \sum_{i=1}^{\text{Num}} w_t^i x_t^i, \\ y_t = \sum_{i=1}^{\text{Num}} w_t^i y_t^i. \end{cases} \tag{10}$$

3.2. IMCL Algorithm Steps. The basic steps of the underground personnel positioning algorithm based on improved Monte Carlo are as Algorithm 1.

4. Experimental and Simulation

In order to verify the superiority of the IMCL algorithm, this article has done a lot of experiments and compared it with the MCL algorithm, MCB algorithm, and MCCB algorithm through MATLAB simulation software. The experimental data are the test data of the actual operation base of the actual mine.

In order to effectively simulate the real roadway environment, the simulation area is set as a long and narrow area of 500 m × 5 m. Unknown nodes are randomly deployed in the simulation area, and beacon nodes are nonuniformly cross-deployed on both sides of the area with known coordinates. Suppose the communication radius of the node is $R_c = 5$ m, the maximum moving speed of the node is $v_{\max} = 5$ m/s, the minimum moving speed is $v_{\min} = 0$ m/s, the number of samples is $N = 50$, and the maximum number of samples is Num = 500.

The pros and cons of the algorithm are evaluated by positioning error, that is,

$$\sigma = \frac{\sqrt{(x - x_{\text{true}})^2 + (y - y_{\text{true}})^2}}{R_c}. \tag{11}$$

4.1. The Effect of Node Moving Speed on Positioning. Through experiments, it is found that the maximum speed of the mobile node changes from 5 m/s to 25 m/s, and the changes in the positioning errors of the four algorithms are shown in Figure 6.

Experiments show that with the increase in $v_{\max}$, the positioning errors of the four positioning algorithms IMCL, MCL, MCB, and MCCB all gradually increase. As $v_{\max}$ increases, the sample sampling area in the algorithm will also become larger, and invalid samples will also increase, resulting in larger positioning errors. At the same time, Figure 6 shows that the error curve does not increase linearly. Because the moving speed of the target node under test becomes larger and the activity range of the node per unit time becomes larger, more beacon node communications can be obtained, and the impossible position samples can be filtered out more effectively, thereby improving positioning accuracy and reducing errors. The error of the MCL method is the largest, mainly because the sampling area is not limited, and it becomes larger as $v_{\max}$ increases, the number of sampling is increasing, and the positioning accuracy is low. The MCB and MCCB algorithms limit the sampling area in a relatively small beacon box, which improves the sampling success rate and reduces positioning errors. The IMCL algorithm further limits the initial sampling area to the most likely range based on the influence of the beacon node and combines the sample weight to improve the positioning accuracy with the lowest error.

4.2. The Effect of Node Moving Speed on Sampling Times. The number of sampling also affects the positioning accuracy. Experiments show that with the change of node $v_{\max}$, the sampling times of MCL increase, and the sampling times of the other three types decrease, as shown in Figure 7. For the MCL method, when the node $v_{\max}$ is relatively small, the sampling range is relatively small, the sampling success rate is relatively high, and the sampling times are relatively

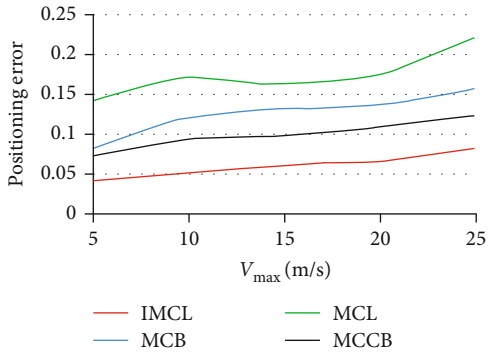


Figure 6: Positioning error of different v_{max}.

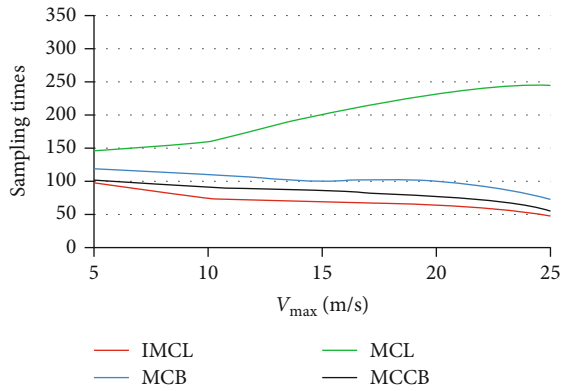


Figure 7: The sampling times of different v_{max}.

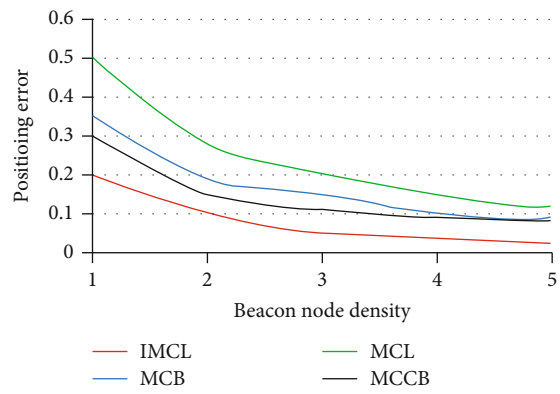


Figure 8: Positioning error under different beacon node densities.

small; as v_{max} increases, the sampling area becomes larger, the sampling success rate is low, and the sampling times increase. In MCB and MCCB, with the increase in v_{max}, the beacon node box becomes smaller, the sampling success rate is high, and the sampling frequency is reduced. The sampling area of IMCL is constructed using the influence of beacon nodes and MCL sampling, which improves the sampling efficiency and success rate, and the number of samplings is relatively small.

4.3. The Effect of Beacon Node Density on Positioning. The beacon node density is defined as the number of beacon nodes within one-hop distance of the node to be tested. As the density of beacon nodes increases, the more observation information the node to be tested receives, and the filtered samples are closer to the posterior probability distribution, which improves the positioning accuracy.

The experimental results are shown in Figure 8. As the density of beacon nodes increases, the positioning errors of the four algorithms have decreased to varying degrees, but the downward trend is gradually flattened. The positioning accuracy of MCB and MCCB is closely related to the density of beacon nodes. The greater the density, the smaller the box of the beacon node and the smaller the positioning error. MCCB is based on the center of mass MCB, so its positioning error is lower than MCB. IMCL is based on the positioning of the nearest beacon node. The greater the density of beacon nodes, the smaller the distance between the unknown node and the nearest beacon node, the greater the signal strength, the more accurate the distance calculation, and the higher the positioning accuracy.

4.4. The Effect of Beacon Node Density on Sampling Times. The experimental results are shown in Figure 9. The sampling times of each algorithm decrease with the increase of the beacon node density. Because of the increase in the density of beacon nodes, the more observation information obtained by the sample, more samples will be retained in the filtering stage, the higher the success rate, and the fewer sampling times, so the changing trend of the sampling times of the four algorithms is roughly the same. Although the sampling times decrease with the increase of the beacon node density, the sampling times of the MCL algorithm are relatively the largest. Similar to the same v_{max}, the sampling area of MCL is fixed, while the sampling areas of the other three algorithms are relatively small. Therefore, when sampling randomly, the probability of success is relatively large, and the number of sampling is naturally relatively low. The sampling area of the MCCB algorithm is reduced on the basis of the MCB, so the sampling times of the MCCB are less than the sampling times of the MCB. The initial sampling area of IMCL is closer to the actual location of the unknown node. As the density of beacon nodes increases, more samples are drawn that meet the filtering conditions, and the number of samples is less.

4.5. The Effect of Roadway Width on Positioning Error and Sampling Times. During the experiment, we changed the width of the roadway network while keeping v_{max} unchanged and repeated the experiment to observe changes in positioning errors and sampling times.

Figure 10 shows that as the width changes from 5 m to 10 m, the positioning errors of various algorithms gradually become larger. Since the beacon nodes are cross-deployed on both sides of the area, the increase of the width causes the longitudinal distance between the beacon nodes to increase, the area of the same length increases, and the number of one-hop or two-hop beacon nodes decreases. The information is less, and the positioning error is large.

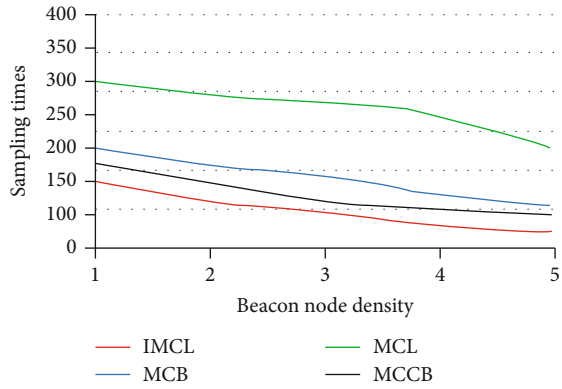


Figure 9: Comparison of sampling times of different beacon node densities.

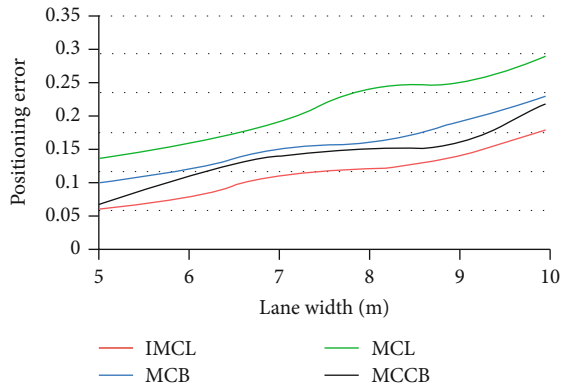


Figure 10: Positioning error under different roadway widths.

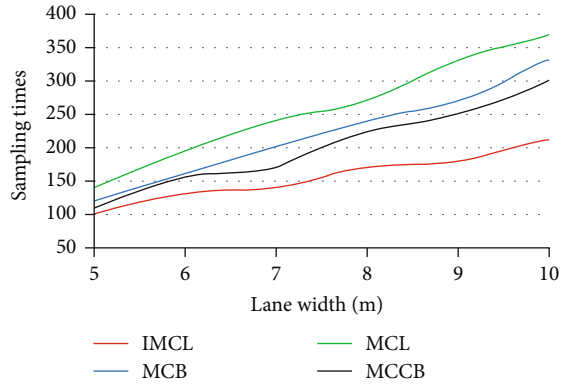


Figure 11: Sampling times under different roadway width.

Figure 11 shows the trend of sampling times of various algorithms under different widths of the simulation area. According to the above analysis, the increase of the width leads to the decrease of one-hop or two-hop beacon nodes. In the filtering stage, relatively more samples are filtered out, and the number of sampling increases.

5. Conclusions

According to the size of RSSI received by the beacon node and combined with MCL sampling, IMCL establishes the sampling area where unknown nodes are most likely to appear, reduces the sampling range of MCL, and improves the sample sampling success rate. The algorithm performs two filtering and screening in the sample filtering stage to retain high-quality samples; in the recollection stage, it uses invalid sample sets to match sampling to reduce the number of recollections and reduce energy consumption. The final sample is weighted to reflect the influence of the sample on the location of unknown nodes and improve the positioning accuracy. Experimental simulations show that under different moving speeds, different beacon node densities, and different roadway widths, the positioning accuracy of this algorithm is higher than other algorithms, and the sampling times are the least.

Data Availability

The data used to support the findings of this study are included within the article.

Consent

Consent is not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This research was funded by the Natural Science Foundation of Fujian Province of China, under grant No 2019J01719; Department of Education, Fujian Province of China, under grant No JT180879; and Ph.D. Research Fund of Chengyi University College, Jimei University, under grant No CK17064.

References

[1] B. Wu, J. Luo, and H. Huang, "Research on comprehensive automation system of southern coal mine," *Industry and Mine Automation*, vol. 40, no. 1, pp. 23–26, 2014.

[2] Y. Liang, "Innovation and development of safety science and technology in coal industry of China," *Safety in Coal Mines*, vol. 46, no. S1, pp. 5–11, 2015.

[3] B. Wu, J. Luo, and C. Yang, "Wireless sensor network minimum beacon set selection algorithm based on tree model," *Neural Computing and Applications*, vol. 30, no. 3, pp. 965–976, 2018.

[4] M. Mansouri, O. Ilham, H. Snoussi, and C. Richard, "Adaptive quantized target tracking in wireless sensor networks," *wireless networks*, vol. 17, no. 7, pp. 1625–1639, 2011.

[5] J. Ma, Q. Hu, B. Song, and S. Zhang, "Underground personnel positioning system based on fingerprint film and track estimation," *Industry and Mine Automation*, vol. 42, no. 5, pp. 19–23, 2016.

[6] R. Jiang and Z. Yang, "An improved centroid localization algorithm based on iterative computation for wireless sensor network," *Acta Physica Sinica*, vol. 65, no. 3, pp. 1–9, 2016.

[7] J. Li, "Application of KJ190 belt conveyor temperature monitoring system in coal mine," *Colliery Mechanical & Electrical Technology*, vol. 3, pp. 107–110, 2014.

[8] J. Ma, Q. Hu, B. Song, and S. Zhang, "Underground personnel positioning system based on fingerprint and dead-reckoning," *Industry and Mine Automation*, vol. 42, no. 5, pp. 19–23, 2016.

[9] P. N. Pathirana, N. Bulusu, A. V. Savkin, and S. Jha, "Node localization using mobile robots in delay-tolerant sensor networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 3, pp. 285–296, 2005.

[10] A. Baggio and K. Langendoen, "Monte Carlo localization for mobile wireless sensor networks," *Ad Hoc Networks*, vol. 6, no. 5, pp. 718–733, 2008.

[11] K. Jiang, L. Chen, and A. Yang, "Monte Carlo localization for an autonomous underwater vehicle with a low-cost sonar," *IOP Conference Series: Earth and Environmental Science*, vol. 440, no. 4, 2020.

[12] S. Thrun, D. Fox, W. Burgard, and S. Thrun, "Robust Monte Carlo localization for mobile robots," *Artificial Intelligence*, vol. 128, no. 1–2, pp. 99–141, 2001.

[13] L. X. Hu and D. Evans, "Localization for mobile sensor networks," in *10th anual International Conference on Mobile Computing and Networking*, pp. 45–57, 2004.

[14] D. Wang, C. Xu, P. Yuan, and D. Huang, "A revised Monte Carlo method for target location with UAV," *Journal of Intelligent & Robotic Systems*, vol. 97, no. 2, pp. 373–386, 2020.

[15] W. Chen, T. Huang, and A. Maalla, "Research on adaptive Monte Carlo location method based on fusion posture estimation," in *2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*IEEE.

[16] M. S. Gupta and K. Kumar, "Application aware networks' resource selection decision making technique using group mobility in vehicular cognitive radio network," *Vehicular Communications*, vol. 26, 2020.

[17] A. Bagwari and G. S. Tomar, "Cooperative spectrum sensing with adaptive double-threshold based energy detector in cognitive radio networks," *Wireless Personal Communications*, vol. 73, no. 3, pp. 1005–1019, 2013.

[18] A. Kumar and K. Kumar, "Relay sharing with DF and AF techniques in NOMA assisted cognitive radio networks," *Physical Communication*, vol. 42, p. 101143, 2020.

[19] M. Qin and R. Zhu, "A Monte Carlo localization method based on differential evolution optimization applied into economic forecasting in mobile wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, 2018.

[20] J. M. Pak and D. W. Kim, "Monte Carlo localization using the filter-based random motion model in wireless sensor networks," *Journal of Platform Technology*, vol. 5, 2017.

[21] Y. Wang, W. Zhang, F. Li et al., "An improved adaptive Monte Carlo localization algorithm fused with ultra wideband sensor," in *2019 IEEE International Conference on Advanced Robotics and its Social Impacts (ARSO)*, 2019.

[22] S. P. Fan, Y. J. Wen, and J. Zhou, "An enhanced Monte Carlo localization algorithm for mobile node in wireless sensor networks," *Applied Mechanics and Materials*, p. 2668, 2013.

[23] G. Z. Qiao, *Design and Key Technology Research of Coal Mine Safety Comprehensive Monitoring System Based on Wireless Sensor Network*, Ph.D., Lanzhou University of Technology, Lanzhou City, Gansu Province, China, 2012.

[24] B. Dil, S. Dulman, and P. Havinga, "Range-based localization in mobile sensor networks," *European Conference on Wireless Sensor Networks*, 2006.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5570685, 13 pages
https://doi.org/10.1155/2021/5570685

Research Article

RBF Neural Network-Based Frequency Band Prediction for Future Frequency Hopping Communications

Shengyan Zhu[1], **Yongjian Wang**[2], **Jianbo Zheng**[3] and **Shupeng Wang**[4]

[1]*Faculty of Quality Management and Inspection & Quarantine Sanjiang Research Institute of Artificial Intelligence & Robotics, Yibin University, China*
[2]*National Computer Network Emergency Response Technical Team, Coordination Center of China, China*
[3]*Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China*
[4]*Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China*

Correspondence should be addressed to Yongjian Wang; wyj@cert.org.cn, Jianbo Zheng; jb.zheng@siat.ac.cn, and Shupeng Wang; wangshupeng@iie.ac.cn

Received 30 January 2021; Revised 22 February 2021; Accepted 15 March 2021; Published 30 March 2021

Academic Editor: Wei Wang

On the basis of the chaotic features of the frequency hopping signal, frequency band prediction for frequency hopping signal can enhance the interference effect of the signal greatly. However, poor prediction accuracy often limits its development in the military field. Therefore, for the sake of enhancing the frequency band prediction accuracy of frequency hopping signal, this paper studies the radial basis function (RBF) neural network frequency hopping signal frequency band prediction model based on the gradient descent method and improved the particle swarm optimization algorithm, respectively. The former uses a step-by-step algorithm to optimize the center value and weight so that the network can find the most suitable initial state. Then, the clustering selection optimization algorithm is employed to optimize the central value. In addition, it optimizes the weight by using a gradient descent method of the optimal learning rate. The latter optimizes the structure of the RBF neural network through the combination of the subtractive clustering algorithm and improved the particle swarm optimization (PSO) algorithm. Simulation results demonstrate that the gradient RBF algorithm model performs better in terms of accuracy, but time efficiency is lower, while the PSO-RBF algorithm has better time efficiency.

1. Introduction

The purpose is to effectively combat the enemy's communication system or protect the security of our army's communication system in the modern military battlefield. In addition to the traditional physical attack accident, we can also attack the enemy's communication system by means of signal interference. Network confrontation has become the main mode of enemy-friend warfare. Its main form of confrontation is exploiting a security flaw in an enemy network system to invade the enemy network and reduce and destroy the use efficiency of the enemy network [1]. Therefore, we need to understand the enemy's communication frequency band, and if we interfere and strike arbitrarily without knowing anything, it will also interfere with our communication system. As a result, it is necessary to predict the communication frequency band of the enemy. Through the prediction of the frequency band of frequency hopping communication, we can understand the enemy communication frequency band and attack effectively. Therefore, the prediction of the frequency hopping signal is the most important work before the specific strike.

In the traditional conventional frequency hopping technology, both sides of the communication carry out synchronous hopping according to the agreed frequency hopping strategy in the process of communication to awaken the

reliable transmission of data information [2]. However, with the increasing intensity of modern electronic warfare, the spectrum resources are becoming more and more limited, which leads to the intensification of mutual interference among all kinds of communication equipment and the continuous upgrading of the interference technology of the interference party. The traditional frequency hopping technology cannot meet the antijamming needs of modern electronic warfare. Therefore, based on conventional frequency hopping, many experts and scholars worldwide raise the communication technology of adaptive frequency hopping (AFH) [3]. The core technology is monitoring the channel in real time, carrying on the link quality analysis (LQA) and feeding the evaluation information back to the other side of the communication. Then, the two sides communicate adaptively at the same time according to a certain strategy. Adaptive frequency hopping technology dramatically enhances the antijamming ability of the communication system. The common adaptive frequency hopping techniques include frequency point adaptive frequency hopping technology, power adaptive frequency hopping technology, frequency hopping rate adaptive frequency hopping technology, and the adaptive frequency hopping technology combined with them.

Nowadays, the existing neural network researches mainly focus on three aspects: network structure and optimization, learning and training algorithm, and practical application [4]. According to the real-time performance of network applications, these researches are separated into two fields: static network research and dynamic network research. The back propagation network (BP network) is the most extensively applied network model. It has a strong biological background because of its excellent input-output mapping characteristics. The BP network has a strong advantage in multivariable function approximation. Although it can achieve global optimization in theory, the optimization easily sinks into local minimum due to the algorithm restriction. The radial basis function network not only has the biological background but also coincides with the function approximation theory and is also suitable for multivariable function approximation. The theoretical basis of the orthogonal polynomial function network is relatively perfect, but for the nonlinear modeling and prediction of complex problems, the magnitude of network nodes often augments rapidly [5]. The advantage of the spline function network is that it only needs local information when learning; so, it has obvious advantages in the parallelism and convergence velocity of the algorithm, but because the subarea network separation in its definition domain is extraordinary complex, thus increasing the difficulty of practical application. The existing dynamic network research mainly focuses on real-time control, which requires the designed network to be simple in structure and fast in convergence. Typical dynamic networks include Hopfield network, adaptive resonance theory (ART) network, and dynamic recursive network [6]. The dynamic network is a single-layer network in the network structure, realizing the system's complex behavior control and simulation with less network structure overhead because of the internal feedback connection. So, it is more suitable for nonlinear dynamic system identification and control and other fields.

In view of the existing network vulnerability analysis methods, the important differences of nodes in the network and the relationship between network nodes are often ignored. The research challenges of prediction frequency are listed as follows: (1) The prediction accuracy of frequency hopping signals with interference is strictly requested in the military. Meanwhile, the applied prediction scheme must have a high operability. (2) Considering that the military scenario has a strict requirement on the timeliness of applications, it is rather a challenge to reduce the running time of the applied algorithm while ensuring accurate decisions are made. (3) Limited spectrum resources aggravate the interference between communication devices. Designing a method overcoming the interference among devices while ensuring network efficiency is necessary. In the network vulnerability analysis, this project considers the communication characteristics of the node itself and the standing and influence of the network node structure from the multidimensional point of view of time and space. Finally, the fragile node is determined to provide a prerequisite for the frequency band selection of interference targets. The main contributions of this paper are summarized as follows:

(1) Compared with the interference to the enemy signal in the military field, this scheme predicts the frequency band of the frequency hopping signal from the point of view of the nature of the frequency hopping signal. It not only enlarges the precision of the interference but also is more convenient in practical operation

(2) The clustering selection optimization algorithm is introduced into the radial basis function neural network (RBF neural network) to predict the frequency band of frequency hopping communication. It is improved based on k-means and has higher accuracy than the k-means algorithm. It provides convenient conditions for the second step prediction and reduces the amount of computation. The concept of the optimal learning rate is introduced into the traditional gradient descent method, which makes the optimization process more efficient and convenient. The optimal learning rate is calculated by the least mean square matrix algorithm, which effectively reduces the running duration of the algorithm

(3) The second scheme applies the improved particle swarm optimization algorithm to the RBF neural network and uses a subtractive clustering algorithm to optimize the number of RBF network centers. The proposed scheme in this paper availably raises the overall efficiency of the network and the accuracy of frequency hopping communication band prediction

The remaining work of this paper is organized as follows: Section 2 introduces the frequency hopping signal prediction scheme based on the RBF neural network proposed, Section 3

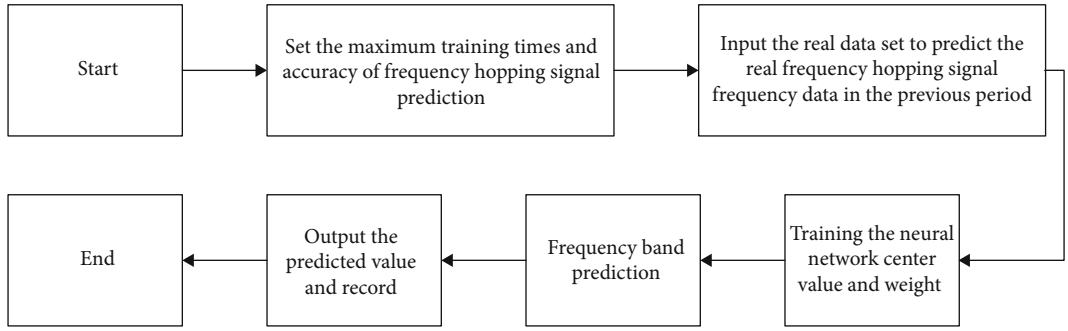


Figure 1: The RBF neural network prediction of frequency hopping signal flow diagram.

introduces the structure of RBF neural network based on the improved particle swarm optimization and subtractive clustering algorithm, Section 4 explains the experimental results, and Section 5 summarizes the full-text work.

2. Frequency Band Prediction Model of Frequency Hopping Signal Based on the Gradient Descent Method

This section proposes an RBF neural network prediction model based on an improved gradient descent method due to the existing problem of low accuracy of prediction models. It combines the clustering algorithm and the optimal learning rate to predict the frequency band of the frequency hopping signal.

2.1. Design Overview. In the RBF neural network prediction model, this scheme regards the current signal frequency and the signal frequency of the next time unit as the input and output in the RBF neural network [7], respectively. This prediction process is adaptive and is carried out gradually. In the specific prediction process, the network is constantly optimizing according to the actual value of the predicted signal. The number of signals in this process always remains the same, that is, when the next signal is optimized, the first signal is removed, and when the next signal is predicted, the first two signals are removed. In this way, a requirement of always keeping the same data optimized in the network is achieved. This approach can reduce the computing time and improve the optimization efficiency of the entire network. The method of real-time prediction of the network structure can greatly improve the accuracy of the network to achieve the experimental goal. Figure 1 shows a schematic diagram of the specific RBF neural network prediction process.

According to Figure 1 above, the RBF neural network's prediction model needs to go through the following steps: First, set the accuracy of the predicted model to achieve higher prediction accuracy than the existing scheme, which is set as 90%. Next, we need to obtain a certain number of frequency hopping signal transmission frequency bands as the training set of the algorithm and then use the generated data set to optimize the structure of the entire network and determine the number, size, and weight of the center. Then, the frequency band prediction of the frequency hopping signal can be realized.

In the prediction process of the RBF neural network, training the central value and weight of the neural network is the key step in this scheme, and it is also the key research link in this section. The training of the central value and weight directly determines the quality of the RBF neural network prediction model. This section proposes a scheme based on a clustering selection optimization algorithm and gradient descent method, which constructs an RBF neural network prediction model to achieve an accurate prediction of frequency hopping signal frequency bands.

The prediction model that had been put forward in this section is divided into two steps in the optimization design of the model [8] according to the gradient-RBF neural network. The first step is to optimize the selection of the central value in the RBF neural network, and the main application is a new cluster selection optimization algorithm. Compared with the commonly used k-means clustering selection optimization algorithm, it will make the size of the center value more accurate, thereby reducing the time required for the second step, improving the work efficiency of the network, and improving the accuracy. The second step is the optimal selection of weights. This section uses the gradient descent method and the least mean square algorithm to convert the learning rate in the gradient descent method into the optimal learning rate so that the efficiency of the entire neural network will be significantly improved, and the signal frequency prediction will be more accurate.

2.2. Selection of the Number of the Center c of the Hidden Layer of the RBF Neural Network. This program introduces the RBF neural network prediction method, which is used to predict the frequency of the frequency hopping signal in the frequency hopping communication. Due to the particularity of the RBF neural network structure, it is necessary to construct a reasonable network structure before predicting the frequency of the frequency hopping signal to achieve the purpose of accurate prediction. Therefore, how to find the optimal structure has become the most important issue in this scheme. The main idea of selecting the center is to make the center point as the center of a certain area, and the number of centers is required to be appropriate.

The pseudorandom code irregularly controls the frequency of the frequency hopping signal, but it may be concentrated in a certain area at different times [9]. Therefore, choosing the center value in these several areas will make

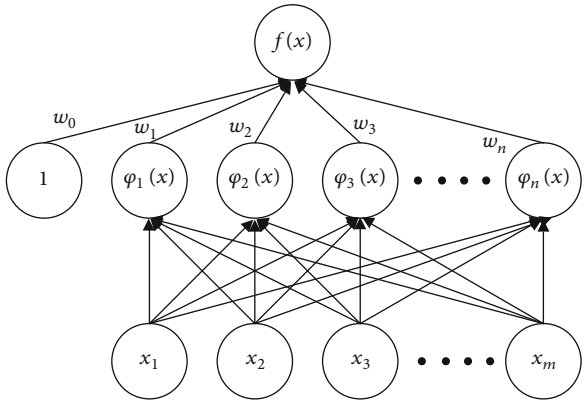


Figure 2: The input-neuron-output diagram of the RBF neural network.

the neural network structure more optimized and simplify the training process during the training process. The following scheme introduces an improved algorithm based on the k-means clustering selection algorithm. This algorithm is more suitable for one-dimensional input of frequency hopping signal frequency, and its calculation time is shorter, and the accuracy is higher.

After comparing the radial basis functions mentioned above, this scheme selects the Gaussian function as the radial basis function of the RBF neural network [10]. Because the Gaussian function decreases monotonously from the center to both sides, its response is locally limited. There will be a local response only at a position close to its center. The Gaussian function is similar to the biological neuron in the RBF neural network and the linear summation weighted by N Gaussian functions as the final output. Thus, the RBF neural network transformed the original nonlinear problem into the current linear problem [5]. The functional equation $f(x) = w_0 + \sum_i^n w_i \varphi(\|x - c_i\|)$ is the mapping equation from input to output, where $c_i (1 \le i \le n)$ denotes the center of the RBF neural network, and norm $\|\bullet\|$ represents the Euclidean distance and $w_i (1 \le i \le n)$ is the weight of the linear summation. x and w_o represent the input vector and deviation, respectively. The Gaussian function is selected as the radial basis function $\varphi(\cdot)$. Figure 2 shows a schematic diagram of its structure.

For the selection of center c, this section introduces an improved cluster selection optimization algorithm based on the k-means clustering algorithm, which has a faster convergence speed than the k-means algorithm, and the accuracy of optimizing the center value is higher.

In this learning method, the center is determined by self-organizing learning, which can be moved. The selection of the center can use the mean clustering selection optimization algorithm. This is an unsupervised learning method, and the center changes with the number of samples. The following are the detailed steps:

(1) Initialize cluster center $c_i (1 \le i \le n)$. If the value of n is increasing, the accuracy is also improved, but the amount of calculation will increase [11]. Select n samples from the input sample $x(1, 2, \cdots, m)$ as cluster centers. (n is given in advance, according to the characteristics of the frequency hopping signal, n is given as 50)

(2) Group the input samples according to the nearest neighbor rule, assign $x(1, 2, \cdots, m)$ to the center c; θ_n is the cluster set of the input samples and meets$d = d_{\min} = \|x_m - c_i\|$, and j is the number of elements in each cluster set

(3) Calculate the average distance $\bar{d} = (1/j)\sum_1^i \|x_j - c_i\|$ of each center c_i and the average distance $\bar{D} = (1/m)\sum_1^m \|x_m - c_i\|$ of the entire system (each input x_m corresponds to its cluster center)

(4) Find c_a corresponding to $\bar{d}_{i\,\max}$ and randomly select another center c'_a, if it is $\bar{D}_i{}' \le \bar{D}_i$, keep c'_a; otherwise, keep c_a

Repeat steps 3 and 4, when the random center c'_a always satisfies $\bar{D}_i{}' \le \bar{D}_i$, and the optimization is completed, and the cluster center c_i is output. The width of the RBF network is $\sigma = d/\sqrt{N}$, where d represents the maximum distance between all centers and N represents the number of centers.

2.3. Improved Gradient Descent Method and Implementation Scheme. After the center optimization is selected, the weights are optimized by the gradient descent method. In this scheme, the Gaussian function and matrix are used to calculate the actual output value $\hat{y}$ of the network, and y, t, and w denote the actual output value of the network, the number of iterations, and the weight, respectively. Herein, $e(t)$ represents the error function, $E(t)$ is the cost function, and $\eta(t)$ is the learning efficiency function.

The update function of the weight is

$$w(t+1) = w(t) - \eta(t)\frac{dE(t)}{dt}. \tag{1}$$

By using the least mean square algorithm, we can obtain

$$\varphi(\|x - c_i\|) = \exp\left(-\frac{1}{2\sigma^2}\|x - c_i\|^2\right) (i = 1, 2, \cdots, n), \tag{2}$$

where n is the center number, and m is the total number of input samples. In this experiment, we assume n is 50.

$$\varphi_{ji} = \varphi\left\|x_j - c_i\right\| \; j = 1, 2 \cdots, m, \tag{3}$$

$$\phi\omega = y\,\omega = (\omega_1, \omega_2, \cdots, \omega_n)^T. \tag{4}$$

The cost function can be obtained by

$$E(t) = \frac{1}{2}\sum_{j=1}^{m}\left[y_j - y\wedge(t)\right]^2 = \frac{1}{2}\sum_{j=1}^{m} e_j{}^2(t). \tag{5}$$

By using the gradient descent method for function $E(t)$,

we can obtain

$$\nabla E(t) = \sum_{j=1}^{m} e_j(t) \frac{\partial e_j(t)}{\partial \omega(t)} = -\sum_{j=1}^{m} e_j(t) \frac{\partial \left[\varphi_j \bullet \omega(t)\right]}{\partial \omega(t)}, \tag{6}$$

$$\nabla E(t) = -\sum_{j=1}^{m} e_j(t) \varphi_j^T. \tag{7}$$

The variable for the tth learning weight is $\Delta\omega(t)$.

$$\Delta\omega(t) = \omega(t+1) - \omega(t), \tag{8}$$

$$\Delta\omega(t) = -\eta(t)\nabla E(t-1) = \eta(t) \sum_{j=1}^{m} e_j(t-1)\varphi_j^T. \tag{9}$$

The optimal learning rate is to use the learning rate in the gradient descent method to achieve the highest efficiency through matrix calculation [12]. The scheme is listed as follows:

$$\Delta e(t) = e(t) - e(t-1) = -\Delta \hat{y}(t) = -\eta(t)\phi\Delta\omega(t) = -\eta(t)\phi\phi^T e(t-1), \tag{10}$$

$$e(t) = e(t-1) - \eta(t)\phi\phi^T e(t-1), \tag{11}$$

$$E(t) = \frac{1}{2} e^T(t)e(t) = \frac{1}{2}\left[e(t-1) - \eta(t)\phi\phi^T e(t-1)\right]^T \bullet \left[e(t-1) - \eta(t)\phi\phi^T e(t-1)\right]. \tag{12}$$

From the above derivation, the optimal learning rate can be obtained:

$$\eta^*(t) = \nabla e(t) = e(t) - e(t+1) = \frac{e^T(t-1)\phi\phi^T e(t-1)}{e^T(t-1)\phi\phi^T\phi\phi^T e(t-1)}. \tag{13}$$

Specific steps are as follows:

(1) We need to get the number of hidden layer nodes n through the above method and use the mean cluster selection algorithm to find the center c and the width σ. Set learning accuracy rmse* and the maximum number of iterations max t

(2) Set the weight of the output layer and hidden layer to $w_0 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ and calculate $E(t) = \text{rmse}$

(3) Let $\eta(t) = \eta^*(t)$ and $w = w + \eta(t)\phi^T e(t-1)$ iterate

(4) If rmse $\geq$ rmse* or $t \leq \max t$, repeat step 3; otherwise, stop and output $\hat{y}$ and t

The above optimization scheme can obtain the center, weight, and width of the RBF neural network through optimization training and learning [13]. When testing the data, the obtained data can continue to be input into the network structure as a sample, and the above steps are repeated. In this way, when the test data is more, the network structure is more accurate, and the measurement result is more accurate.

3. Frequency Band Prediction Model of Frequency Hopping Signal Based on the RBF Neural Network Based on Improved Particle Swarm Optimization (PSO-RBF Algorithm)

In view of the limitations of the above schemes in the military field and relatively speaking, the algorithm simulation time is long, and a global prediction scheme is proposed to predict the frequency band of frequency hopping signals in this section. This scheme uses the subtractive clustering algorithm and improved the particle swarm optimization (PSO) algorithm. We propose an RBF neural network prediction model based on improved PSO and optimize the center value and weight of the network to achieve the purpose of frequency hopping signal frequency band prediction.

3.1. Design Overview. In the previous section, the center and weight of the hidden layer in the RBF neural network are optimized by the step-by-step method. To improve the accuracy of the scheme as a whole, the accuracy of each step is high, which leads to too many iterations. Therefore, this section applies the improved PSO algorithm to the RBF neural network to predict the signal frequency band in the frequency hopping communication [14]. The prediction model proposed in this section optimizes the structure of the RBF neural network through real data, compares the predicted value with real value, and reoptimizes the adaptive prediction process of the network mechanism with real value as the training sample. Figure 3 shows a schematic diagram of the specific forecasting process.

Figure 3 above shows the prediction diagram, which is similar to the prediction process mentioned in Section 2. Some improvements are made on the optimization of the center, width, and weight of the RBF neural network. Data flow in this prediction process represents the real FM signal transmission frequency. The real value is taken as the training sample of the prediction model, which is optimized by the scheme in this section. By constantly adjusting the parameters, we can finally achieve the purpose of optimizing the RBF neural network construction.

In this scheme, the improved PSO algorithm is applied to the RBF neural network to predict the frequency band of the frequency hopping signal model. PSO is a random search algorithm that simulates biological activities in nature and is a branch of evolutionary computing. It finds the optimal solution through the cooperation mechanism in the population and is extensively used in various engineering optimization problems. The RBF neural network prediction model applies the improved PSO algorithm, which regards the center, width, and weight of the RBF neural network as a three-dimensional optimization structure, and optimizes the prediction model continuously through the parameters and

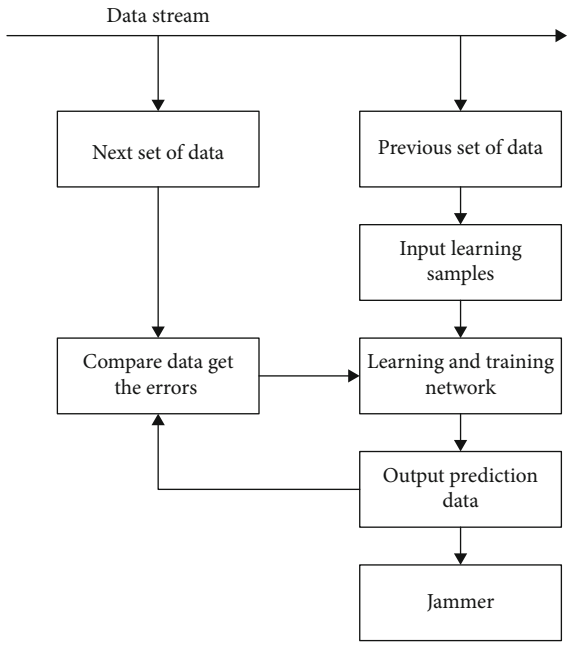


Figure 3: The data processing flow diagram of the prediction system.

fitness function of the particle swarm [15]. The step of optimizing the speed and position of particles is the most important link in this scheme. Combined with the advantages of the above two, the PSO algorithm is used to optimize the central parameters of the basis function. While using the subtractive clustering algorithm to determine the number of centers of the radial basis function, the PSO algorithm is used to adjust the connection weight between the hidden layer and the output layer of the network to achieve the overall optimization effect of the RBF neural network [16].

3.2. The Center Number of the RBF Neural Network Is Determined by the Subtractive Clustering Algorithm. The Gaussian function constitutes the vector basis, and several vector bases constitute the hidden layer of the RBF neural network [17], which is a method of approximating nonlinear functions by linear functions. In this section, the radial basis function used in the hidden layer of the RBF neural network is the Gaussian function, and the transmission frequency of frequency hopping signal in frequency hopping communication is the nonlinear function. In this scheme, the accuracy of the prediction model is directly determined by the number of radial basis functions in the hidden layer of the neural network. Under normal circumstances, the more basic functions, the higher the accuracy of the algorithm and then the more iterations of the algorithm [18]. Therefore, the key to this prediction model is that the count of hidden layer centers needs to be determined.

The subtractive clustering algorithm is a relatively effective clustering method to determine the number of basis function centers. The calculation principle of the subtractive clustering algorithm is to select the first central point from the position with the highest density in a group, get the first central point, subtract this point and the points near this point, and then repeat this process, and so on; until the current highest density is met. The specific algorithm is as follows.

Consider normalizing the data into p data points $(x_1, x_2, \cdots, x_p)$ in an n-dimensional unit of hyperspace [19], the following formula gives the density at data point x_i:

$$D_i = \sum_{i=1}^{p} \exp\left[-\frac{\left\|x_i - y_j\right\|^2}{\gamma_a/2}\right]. \tag{14}$$

After calculating the density of each data point, we chose the point with the highest density as the first cluster center and denoted its density as D_c. The updated formula for the density of each data point is

$$D_i = D_i - \sum_{i=1}^{p} \exp\left[-\frac{\|x_i - x_{c1}\|^2}{\gamma_a/2}\right]. \tag{15}$$

After the density of each point is updated, we select the next clustering center and modify its density again [20]. This process is repeated until the current maximum density D is far less than the initial maximum density, and the formula is as follows:

$$\frac{D_{\max}}{D_{c1}} < \lambda. \tag{16}$$

Thus, the clustering ends, and the number of clustering centers obtained is equal to the number of centers of the basis function.

3.3. Implementation of Improved PSO in the RBF Neural Network

3.3.1. Fitness Selection Based on Frequency Hopping Signal. In this scheme, the fitness function selects the error function of the whole system, which can show the quality of the system more intuitively and has more intuitive significance for the optimization of the internal structure of the system.

In the PSO algorithm, a particle corresponds to a feasible solution. First, we code the particle, which includes the center value and width of the basis function, particle velocity, and fitness [21]. According to the subtractive clustering algorithm, suppose m centers are determined and each center is k-dimensional, then the position of the particle is $m \times (k + 1)$-dimensional, the velocity of it is also $m \times (k + 1)$-dimensional, σ_i represents the width of the i basis function, and f_i is the fitness of the i individual. Formulas (17) and (18) show the fitness function:

$$f_i = \frac{1}{R_i}, \tag{17}$$

$$R_i = \frac{1}{M}\sum_{k=1}^{m} (y_k - y\wedge_k)^2. \tag{18}$$

Table 1: Initialization values of PSO-RBF algorithm.

Parameter	Value
Number of iterations G	20
Particle dimension n	15
Population size m	20
Algorithm parameters w	0.1
Algorithm parameters $c1$	2
Algorithm parameters $c2$	2

3.3.2. Selection of the Update Function of Improved PSO. A random particle swarm is generated by the PSO algorithm, and each particle is given a random velocity. During the flight, the speed of the particles is dynamically adjusted by the pulling force of their own and companions' flight experience, and the whole group can fly to a better search area [22]. Suppose that the search space is N-dimensional, the total number of particles is Z, the position of the i particle in the N-dimensional space is expressed as $x_i = (x_{i1}, x_{i2}, \cdots, x_{\mathrm{iN}})$, and the flight speed is expressed as $v_i = (v_{i1}, v_{i2}, \cdots, v_{\mathrm{iN}})$. The optimized objective function assigns a fitness value to each particle. Each particle is searched in the solution space immediately following the current optimal particle. In iterations, each process is not completely random, and if we find a better solution, we can use it as a basis for finding the next solution. Each time the particle updates itself, it tracks two "extremes" in the iteration: one is the individual extreme point [23], that is, the best solution found by the particle itself, whose position is expressed as pbest; the other is the global extreme point, which is the best solution found by the whole population at present, and its position is expressed as gbest. When the two best solutions are found, the velocity and position of particle will be updated according to formula (19):

$$v_{id}^{k+1} = w v_{id}^{k} + c_1 \, \mathrm{rand}_1^{k} \left(\mathrm{pbest}_{id}^{k} - x_{id}^{k} \right) + c_2 \, \mathrm{rand}_2^{k} \left(\mathrm{gbest}_{id}^{k} - x_{id}^{k} \right), \tag{19}$$

where v_{id}^{k} denotes the d-dimensional component of the flight velocity in the k iteration of the i particle, and x_{id}^{k} represents the d-dimensional component of the position in the k iteration of the i particle [24]. pbest_{id}^{k} denotes the position of the individual extreme point of the particle i in the d-dimension. gbest_{id}^{k} denotes the position of the global extreme point of the whole swarm in the d-dimension. w stands for inertia weight, and c_1 and c_2 are learning factors. The maximum step size of the global best particle and the individual best particle flying in the direction of the global best particle and the individual best particle is adjusted by them severally. General order$c_1 = c_2 = 2$andrand 1are stochastic numbers among [0, 1].

Different from the application of the traditional particle swarm optimization algorithm, the prediction model of the RBF neural network regards the radial basis function in the hidden layer of the neural network as a particle. The center value, width, and weight are regarded as particles that need to be searched in the three-dimensional solution space. Therefore, the value of d in this scheme is 3. pbest_{id}^{k} represents the best position of the predicted value of this time in the training sample, and gbest_{id}^{k} represents the best position of the predicted value of all training samples up to the current time. In this training model [25], each radial basis function has a memory function and records its center value and width, as well as its weight connected to the output layer.

3.3.3. Algorithm Step. This algorithm optimizes the RBF neural network, using subtractive clustering algorithm and improved the particle swarm optimization algorithm [26]. The following are the specific steps:

(1) Input the transmission frequency bands of a certain number of real frequency hopping signals (2000-100) as set training samples optimized by the RBF neural network

(2) According to the subtractive clustering algorithm mentioned above, we perform clustering analysis on the samples, so that the number of centers can be determined

(3) Initialize the particle swarm, that is, we set the parameters of the PSO algorithm, as well as the initial position of the particles. The neurons in the hidden layer of the RBF neural network are represented by the particles here

(4) If the fitness of each particle is better than that of the best position it has ever experienced, then update the optimal position pbest_{id}^{k} of the particle

(5) If the fitness of each particle is better than that of the best position experienced by the particle population, then update the global optimal position gbest_{id}^{k} of the particle

(6) Modify the velocity and position of particles

(7) Continue the cycle from steps 4 to 6 and end the cycle when the result meets the calculation requirements

(8) Decode the best position experienced by the particle swarm and use the decoded value as the structural parameter of the RBF neural network. Then, learn from the network

4. Performance Analysis

In this experiment, 2000 data generated by MATLAB are used as training samples and 100 as test samples. The initialization values of the PSO-RBF algorithm are shown in Table 1.

Figures 4 and 5 compare between the real and predicted values obtained when the training set of different methods is 2000 data. There are 100 test sets, and it can be seen from the simulation figure that the distance between the real value

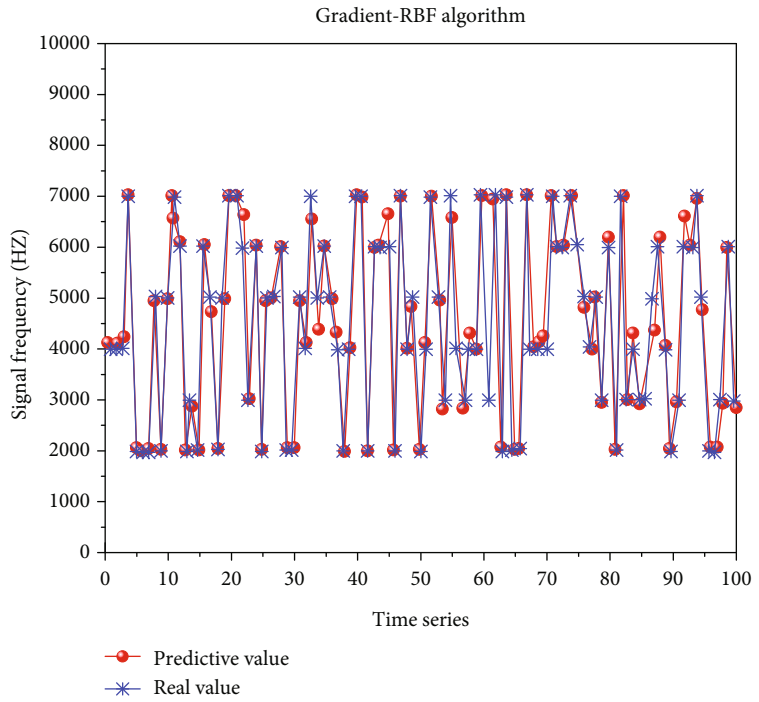


Figure 4: The error diagram of gradient-RBF prediction results.

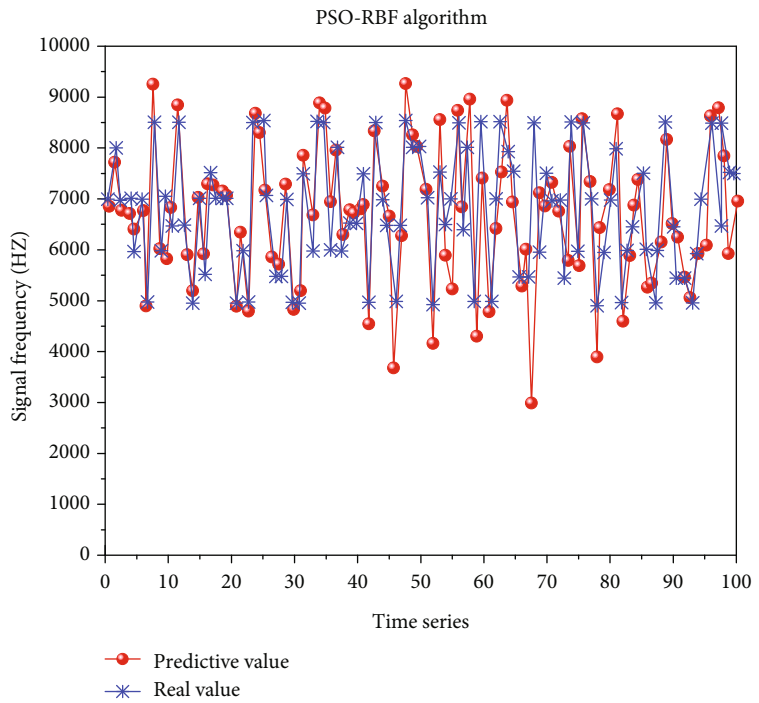


Figure 5: The error diagram of prediction result of PSO-RBF algorithms.

and the predictive value of most of the vast majority data is small, indicating that the error is small.

Because the scheme of this paper is mainly used to attack the enemy's communication system in the military field, therefore, the demand for time efficiency is high. In the first scheme, the RBF neural network based on the gradient descent method utilizes 100 to 2000 training data to analyze the impact of time efficiency and accuracy in detail. The details are shown in Figures 6 and 7 below.

Figure 7 depicts the changing trend between the gradient-RBF algorithm and the traditional correlation-based algorithm model prediction accuracy and the simulation time

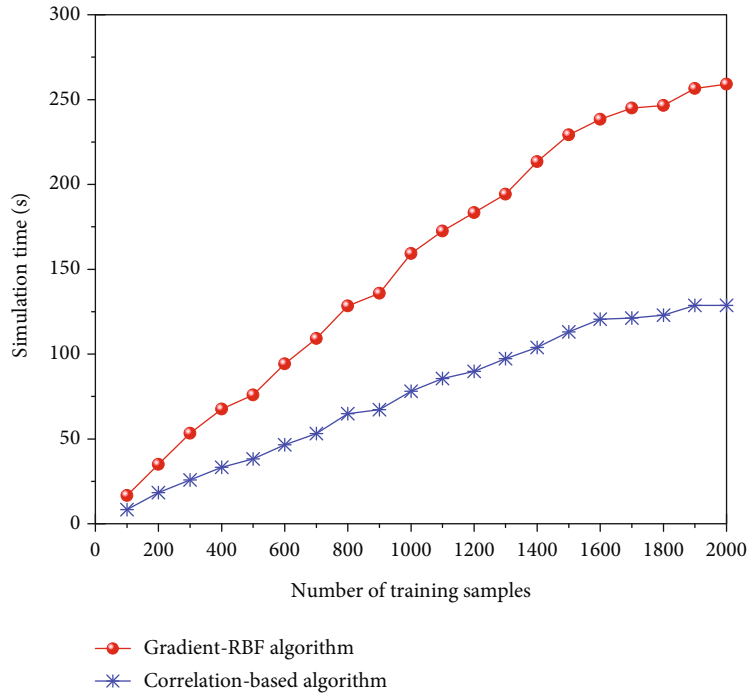


Figure 6: The relation diagram between the number of training samples and time efficiency.

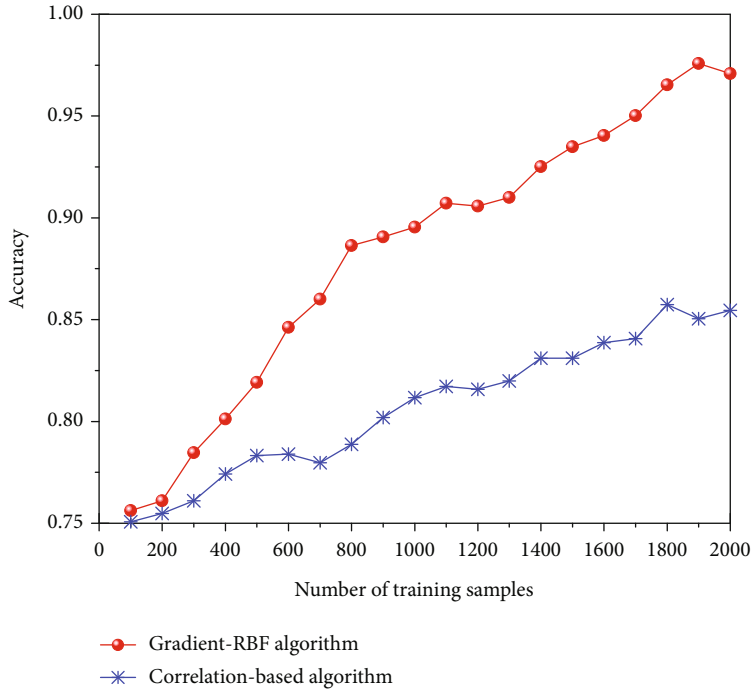


Figure 7: The accuracy comparison of the scheme.

of the algorithm. We can obtain from the figure that as the number of training sets increases, the simulation time curve of the two prediction models is tortuous. The simulation time of the gradient-RBF algorithm is getting longer. It can also be said that the simulation time is proportional to the number of training set samples. Compared with the simulation results of the two algorithms in the figure, the simulation time of the prediction model based on the correlation degree algorithm is significantly lower than that of the gradient-RBF algorithm. The cause of this trend is that the massive training samples and the massive iterations of the algorithm also increase, resulting in longer simulation times.

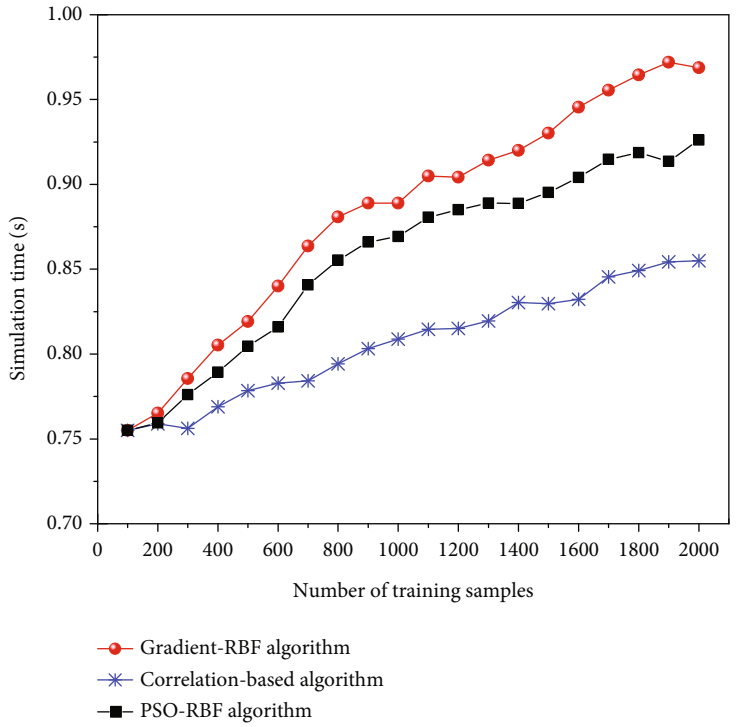


Figure 8: The relation diagram between the number of training samples and accuracy.

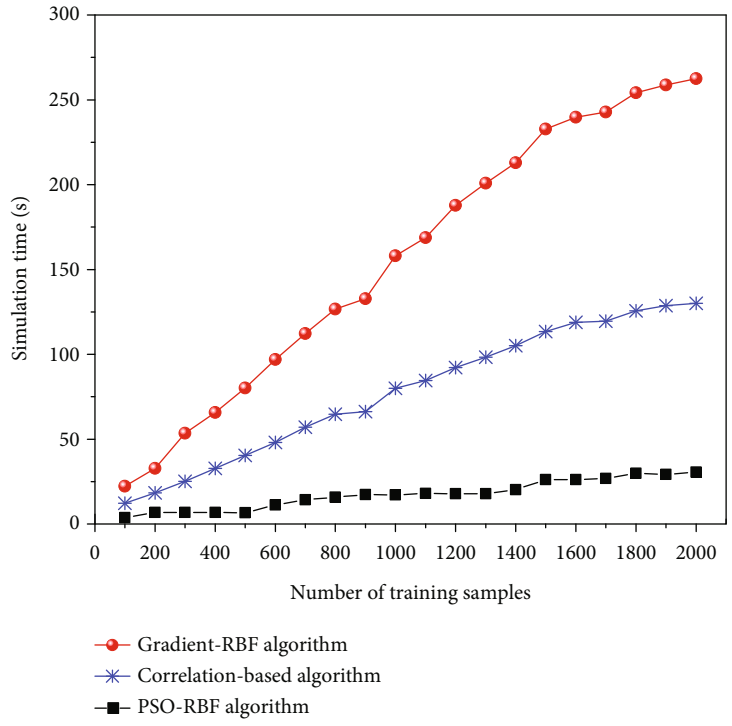


Figure 9: the relation diagram between the number of training samples and time efficiency.

Figure 8 shows the yellow line in the figure shows the changing trend between model prediction accuracy and accuracy in the PSO-RBF algorithm. Through the comparison of the three lines, we can see that the accuracy of the gradient-RBF algorithm and the PSO-RBF algorithm is obviously better than the traditional correlation-based algorithm. As the number of training sets increases, the accuracy curve of the prediction model rises tortuously, and the accuracy

of the algorithm gradually increases, up to 92%. This indicates that the accuracy is proportional to the number of training set samples. The reason for this trend is that the number of training set samples increases, the number of iterations of algorithms also increases, and the center value and weight are better optimized, thus improving the accuracy of the algorithm. In this paper, we make use of the subtractive clustering algorithm to determine the number of centers of the radial basis function of the RBF neural network and the center value, width, and weight between the hidden layer and the output layer is optimized by the PSO algorithm. The experimental results illustrate that the optimization of the neural network by particle swarm optimization is effective.

Figure 9 shows that the yellow line in the figure indicates the changing trend between the model prediction accuracy of the PSO-RBF algorithm and the simulation time of the algorithm. Through the comparison of the simulation time of the three schemes, we can obtain that the simulation time curve of the prediction model rises tortuously as the number of training sets increases, and the simulation time of the algorithm becomes longer gradually. In other words, the simulation time is proportional to the number of training set samples. This is due to the number of training samples increasing, and the number of iterations of the algorithm also increasing, thus making the simulation time longer. The PSO-RBF algorithm of this scheme adopts the global optimization prediction model; so, the simulation time is much less than the other two schemes.

We select four specific indicators for systematic analysis to analyze the performance of the proposed scheme and the other existing schemes more clearly, including mean absolute error (MAE), mean absolute error rate (MAER), root mean square error (RMSE), and symmetrical mean absolute percentage error (SMAPE). The specific formulas are as follows. Through these indicators to judge the relationship between the final output and the actual value, the specific criteria are as follows:

(1) Mean absolute error (MAE) represents the actual situation of the prediction error

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}|X_t - F_t|, \tag{20}$$

where X_t is the actual value, F_t is the predicted value, and n is the total number of samples

(2) Mean absolute percentage error (MAPE) is a criterion for describing predicted and actual error values

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{X_t - F_t}{X_t}\right|, \tag{21}$$

for MAPE, the closer the value is to 0, the better the prediction effect is

Table 2: The quantitative analysis of 1000 training sets prediction results.

Algorithm	MAE (Hz)	MAPE	RMSE (Hz)	SMAPE
Gradient-RBF algorithm	105	0.0361	11.7	0.0722
Correlation-based algorithm	656	0.1452	24.6	0.01285
PSO-RBF algorithm	458	0.0932	22.7	0.0872

Table 3: The error analysis of two schemes.

Algorithm	Simulation time (s)
Gradient-RBF algorithm	146
Correlation-based algorithm	82

(3) Root mean squared error (RMSE) is the square root of the mean square difference between the prediction data and the original data

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_t - F_t)^2} \tag{22}$$

(4) Symmetrical mean absolute percentage (SMAPE) is a criterion for describing predicted and actual errors, which is expressed as

$$\mathrm{SMAPE} = \frac{2}{n}\sum_{t=1}^{n}\frac{|X_t - F_t|}{X_t + F_t}, \tag{23}$$

and the range of value is [0,2). The closer the value is to 0, the better the prediction effect and the smaller the error.

The real and predicted values of the test results of 1000 training in this scheme are brought into the above formulas. The result data can be obtained as follows.

In Table 2, SMAPE can clearly show the error between the real value and the prediction value of the system. The two schemes proposed in this paper are only slightly insufficient in accuracy compared with the gradient-RBF algorithm mentioned above when the data selected by the gradient-RBF algorithm is 1000 training sets.

This scheme compares the simulation running time of the two schemes through the specific data of Table 3, which directly shows the time defect of the gradient-RBF algorithm. The reason for this problem is that the number of iterations of the algorithm is too many.

It can be obtained from Table 4 that in the training process of the gradient algorithm, each iteration reaches the highest number of iterations, and the corresponding time is longer, while the prediction result is higher than that of the improved particle swarm algorithm. However, after meeting the minimum accuracy requirement, the improved particle swarm optimization algorithm does not continue to iterate.

Table 4: The error-index analysis of two schemes of 2000 training sets.

Algorithm	Gradient-RBF algorithm	PSO-RBF algorithm
Time performance (s)	256	28
Accuracy	98%	92%

It optimizes the center, width, and weight of the RBF neural network from the global search. This scheme is much higher in time efficiency than the first scheme. As long as the parameter adjustment work is meticulous and the parameter adjustment is reasonable, the accuracy of the gradient algorithm can be achieved.

5. Conclusions

This paper analyzes and compares the features of related algorithms with their optimization algorithms. We apply each method to the frequency band prediction process of frequency hopping signals, compare the pros and cons of related algorithms, and propose two kinds of prediction schemes. The first is to first predict the center c in the RBF neural network and use the cluster selection optimization algorithm. After optimizing the center c, using a gradient descent method and least mean square algorithm to calculate optimal learning rate, so as to improve the efficiency of the entire system. By analyzing simulation results, it is found that the accuracy of this scheme is remarkably progressed compared with a traditional correlation-based algorithm, but time efficiency is slightly longer. The second scheme is proposed in response to the shortcoming of the long computing time of the first scheme. It applies the improved particle swarm algorithm to the RBF neural network and uses the reduced clustering algorithm to calculate the number of RBF neural network centers. Then, the proposed scheme is used to directly optimize the center and weight for predicting the frequency problem of the frequency hopping signal of the nonlinear function in the RBF neural network. After the error analysis and comparison of the two schemes, it is found that using the second scheme to enhance the particle swarm optimization algorithm in the RBF neural network can effectively increase the time efficiency and make system accuracy above 90%. In this way, it realizes the prediction of the nonlinear function, that is, the prediction of the frequency hopping signal, which makes it operable to attack the enemy's communication system in the specific military field. For our future work, we will consider blockchain-based communications and high-efficient learning solutions in our framework, such as [27, 28].

Data Availability

The dataset used in this paper is generated by MATLAB for evaluations.

Conflicts of Interest

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

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61931019, Shenzhen Technology Project (JCYJ20180302145648171), and the Fundamental Research Funds for the Central Universities (lzujbky-2020-kb25).

References

[1] W. Wang, F. Xia, H. Nie et al., "Vehicle trajectory clustering based on dynamic representation learning of internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020.

[2] J. Zhang, B. Wei, F. Wu et al., "Gate-ID: WiFi-based human identification irrespective of walking directions in smart home," *IEEE Internet of Things Journal*, 2020.

[3] C. Cormio and K. R. Chowdhury, "Common control channel design for cognitive radio wireless ad hoc networks using adaptive frequency hopping," *Ad Hoc Networks*, vol. 8, no. 4, pp. 430–438, 2010.

[4] W. Wang, X. Zhao, Z. Gong, Z. Chen, N. Zhang, and W. Wei, "An attention-based deep learning framework for trip destination prediction of sharing bike," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2020.

[5] W. Wang, T. Tang, F. Xia, Z. Gong, Z. Chen, and H. Liu, "Collaborative filtering with network representation learning for citation recommendation," *IEEE Transactions on Big Data*, pp. 1–14, 2020.

[6] Z. Ning, P. Dong, X. Wang et al., "Distributed and dynamic service placement in pervasive edge computing networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 6, pp. 1277–1292, 2021.

[7] S. Gao, J. Sun, and X. Gao, "Soft-sensor modeling of rectification of vinyl chloride based on improved PSO-RBF neural network," in *2012 24th Chinese Control and Decision Conference (CCDC)*, Taiyuan, China, 2012.

[8] M. Er, S. Wu, J. Lu, and H. L. Toh, "Face recognition with radial basis function (RBF) neural networks," *IEEE Transactions on Neural Networks*, vol. 13, no. 3, pp. 697–710, 2002.

[9] J. Zhang, F. Wu, B. Wei et al., "Data augmentation and dense-LSTM for human activity recognition using WiFi signal," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4628–4641, 2021.

[10] M. Xu, H. Chen, and L. Duan, "A combined training algorithm for RBF neural network based on particle swarm optimization and gradient descent," in *2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*, Liuzhou, China, 2020.

[11] X. Wang, Z. Ning, and S. Guo, "Multi-agent imitation learning for pervasive edge computing: a decentralized computation offloading algorithm," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 2, pp. 411–425, 2021.

[12] W. Wang, F. Xia, J. Wu, Z. Gong, H. Tong, and B. D. Davision, "Scholar2vec: vector representation of scholars for lifetime collaborator prediction," *ACM Transactions on Knowledge Discovery from Data*, pp. 1–20, 2020.

[13] X. Wang, Z. Ning, S. Guo, M. Wen, and V. Poor, "Minimizing the age-of-critical-information: an imitation learning-based scheduling approach under partial observations," *IEEE Transactions on Mobile Computing*, p. 1, 2021.

[14] J. Y. Chen, Z. Qin, and J. Jia, "A PSO-based subtractive clustering technique for designing RBF neural networks," in *IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, pp. 2047–2052, Hong Kong, China, 2008.

[15] Z. Ning, P. Dong, X. Wang et al., "Partial computation offloading and adaptive task scheduling for 5G-enabled vehicular networks," *IEEE Transactions on Mobile Computing*, p. 1, 2020.

[16] B. Zhang, J. Y. Wang, and S. L. Zhang, "A new PSO-RBF model for groundwater quality assessment," *Advanced Materials Research*, vol. 463-464, pp. 922–925, 2012.

[17] Z. Ning, P. Dong, X. Wang et al., "Mobile edge computing enabled 5G health monitoring for internet of medical things: a decentralized game theoretic approach," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 2, pp. 463–478, 2021.

[18] D. Xiang-Jun and W. Yan-Qin, "RBF neural network identifier based on optimal selection cluster algorithm and PSO algorithm and its application," in *2011 Fourth International Conference on Intelligent Computation Technology and Automation*, pp. 884–887, Shenzhen, China, 2011.

[19] M. L. Zhang, "Ml-rbf: RBF neural networks for multi-label learning," *Neural Processing Letters*, vol. 29, no. 2, pp. 61–74, 2009.

[20] Z. Ning, S. Sun, X. Wang et al., "Intelligent resource allocation in mobile blockchain for privacy and security transactions: a deep reinforcement learning based approach," *Science China Information Sciences*, 2020.

[21] Y. Zhong, X. Huang, P. Meng, and F. Li, "PSO-RBF neural network PID control algorithm of electric gas pressure regulator," *Abstract and Applied Analysis*, vol. 2014, Article ID 731368, 7 pages, 2014.

[22] Z. Ning, K. Zhang, X. Wang et al., "Intelligent edge computing in Internet of Vehicles: a joint computation offloading and caching solution," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–14, 2020.

[23] G. Wenxian, W. Hongxiang, X. Jianxin, and Z. Yunfeng, "RBF neural network model based on improved PSO for predicting river runoff," in *International Conference on Intelligent Computation Technology & Automation, IEEE Computer Society*, Changsha, China, 2010.

[24] Z. Ning, R. Y. K. Kwok, K. Zhang et al., "Joint computing and caching in 5G-envisioned internet of vehicles: a deep reinforcement learning based traffic control system," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2020.

[25] N. Zhang, "Urban stormwater runoff prediction using recurrent neural networks," in *Advances in Neural Networks-ISNN 2011-8th International Symposium on Neural Networks, ISNN 2011*, Guilin, China, 2011May 29-June 1, 2011, Proceedings, Part I. DBLP.

[26] J. Y. Chen and Z. Qin, *Training RBF neural networks with PSO and improved subtractive clustering algorithms*, pp. 1148–1155, Neural Information Processing, Berlin, Heidelberg, 2006.

[27] Z. Ning, S. Sun, X. Wang et al., *Intelligent Resource Allocation in Mobile Blockchain for Privacy and Security Transactions: A Deep Reinforcement Learning Based Approach*, SCIENCE CHINA Information Sciences, 2020.

[28] X. Wang, Z. Ning, S. Guo, and L. Wang, "Imitation learning enabled task scheduling for online vehicular edge computing," *IEEE Transactions on Mobile Computing*, p. 1, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9896460, 1 page
https://doi.org/10.1155/2023/9896460

Retraction

Retracted: Mathematical Modeling and Simulation of Wireless Sensor Network Coverage Problem

Wireless Communications and Mobile Computing

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

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

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

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

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

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

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

References

[1] J. Guo, "Mathematical Modeling and Simulation of Wireless Sensor Network Coverage Problem," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5537004, 12 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5537004, 12 pages
https://doi.org/10.1155/2021/5537004

Research Article

Mathematical Modeling and Simulation of Wireless Sensor Network Coverage Problem

Jianxia Guo

School of Mathematics and Statistics, Xinxiang University, Xinxiang, Henan 453003, China

Correspondence should be addressed to Jianxia Guo; guojianxia@xxu.edu.cn

Received 6 February 2021; Revised 4 March 2021; Accepted 10 March 2021; Published 19 March 2021

Academic Editor: Wei Wang

The research on wireless sensor networks has achieved a lot in recent years and some of the results have been put into practical applications, but with the increasing demand and requirements for wireless sensor networks, many old and new problems need to be solved urgently. In this paper, a data topology optimization algorithm based on local tree reconstruction for heterogeneous wireless sensor networks is proposed for data transmission in wireless sensor networks that are easily affected by external instabilities. This heterogeneous network can accomplish better data transmission; firstly, the nodes are divided into different layers according to the hop count of nodes in the network, and a certain proportion of relay nodes are selected for different layer nodes; then, different initial energy is set for different layer nodes, and since the data packets of different nodes have different sizes, the corresponding data aggregation coefficients are used in this paper according to the actual data requirements of the network during data transmission; finally, the topology of the tree is dynamically updated in real time during the operation of the network to extend the lifetime of the nodes. The simulation results verify that the proposed heterogeneous network topology evolution algorithm effectively extends the network lifetime and improves the utilization of nodes. This paper establishes a modified least-squares target localization model to achieve accurate 3D localization of targets in real scenes and proposes an optimal base station node selection strategy based on spectral clustering using the location distribution information of base station nodes in space. The simulation results show that the error of the terminal 3D coordinates calculated by the proposed algorithm is smaller than the real coordinates, and the error is smaller than other existing algorithms with the same simulation data.

1. Introduction

Wireless sensor networks (WSNs) consist of a large number of inexpensive miniature sensor nodes deployed in a monitoring area, forming a multihop self-organizing network system through wireless communication, so that the network setup is flexible, the location of the devices can be changed at any time, and it can also be connected to the Internet by wired or wireless means. The development of this technology is due to the rapid development of microelectromechanical systems, system-on-a-chip, wireless communication, and low-power embedded technologies. The purpose of this technology is to collaboratively sense, collect, and process information from sensed objects in the network coverage area and send it to an observer. Sensors, sensed objects, and observers constitute the three elements of a wireless sensor network [1]. The dense deployment of nodes in the network enables quantitative data collection, data aggregation and processing, and data transmission, and the technology has played a pivotal role in the rapid development of the world [2]. Wireless sensor networks are a frontier hot research area involving multidisciplinary intersection and have a very promising application [3].

Wireless sensor network as a kind of big data transmission network in the context of the new era is gradually changing the human living conditions, convenient human life in all aspects; wireless sensor network has an important market value and research value. Government agencies and research units around the world are continuously increasing their research efforts [4]. For example, Jingdong Research Institute has invested a lot of money in the drone delivery business so that the delivery can be delivered to customers quickly and

precisely, which requires precise positioning between drones and customers, and this positioning communication is done by self-organized communication between sensor nodes. The maturation of this technology will greatly facilitate people's daily life [5]. Positioning is not limited to the commonly understood positioning: two-dimensional latitude and longitude positioning, three-dimensional positioning, but has gradually developed to a more practical direction of more accurate physical positioning, and the accuracy, fast convergence, low algorithm complexity, and other indicators of positioning algorithms are the main indicators to measure the merits of a positioning algorithm [6]. Research on the covert communication technology of sparse sensor networks directed broadcast and undirected interest diffusion, based on minimizing information replication, to meet the ability of fast data transmission, secure data exchange, subject security authentication. Further research on the opportunity network communication capability in specific scenarios achieved the information back transmission capability under sparse network conditions [7].

In wireless sensor networks, node location deployment, topology control of the network, network life maximization optimization algorithms, different routing protocols, for precise target location, etc., are research topics of great interest to researchers.

2. Related Studies

Singh et al. proposed a new hybrid optimization method that combines focal plane array (FPA) with chaotic and acoustic search algorithms to solve the Sudoku puzzle by improving the search accuracy of the algorithm [8]. Cayirpunar et al. proposed a hybrid FPA algorithm whose unique feature is a path relinking-based strategy that replaces the local and global pollination operations in FPA and its ability to accomplish the task of giving elderly people with the premise of providing healthy nutritious meals, which can provide a solution to improve its execution time and adaptive value [9]. Considering the characteristics of limited node resources, limited communication bandwidth, limited transmission capacity of network information, many redundant network transmissions, relatively low throughput rate, and poor real-time performance, once some nodes in the network die, it will cause incomplete information monitoring transmission and data errors in the relevant area, and the base station will then make wrong decisions based on the incomplete and wrong information transmitted to it by the relay nodes [10]. The security of data transmission in wireless sensor networks, the antidestructiveness of network topology, the balance of energy consumption of nodes, and the scalability are the main indicators of a good or bad network, which may lead to a link failure or other unreliable links due to the premature death of some nodes in the transmission process, which will lead to an increase in network energy consumption and make more nodes depleted prematurely [11]. This makes the network vulnerable to various attacks and reduces the robustness of the network data transmission path. To address the routing security issue, malicious attack nodes in the network are detected and excluded from communication routing as early as possible to ensure the accuracy of the data and thus improve the robustness of the network data transmission path [12].

Verma et al. apply Good Working Order (GWO) to optimize the influence maximization problem for social networks [13]. The problem is first formulated as an optimization model, and then, GWO is used to optimize the model. Experiments show that GWO has better performance than the latest influence maximization algorithms and has the advantage of less computation time than other metaheuristics [14]. Data aggregation is the operation of aggregating the aggregated data according to certain criteria of classification and aggregation, using any means, which can be considered as centralized to distributed diffusion. There are many types of centralized data, and if all of them are packaged and transmitted out, it is not easy to achieve the purpose, but also consumes more energy [15]. Classifying and compressing the data to get the desired data greatly reduces the energy consumption during network transmission and removes the redundant data. In recent years, the research on aggregation technology has achieved many results, and the development of data aggregation is becoming increasingly rapid, and with the advent of the era of big data, the importance of data aggregation will become increasingly prominent. To obtain good optimization accuracy, Kumar et al. proposed a hybrid algorithm combining GWO and Dragonfly algorithm, and the new algorithm combines the advantages of GWO in local exploitation and the global exploration capability of Dragonfly algorithm [16]. The hybrid algorithm was simulated on an IEEE 30 bus system and verified to be more effective than other algorithms in reducing cost and minimizing power consumption. Shankar et al. added a chaos mapping strategy to enrich the population diversity in the population initialization phase of GWO to enhance the global exploration capability of the algorithm and named it chaos enhanced grey wolf optimization (CEGWO) [17]. Later, CEGWO was applied to an extreme learning machine to identify patients with paraquat poisoning.

The degree of the node has a great impact on the load of the node, i.e., the number of child nodes, and the larger the degree of the node, the more energy it needs to consume, requiring more power.

(1) In this case, the nodes need to be optimized for power control. Many algorithms often take the means to dynamically adjust the power of each node to transmit the received data, the same need to establish routing tables to facilitate control analysis, for the node degree of power control, if updated too quickly, the complexity of the algorithm is high, but affect the transmission performance of the network, usually take the means to set a fault-tolerance threshold to limit according to the routing power table

(2) The power control algorithm of the control node degree has high local convergence and can achieve accurate results quickly and dynamically, which is a better means of controlling the use of power

(3) The disadvantages of this algorithm are poor globalization, high energy consumption, and the tendency to have bottleneck nodes

3. Wireless Sensor Network Coverage Modeling and Simulation

3.1. Wireless Sensor Network Coverage Modeling Optimization Algorithm Design. Different WSN applications require different levels of coverage within the monitoring area. Therefore, the coverage requirements vary by application scenario, and that key point must be considered first when developing a deployment strategy. Simply put, the degree of coverage or the size of the area being monitored within a given monitoring area is related to the number of nodes scheduled for sensing tasks [18–20]. Therefore, when deploying a network coverage solution, the nature of the application scenario and some other factors must be considered, such as whether the functionality of the sensor nodes meets the requirements of that deployment. Therefore, synthesizing the variability of deployment tasks, the WSN coverage optimization problem can be classified into stochastic coverage and deterministic coverage if classified by the deployment method. If classified by coverage area type, it can be divided into point coverage, fence coverage, and area coverage [21, 22].

The virtual force method (VFM) is a cluster-based approach in which randomly distributed sensor nodes form clusters based on their random physical locations, and one of them is selected as the cluster head to manage the other nodes [23–25]. VFM relies on sensor mobility, which uses virtual repulsive and attractive forces to force nodes to move away or to achieve coverage optimization. If two nodes are far away (greater than a preset distance threshold), they will exert an attraction on each other; if they are too close (less than a preset distance threshold), they will repel each other to increase the coverage. The sensors will keep moving, and when the repulsive and attractive forces are equal, they eventually cancel each other and reach a state of global equilibrium [26–28]. The coverage optimization problem is categorized according to the type of coverage area, which can include point coverage, fence coverage, and area coverage, while each coverage type is used in different WSN applications depending on its deployment characteristics. Figure 1 shows the classification diagram.

The purpose of point coverage is to require monitoring a set of target points in a known area or location, as shown in Figure 1. The point coverage scheme focuses on determining the exact location of the sensor nodes, with a limited number of sensing nodes to ensure effective coverage of the fixed points (target points). Usually, point coverage can be considered as a special case of area coverage problem if the numbering of sensor nodes is not considered. Fence coverage refers to sensing nodes monitoring and tracking the motion trajectory of an event or a moving target. It is widely used in some deployment scenarios where the main goal is to monitor the boundary, such as monitoring intruders who cross the boundary or penetrate the protected area. The basic requirement is to form a sensor fence such that the area covered by sensors in the fence has continuous isolation against intrusion detection, as shown in Figure 1. In fence coverage, it is usually considered that the intruder penetrates the fence sensing network with the minimum possibility of not being detected. The main goal in area coverage is to cover the entire monitoring area or to monitor all points in that space, as shown in Figure 1. The WSN area coverage requires being able to obtain real-time changes in the data in the area where it is located, and area coverage is often full coverage; the entire area is covered by the WSN. Full coverage means that each point in the area is sensed by at least one sensor node, and it helps to achieve static deployment of sensor nodes, thus maximizing network coverage. The main source of data for coverage area classification is literature [14]. Typically, a minimum number of sensor nodes are deployed in the monitoring area to achieve full coverage of the area.

The Boolean sensing model has become the simplest and most used model for WSN sense because it does not consider the uncertainties in node monitoring and the weakening of physical signals. A monitoring point m is covered or sensed m_j if it lies within the sensing range of the sensor node st. P in the figure is the sensing radius of the node. S_j sensing area is defined as a disk with s, as the center and a radius of P. The probability represented a monitoring point m.

$$P_{\text{cov}}(s_i, m_j) = \begin{cases} 0, & \text{if } d(s_i, m_j) \le r^P, \\ 1, & \text{otherwise.} \end{cases} \tag{1}$$

Then, the Euclidean distance between two nodes is

$$d(s_i, m_i) = \sqrt{(x_i - x_i)^3 + (y_i - y_i)^3 + (z_i - z_i)^3}. \tag{2}$$

Then, the Euclidean distance between two nodes can be expressed as

$$d(s_i, m_j) = \sqrt{(x_i - x_j)^3 + (y_i - y_j)^3 + (z_i - z_j)^3},$$

$$P_{\text{co}}(s_i, m_j) = \begin{cases} 0, & d(s_i, m_j) \le r_1, \\ e^{-p(d(s_i, m_j) - r_j)^{ij}}, & r_1 < d(s_i, m_j) \le r^P, \\ 1, & d(s_i, m_j) \ge r^P. \end{cases} \tag{3}$$

Sensor node clustering is an effective method of topology control that maximizes the energy utilization of the network. Several clustering protocols have been used in various WSN applications. However, most of these protocols focus only on selecting the optimal set of cluster heads to reduce or balance the energy consumption of a given network, while neglecting how to efficiently cover the network area. To this

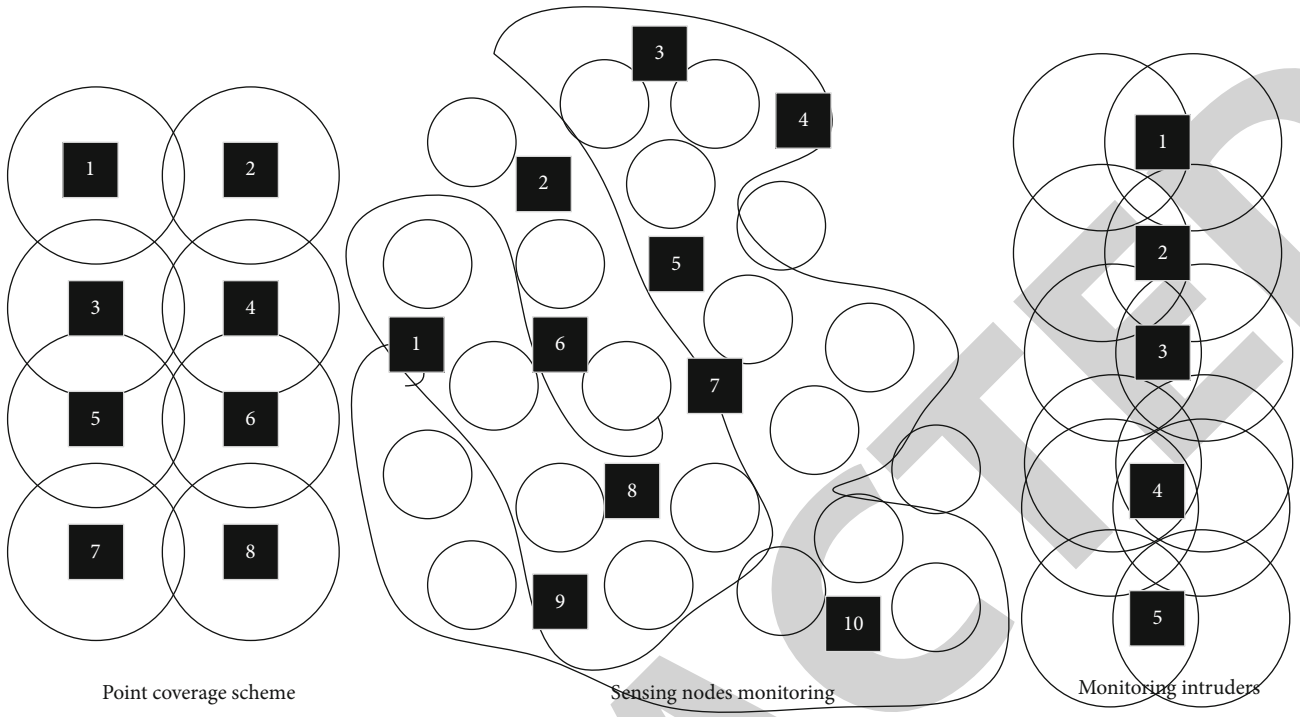


Figure 1: Classification of coverage area types.

end, the literature [74] describes a commonly used energy consumption model, namely, Equations (4) and (5).

$$E_{\mathrm{TX}}(k_1, d) = \begin{cases} \left(\mathrm{E}_{\mathrm{ele}} + \varepsilon_{\mathrm{fs}} \times d^3\right) \times k_1, & d \le d_0, \\ \left(\mathrm{E}_{\mathrm{ele}} + \varepsilon_{\mathrm{fs}} \times d^5\right) \times k_1, & d > d_0, \end{cases} \tag{4}$$

$$E_{\mathrm{TX}}(k) = \mathrm{E}^3{}_{\mathrm{ele}} \times k_1. \tag{5}$$

From the above equation, the model is an ideal state model and lacks the consideration of time on energy consumption. To make the discrete radio model more accurate for calculating WSN power consumption and to determine which links between sensor nodes are available for transmission, the energy consumption model based on the wireless transceiver data sheet can be expressed as

$$E_i = \sum_{\mathrm{state}j} P_{\mathrm{state}j} \times k_1 + \sum E_{\mathrm{state}j}. \tag{6}$$

The coverage optimization problem for WSNs is a typical NP-hard problem, and such problems are difficult to obtain optimal solutions. Therefore, many researchers have proposed various coverage optimization methods by combining the characteristics of sensor deployment, including virtual force-based optimization methods, grid methods, Delaunay triangulation, Voronoi diagram methods, and swarm intelligence optimization algorithms. The grid method is usually used for predefined deployment methods in which the sensor nodes need to be precisely placed on the specified grid points. The network deployed by this method can improve the network coverage and connectivity to a certain extent. The Kano model is based on the analysis of the impact of user needs on user satisfaction and reflects the nonlinear relationship between user demand satisfaction and user satisfaction. The ordinate represents user satisfaction. The higher the upward, the more satisfied, the lower the more dissatisfied; the abscissa indicates the degree of existence of a certain demand. The more to the right, the higher the degree of existence, and the more to the left, the lower the degree of existence. There are three commonly used grid types: triangular grid, square grid, and hexagonal grid. Among them, the triangular grid is the most outstanding because it has the smallest overlapping area. Therefore, this grid requires the least number of sensors, while the hexagonal grid is the worst. In addition to the grid type, the size of the grid also has a significant impact on the network coverage, so the size of the grid needs to be chosen based on the density of the WSN. For high-density networks, a small-sized grid helps to reduce coverage voids, thus improving the stability of the network. However, in sparse networks, a large grid size is more suitable, because it can minimize the redundant coverage of the network, thus ensuring that the node's sensory capabilities are fully utilized.

The control of energy consumption is particularly important to the WSN life cycle. Under the same conditions, the smaller the energy consumption of WSN, the longer the life cycle. Since the energy consumption model described nodes and does not consider the real-time changes of the whole network deployment, this chapter proposes a new energy consumption model in conjunction with reality. The energy consumption in this chapter considers three components: first, the energy consumption to complete the sensing (radiation) task; second, the energy consumption when sending data and receiving data during wireless communication;

and third, the energy consumed by the nodes to move the distance after optimized deployment.

The first part of the energy model consumed by the sensing task uses a model in which the energy consumption of the node E_a is proportional:

$$E_\alpha = \eta \cdot \sum_{i=1}^{m} \left(r_i^p\right)^2. \tag{7}$$

When sensor nodes are deployed in a forest environment, the greater the coverage of the network, the better the deployment. However, since the node batteries are not rechargeable, energy consumption is crucial to the network life cycle, and the smaller the energy loss, the longer the network life cycle. Also, when the sensing range of sensor nodes is mostly within the monitoring area, the coverage area of nodes can be maximized. Therefore, the smaller the radiation range of nodes is outside the monitoring area and within the obstacles, the more effectively the network voids can be reduced and the second deployment of the network can be facilitated.

$$\begin{cases} f_1(I) = \min\left(C_r(I)\right) \\ \text{s.t.} \quad A(s_i) \in H_2. \end{cases} \tag{8}$$

Using the angle of arrival to measure the location of the target source is now a frequently used high-precision method, using signal receiving equipment base station to measure the angle of arrival of the signal, and finally using the mathematical triangulation method can calculate the precise three-dimensional location of the target. The cost of this measurement means is relatively high and requires very special hardware equipment with special features; this paper focuses on wireless sensor nodes, nodes with small functions many times do not have such ranging and positioning capabilities, but other high-end equipment can.

3.2. Modeling and Simulation Experimental Design. Within an urban area, the mobile wireless sensor creator will initially arrange the sensors to meet certain coverage criteria. The initial location of these sensors can be limited to a certain range, which is assumed to be a grid of a specific size. This means that the location of the sensors can be roughly determined within the bounded range. These mobile sensors can be moved to accomplish certain data acquisition tasks. The location information of the sensors will not be known (or probabilistically known) after the movement. Also, the network creator needs to determine the location information of all sensors again to perform sensor movement control. Therefore, high accuracy positioning plays a very important role in mobile wireless sensor networks. As shown in Figure 2, a generic framework can be used for swarm aware-assisted localization. This framework is multiround repeatable for localization and network construction. However, sensors without GPS modules are widely used devices today.

First, having sensors with GPS modules in the network will increase the overall cost of the network. Second, GPS modules will continually consume sensor energy, but network builders want to minimize the energy consumption of sensors that can only be powered by batteries. Third, GPS modules are also fragile and difficult to maintain. Cluster aware-assisted positioning becomes a viable alternative. When a participant positioned with swarm intelligence perception approaches a sensor, a communication connection is quickly established between them. The participant mobilizes the computing power of the smart device to quickly compute the GPS location of the sensor. The location computation is based on the participant's smart device's own GPS location information and the relative position between the device and the sensor, which can be aided by some local positioning techniques. At least three anchor points are required for range-based positioning techniques. This is also feasible because participants can perform three different location-based localization calculations within the sensor communication range. This process is given in Figure 2. Two special examples for localization are if the sensor is located on a common edge of two grids or at a vertex of the grid. This means that this sensor is associated with two grids as in Figure 2 or with four grids as in Figure 2. To compute the locations of all sensors, group-wise perception-assisted localization should recruit enough competent participants. This goal of enough is achieved thanks to the sufficiently dense population in the city and the pervasiveness of smart devices.

Competent participants are those smart device carriers who can pass through sensor(s). However, how to recruit these participants remains the main challenge for swarm intelligence-aided localization. Nowadays, people widely use navigation systems. People use the navigation to plan a path before going to a destination. Naturally, these paths, which are chosen by many people, can be imported into a group intelligence-aided location control center to help the system select candidate participants. The selection is mainly based on whether the planned path of the candidate participant can assist GPS positioning. In our model, the target area is divided into a square grid, and the sides of the grid are bounded by the communication range of the sensor nodes. If the communication range of a sensor node is represented by a circular region of radius Rc, the side length of the grid must be no larger than Rc. The leader in this field is Amazon's Lambda, which allows rapid deployment of code written in python, JavaScript, and Java. The Lambda function can be a script or a complex application that depends on other services and I/O. They can be called manually or triggered by events generated by other Amazon services, such as S3. When used with Pathway, it can be used to deploy the entire microservice implementation in a zero-infrastructure environment. Other mainstream cloud platforms have also made strides into this field, such as Microsoft's Azure Functions and Google's Cloud Functions. This requirement ensures that a sensor in this grid can communicate with a participant at any location in this grid. Since the sensor's approximate localization information is known, the sensor can be identified as being positioned within a certain grid, i.e., associated with this grid. Also, the planned path trajectory provided by the navigation system for the candidate participant can be associated with a series of grids. We assume that the candidate participant can assist in the

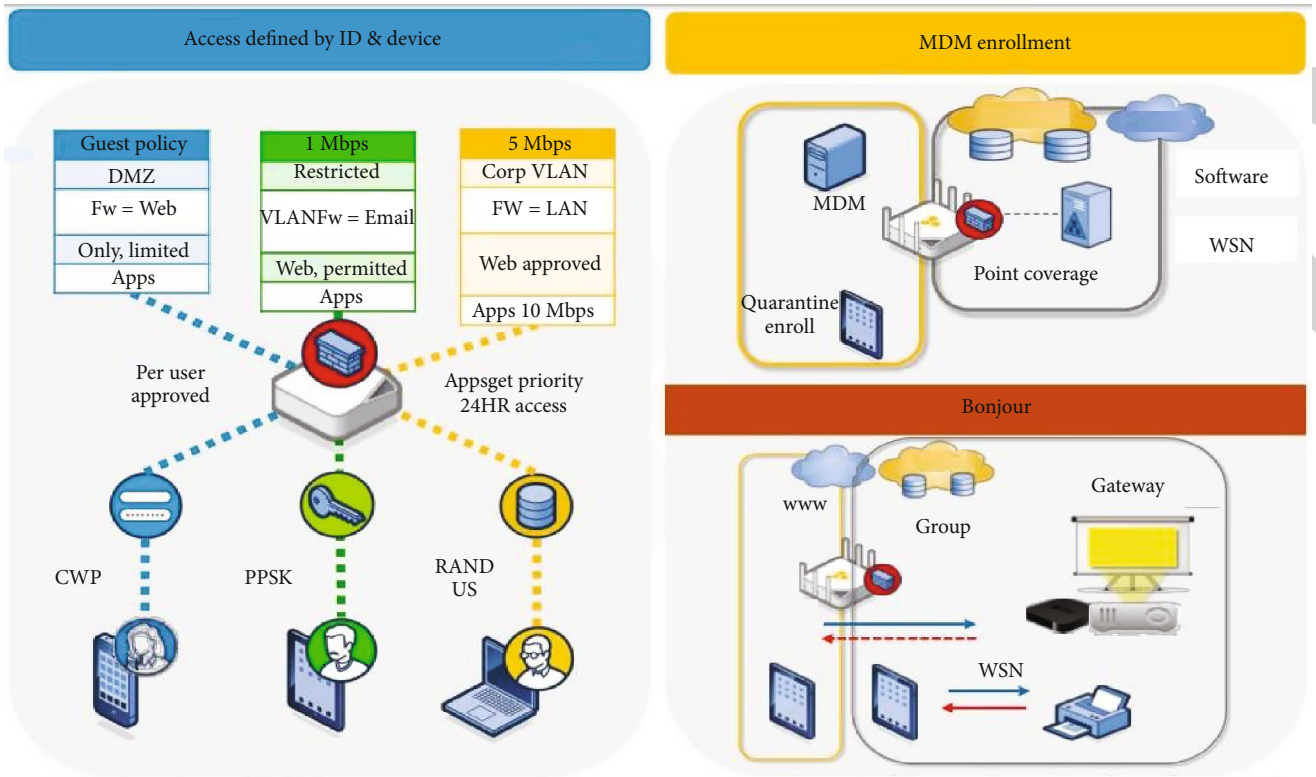


Figure 2: Deployed WSN.

localization of the sensors in the grid associated with her trajectory.

To test the efficiency of the algorithm under different data scenarios, we applied the algorithm to five real datasets from CRAWDAD. These data collected daily GPS trajectory logs from five different locations. This trajectory log records the GPS location information of the participants every 10 seconds. We divided the map of these places into a square grid with a side length of 5. The basic statistics of these data are listed in Table 1.

The main purpose of the network is to transmit the collected data and finally to the base station node, the simple network requirement is that all nodes in the network do not perform any processing of the data directly to their next-hop node that is the parent node, which ensures the correctness of the data, but the fault tolerance of the data is not high and the relay parent node because of the processing of a lot of data, energy will be quickly depleted, the network will affect the connectivity of the network by prematurely appearing dead nodes. To avoid these shortcomings, data aggregation is often adopted in this paper, and the data aggregation technique and its advantages and shortcomings have been described specifically. Since the data collected by the nodes in close spatial proximity have strong similarity, the nodes' functions are firstly preset when building the network in common networks. The compressed data may contain various types, and the next-hop node receives the packet and decompresses it and then takes the same way of compression to continue transmission until the Sink node, but the requirements for data in many networks are very high. Due to the rapid development of big data, the information content of the data is increasingly required, and the data collected by the nodes may contain multiple types, so if we take the simple approach of compressing the data into a packet will inevitably lead to the loss of data, which is not worth the loss and the data will lose the meaning of network monitoring if its true value cannot be obtained, so the appropriate data should be selected for different topologies' aggregation ratio.

The coverage vulnerability problem is an important research problem in wireless sensor networks. This chapter introduces the specifics of the mobile node-based vulnerability repair algorithm for wireless sensor networks, starting with the network model of the algorithm, the problem description, and analysis of the algorithm to introduce the prerequisites for the implementation of the algorithm. After that, the relevant terms involved in the algorithm and the basic geometric theoretical knowledge required are introduced to pave the way for the introduction of the content of the algorithm. Finally, the specific stages and steps related to the implementation of the algorithm are detailed, the criteria for the algorithm to select the repair nodes, and how to transform these criteria into operational variables utilizing mathematical formulas, in preparation for the simulation experiments later.

4. Analysis of Results

4.1. Path Success Rate. In this paper, we consider how to circumvent the influence of malicious nodes when nodes select the next-hop node to establish a path, that is, not to select a

Table 1: Five GPS track data information.

Dataset	Number of participants	Area width	Area height	Grid scale
KAIST	98	10000	9000	2000∗2200
NCSU	36	11000	10000	2200∗1700
New York	68	20000	18000	2500∗2300
Orlando	39	15000	14000	1800∗2500
State Fair	38	12410	11000	2600∗2100

malicious node as a relay node. In this paper, we use a Bayesian detection algorithm to first detect the attacking nodes in the network. Figure 3 shows the Bayesian detection simulation; there are 100 nodes in the simulation network; in the figure, it is given that as the number of malicious nodes in the network increases in the Bayesian detection rate graph, it can be seen that with the increase of malicious nodes detection rate decreases rapidly; this is because the proportion of normal nodes and malicious nodes is considered; when the malicious nodes reach 50% of the network, detection rate is 0, so we can see the malicious nodes can cause a lot of damage to the network.

Simulation in Figure 4 shows the comparison of the path success rate between the proposed algorithm ROACO and the classical algorithm ACO, and ROACO is significantly better than ACO in terms of path success rate. This is because the proposed algorithm does not select the malicious nodes detected by the Bayesian algorithm as the connected nodes in the node path selection, which can effectively avoid the malicious nodes from joining the link and destroying the link, thus having stronger connectivity and better protection of data transmission to the Sink node, making the data transmission path of the network more robust.

This is because if the network is updated too frequently (less than the optimal value of 90), the nodes in the network will consume a large portion of energy for data computation, resulting in the shortening of the overall network lifetime; if the network topology is updated slowly, some nodes in the network that bear too much data forwarding will consume energy rapidly, resulting in the premature appearance of dead nodes and thus affecting the network lifetime. In the simulation, the optimal number of updates is firstly selected and fixed, and the topology evolution is performed in the optimal environment to maximize the utilization rate of the network. It can be seen from the figure that the lifetime of the network decreases more than twice when the number of updates is more than 110 or less than 70, which also shows that the energy consumption of data computation and the energy consumption of data transmission have a great impact on the lifetime of the network.

The number of surviving nodes reflects the variation of the maximum connected subgraph in the network. It can be seen from Figure 5 that the algorithm DA-LTRA effectively improves the network lifetime and delays the death of the first node in the network by 3000 rounds, because DA-LTRA is better than the DADAT algorithm in the case of the same heterogeneous network in terms of initial node energy setting, selection of relay nodes. This is because DA-LTRA is better than DADAT in terms of node initial energy setting, relay node selection, and local tree reconstruction technique in the later tree maintenance stage to balance the load of the network.

The tree topology has a strong resistance to destruction, and the maximum connectivity subgraph is generally the basis for studying the topology properties. The tree topology optimization algorithm constructed in this section is based on the transmission of heterogeneous data, the transmission of heterogeneous data has higher requirements on the network topology, and it is difficult for the general homogeneous network to meet the quality of service requirements for heterogeneous data transmission, so the strategy adopted in this section is to use heterogeneous nodes to complete the transmission of these different types of data, such a transmission method can meet both the real-time data transmission and the data accuracy requirements. In the process of topology optimization, the thesis takes the approach of adjusting the structure of the network through three major steps, firstly, relay node selection, secondly, heterogeneous energy setting, and finally, dynamic local tree structure adjustment; after the above three steps optimization means the network has a strong topology. Preprocessing the training dataset, in the data preparation stage, the data in each category is segmented, punctuation and adverbs are removed, and the files in each category are read into a large file, so that there is only one file at the end of each category. Containing all the files at the beginning, the files processed under mahout must be in sequence file format, and textile needs to be converted to sequence file.

4.2. Simulation Results. For each simulated experiment, we ran algorithms 4 and 5 in the least participant group-wise perception-assisted localization and time-efficient group-wise perception-assisted localization 100 times, respectively, to calculate the average experimental effect. We set the time at which a participant enters the grid to the time at which the sensor in that grid is assisted to be localized. Also, we need to do more preprocessing, such as removing sensors that are not in contact with any participant trajectory and completing all missing sensor-participant pairs for the assisted localization time. All algorithms, preprocessing, and simulation experiments were implemented in Python.

We present in Figure 6 a comparison of least participant group-wise perception-aided localization and time-efficient effect group-wise perception-aided localization performed on different datasets. For each experiment, 100 sensor target nodes were randomly selected. The bar chart corresponds to the left axis, which indicates the number of selected participants. The line plot corresponds to the right axis which indicates the localization completion time. The least participant algorithm has better results than the time-efficient algorithm in the bar chart. It is particularly noteworthy that the effect is more pronounced in the KAIST dataset. This is because there

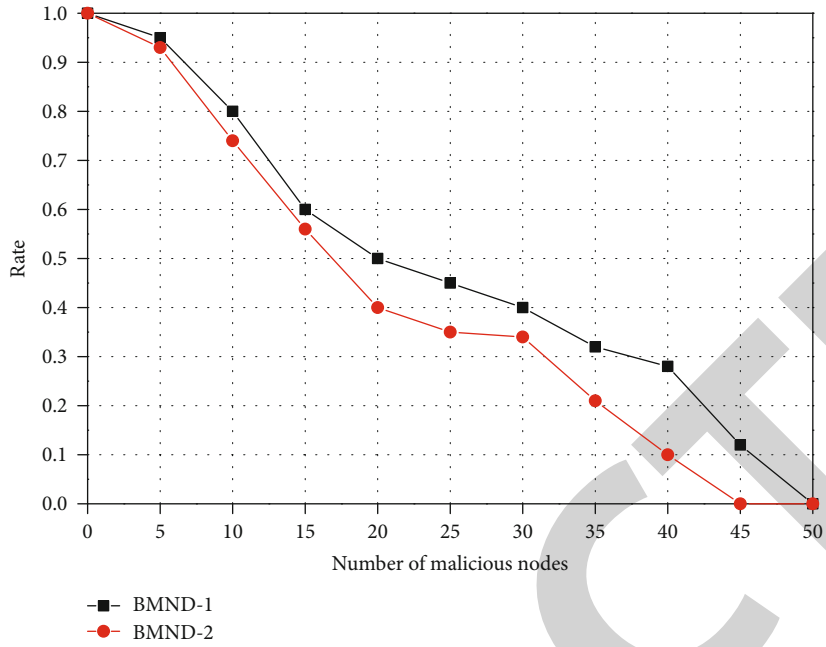


Figure 3: Bayesian detection.

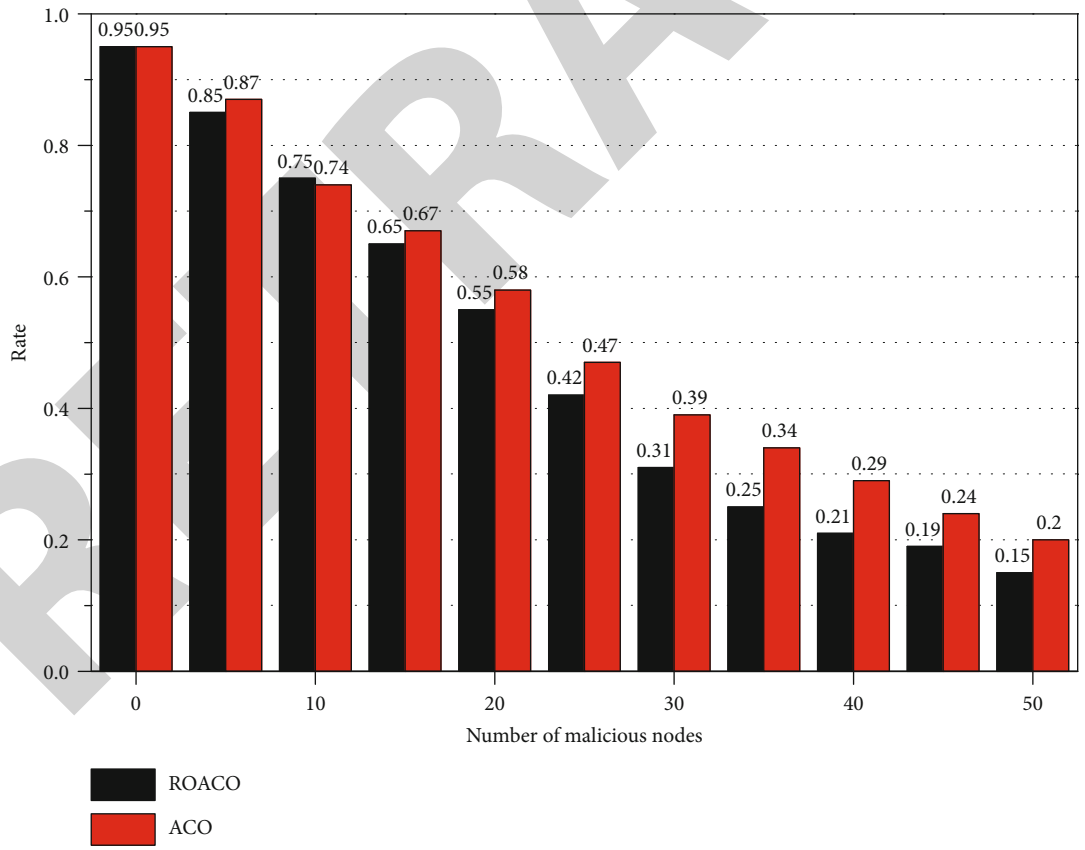


Figure 4: Comparison of detection rates of different algorithms.

are enough candidate participants in the KAIST dataset, which allows the time-efficient algorithm to have sufficiently many choices to shorten the participant localization completion time, but it also allows the number of recruited participants to grow rapidly. In the line graph, the gap in effectiveness between the two algorithms then drops

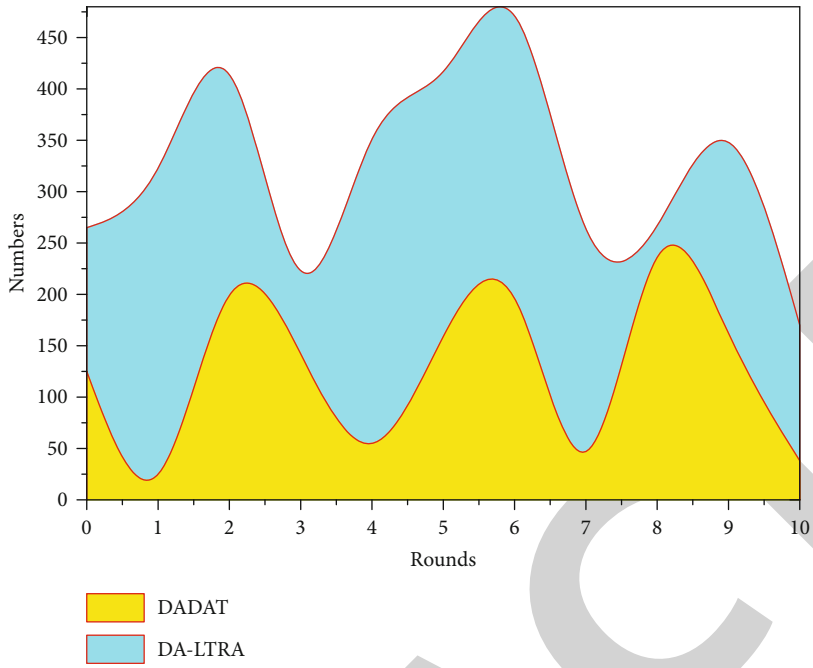


Figure 5: Comparison of the number of surviving nodes of different algorithms.

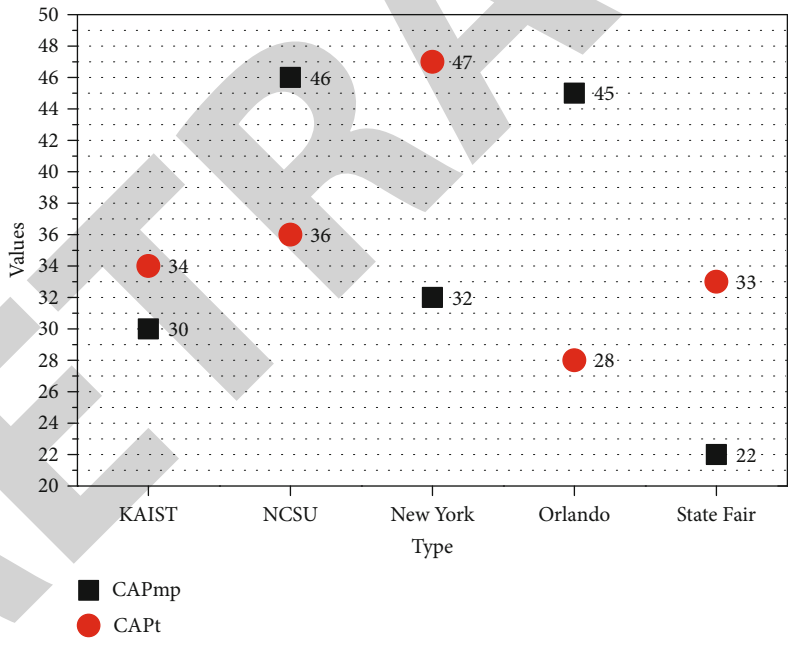


Figure 6: Comparison of experimental performance of CAPmp and CAPt.

dramatically because the data sampling time is limited to a small interval. For the last three datasets, New York, Orlando, and State Fair, the time-efficient algorithm does not improve the localization completion time much because there is only a small amount of trajectory data.

In this section, we consider a framework for group-wise sensing-assisted GPS localization for wireless sensor networks and propose two recruitment participant optimization objectives, minimum participants and most time-efficient, respectively. We portray both problems as integer linear programming problems and propose optimization objectives for submodular cost set functions. An algorithm based on greedy ideas is proposed to solve both problems, and the correctness of this algorithm is demonstrated and compared in experiments. The main contribution of our work in this section is that we propose a multiround implementation framework to show how swarm-wise perception can assist in GPS accurate positioning of sensor networks. The technical route and application problem scenario of this work is novel. We define the Crowdsensing Aided Positioning with minimum

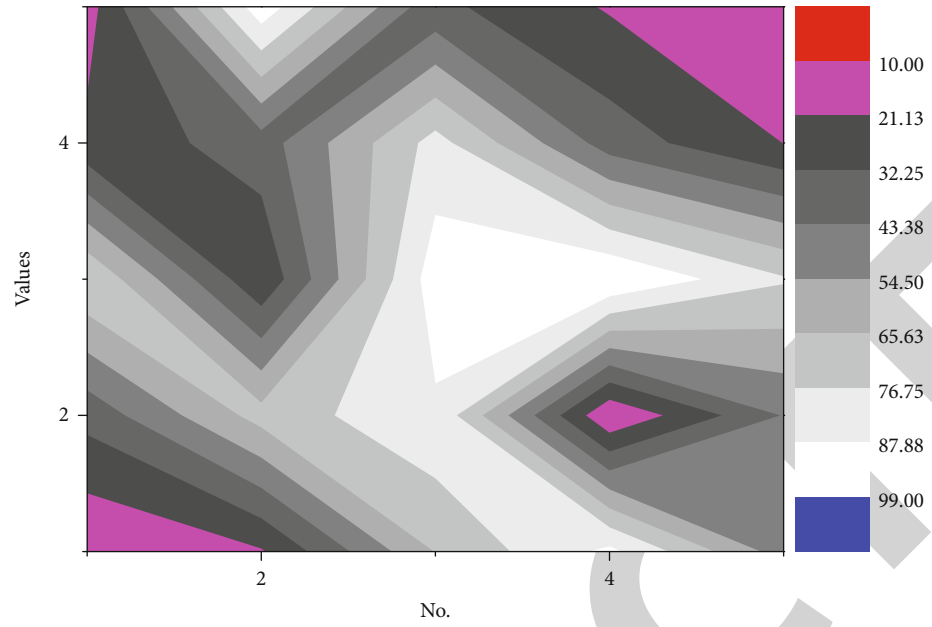


Figure 7: Simulation experimental results with different variance value settings.

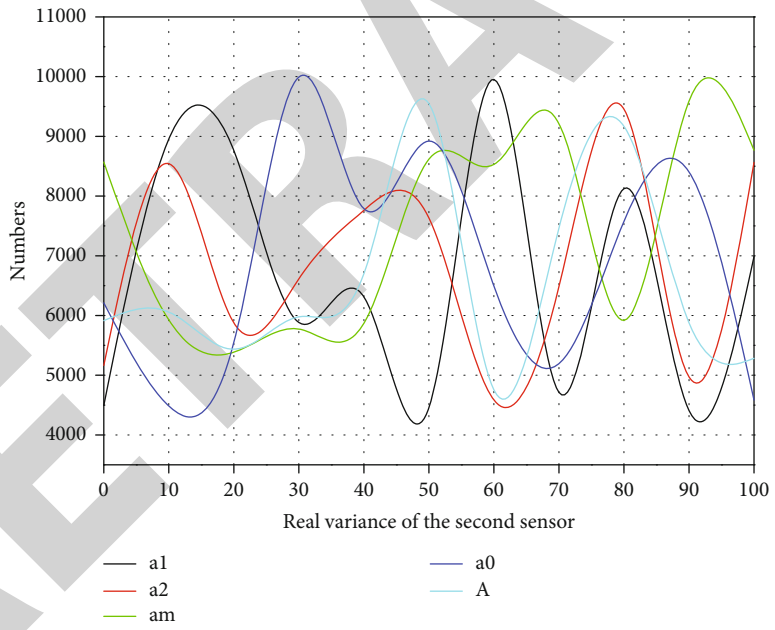


Figure 8: Experimental results of the variation of accuracy results with the variance of another sensor.

participants' problem as a problem called Crowdsensing Aided Positioning with minimum participants (CAPmp) in the context of integer linear programming. And we reduce this problem to the classical ensemble coverage problem. We further propose the Crowdsensing Aided Positioning Timely (CAPt) problem. To solve the above two problems, we propose an approximate greedy algorithm based on submodular functions and analyze it theoretically. We also fully conduct experiments on the above algorithm on real data. The effectiveness of the proposed algorithm is verified under different parameter settings.

For our simulation experiments, we set $n = 100$ and $\mu = 0$. We repeated the experiment 10000 times and averaged the results to evaluate the effectiveness of our method. The results are summarized in Figure 7.

Among them is the number of experiments in which it is more accurate than the sensor in all experiments. From the experimental results, we can say that the results of our method are better, i.e., more accurate, than at least one of these sensors. Also, we can say that the accuracy is related to the measurement variance of the sensor or more specifically to it. Afterward, we performed more experiments with

the setup and changed the value so that it varied in an integer from 1 to 100, and we redid the above experiments in this way. The results of the experiment are shown in Figure 8. We can see that as the increment, the accuracy results of our method are further better than the results of the maximum likelihood estimation, while the results are progressively more accurate than the estimates of both sensors and better than the results of the sensor with small variance, but decreasing and converging to 50%. At the same time, the cases that are more inaccurate than both sensors converge to zero.

We present a study of the fusion of wireless sensor networks and group-wise sensing networks. Firstly, a framework for group-wise sensing networks to assist GPS localization of wireless sensor networks is considered, which provides the foundation for the location information required for the mobility problem of mobile wireless sensor networks in Chapter 3. The experimental performance tests of CAPmp and CAPt are tested in accordance with common standards. Secondly, as inspired by the opportunity network framework based on location information, we further propose to combine group-wise sensing networks and wireless sensor networks to implement a data fusion framework. The basic idea is that through data fusion calibration, group-wise sensing participants obtained the data trustworthiness attributes of group-wise sensing network participants: data perception accuracy and estimation reliability in the framework of opportunity networks through sensor networks. After proposing a simple single-value accuracy model, a confidence interval accuracy model is proposed in this chapter. Confidence interval-based models are generic models and are introduced mainly to introduce the completeness of probabilistic statistical models, which can be transformed into single-value accuracy models by methods such as sampling. Using the single-value accuracy model, three calibration methods are proposed and validated, namely, data calibration by sensors in collaboration with participants to be tested, data calibration by qualified participants in collaboration with candidate participants, and multiple iterations of data calibration.

5. Conclusion

Because the data of the network often face some unstable external factors that influence the process of transmission, these influences many times will reduce the efficiency of data transmission. A safe and reliable data transmission model is established for the complex network environment, which combines the actual requirements to ensure safe and reliable data transmission. The model considers energy, node distance, data redundancy, and link security in the node path selection method and presents an intelligent, secure, efficient, and robust data transmission path robust optimization algorithm based on ant colony algorithm. The algorithm is a fast converging global optimization algorithm with stronger robustness. The heterogeneous network topology optimization algorithm proposed in this paper can maximize the heterogeneous data transmission requirements. In Chapter 4, the nodes in the wireless sensor network are heterogeneous, and these nodes have different functions and different initial energies. The implementation of the algorithm is specifically divided into three steps: selecting relay nodes in different layers according to the proportion, setting different initial energies for the nodes in different layers, and local tree reconstruction, and finally, the heterogeneous network topology is established through the adjustment of the three steps. Under the same simulation conditions and simulation parameters, the simulation results of the algorithm proposed in this paper have higher performance and can effectively reduce the energy consumption of nodes in the network.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] L. Wang, W. Wu, J. Qi, and Z. Jia, "Wireless sensor network coverage optimization based on whale group algorithm," *Computer Science and Information Systems*, vol. 15, no. 3, pp. 569–583, 2018.

[2] G. P. Gupta and S. Jha, "Biogeography-based optimization scheme for solving the coverage and connected node placement problem for wireless sensor networks," *Wireless Networks*, vol. 25, no. 6, pp. 3167–3177, 2019.

[3] C. K. Ng, C. H. Wu, W. H. Ip, and K. L. Yung, "A smart bat algorithm for wireless sensor network deployment in 3-D environment," *IEEE Communications Letters*, vol. 22, no. 10, pp. 2120–2123, 2018.

[4] M. Hammoudeh, F. al-Fayez, H. Lloyd et al., "A wireless sensor network border monitoring system: deployment issues and routing protocols," *IEEE Sensors Journal*, vol. 17, no. 8, pp. 2572–2582, 2017.

[5] R. E. Mohamed, A. I. Saleh, M. Abdelrazzak, and A. S. Samra, "Survey on wireless sensor network applications and energy efficient routing protocols," *Wireless Personal Communications*, vol. 101, no. 2, pp. 1019–1055, 2018.

[6] M. E. Keskin, "A column generation heuristic for optimal wireless sensor network design with mobile sinks," *European Journal of Operational Research*, vol. 260, no. 1, pp. 291–304, 2017.

[7] R. Priyadarshi, B. Gupta, and A. Anurag, "Wireless sensor networks deployment: a result oriented analysis," *Wireless Personal Communications*, vol. 113, no. 2, pp. 843–866, 2020.

[8] S. P. Singh and S. C. Sharma, "Implementation of a PSO based improved localization algorithm for wireless sensor networks," *IETE Journal of Research*, vol. 65, no. 4, pp. 502–514, 2018.

[9] O. Cayirpunar, B. Tavli, E. Kadioglu-Urtis, and S. Uludag, "Optimal mobility patterns of multiple base stations for wireless sensor network lifetime maximization," *IEEE Sensors Journal*, vol. 17, no. 21, pp. 7177–7188, 2017.

[10] X. Song, Y. Gong, D. Jin, Q. Li, and H. Jing, "Coverage hole recovery algorithm based on molecule model in heterogeneous

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6654349, 10 pages
https://doi.org/10.1155/2021/6654349

Research Article

Fusing Node Embeddings and Incomplete Attributes by Complement-Based Concatenation

Zheng Wang,[1,2] Yuexin Wu,[3] Yang Bao,[3] Jing Yu,[3] and Xiaohui Wang[4]

[1]*Department of Computer Science and Technology, University of Science and Technology Beijing, Beijing, China*
[2]*Department of Computer and Information Science, University of Macau, Macao, China*
[3]*School of Software, Tsinghua University, Beijing, China*
[4]*School of Mechanical Engineering, University of Science and Technology Beijing, Beijing, China*

Correspondence should be addressed to Xiaohui Wang; wangxh14@ustb.edu.cn

Received 27 December 2020; Revised 22 January 2021; Accepted 8 February 2021; Published 25 February 2021

Academic Editor: Wei Wang

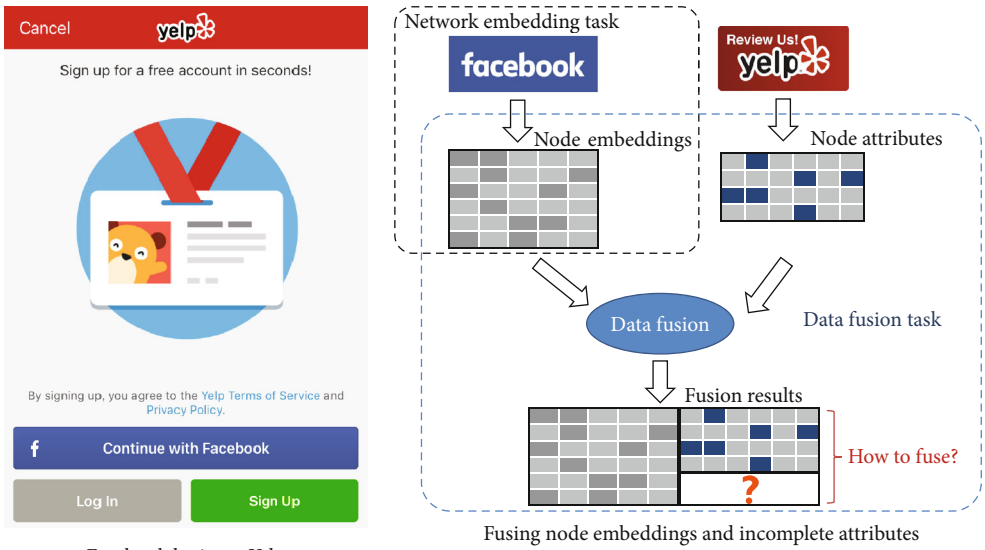
Network embedding that learns representations of network nodes plays a critical role in network analysis, since it enables many downstream learning tasks. Although various network embedding methods have been proposed, they are mainly designed for a single network scenario. This paper considers a "multiple network" scenario by studying the problem of fusing the node embeddings and incomplete attributes from two different networks. To address this problem, we propose to complement the incomplete attributes, so as to conduct data fusion via concatenation. Specifically, we first propose a simple inductive method, in which attributes are defined as a parametric function of the given node embedding vectors. We then propose its transductive variant by adaptively learning an adjacency graph to approximate the original network structure. Additionally, we also provide a light version of this transductive variant. Experimental results on four datasets demonstrate the superiority of our methods.

1. Introduction

Social network sites (SNSs, also commonly referred as social networking services) are online platforms which provide users with various features to facilitate digital social interaction and information sharing [1, 2]. Over three billion users are currently active on various SNSs (like Facebook, Twitter, and QQ), spending on average two hours daily. These wide and active SNSs naturally form an important part of the digital economy, making social network analysis [3, 4] become a hot research topic over the years.

Recently, network embedding [5], as a fundamental problem in network analysis, has aroused considerable research interest. Network embedding learns low-dimensional vector representations for network nodes. The learned vectorized representations, which preserve certain structural and content information of networks, can be easily combined with off-the-shelf learning algorithms for many social network analysis tasks such as node classification [6], link prediction [7], and diffusion prediction [8].

1.1. Problem. Although various network embedding methods have been proposed, they mainly focus on a single network scenario. In the era of big data, the related information from different networks should be fused together to facilitate applications. In this paper, we consider a "multiple network"



Facebook login on Yelp

(a)

(b)

Figure 1: Illustration of the studied problem.

scenario by studying the problem of fusing the node embeddings and incompleted attributes provided by two different networks.

As illustrated in Figure 1, this problem has practical importance. Imagine that you use Yelp (see Figure 1(a)), a popular review app, and try to get in your account there. Yelp allows you to sign in using your Facebook account. In addition, as the node (user) embeddings not only preserve certain characteristics of networks but also protect users' privacy [9], Facebook may provide these embeddings to Yelp to facilitate its applications, e.g., cold-start recommendation. More importantly, as some Yelp users begin to write reviews, a very practical problem would arise: is it possible to fuse the original node embeddings provided by Facebook and the reviews provided by Yelp to get new user embeddings (illustrated in Figure 1(b))?

1.2. Challenge and Solution. Certainly, one fundamental challenge is the incompleteness of attributes, i.e., only a small part of nodes are further provided with attributes. This challenge is very common. As reported [10], the distribution of user activity tends to be long-tailed, suggesting most social media contents (like the reviews on Yelp) are actually written by a few active users. To address this, we propose to complement the incomplete attributes by defining attributes as a parametric function of the given node embedding vectors. This complement enables us to conduct data fusion via concatenation (illustrated in the bottom right corner of Figure 1(b)).

To obtain high-quality fusion results, we further propose a transductive method by adaptively learning an adjacency graph to approximate the original network structure. In particular, the adjacency graph is learned by jointly considering the given node embeddings and attribute knowledge. Additionally, we also provide a light version of the proposed transductive method. Specifically, for each node, this light version reduces its neighbor candidate set for efficient adjacency graph learning. We then conduct extensive experiments to verify the effectiveness of our methods.

In summary, our main contributions are as follows:

(i) We study the problem of fusing node embeddings and incomplete attributes from two different networks. To our best knowledge, little work has addressed this problem

(ii) We propose a very simple and effective inductive method based on the idea of attribute complement

(iii) We further propose a transductive method POINTS and its light version POINTS*, both of which could obtain superior performance

(iv) The remainder of the paper is organized as follows. We review related work in Sect. 2 and formalize the problem in Sect. 3. We present our method in Sect. 4 and then provide some discussion in Sect. 6. We conduct experiments in Sect. 7. Finally, we end with a conclusion in Sect. 8

2. Related Work

2.1. Network Embedding. Over the past few years, there has been a lot of interest in learning useful node embeddings (i.e., features) from large-scale networks automatically [5]. A representative work is DeepWalk [6] which performs random walk on a network to generate node sequences and then perform the skip-gram algorithm [11] on those sequences to achieve the embedding. Another well-known work is LINE which preserves both first-order proximity (i.e., the similarity between linked nodes) and second-order proximity (i.e., the similarity between the nodes with shared neighbors) of a network. In addition, researchers have also proposed some deep learning-based embedding models, such as SDNE [12] and

GraphGAN [13]. Recently, lots of studies consider the network embedding with side information, such as node attributes. For example, by proving DeepWalk is equivalent to matrix factorization, the work in [14] presents text-associated DeepWalk (TADW). GraphSAGE [15] employs graph convolutional networks [16] to aggregate features among node neighborhood for network embedding. RSDNE [17] and RECT [18] further consider the problem of zero-shot graph embedding, i.e., the completely imbalanced label setting.

2.2. Data Fusion. Data fusion is the study of efficient methods for automatically transforming information from different sources and different points in time into a representation that provides effective support for human or intelligent systems. Data fusion has proved useful in many disciplines, as discussed in [19, 20]. For example, in bioinformatics, jointly analyzing multiple datasets describing different organisms improves the understanding of biological processes [21]. In information retrieval, fusing the retrieval results from multiple search engines would significantly improve the retrieval performance [22]. In biometric recognition systems, feature fusion could greatly improve the recognition performance [23]. We refer to [24, 25] for a comprehensive survey.

However, little previous work considers the fusion of incomplete data or network embedding data. Our work fills this gap.

3. Problem Statement

The studied problem is defined as follows. We are given the node embeddings of a network $U \in \mathbb{R}^{n\times d}$, where n is the node number and the i-th row of U (denoted as u_i) is a d-dimensional embedding vector of node i. On the other hand, another network further provides the attributes of l ($l < n$) nodes: $\mathscr{L}_A = \{(u_1, a_1), \cdots, (u_l, a_l)\}$, where $a_i \in \mathbb{R}^{1\times m}$ is the attribute vector, and m is the attribute feature number. Our goal is to fuse the given node embeddings and those incomplete attributes, so as to get the updated embeddings for all nodes. Note that different from existing network embedding methods, the original network structure is unknown in our problem.

4. The Proposed Method

4.1. Fusion via Attribute Complement. Since only a small part of nodes are further provided with attributes, we cannot directly fuse node embeddings and attributes. To address this problem, we adopt a very simple complement strategy: predicting the nonexist attributes. In particular, for each node i which is further provided with attributes, we assume that its node embedding u_i should have the ability to generate its attribute vector a_i. The optimal generation function f can be obtained by solving the following minimization problem:

$$\min_f \sum_{i\in\mathscr{L}_A} \ell(f(u_i), a_i), \tag{1}$$

where ℓ is a loss function that measures the reconstruction error, such as squared loss or hinge loss.

By solving the problem in Eq. (1), we can obtain the generation function f. Then, for a node i with no attributes, we can predict its attributes by applying $f(u_i)$. This complement enables us to conduct data fusion via concatenation. More details and discussion about the concatenation strategy can be found in Sect. 6.2.

4.2. Transductive Attribute Prediction. The method formulated in Eq. (1) is inductive. In this section, we present a transductive method. Generally, transductive methods, which leverage the test data for model training, perform better than inductive methods [26]. For network embedding, classical transductive methods exploit all network nodes by preserving the inherent network structure in the embedding space, i.e., connected nodes tend to have similar embeddings [27, 28]. Although the original network structure is unknown, one can simply build a sparse adjacency graph (We use the term "graph" to describe the recovered network structure, as to avoid ambiguity with the original network structure.) S to approximate it, i.e., $S_{ij} = 1$ when node j is the k-nearest neighbors of node i in the given node embedding space, otherwise $S_{ij} = 0$. This approximation can capture the intuition of transductive learning by the following cost term:

$$\begin{aligned} &\min_Y \sum_{i,j} S_{ij} Dist\left(y_i, y_j\right), \\ &\text{s.t.} \forall i \in \mathscr{L}_A, y_i = a_i, \end{aligned} \tag{2}$$

where $Dist(\cdot,\cdot)$ is a distance function, and $y_i \in \mathbb{R}^{1\times m}$ (the i-th row of matrix $Y \in \mathbb{R}^{n\times m}$) is the predicted attribute vector of node i. The imposed constraint ensures the predicted attributes to be consistent with the known attributes.

The adjacency graph plays a crucial role in this kind of graph-based transductive learning methods [29, 30]. However, the matrix S in Eq. (2) might not be the optimal adjacency graph. On the one hand, the original network information is only approximately described by the given node embeddings (i.e., $u_{i=1,\cdots,n}$) which S is built from. On the other hand, the construction of S ignores the attribute information, i.e., similar (dissimilar) attributes indicate similarity (dissimilarity) between different nodes. In this paper, we solve this problem in an adaptive way. Specifically, we propose to learn S by jointly considering the given node embeddings and attribute knowledge. This yields the following cost term:

$$\begin{aligned} \min_{Y,S} \ &\frac{\alpha}{2}\sum_{i,j} S_{ij} Dist\left(y_i, y_j\right) + \frac{\beta}{2}\sum_{i,j} S_{ij} Dist\left(u_i, u_j\right), \\ \text{s.t.} \quad &\forall i \in \mathscr{L}_A, y_i = a_i \\ &\forall i, \ \left\|s_i'\right\|_0 = s_i'\mathbf{1} = k, \ S_{ii} = 0 \end{aligned}, \tag{3}$$

where $s_i \in \mathbb{R}^{n\times 1}$ is a vector with the j-th element as S_{ij} (i.e., s_i' is the row vector of matrix S), 1 denotes a column vector with

all entries equal to one, and α and β are two adjustable parameters. Intuitively, the first and second term of Eq. (3) measure how well the adjacency graph fits the attributes and the given node embeddings, respectively.

The unified model: POINTS with learning the attribute generation function (Eq. (1)) and adjacency graph (Eq. (3)), the proposed method is to solve the following optimization problem:

$$\begin{aligned}\min_{f,Y,S} \mathcal{J} = & \sum_{i\in\mathcal{L}_A} \ell(f(u_i), y_i) + \frac{\alpha}{2}\sum_{i,j} S_{ij} Dist\left(y_i, y_j\right) \\ & + \frac{\beta}{2}\sum_{i,j} S_{ij} Dist\left(u_i, u_j\right), \\ \text{s.t.} \quad & \forall i \in \mathcal{L}_A, y_i = a_i \\ & \forall i,\ \left\|s_i'\right\|_0 = s_i'\mathbf{1} = k,\ S_{ii} = 0\end{aligned} \tag{4}$$

Since the key idea of this method is to learn the adjacency graph adaptively, we term our method as adaPtively netwOrk embeddIng aNd aTtribute fuSion (POINTS).

A Light Version of POINTS: for each node i, to learn its optimal neighbors, POINTS needs to consider all nodes. This is very inefficient, as the network may be extremely large (some theoretical analysis can be found in Sect. 6.3). Therefore, we give a light version of POINTS (denoted as POINTS*). In particular, we propose to build a candidate neighbor set (denoted as $\mathcal{N}_{k^*}(i)$) for each node, where k^* ($k < k^* \ll n$) is the candidate neighbor number. Based on this idea, the light version POINTS* is to solve the following optimization problem:

$$\begin{aligned}\min_{f,Y,S} \mathcal{J}^* = & \sum_{i\in\mathcal{L}_A} \ell(f(u_i), y_i) + \frac{\alpha}{2}\sum_{i,j} S_{ij} Dist\left(y_i, y_j\right) \\ & + \frac{\beta}{2}\sum_{i,j} S_{ij} Dist\left(u_i, u_j\right), \\ \text{s.t.} \quad & \forall i \in \mathcal{L}_A, y_i = a_i \\ & \forall i,\ \left\|s_i'\right\|_0 = s_i'\mathbf{1} = k,\ S_{ii} = 0 \\ & \forall i,\ \text{if } j \notin \mathcal{N}_{k^*}(i),\ S_{ij} = 0\end{aligned} \tag{5}$$

5. Optimization

The objective functions of POINTS (i.e., Eq. (4)) and POINTS* (i.e., Eq. (5)) both contain 0/1 constraints, which might be difficult to solve by the conventional optimization tools. In this section, we develop efficient solutions for these two problems.

5.1. Optimization for POINTS. Before deriving the optimization algorithm, we need to specify the choice of functions in Eq. (4). For simplicity, we choose a linear model for f, i.e., $f(u_i) = u_i W$, where $W \in \mathbb{R}^{d\times m}$ is the model parameter matrix. In addition, we adopt squared loss for ℓ and squared Euclidean distance for $Dist(\cdot,\cdot)$. As such, we can update the variables in Eq. (4) iteratively, as follows.

Update rule of W and Y: by fixing the other variables, the partial derivative of $\mathcal{J}$ w.r.t. W is

$$\frac{\partial \mathcal{J}}{\partial W} = 2\left(U'UW - U'Y\right). \tag{6}$$

Therefore, we can update W as $W \leftarrow W - \eta(\partial\mathcal{J}/\partial U)$, where η is the learning rate.

When the other variables are fixed, we can obtain the partial derivative of $\mathcal{J}$ w.r.t. Y as

$$\frac{\partial \mathcal{J}}{\partial Y} = 2(-UW + Y) + 2\alpha\Delta Y, \tag{7}$$

where $\Delta = D - (S + S')/2$, and D is a diagonal matrix whose i-th diagonal element is $\sum_j (S_{ij} + S_{ji})/2$. Then, we can update Y as $Y \leftarrow Y - \eta(\partial\mathcal{J}/\partial Y)$. After that, for each node i with given attributes a_i, we adjust its predicted attributes as $y_i = a_i$, so as to satisfy the constraint in Eq. (4).

Update rule of S: when the other variables are fixed, the original optimization problem reduces to

$$\begin{aligned}& min_S \alpha\sum_{i,j} S_{ij}\left\|y_i - y_j\right\|_2^2 + \beta\sum_{i,j} S_{ij}\left\|u_i - u_j\right\|_2^2 \\ & \text{s.t.}\forall i,\ \left\|s_i'\right\|_0 = s_i'\mathbf{1} = k,\ S_{ii} = 0.\end{aligned} \tag{8}$$

As problem (8) is independent between different i, we can instead to solve n decoupled subproblems:

$$\begin{aligned}& \min_{s_i,\forall i} \sum_{j=1}^{n} \alpha\left\|y_i - y_j\right\|_F^2 S_{ij} + \beta\left\|u_i - u_j\right\|_F^2 S_{ij} \\ & \text{s.t.}\forall i,\ \left\|s_i'\right\|_0 = s_i'\mathbf{1} = k,\ S_{ii} = 0.\end{aligned} \tag{9}$$

The optimal solution of problem (9) is (proved in Sect. 6.1)

$$S_{ij} = \begin{cases} 1, & \text{if } j \in \mathcal{N}_k^{UA}(i); \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

where set $\mathcal{N}_k^{UA}(i)$ contains the top-k nearest nodes to i in the network "embedding-attribute" space, where the distance between node i and j is defined as $\alpha\|y_i - y_j\|_2^2 + \beta\|u_i - u_j\|_2^2$..

We can iteratively update these three variables until convergence to obtain the final solution. After that, as discussed in Sect. 6.2, we can get the final fusion results by concatenation. For clarity, we summarize the complete fusion procedure in Alg. 1.

5.2. Optimization for POINTS.* The optimization approach of POINTS* is very similar to that of POINTS in Sect. 5.1. The only difference is that when updating S as other variables are fixed, we only need to sort the nodes in (its neighbor candidate set) $\mathcal{N}_{k^*}(i)$ to get the top-k nearest neighbors in the

Input: The given node embeddings U, the attribute information, learning rate η, and parameters α and β;
Output: The final fusion results;
1: Initialize W, Y and S;
2: **repeat**
3: Update W as $W \leftarrow W - \eta(\partial \mathscr{J}/\partial U)$.
4: Update Y as $Y \leftarrow Y - \eta(\partial \mathscr{J}/\partial Y)$;
5: Set $y_i = a_i$; $\forall i \in \mathscr{L}_A$;
6: Update S by solving problem (8);
7: **until** Convergence or a certain iteration;
8: Obtain the final fusion results via concatenation, as discussed in sect. 6.2
9: **return** The final fusion results.

Algorithm 1. POINTS.

network embedding attribute space, so as to get the optimal solution of S.

6. Algorithm Analysis

6.1. Optimization Algorithm Solving Problem (9)

Theorem 1. *The optimal solution of problem (9) is Eq. (10).*

Proof. By contradiction, suppose node i has gotten its optimal neighbor set $\mathscr{N}_k^{UA}$ which contains a node p not in i's top-k nearest nodes in the "node-attribute" space. For convenience, we use $\Psi(i, j)$ to denote the distance between nodes i and j in this space, i.e., $\Psi(i, j) = \alpha\|y_i - y_j\|_2^2 + \beta\|u_i - u_j\|_2^2$. As such, there must exist a node $q \notin \mathscr{N}_k^{UA}$ which is one of i's top-k nearest nodes in this space. Then, we get $\Psi(i, p) > \Psi(i, q)$. Considering our minimization problem (i.e., Eq. (9)), this inequation leads

$$\sum_{j \in \mathscr{N}_k} \Psi(i, j) > \sum_{j \in \{\mathscr{N}_k + q\} \setminus p} \Psi(i, j). \tag{11}$$

This indicates that $\{\mathscr{N}_k^{UA} + q\} \setminus p$ is a better optimal solution than $\mathscr{N}_k^{UA}$, a contradiction.

Actually, we can generalize the above proof to a more general case.

Theorem 2. *Suppose node i is close to j than node p. If the chosen distance function $Dist(\cdot, \cdot)$ satisfies $Dist(i, j) < Dist(i, p)$, the optimal solution of problem (9) is Eq. (10) (which adopts the distance function $Dist(\cdot, \cdot)$).*

Proof. This conclusion can be proved by replacing the squared Euclid distance function in the proof of Theorem 1 by $Dist(\cdot, \cdot)$.

6.2. Fusion Strategy. In this part, we will discuss how to conduct data fusion, based on the proposed attribute complement methods. The inductive method (described in Sect. 4.1) would learn a generation function f. Then, for each node i, we can predict its attribute vector as $y_i = f(u_i)$. For those two transductive methods (described in Sect. 4.2), we will directly obtain the predicted attribute vectors $y_{i=1,\cdots,n}$. As such, the attributes are completed for fusion. Specifically, we adopt a "trick" concatenation strategy: (1) if node i has no attributes, we obtain its final fusion vector by concatenating u_i and the predicted attribute vector y_i; (2) if node i has attributes, we obtain its final fusion vector by concatenating u_i and a_i. The principle of this trick is that the given attributes are always more stable and accurate than the predicted attributes for node description.

Table 1: The statistics of datasets.

Name	Citeseer	Cora	Wiki	Pubmed
#nodes	3,312	2708	2,405	19,717
#edges	4,732	5429	17,981	44,338
#classes	6	7	17	3
#attributes	3,703	1433	4,973	500

6.3. Time Complexity. The time complexity of Alg. 1 is as below. The complexity for updating W is $O(d^2n + d^2m + dmn)$. The complexity for updating Y is $O(nnz(\Delta)m + dmn)$, where $nnz(\cdot)$ is the number of nonzeros of a matrix. The complexity for updating S is $O(n^2 \log n)$, because for each node, we have to calculate its top-k nearest neighbors. As $d, m \ll n$ is linear with n and $nnz(\Delta)$ is linear with n, the overall complexity of POINTS is $O(\tau(n^2 \log n))$, where τ is the number of iterations to converge.

For the light version, i.e., POINTS*, the complexity of updating S becomes $O(nk^* \log k^*)$, and all others remain the same. Hence, since $k^* \ll n$, the overall complexity becomes $O(\tau(d^2n + dmn + nk^* \log k^*))$. As our method usually converges fast ($\tau \leq 20$ in our experiments) and $d, m \ll n$, the complexity of POINTS* is linear to the number of nodes.

7. Experiments

Datasets: we conduct our experiments on four widely used citation network datasets: Citeseer [31], Cora [31], Wiki [32], and Pubmed [32]. In these networks, nodes are documents, and edges denote the citation relationship between them. Node attributes (i.e., features) are the bag-of-word representations of documents. The statistic of these networks is shown in Table 1.

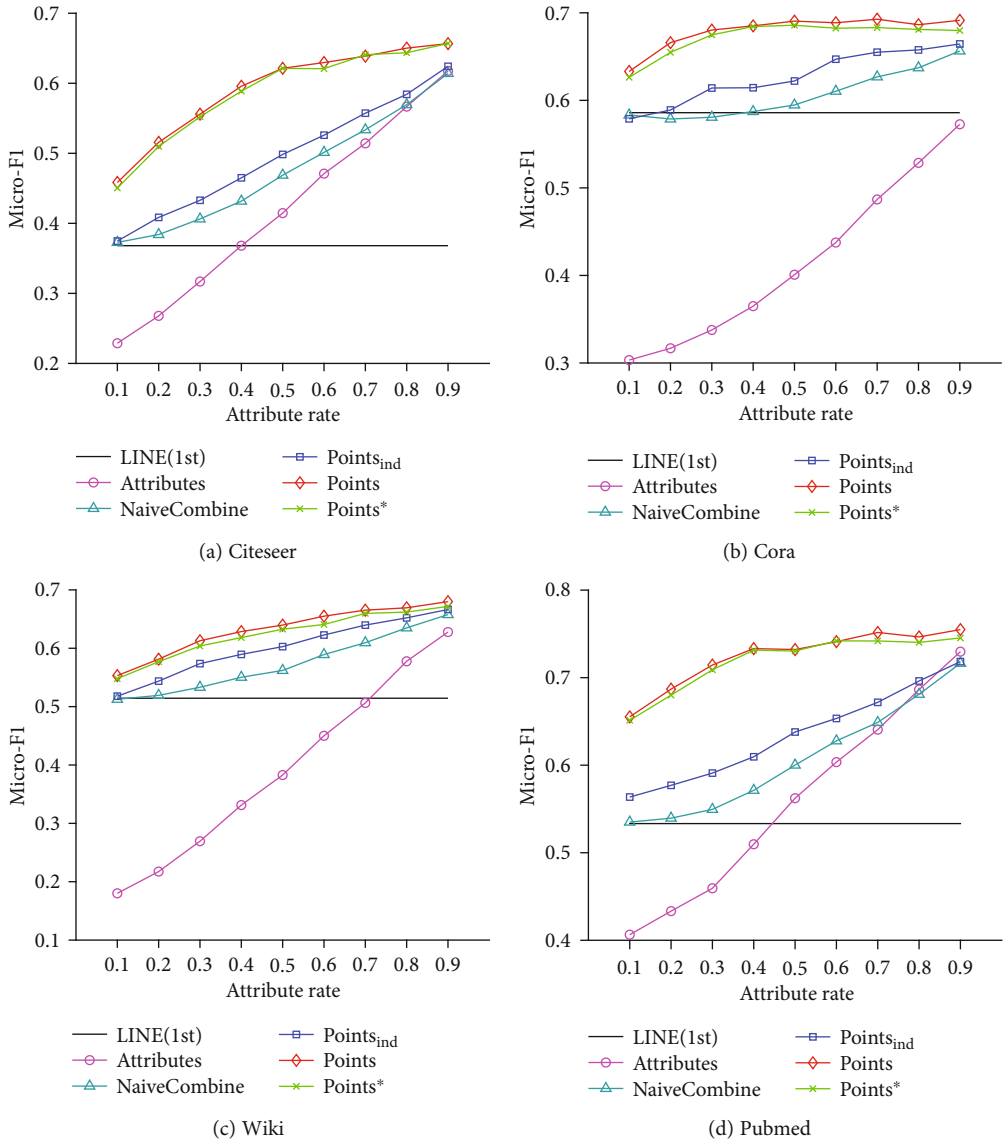


Figure 2: Node classification results (Micro-F1).

Experimental setting: as illustrated in Figure 1, for each dataset, we first obtain the original node embeddings and then provide some nodes with attributes for data fusion, so as to simulate fusing data from two different networks. Specifically, we first obtain the original node embeddings by the famous network embedding method LINE. We adopt its first-order proximity version LINE (1st). Besides, we also try other network embedding methods in Sect. 7.3. After obtaining the original node embeddings, we randomly select some nodes and provide them with attributes. At last, we employ different fusion methods to obtain the final fusion results.

Baseline methods: since this incomplete data fusion problem has not been previously studied, there is no natural baseline to compare with. We thus compare our methods with those methods which directly fuse the original given node embeddings and attributes. We list them as follows:

(1) LINE(1st). We adopt LINE(1st) to obtain the original node embeddings. This method neglects the incomplete node attributes. Note: in Sect. 7.3, we also try more network embedding methods

(2) Attributes. We use the zero-padded attributes as fusion results. This method neglects the given node embeddings

(3) NaiveCombine. We simply concatenate the vectors from the given node embeddings and the zero-padded attributes

For our method, we test its three different versions: POINTS$_{ind}$ (the inductive version formulated in Eq. (1)), POINTS (the full transductive version formulated in Eq. (4)), and POINTS* (the light version formulated in Eq. (5)).

Parameters: we follow the suggestion of LINE to set the embedding dimension to 128. In addition, following [14],

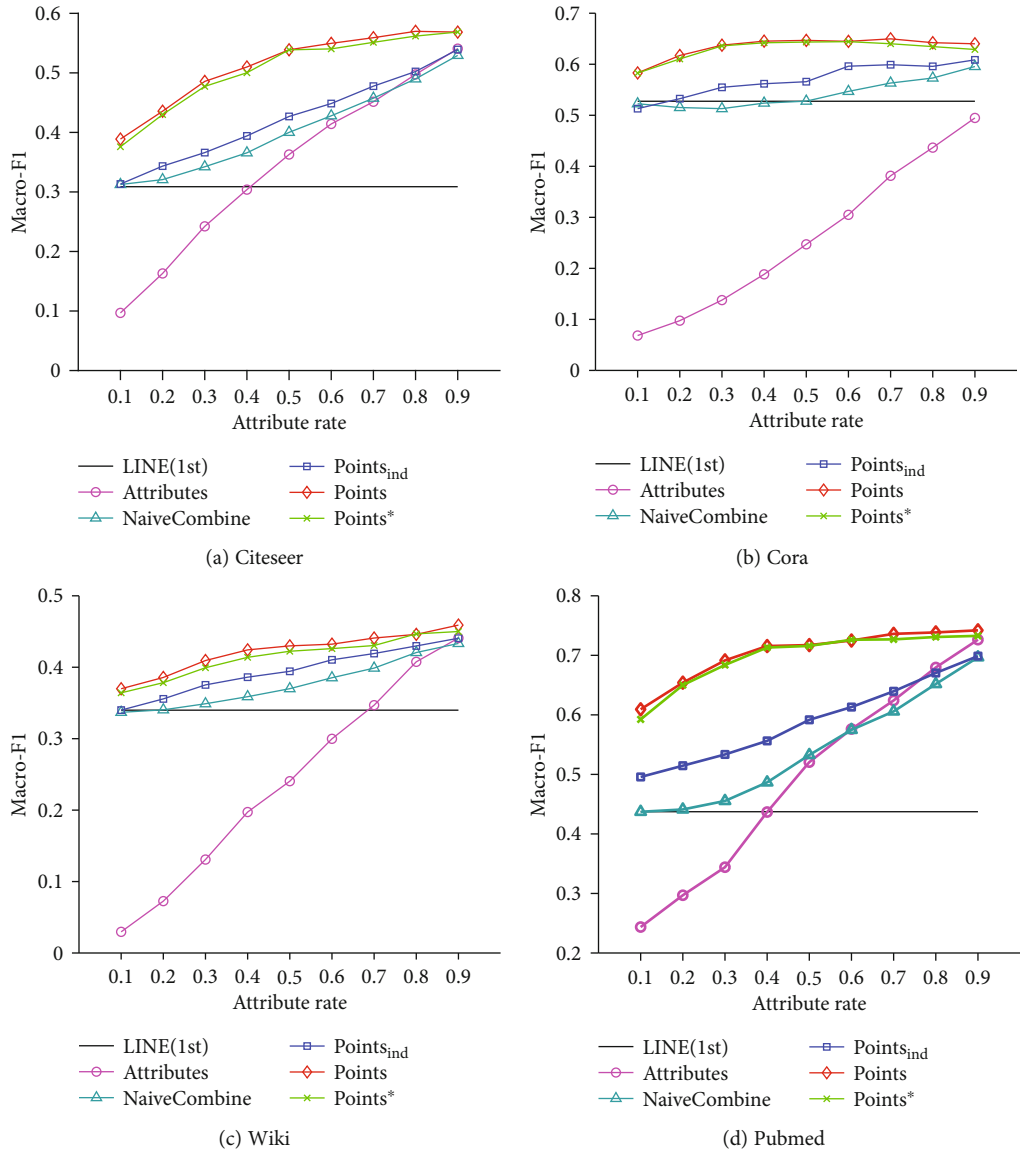


Figure 3: Node classification results (Macro-F1).

we reduce the dimension of attributes by applying SVD decomposition on the original text features. For simplicity, we also reduce this dimension to 128. In the proposed methods POINTS and POINTS*, we fix parameters $\alpha = 1$ and $\beta = 1$ throughout our experiments, although adjusting them would yield better results. Besides, we simply set the neighbor number $k = 5$ like most graph-based transductive methods [33] and set the candidate number $m = 20k$ for POINTS∗.

7.1. Node Classification. Following [6], we train one-vs-rest logistic regression classifiers to evaluated the fusion (i.e., the updated embeddings) quality. Specifically, for Citeseer, Cora, and Wiki, we fix the label rate in the classifiers to 10%. Since Pubmed is a much larger dataset with fewer classes, we follow [34] to set the percentage of labeled data to 1%. In addition, we increase the rate of nodes with attributes from 10% to 90% on all datasets. Following [28], before evaluation, we normalize all representation vectors to unit length for a fair comparison. Figures 2 and 3 show the classification performance measured by micro-F1 and macro-F1 [35], respectively. We can draw the following three conclusions from these results.

Firstly, all our methods (including POINTS$_{ind}$, POINTS, and POINTS*) outperform baseline methods significantly. For example, on Citeseer with 50% attributes, POINTS$_{ind}$, which performs worst in the proposed three methods, still outperforms LINE(1st) by 13%, attributes by 8%, and NaiveCombine by 3%. Additionally, the improvements of our two transductive methods POINTS and POINTS* are more remarkable. These results clearly demonstrate the effectiveness of our complement strategy.

Secondly, the proposed two transductive methods (i.e., POINTS and POINTS*) consistently outperform our inductive method POINTS$_{ind}$. Especially on Citeseer, Cora, and Pubmed, these two transductive methods generally

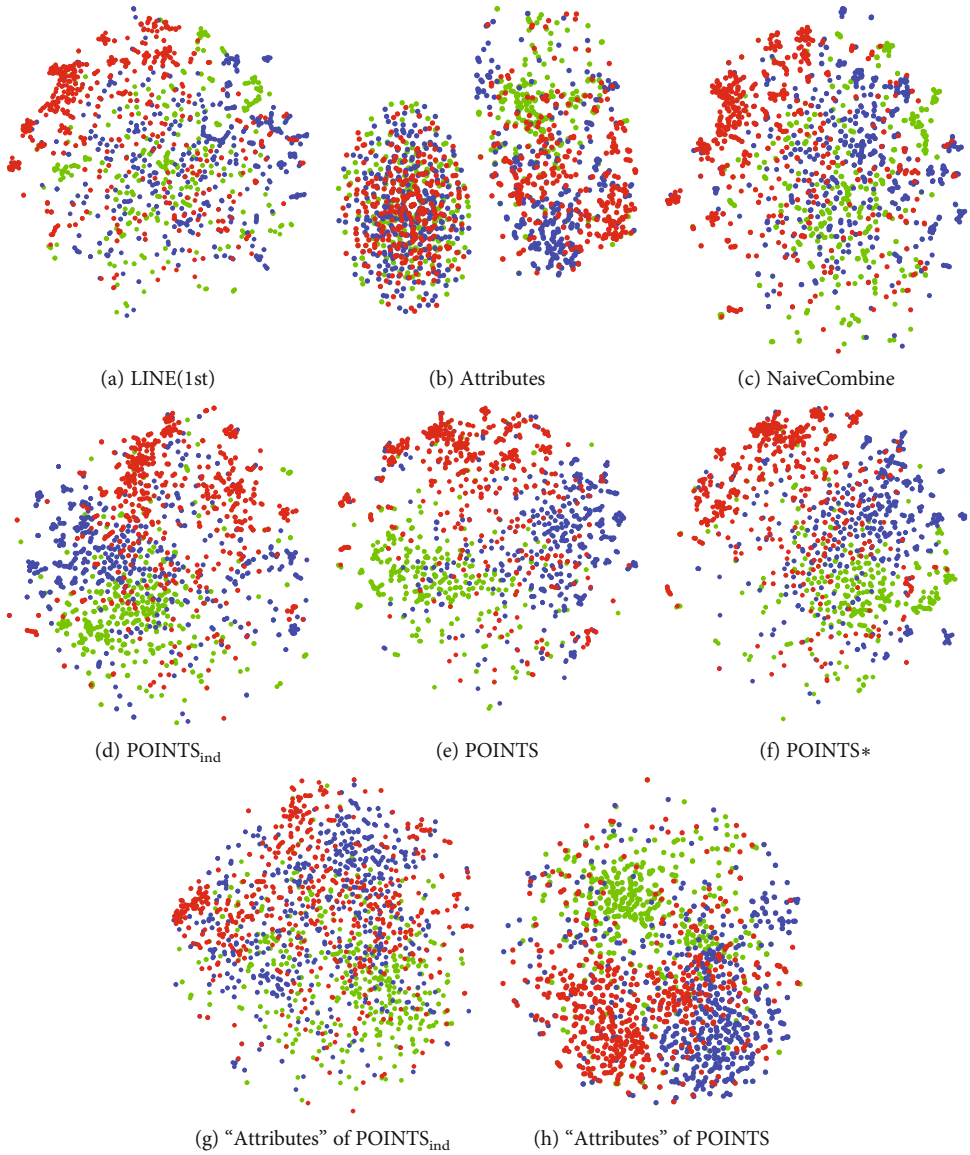


Figure 4: Visualization on Citeseer with 50% attributes. Each point stands for one document. Color of a point indicates the document category. The red indicates the topic of IR, the blue indicates the topic of DB, and the green indicates the topic of HCI. Note: the "Attributes" in (g) and (h) stands for the predicted attributes.

outperform POINTS$_{ind}$ by 5-12%. On the other hand, we also find that the improvement becomes less significant on Wiki. We conjecture that it may be hard to recover its original network structure from the given node embeddings and attributes. More specifically, this might be because Wiki (whose edge num is eight times greater than node number) is much denser than the other three datasets.

Thirdly, the light version POINTS* is comparable to POINTS on all datasets. This indicates that we can reduce the neighbor candidate set size for efficient transductive learning.

7.2. Visualization. Following [28], we use t-SNE package [36] to visualize the final node representations obtained by different fusion methods. Without loss of generality, we choose the first dataset Citeseer and test the case with 50% attributes. Similar to [28], for a clear comparison, we visualize the nodes from three different research fields: IR, DB, and HCI. Figure 4 shows the visualization results.

As shown in Figures 4(a)–4(c), the visualization results of the compared three baselines are not very meaningful, in which the points belonging to different categories are heavily mixed with each other. This is due to the fact that all these baselines cannot sufficiently utilize the incomplete attributes. In contrast, as shown in Figures 4(d)–4(f), the results of our three methods are much better (nodes with same colors are distributed closer). In addition, compared to our inductive method POINTS$_{ind}$, our two transductive methods POINTS and POINTS* show more meaningful layout. Specifically, the blue points in POINTS$_{ind}$ are partly separated by the red points, while these two types of points in POINTS and POINTS* are less mixed with each other. To clarify the reason, we further visualize the predicted attributes of POINTS$_{ind}$ and

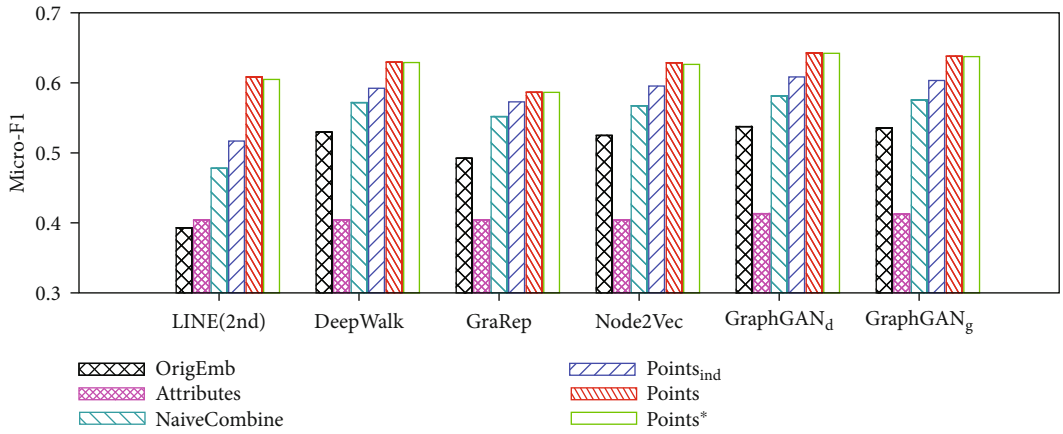


FIGURE 5: Average node classification performance (micro-F1) of more network embedding methods on Citeseer. LINE(2nd) [28] preserves the second-order proximity of a network to learn node embeddings. DeepWalk [6] learns node embeddings by simulating uniform random walks on networks. GraRep [37] generates node embeddings by computing high-order proximity and uses the SVD to reduce their dimensionality. Node2Vec [7] adopts a biased random walk strategy based on DeepWalk to efficiently explore neighborhood architecture. GraphGAN [13] is a newly proposed deep method which employs adversarial training in a minimax game to learn node embeddings. In this method, its discriminator function (denoted as GraphGAN$_d$) and generator function (denoted as GraphGAN$_g$) could separately obtain network embedding results.

POINTS in Figures 4(g) and 4(h), respectively. We can clearly find that POINTS could obtain high-quality attributes, which explains the superiority of our transductive methods. [21].

7.3. More Network Embedding Baselines. We evaluate the performance of our methods based on more network embedding methods. In particular, we further test another five network embedding methods as follows:

Without loss of generality, we fix the label rate to 10% and choose 50% nodes with attributes. For convenience, we use "OrigEmb" to denote the original node embeddings obtained by various network embedding methods.

Figure 5 shows the performance on Citeseer. We can clearly find that our methods (including POINTS$_{ind}$, POINTS, and POINTS*) consistently outperform baselines by a large margin. On the other hand, the light version POINTS* could always achieve similar accuracy as its full version POINTS. Taken together, all these observations clearly indicate the effectiveness of our methods.

8. Conclusion

This paper investigates the problem of fusing node embeddings and incompleted attributes provided by two different networks. We develop both inductive and transductive variants of our method. Additionally, we also provide an efficient light version of our transductive variant. Extensive experiments have demonstrated the effectiveness of our methods. In the further, we would extend our method to fuse more types of related information from more different networks and resources.

Data Availability

The datasets used in this paper can be found at https://linqs.soe.ucsc.edu/data.

Conflicts of Interest

The author(s) declare(s) that they have no conflicts of interest.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China (No. 61902020) and Macao Youth Scholars Program (AM201912).

References

[1] X. Kong, S. Tong, H. Gao et al., "Mobile edge cooperation optimization for wearable internet of things: a network representation-based framework," *IEEE Transactions on Industrial Informatics*, p. 1, 2020.

[2] W. Wang, F. Xia, H. Nie et al., "Vehicle trajectory clustering based on dynamic representation learning of internet of vehicles," in *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[3] C. Wang, C. Wang, Z. Wang, X. Ye, J. X. Yu, and B. Wang, "Deepdirect: learning directions of social ties with edge-based network embedding," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 12, pp. 2277–2291, 2018.

[4] W. Wang, X. Zhao, Z. Gong, Z. Chen, N. Zhang, and W. Wei, "An attention-based deep learning framework for trip destination prediction of sharing bike," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2020.

[5] P. Cui, X. Wang, J. Pei, and W. Zhu, "A survey on network embedding," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 5, pp. 833–852, 2018.

[6] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: online learning of social representations," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 701–710, New York, New York, USA, 2014.

[7] A. Grover and J. Leskovec, "node2vec: scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD*

international conference on Knowledge discovery and data mining, pp. 855–864, San Francisco California USA, 2016.

[8] S. Bourigault, S. Lamprier, and P. Gallinari, "Representation learning for information diffusion through social networks: an embedded cascade model," in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pp. 573–582, San Francisco California USA, 2016.

[9] A. Shakimov, H. Lim, R. Cáceres et al., "Visa-vis: privacy-preserving online social networking via virtual individual servers," in *Proceedings of the 3rd International Conference on Communication Systems and Networks*, pp. 1–10, NW Washington DC, United States, 2011.

[10] X. Cheng, H. Li, and J. Liu, "Video sharing propagation in social networks:measurement, modeling, and analysis," in *INFOCOM, 2013 Proceedings IEEE*, pp. 45–49, Turin, Italy, 2013.

[11] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *International Conference on Machine Learning*, pp. 1188–1196, Beijing, China, 2014.

[12] D. Wang, P. Cui, and W. Zhu, "Structural deep network embedding," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1225–1234, San Francisco California USA, 2016.

[13] H. Wang, J. Wang, J. Wang et al., "GraphGAN: graph representation learning with generative adversarial nets," in *In Thirty-Second AAAI Conference on Artifificial Intelligence*, pp. 2508–2515, New Orleans, Louisiana, USA, 2018.

[14] C. Yang, Z. Liu, D. Zhao, M. Sun, and E. Y. Chang, "Network representation learning with rich text information," in *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pp. 2111–2117, Buenos Aires, Argentina, 2015.

[15] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Advances in Neural Information Processing Systems*, pp. 1024–1034, Curran Associates Inc., 57 Morehouse Lane, Red Hook, NY, United States, 2017.

[16] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *International Conference on Learning Representations (ICLR)*, Toulon, France, 2017.

[17] Z. Wang, X. Ye, C. Wang, Y. Wu, C. Wang, and K. Liang, "RSDNE: Exploring relaxed similarity and dissimilarity from completely-imbalanced labels for network embedding," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, New Orleans, Louisiana, USA, 2018.

[18] Z. Wang, X. Ye, C. Wang, J. Cui, and P. Yu, "Network embedding with completely imbalanced labels," *IEEE Transactions on Knowledge and Data Engineering*, p. 1, 2020.

[19] A. Zhang, S. Song, and J. Wang, "Sequential data cleaning: a statistical approach," in *Proceedings of the 2016 International Conference on Management of Data*, pp. 909–924, San Francisco, California, USA, 2016.

[20] A. Zhang, S. Song, J. Wang, and P. S. Yu, "Time series data cleaning," *Proceedings of the VLDB Endowment*, vol. 10, no. 10, pp. 1046–1057, 2017.

[21] O. Alter, P. O. Brown, and D. Botstein, "Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms," *Proceedings of the National Academy of Sciences*, vol. 100, no. 6, pp. 3351–3356, 2003.

[22] D. F. Hsu and I. Taksa, "Comparing rank and score combination methods for data fusion in information retrieval," *Information Retrieval*, vol. 8, no. 3, pp. 449–480, 2005.

[23] Y. Chen, J. Yang, C. Wang, and N. Liu, "Multimodal biometrics recognition based on local fusion visual features and variational bayesian extreme learning machine," *Expert Systems with Applications*, vol. 64, pp. 93–103, 2016.

[24] J. Bleiholder and F. Naumann, "Data fusion," *ACM Computing Surveys*, vol. 41, no. 1, pp. 1–41, 2009.

[25] J. J. Clark and A. L. Yuille, *Data fusion for sensory information processing systems*, vol. 105, Springer Science & Business Media, 2013.

[26] T. Joachims, "Transductive inference for text classification using support vector machines," in *International Conference on Machine Learning*, vol. 99, pp. 200–209, Bled, Slovenia, 1999.

[27] Y. Jacob, L. Denoyer, and P. Gallinari, "Learning latent representations of nodes for classifying in heterogeneous social networks," in *Proceedings of the 7th ACM international conference on Web search and data mining*, pp. 373–382, New York, New York, USA, 2014.

[28] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, "Line: Large-scale information network embedding," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 1067–1077, Florence, Italy, 2015.

[29] F. Wang and C. Zhang, "Label propagation through linear neighborhoods," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 1, pp. 55–67, 2008.

[30] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency," in *Advances in neural information processing systems*, pp. 321–328, Vancouver, British Columbia, Canada, 2004.

[31] A. K. McCallum, K. Nigam, J. Rennie, and K. Seymore, "Automating the construction of internet portals with machine learning," *Information Retrieval*, vol. 3, no. 2, pp. 127–163, 2000.

[32] P. Sen, G. Namata, M. Bilgic, L. Getoor, B. Galligher, and T. Eliassi-Rad, "Collective classification in network data," *AI Magazine*, vol. 29, no. 3, p. 93, 2008.

[33] X. Zhu, *Semi-supervised learning literature survey*, vol. 2, no. 3, 2006Computer Science, University of Wisconsin-Madison, 2006.

[34] Q. Li, Z. Han, and X. M. Wu, "Deeper insights into graph convolutional networks for semi-supervised learning," 2018, https://arxiv.org/abs/1801.07606.

[35] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine learning: an artificial intelligence approach*, Springer Science & Business Media, 2013.

[36] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.

[37] S. Cao, W. Lu, and Q. Xu, "Grarep: learning graph representations with global structural information," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pp. 891–900, Melbourne, Australia, 2015.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6618980, 11 pages
https://doi.org/10.1155/2021/6618980

Research Article

A Cyber Physical System Crowdsourcing Inference Method Based on Tempering: An Advancement in Artificial Intelligence Algorithms

Jia Liu[1], **Mingchu Li**[1], **William C. Tang**[2] **and Sardar M. N. Islam**[3]

[1]*School of Software Technology and Key Laboratory for Ubiquitous Network and Service Software, Dalian University of Technology, Dalian 116620, China*
[2]*Department of Biomedical Engineering, University of California, Irvine 92697-2715, USA*
[3]*Institute for Sustainable Industries and Liveable Cities, Victoria University, Melbourne 14428, Australia*

Correspondence should be addressed to Jia Liu; jialiudlut@gmail.com

Received 19 December 2020; Revised 24 January 2021; Accepted 6 February 2021; Published 19 February 2021

Academic Editor: Wei Wang

Activity selection is critical for the smart environment and Cyber-Physical Systems (CPSs) that can provide timely and intelligent services, especially as the number of connected devices is increasing at an unprecedented speed. As it is important to collect labels by various agents in the CPSs, crowdsourcing inference algorithms are designed to help acquire accurate labels that involve high-level knowledge. However, there are some limitations in the algorithm in the existing literature such as incurring extra budget for the existing algorithms, inability to scale appropriately, requiring the knowledge of prior distribution, difficulties to implement these algorithms, or generating local optima. In this paper, we provide a crowdsourcing inference method with variational tempering that obtains ground truth as well as considers both the reliability of workers and the difficulty level of the tasks and ensure a local optimum. The numerical experiments of the real-world data indicate that our novel variational tempering inference algorithm performs better than the existing advancing algorithms. Therefore, this paper provides a new efficient algorithm in CPSs and machine learning, and thus, it makes a new contribution to the literature.

1. Introduction

1.1. Research Background. Crowdsourcing refers to the practice whereby a company or organization outsources the tasks that used to be performed by employees to the crowds of nonspecific Internet agents. The wide use of the Internet enables crowdsourcing to make use of the wisdom of crowds as a cheap, fast, and convenient method. Therefore, crowdsourcing is much more powerful than traditional methods, especially in the fields of computer vision, natural language processing, environmental protection [1, 2], etc.

In recent years, the development of embedded computing and wireless communications has enabled the CPSs to become an important research and industrial field. A CPS is often designed as a cooperative network, consisting of sensors, actuators, and controllers [3–6]. In many applications, such as object identification, traffic management, and smart health, CPSs need to extract information and process data in a large scale, a task that is difficult or unable to accomplish at one single device [7, 8]. Therefore, crowdsourcing is important for CPSs to combat the above limitations [9–11].

However, existing CPS crowdsourcing also has disadvantages: for instance, (1) sometimes, agents can be unreliable, or there may be attackers or spammers who aim to corrupt the system; therefore, it is intuitive to assign the same tasks to multiple workers; (2) the difficulty levels of the tasks can vary a lot that if the workers are not properly rewarded, they may choose only the easy tasks, thus degrading the system; (3) it is also critical to allocate tasks to the right kind of workers in order to improve the efficiency and quality of the answers. However, when there is no longer a need to improve the quality of the collected data, we need to improve the computing algorithm to better extract the useful information from the data. Therefore, it is very important to implement

inference algorithms to extract information from the data and deal with the above three disadvantages, the procedure of which is called crowdsourcing inference problem.

1.2. Related Work. In many fields, such as human pose estimation and smart medical, crowdsourcing inference has become a useful and cost-effective method to denote a large quantity of data. The existing crowdsourcing inference algorithms can be classified into four categories, including the weighted majority voting (WMV) method, the statistical method, new kind of inference tools, and variational method.

The first way to obtain ground truth is by majority voting (MV) or WMV, assigning different workers' different weight according to their expertise or reliability [12–15]. The limitation of MV is that it regards all workers to be equally and easily subject to attacks. In contrast to MV, WMV is intuitively more efficient, and all other inference algorithms can be treated as a kind of WMV, so it is meaningless to discuss the limitations of WMV since the situation varies accordingly. In general, the weight of workers should be assigned according to their reliability; therefore, some researchers use qualification test or hidden test to determine the weight of workers [16]. However, this approach is highly dependent on a rich budget and the cooperation of workers.

Another way is to use a statistical method, for example, maximum likelihood estimation (MLE) and expectation maximization (EM). Expectation maximization (EM) is a well-established way to compute the hyperparameters and ground truth in the process of implementing MLE. EM first makes some initial guess on the hyperparameters, then computes expectations under these hyperparameters, and repeats this process until convergence is reached. This process can be treated as an iterative decoding process, so EM is convenient and easy to understand [16–18]. However, EM may result in local optima and is challenging to scale.

Therefore, some researchers turn to a new kind of inference tools, including the widely used back-propagation (BP) and mean filed method (MF) [19]. These algorithms can be effective, as well as can guarantee local optimality and sometimes can guarantee global optimality. However, standard BP requires the knowledge of prior distribution, and in reality, it is difficult to implement due to the process of passing messages in the form of sufficient statistics. MF is the approximation approach that maximizes energy functional over the approximate distribution, which is easy to compute. However, MF may lose valuable information and fail to capture the dependency property of posterior distribution.

An efficient way to avoid the above disadvantages is to use variational method. For example, Liu et al.'s [20] variational inference algorithm (VMP) performs extremely well with good worker distribution prior. In tasks of crowdsourcing inference, Hoffman et al. [21] have formulated it into a stochastic variational inference problem. Liu et al. [22] also formulated the crowdsourcing inference problem into a variational inference model. Cai et al. [23] proposed a crowdsourcing prediction algorithm using variation inference. In summary, variational inference (VI) translates inference problem into an optimization problem with multiple local optimum. However, using the tempering technique, we can curb this to some extent. Intuitively, VI objective uses high-entropy variational distribution to replace data-proper variational distribution. By tempering, we can punish low-entropy distributions and loose this restraint gradually to reach distributions that fit the data.

1.3. Motivation. In view of the problems of the existing literature in this area of CPS crowdsourcing inference algorithms discussed above, the main objectives of this paper are to model the CPS crowdsourcing inference problem into a message passing algorithm, i.e., our novel variational tempering inference algorithm (VTI), and to study the performance error bound of this algorithm.

As a popular sampling as well as computing method, Markov Chain Monte Carlo (MCMC) has proved its efficiency in the field of CPSs and big data, due to its ability to deal with large amount of data, and simplifies the gradient computation [24]. Variational tempering (VT) is an extension of MCMC. Therefore, in VTI, we introduce global temperature. We then implement the traditional VI method to compute the gradient of it in each iteration and further update it in the next iteration, until the algorithm reaches convergence, which is also similar to the way that we treat tasks' answers as a distribution as well as worker reliability.

After researching into the existing works, we borrow the idea of VT and solve the problem of crowdsourcing inference by considering worker reliability and task difficulty. There are few papers in the literature considering task difficulty level. The reason we are considering task difficulty level in this paper is that we can assign tasks in a more efficient way [25], infer the ground truth more accurately [26, 27], and reward workers by means of incentives [28] etc. Therefore, the main objective of this paper is to obtain task ground truth by making the best use of worker reliability and task difficulty. The specific tasks we will perform in this paper are the following: (1) Model the crowdsourcing inference problem into a message passing algorithm. (2) Study the performance error bound of our novel variational tempering inference algorithm (VTI). (3) Use real data to simulate the probability error of VTI.

1.4. Contributions. The main contribution of this paper is a new algorithm to CPS crowdsourcing inference. Based on variational analysis and the property of exponential family, a kind of probabilistic graphical model, we try to solve crowdsourcing inference problem in CPSs. First, we infer the ground truth by considering worker reliability. Being different from previous works, we then hypothesize that tasks have different difficulty levels and use this as a parameter to eliminate the same worker constraint.

Therefore, the contributions of the paper are listed as follows:

(1) Formulate the CPS crowdsourcing inference problem into a message passing algorithm

(2) Study the performance error bound of the proposed variational tempering inference algorithm (VTI)

(3) Use real data to simulate the probability error of VTI. In general cases, VTI performs better than other algorithms

In this paper, we first review CPS crowdsourcing inference and variational tempering in Section 1. We then define the background assumptions of the inference model in Section 2. Thereafter, we derive our algorithm stepwise and show the procedure of how we obtain the results in Sections 3 and 4. We show the results of our algorithm in comparison with others in Section 5. Our VTI algorithm requires less constraint on worker prior and task difficulty level and performs better than the existing inference algorithms. In the end, Section 6 presents the conclusion of this study.

2. Preliminaries

In this part, we describe the mathematical notations of VTI. Assuming in the crowdsourcing network, there are M workers and N tasks. Suppose that each task has the value of $\{\pm 1\}$, and we use z_i to denote the true label of it, i.e., $z_i \in \{\pm 1\}$. We then denote $\mathcal{N}_j$ as worker j's neighbor set, i.e., the indexes of tasks that worker j takes. Likewise, $\mathcal{M}_i$ task i's neighbor set, i.e., the indexes of workers that task i is assigned to. $L \in \{0, \pm 1\}^{N \times M}$ is the label matrix, and $L_{ij} \in \{\pm 1\}$ is worker j's answer or label to task i, and $L_{ij} = 0$ means there is no link between worker j and task i. The goal of our VTI is to get the ground truth and difficulty level of task i as well as the reliability of worker j.

We use $q_j = \text{prob}[L_{ij} = z_i], i \in \mathcal{N}_i$ to denote the reliability of worker j and $d_i = \text{prob}[L_{ij} = z_i], j \in \mathcal{M}_j$ to denote the difficulty level of task i. Note that q and d are two separate factors, so the calculation of q does not involve d, and the calculation of d does not involve q. However, they both influence the label of each task. In addition, when $q_j \approx 1$, worker j produces 100% right answers. $q_j \approx 1/2$ can be tricky since the worker may be a spammer or just submits random answers that contain no useful information. However, if $q_j \leq 1/2$, the worker's answers can be made use of by just reverting its answer matrix. Further, we assume each worker is independent, and its reliability is drawn from the same distribution with hyperparameter θ_1. We assume different tasks are also independent, drawn from the distribution with hyperparameter θ_2.

It is reasonable to study cases when there are fewer spammers and attackers in the system; therefore, we assume $E[q_j \mid \theta_1] \geq 1/2$. We use beta prior for worker reliability $p(q_j \mid \theta_1) \propto q_j^{\alpha_1 - 1}(1 - q_j)^{\beta_1 - 1}$ [29].

In Liu et al.'s work [22], the authors treat the crowdsourcing issue as a graphical inference problem, of which they first use Bayesian theory to reduce the objective function and then use variational inference to compute

$$\begin{aligned} p(z, q, d \mid L, \theta) &\propto \prod_{j \in [M]} p\left(q_j \mid \theta_1\right) \prod_{i \in [N]} p(d_i \mid \theta_2) \prod_{i \in \mathcal{N}_j} p\left(L_{ij} \mid z_i, q_j, d_i\right) \\ &= \prod_{j \in [M]} p\left(q_j \mid \theta_1\right) q_j^{c_j} \left(1 - q_j\right)^{\gamma_j - c_j} \prod_{i \in [N]} p(d_i \mid \theta_2), \end{aligned} \tag{1}$$

where $\gamma_j = |\mathcal{N}_j|$ is the number of answers given by worker j and $c_j \coloneqq \sum_{i \in \mathcal{N}_j} \mathbb{I}[L_{ij} = z_i]$ is the number of correct answers worker j gives. It follows that the value of z with the minimum error probability is

$$\begin{aligned} \widehat{z}_i &= \text{argmax}_{z_i} p(z_i \mid L, \theta) \text{ where } p(z_i \mid L, \theta) \\ &= \sum_{z_{[N] \setminus i}} \int_d \int_q p(z, q, d \mid L, \theta) dq dd \end{aligned} \tag{2}$$

Note that variable z is discrete and q is continuous, and $\theta = \{\theta_1, \theta_2\}$. Because the joint probability depends on both q and z, it is intractable to calculate $p(z, q \mid L, \theta)$ directly. We can take a detour and take the integration over q_j, resulting in a marginal posterior over z.

$$\begin{aligned} p(z \mid L, \theta) &= \int_d \int_q p(z, q, d \mid L, \theta) dq dd \\ &= \prod_{j \in [M]} \int_0^1 \int_0^1 p\left(q_j \mid \theta_1\right) p(d_i \mid \theta_2) q_j^{c_j} \left(1 - q_j\right)^{\gamma_j - c_j} dq dd \\ &\overset{def}{=} \prod_{j \in [M]} \psi_j\left(z_{\mathcal{N}_j}\right) \prod_{i \in [N]} \varphi_i\left(z_{\mathcal{N}_i}\right) \end{aligned} \tag{3}$$

of which $\psi_j(z_{\mathcal{N}_j})$ is the local factor of worker j, i.e., the summation of all tasks taken by worker j; $\varphi_i(z_{\mathcal{N}_i})$ is the local factor of task i, i.e., the summation of all workers take task i. Here, we also assume that ψ_j on θ_1 and L are independent, and φ_i on θ_2 and L are independent.

Further, we use Markov random field assumption to model $p(z \mid L, \theta)$ and treat the task assignment graph as a factor graph, of which variable nodes are task nodes, and factor nodes are worker nodes [30, 31]. We use $m_{i \to j}$ to denote the messages passing from tasks to workers and $m_{j \to i}$ to denote the messages passing from workers to tasks. Messages are initialized as independent Gaussian distribution or other suitable distribution. We then update them iteratively until convergence. To put it in a simple way, the messages are equivalent to the ground truth combined with all relevant factors. Therefore, $b_i(z_i)$ is the belief of the tasks, i.e., the intermediate estimated ground truth of the tasks. Therefore, the updating equations can be formulated as follows:

$$\begin{aligned}
&\text{From tasks to workers}: \quad m_{i\to j}^{t+1}(z_i) \propto \sum_{z_{\mathcal{N}_i\setminus j}} \varphi_i\left(z_{\mathcal{N}_i}\right) \prod_{i'\in\mathcal{N}_{i\setminus j}} m_{j'\to i}^{t}(z_i),\\
&\text{From workers to tasks}: \quad m_{j\to i}^{t+1}(z_i) \propto \sum_{z_{\mathcal{N}_j\setminus i}} \psi_j\left(z_{\mathcal{N}_j}\right) \prod_{i'\in\mathcal{N}_{j\setminus i}} m_{i'\to j}^{t+1}(z_{i'}),\\
&\text{Calculating the beliefs}: \quad b_i^{t+1}(z_i) \propto \prod_{j\in\mathcal{M}_i} m_{j\to i}^{t+1}(z_i).
\end{aligned} \tag{4}$$

Note that tasks can have binary or multiple dimension labels, and in this case, $m_{i\to j}$ or $m_{j\to i}$ can instead be vectors.

3. System Model

In this section, we introduce our VTI and then form its derivations. We use label aggregation problem as an example to study crowdsourcing problem, and since others can be transformed to a label model easily, we use label aggregation to show VTI algorithm. Intuitively, we use log error rate to indicate the performance of our algorithm and the comparison criteria between existing inference algorithms. Further, the convergence rate and speed are also important criteria for inference problems.

In VTI, we add the task difficulty d to improve inference efficiency. At first, we can use d to help compute worker reliability q and then use both d and q to predict ground truth z. In general, there are three advantages when task requestors take into account of the task difficulty level. First, worker reliability parameter will be more valid if task difficulty is included to compute it. The reason is simple: more difficult tasks require workers with more expertise [32, 33]. Second, task difficulty is useful when assigning tasks and can reduce the task time and budget to some extent. Third, it is useful to consider task difficulty when compensating workers, which in turn can sustain the crowdsourcing system. Surely, it is important to consider cases where workers have no incentive to complete a task as soon as possible and may always submit completed tasks in the last minute. Some workers might even submit a random label in a short time. Therefore, a good difficulty level indicator should be able to deal with those situations.

First, we denote worker j's reliability as $q = \{q_j : j \in [M]\}$, task difficulty as $d = \{d_i : i \in [N]\}$, and the true answer of tasks $z = \{z_i : i \in [N]\}$. Therefore, the joint probability of worker reliability, task difficulty, ground truth, and conditioned on the answer matrix L and θ is as Equation (1).

3.1. Assumptions and ELBO Formation. In VTI, we use a three-layered graphical model to show the causal relationship of the parameters. In this model, all data points share global variables, while each local hidden variable belongs to each data point. As defined above, we denote $L = L_{1:N}$ as observed variables (label vector of the tasks), $z = z_{1:N}$ as local hidden values, and θ as global hidden values. The joint density of the model is

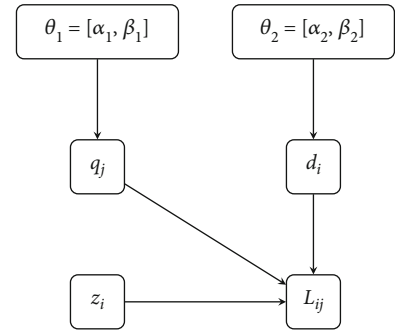


Figure 1: Probability graphical model of VTI.

$$p(q, d, z, L) = p(q, d \mid \theta) \prod_{i=1}^{N} p(z_i, L_i \mid d, q) \tag{5}$$

Figure 1 illustrates the graphical model of VTI.

Computing the posterior inference is the primary task of variational inference. Therefore, our main task is to estimate the ground truth of every task according to collected answers, which is $p(\theta, z \mid L)$, the conditional probability of hidden variables (ground truth), and global hidden variables (data distribution), given observations (collected answers from the crowdsourcing system). However, it is often the case that this conditional probability is difficult or computing-intensive to solve. So it is intuitive for us to resort to approximation methods.

VI is the method of using a family of distributions that is easy to compute to approximate the desired objective posterior distribution. The direct way is to minimize the KL divergence between the posterior distribution and the approximate distribution, which is also a function of the hyperparameters. Thus, the computing process is to solve the optimum of a functional, which is called variational analysis. In the domain of machine learning (ML), this is called the procedure of solving evidence lower bound (ELBO) in terms of a variational parameter.

$$\mathcal{L}(v) = \mathbb{E}_q[\log p(q, d, z, L)] - \mathbb{E}_q[\log q(q, d, z \mid v)] \tag{6}$$

of which $q = q(q, d, z \mid v)$ is the variational distribution we choose to approximate the true joint distribution $p(q, d, z \mid L)$. The detailed derivation of Equation (6) can be found in Appendix.

Assuming that the Bayesian network of our crowdsourcing inference model is fully factored, i.e., each worker and task are independent, and using the chain rule to extend the joint probability, we have

$$q(q, d, z \mid v) = q(q \mid \lambda_1) \prod q(z_i \mid \phi_i) q(d \mid \lambda_2) \tag{7}$$

of which v is the general notation of the variational parameter, λ_1 is the global variational variable for q, λ_2 is the global variational variable for d, $\lambda = (\lambda_1, \lambda_2)$, and ϕ is local variational variable. VI optimizes Equation (2) using a gradient

or coordinate ascent. To better find local optimum, we use variational inference with tempering.

As mentioned before, the procedure of tempering can promote the performance of the optimization process by curbing multioptimum. Therefore, we apply it in our VTI. Using $T \geq 1$ to denote the global temperature, the joint probability becomes

$$p(q, d, z, L \mid T) = \frac{p(z, L \mid q, d)^{1/T} p(q \mid \theta_1) p(d \mid \theta_2)}{c(T)} \tag{8}$$

of which

$$c(T) = \int p(z, L \mid q, d)^{1/T} p(q \mid \theta_1) p(d \mid \theta_2) dL dz \mathrm{d}q dd. \tag{9}$$

Applying Equation (8) in (6), we reach the tempered ELBO.

$$\begin{aligned} \mathscr{L}_A(\lambda, \phi; T) = {} & \mathbb{E}_q\left[\log p(q \mid \theta_1) + \mathbb{E}_q[\log p(d \mid \theta_2)\right] \\ & - \mathbb{E}_q[\log q(q, d \mid \lambda)] \\ & + \sum_{i=1}^{N} \frac{\mathbb{E}_q[\log p(L_i, z_i \mid q, d)]}{T} \\ & - \mathbb{E}_q[\log p(z_i \mid \phi_i)]. \end{aligned} \tag{10}$$

Note that the standard tempering method starts from high temperature and gradually reduces it, taking little account of the influence of data at first, reaching a high-entropy solution, and then slowly makes use of the data or evidence, and reaching a distribution more resembling the data.

Now, we discuss why tempering is effective. The 1st and 3rd terms of Equation (10) are the expected log prior and the log-likelihood. By maximizing these terms, it can result in hidden variables that can better explain the data. The 2nd and 4th terms combined are the variational distribution entropy. Since entropy is convex, it can act as a regularizer which distributes into a hidden variable's configuration. With smaller $1/T$, we obtain a small likelihood of ground truth and smooth entropy distribution. By decreasing T gradually, our variational distribution puts more and more weight on observed data. According to Mandt et al. [34], there are two types of tempering method, one method tempers both likelihood and prior, and the other method tempers only the likelihood. The latter has no problem of gradient stuck during iteration.

This paper discusses gradient-based tempering. As similar to a regular optimization problem, if the tempering is too slow, the algorithm will converge too fast and never reach the global optimum. If we temper too fast, on the other hand, we may skip the global optimum. Since it is difficult to find the proper schedule, we put our emphasis on learning data sequence adaptively. We use variational tempering, a method learning temperature schedule. In variational tempering, we introduce temperature as an auxiliary variable. Therefore, if the temperature is 1, it will degenerate to the original variational inference. We consider finite discrete temperature, for example, $1 = T_1 \leq T_2 \leq \cdots \leq T_M$, which allows us to compute partition function beforehand.

3.2. Objective Function. It is easy to derive that $p(L, z, q, d, y) = p(L, z, q, d \mid y)p(y)$. Given $p(y) = \prod_{m=1}^{M} \pi_m^{y_m}$, the tempered joint of task difficulty, worker reliability, ground truth and worker labels can be written as

$$p(L, z, q, d \mid y) = \frac{p(q, d)}{c(T_y)} \prod_{i=1}^{N} p(L_i, z_i \mid q, d)^{1/T_y} \tag{11}$$

of which T_y is a temperature related to y. Thus, the model becomes

$$p(L, z, q, d, y) = p(q, d) \prod_{m=1}^{M} \left(\frac{\pi_m}{c(T_m)} \prod_{i=1}^{N} p(L_i, z_i \mid q, d)^{1/T_m} \right)^{y_m}. \tag{12}$$

The objective of VTI is to maximize the above joint probability. By further using the chain rule of Bayesian theory, we obtain

$$p(z, q, d, y \mid \psi, \lambda, \gamma) = q(z \mid \psi) q(q, d \mid \lambda) q(y \mid \gamma) \tag{13}$$

of which r denotes the variational parameter for the temperature. Therefore, we obtain the tempered ELBO of VTI.

$$\begin{aligned} \mathscr{L}_T(\lambda, \phi; T) = {} & \mathbb{E}_q[\log p(q)] + \mathbb{E}_q[\log p(d)] + \mathbb{E}_q[\log p(y)] \\ & - \mathbb{E}_q[\log q(q)] - \mathbb{E}_q[\log q(d)] \\ & - \mathbb{E}_q\left[\frac{1}{T_y}\right] \sum_{i=1}^{N} \left(\mathbb{E}_q[\log p(L_i, z_i \mid q, d)]\right) \\ & - \mathbb{E}_q\left[\log C(T_y)\right] \\ & - \sum_{i=1}^{N} \mathbb{E}_q[\log q(z_i)] - \mathbb{E}_q[\log q(y)]. \end{aligned} \tag{14}$$

Therefore, we can maximize this tempered ELBO, instead of optimizing the objective function directly, to obtain ground truth z and hidden parameters q, d.

3.3. Local Temperature. In real practice, often the data would arrive one at a time, or sometimes it is difficult to compute a huge amount of data point at once, so we use the incremental method, which computes each data point as soon as it arrives. This method can simplify the computation and show the influence of each data point to the whole network. Therefore, we can treat the potential or log value of each data point as the local factor. And instead of computing the global temperature at once, we can calculate their local temperature in a similar way. The algorithm is also more scalable to future data. Just like heat flows through the network with thermal conductivity, we can treat the potential or temperature of each local data that passes through the data network. Using

t_i to denote the local temperature, the joint probability of task labels, ground truth, hyperparameter, and temperature becomes

$$p(L, z, q, d, t) \propto p(q, d) \prod_{i=1}^{N} \left[p(L_i, z_i \mid q, d)^{1/t_i} p(t_i) \right]. \quad (15)$$

We can see from the above equation that adding the temperature parameter reduces the weight of the global hidden value and data, i.e., the weight of the ground truth and the task labels. In this way, the resulted distribution of the tasks will have higher entropy. Note that we no longer have to calculate the partition function since there is a $\propto$ in the equation. By adjusting the local temperature, we can obtain a more desired outcome. For instance, outliers can be best demonstrated by higher temperature, while lower temperature can give us the main situation of the data.

4. VTI Algorithms

In this section, we introduce the main part of VTI. We assume that the prior distributions, i.e., the worker prior and the task prior, and the distribution of parameters of the task prior and the worker prior are all from the exponential family [21], which has the following formulation.

$$p(q, d \mid \theta) = h(q, d) \exp \left\{ \theta^T t(q, d) - a_g(\theta) \right\},$$

$$p(z_i, L_i \mid q, d) = h(z_i, L_i) \exp \left\{ (q, d)^T t(z_i, L_i) - a_l(q, d) \right\}, \quad (16)$$

of which $t(q, d)$ and $t(z_i, L_i)$ are sufficient statistics of global and local data points, of which $a_g(\theta)$ and $a_l(q, d)$ are corresponding log normalizers.

Treating the objective function as a function of the global variational parameter, we obtain

$$\begin{aligned} \mathscr{L}_T(\lambda, \phi; T) &= \mathscr{L}_T(\lambda, \phi(\lambda, T); T), \\ \phi(\lambda, T) &= \operatorname{argmax}_{\phi_{ij}} \mathscr{L}_T\left(\lambda, \phi_{ij}; T\right). \end{aligned} \quad (17)$$

According to Hoffman, the tempered ELBO is

$$\begin{aligned} \mathscr{L}_T(\lambda, \phi; T) = \mathbb{E}_q \left[\frac{1}{T_y} \right] \mathbb{E}_q \left[\eta_g(L, z, \theta) \right]^T \nabla_\lambda a_g(\lambda) \\ - \lambda^T \nabla_\lambda a_g(\lambda) + a_g(\lambda) \\ - \mathbb{E}_q \left[\log C\left(T_y\right) \right] - \mathbb{E}_q [\log q(y)]. \end{aligned} \quad (18)$$

4.1. Take-Home Equation. According to Hoffman, the tempered ELBO's natural gradient with respect to the global variational parameter is

$$\begin{aligned} \nabla_\lambda \phi(\lambda, T) &= \mathbb{E}_q \left[\eta_g(L, z, \theta) \right] - \lambda, \\ \widehat{\lambda} &= \mathbb{E}_q \left[\eta_g(L, z, \theta) \right], \end{aligned} \quad (19)$$

of which $\eta_g(L, z, \theta) = (\alpha_1 + \sum_{i=1}^{N} t(z_i, L_i), \beta_1 + N)$.

$$\lambda_{t+1} = \lambda_t + \rho \left(\widehat{\lambda} - \lambda_t \right). \quad (20)$$

In the first step, we estimate $\widehat{\lambda}$ and then update it with λ_t with decreasing learning rate ρ_t. By dividing the expectation of the sufficient statistics, we can reduce the variance of the gradients.

We then optimize tempered ELBO. And the local variational update, i.e., the main body of VTI, is

$$\phi_{ij} = \frac{1}{T} \mathbb{E}_q \left[\eta_l \left(z_{i,-j}, L_i, q, d \right) \right], \quad (21)$$

of which, η_l is the local variational parameter and is also the natural parameter of the exponential family distribution $p(z_{ij} \mid L_i, z_{i,-j}, q, d)$. Therefore, we can compute η_l by constructing the exponential family distribution $p(z_{ij} \mid L_i, z_{i,-j}, q, d)$, since $L_i, z_{i,-j}$ is known, and q, d is initialized.

Therefore, VTI can be summarized in Algorithm 1.

Initialize $\lambda_0, T, q_0, d_0, L, \phi_{ij}, \rho_0$
Output: z, q, d
for each label L_{ij} do
  Update local variational parameter:
  $\phi_{ij} = (1/T)\mathbb{E}_q[\eta_l(z_{i,-j}, L_i, q, d)]$.
end for
Update global variational parameter:
$\widehat{\lambda} = \mathbb{E}_q[\eta_g(L, z, \theta)]$
$\lambda_{t+1} = \lambda_t + \rho(\widehat{\lambda} - \lambda_t)$
until convergence
Output: $q_j = \text{prob}[L_{ij} = z_i], i \in \mathscr{N}_j$
  $d_i = \text{prob}[L_{ij} = z_i], j \in \mathscr{M}_i$
  $z_i = \text{sign}\ [q_j L_{ij}], j \in \mathscr{M}_i$

Algorithm 1. VTI.

4.2. Analysis. In this part, we analyze the performance guarantee of VTI. To allocate the tasks to different workers, we use a bipartite graph algorithm, of which the left degree of the graph l refers to how many workers are needed for one task, and the right degree of the graph r refers to how many tasks are allocated to one worker at one time period. Intuitively, the error rate e is dependent on l, r, q, and d. By the definition of q_j, we can see that the effective information provided by worker j can be measured by $g_j = |2q_j - 1|$; therefore, the expectation of effective information provided by one worker is $g_1 = \mathbb{E}[|2q_j - 1|]$. Let $\widehat{l} = l - 1$ and $\widehat{r} = r - 1$, u_1

$= \mathbb{E}[2q_j - 1]$, $u_2 = \mathbb{E}[2d_i - 1]$, $u = u_1/u_2$, $g_2 = \mathbb{E}[|2d_i - 1|]$, and $g = g_1/g_2$. Define $\rho_k^2 \equiv 2g/u^2(g^2 l \wedge r \wedge)^{k-1} + (3 + 1/g\widehat{lr})1 - (1/g^2 l \wedge r \wedge)^{k-1}/1 - (1/g^2\widehat{lr})$.

For $g^2\widehat{lr} > 1$, let $\rho_\infty^2 \equiv \lim_{k\to\infty}\rho_k^2$ such that

$$\rho_\infty^2 = \frac{\left(3 + \left(1/g\widehat{lr}\right)\right)g^2\widehat{lr}}{\left(g^2\widehat{lr} - 1\right)}. \tag{22}$$

We can then arrive at the following error bound.

Theorem 1. *Suppose n tasks are distributed to m workers, forming a l, r bipartite graph. If the worker reliability distribution and task difficulty distribution satisfy $u_1 > 0$, $u_2 > 0$, and $g > 1/(\widehat{lr})$, the k-th estimates of VTI can achieve*

$$\limsup_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} \mathbb{P}\left(z_i \neq \widehat{z}_i\left(\left\{L_{ij}\right\}_{(i,j)\in E}\right)\right) \geq e^{-\lg/\left(2\sigma_k^2\right)}. \tag{23}$$

Theorem 2. *According to Theorem 1, we assume n tasks are distributed to m workers, forming a l, r bipartite graph. If the worker reliability distribution and task difficulty distribution satisfy $u_1 > 0$, $u_2 > 0$, and $g > 1/(\widehat{lr})$, the k-th estimates of VTI can achieve*

$$\limsup_{k\to\infty}\ \limsup_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} \mathbb{P}\left(z_i \neq \widehat{z}_i\left(\left\{L_{ij}\right\}_{(i,j)\in E}\right)\right) \geq e^{-\lg/\left(2\rho_\infty^2\right)}. \tag{24}$$

The implication of this theorem is that the error rate is upper bounded by $e^{-\lg/(2\sigma_\infty^2)}$. Notice that the temperature parameter only affects the convergence and has no effect on error bound. We present the proof of Theorem 1 in Appendix.

4.3. Discussion. We developed a variational inference algorithm with tempering, which is scalable, and can analyze large amount of data by complex probabilistic graphical models. The main idea is to optimize the variational objective using stochastic optimization. By repeatedly sampling the data, it can estimate the natural gradient with noise. By introducing tempering into the objective function, we can simplify the computation complexity and improve convergence rate of the iteration. With VTI, we can easily apply label aggregation modeling to various kinds of pattern recognition and classification problem. Furthermore, VTI can be generalized to more settings as it opens the door for several research directions. Here, we analyze VTI's complexity and how to interpret its results.

First, VTI has the same runtime with majority voting, which is $O(ml)$.

Second, the assumption of conjugate priors is necessary. Conjugate prior plays an important role in variational inference. It enables the form of the posterior probability to resemble the prior probability, which simplifies Bayesian analysis to a large extent. Taking VTI for example, the worker and task prior are Gaussian, and their posteriors are also Gaussian. Using conjugate prior assumption, the inference model can be easily extended to different kinds of distribution that belong to the exponential family.

Third, the mean-field assumption is also necessary. By approximating the true parameters of the prior distribution with a distribution that is not restricted, we can reduce computation and solve complex probabilistic graphical models.

Fourth, our VTI can establish good performance under different distributions of q_j and d_i. Furthermore, VTI is robust to different initialization, i.e., almost every starting point can result in a unique solution. The reason for this robustness is that VTI computes the energy of the objective function and will surely decrease with each iteration.

5. Experiments and Results

In this section, we will assess seven advancing algorithms on real datasets for the experiments of our algorithm. First, the experimental setup is introduced, and then, the convergence performance of the collected crowdsourcing inference algorithm is analyzed. Finally, we make comparison of the error rate of VTI with the existing algorithms.

5.1. Experimental Setup. We use Amazon Mechanical Turk datasets (Data: https://github.com/musegoduci/variational-inference-for-crowdsourcing) to show the comparative results of the inference algorithms. The Amazon Mechanical Turk is a binary dataset, of which 0 means a worker identifies no target in a task, and 1 means the worker identifies target in a task. Note that the data need to be centralized before processing. We use bipartite assignment graph algorithm to generate the initial answer matrix for the simulation. Some notations are listed in Table 1. It should be noted that KOS (Karger, Oh, and Shah's iterative learning algorithm for crowdsourcing inference [29]) is a message passing-based algorithm, and we first identify the reliability of each worker by some data sample and then implement MLE to estimate the test sample's ground truth. However, BP, VMP, NMP (Liu et al.'s novel crowdsourcing inference method [22]), and VTI assume a probability distribution of worker prior and task prior and then solve the hyperparameters of the distribution, leading to a MLE estimation of the ground truth.

5.2. Convergence. In this part, we analyze the convergence behavior of VTI. Note that the overall executing time for an inference algorithm depends on how long it takes to iterate once, as well as on how many times it will iterate. Since the normalizing constants can be precomputed in advance, the executing time of one iterate depends only on the calculation of the natural gradient, which gives faster convergence than others. Therefore, VTI can converge in seconds. Furthermore, we use subsampling to obtain the natural gradient of the data, which is computationally economical in each iteration. Therefore, VTI possesses advantageous computational performance.

Table 1: Notations and settings.

Parameter	Commentate	Value
l	Degree of tasks	5
γ	Degree of workers	5
N	Number of tasks	1000
M	Number of workers	1000
L_{ij}	Worker j's label toward task i	
q_j	Worker j's reliability	
d_i	Task i's difficulty	
z_i	Task i's ground truth	
θ_1	Hyperparameter of worker reliability's distribution	
θ_2	Hyperparameter of task difficulty's distribution	
$[\alpha_1, \beta_1]$	Worker prior	[0.5,1]
$[\alpha_2, \beta_2]$	Task prior	[0.4,0.6]
ε	Convergence tolerance	10^{-6}

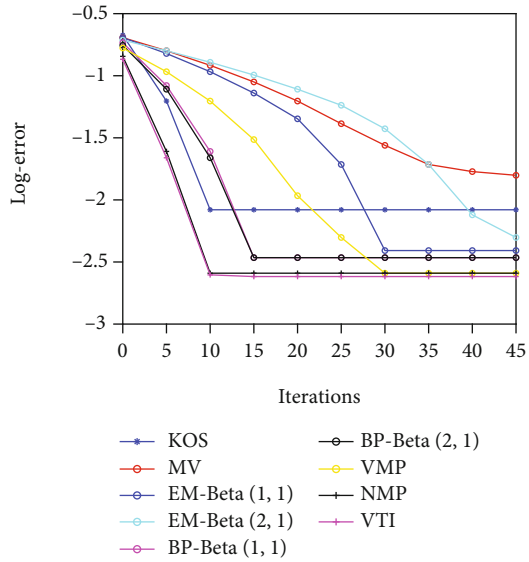


Figure 2: Convergence rate comparison.

From Figure 2, we can show that our VTI has a faster converge rate than KOS, MV, EM, and VMP. It is straight forward to know that this phenomenon is inherent. The reason is that VTI computes both the task difficulty and worker reliability simultaneously and reduces error in both directions. Furthermore, how fast KOS can converge depends on whether and how the ground truth is centered. If the degree of the bipartite graph is low, BP may not even converge. VMP and NMP all have rather good convergence performance because they, like VTI, are both based on the mean-field variational method.

5.3. Error Rate. In this research, we make comparison of VTI with MV, EM, KOS, VMP, and NMP.

In Figure 3, we simulate the performance of VTI with fixed right degree. Intuitively, the log error rate is positively correlated to the left degree, i.e., the quantity of tasks each worker takes. From the figure, we find that VTI has better performance under these settings.

In Figure 4, we simulate the performance of VTI with fixed left degree. Intuitively, the log error rate is positively correlated to the right degree, i.e., the quantity of workers each task takes. From the figure, we find that VTI has better performance under these settings. In Figure 5, we simulate the performance of VTI with fixed left and right degrees. Intuitively, the log error rate is positively correlated to both the degrees. From the figure, we find that VTI has better performance under these settings.

To sum up, it can be seen from Figures 3–5 that our VTI can achieve a slightly higher error rate, which proves that the inference algorithm often performs better when considering both the reliability of workers and task difficulty. And this algorithm can also ensure finding a local optimum.

6. Conclusion

In this paper, we solve the problem of cyber physical system crowdsourcing inference using variational inference with tempering. Our VTI considers not only the worker reliability, but also the task difficulty, which makes VTI adaptable to more complex probabilistic graphical model. With the tempering procedure, the iteration process can reach a smoother local optimum and better demonstrate the relationship of balancing the influence of single data point and global variational parameter. The results are promising in reality and insightful in the field of statistical inference. Therefore, the paper contributes to the research of finding efficient

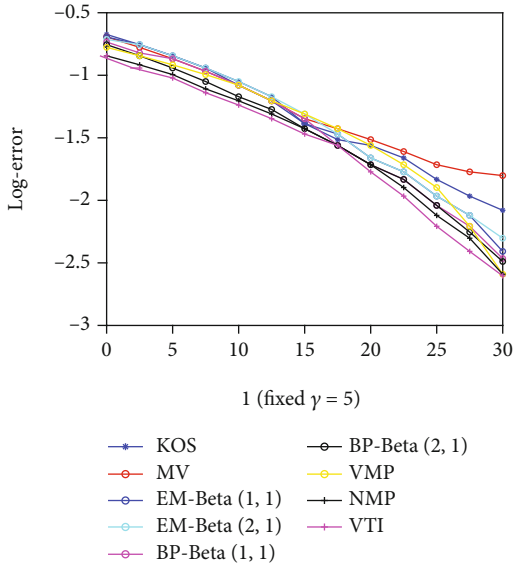


Figure 3: Comparison with fixed number of tasks per worker.

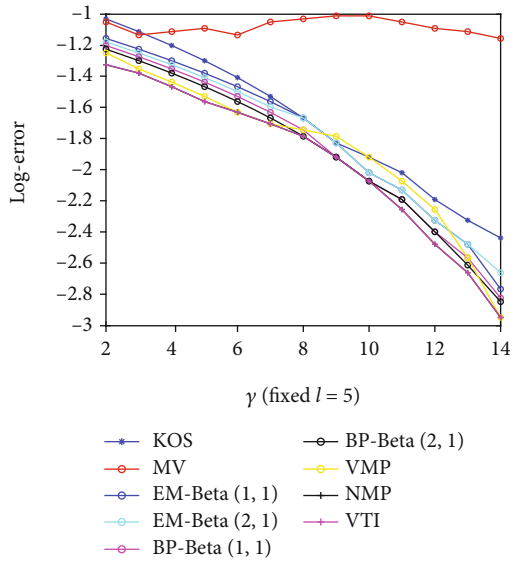


Figure 4: Comparison with fixed number of workers per task.

algorithms in CPSs. Promising directions in future include formulating the inference problem in a more general background and exploring the connection of variational inference with all other inference and estimation algorithms.

7. Limitations of VTI

VTI suffers from a few limitations just like other variational inference algorithms. First, VTI relies on the assumption of the worker prior and difficulty level prior. It is applicable only to CCEF model. Second, VTI assumes that each worker has the same reliability towards different kinds of questions. Although it is capable of dealing with cases when different tasks are categorized into different

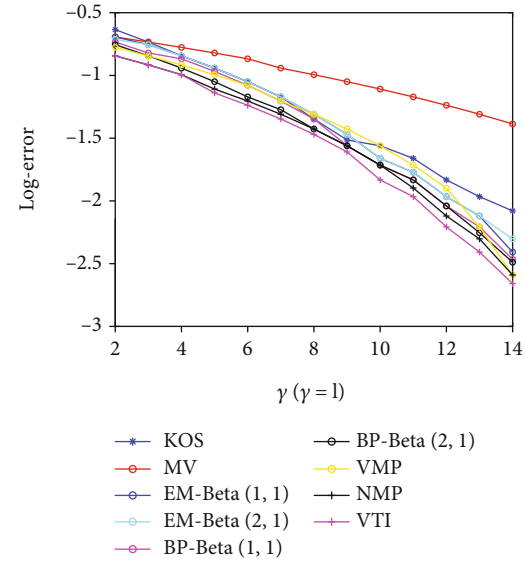


Figure 5: Comparison with the same worker and task degree.

classes, VTI can become extremely complex doing so. Third, VTI has the same cold start issue with other variational inference algorithms. In this paper, we initialize worker reliability and task difficulty uniformly and compute it batch by batch. However, it is also difficult to achieve a computable bipartite graph if in practice, the answers of different tasks arrive at different rate.

Appendix

A. Derivation of Equation (6)

In this part, we show how to obtain Equation (6). Jensen's inequality is used in the 3rd step.

$$\begin{aligned}\log p(L) &= \log \int p(q,d,z,L)\,dzdqdd \\ &= \log \int p(q,d,z,L)\frac{q(q,d,z \mid v)}{q(q,d,z \mid v)}\,dzdqdd \\ &= \log\left(\mathbb{E}_q\left[\frac{p(q,d,z,L)}{q(q,d,z \mid v)}\right]\right) \\ &\geq \mathbb{E}_q[\log p(q,d,z,L)] - \mathbb{E}_q[\log q(q,d,z \mid v)] \triangleq \mathscr{L}(v).\end{aligned} \tag{A.1}$$

B. Proof of Theorem 1

Proof. Because of the local nature of crowdsourcing, we can treat it as a repetition coding and analyze its bit error bound in a similar way. In this case, we repeat each bit r times. By symmetry, we can assume all t_i's are +1. For any of the m tasks, $(1/m)\sum_{i\in[m]}\mathbb{P}(t_i \neq \widehat{t}_i) \leq \mathbb{P}(z_i^k \leq 0)$, where z_i^k is the estimated answer after k-th iteration. Modeling the probabilistic graphical model of VTI as a bipartite graph and each node with its neighbor can be

treated as a regular tree. Intuitively, we have the following equality:

$$\lim_{m\to\infty} \mathbb{P}\left(z_i^k \le 0\right) = \mathbb{P}\left(z\wedge^k \le 0\right). \tag{B.1}$$

In order to analyze the performance guarantee of VTI, we can borrow the idea in Chernoff bound [35]. It follows that $P(x\wedge^k \ge 0) \le \mathbb{E}[e^{\lambda x\wedge^k}]$; therefore, $1 - P(x\wedge^k \ge 0) \ge \mathbb{E}[e^{\lambda x\wedge^k}]$, i.e., $P(x\wedge^k \le 0) \ge \mathbb{E}[e^{\lambda x\wedge^k}]$. Define $\sigma_k^2 \equiv 2(l\wedge r\wedge)^k + u^2 l\wedge^3 \widehat{r}(3g\widehat{r}+1)((g^2 l\wedge r\wedge)^{2k-4})(1-(1/g^2 l\wedge r\wedge)^{k-1})/(1-(1/g^2\widehat{lr}))$, and $\mu_k \equiv u\widehat{l}(g^2 l\wedge r\wedge)^{k-1}$, and variable z with mean μ and variance σ will be a sub-Gaussian distribution, and $\mathbb{E}[e^{\lambda} z] \le e^{\mu\lambda+(1/2)\sigma^2\lambda^2}$. Therefore,

$$P\left(x\wedge^k \le 0\right) \le \mathbb{E}\left[e^{\lambda} x\wedge^k\right] \le e^{-l\mu_k^2/\left(2l\wedge\sigma_k^2\right)}. \tag{B.2}$$

This, finishes the proof.

Data Availability

The data is available at https://github.com/musegoduci/variational-inference-for-crowdsourcing.

Conflicts of Interest

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

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61802097 and 61877007), the Fundamental Research Funds for the Central Universities (Grant No. DUT20GJ205), the Project of Qianjiang Talent (Grant No. QJD1802020), and the Key Research & Development Plan of Zhejiang Province (Grant No. 2019C01012).

References

[1] A. Moayedikia, W. Yeoh, K. L. Ong, and Y. L. Boo, "Framework and literature analysis for crowdsourcing's answer aggregation," *Journal of Computer Information Systems*, vol. 60, no. 1, pp. 49–60, 2020.

[2] A. J. Porter, P. Tuertscher, and M. Huysman, "Saving our oceans: scaling the impact of robust action through crowdsourcing," *Journal of Management Studies*, vol. 57, no. 2, pp. 246–286, 2020.

[3] X. Han, G. Shen, X. Yang, and X. Kong, "Congestion recognition for hybrid urban road systems via digraph convolutional network," *Transportation Research Part C: Emerging Technologies*, vol. 121, p. 102877, 2020.

[4] Y. J. Hsu, "Crowdsourcing agents for smart IoT," in *HAI '14: Proceedings of the second international conference on Human-agent interaction*, Tsukuba Japan, 2014.

[5] W. Wang, J. Liu, Z. Yang, X. Kong, and F. Xia, "Sustainable collaborator recommendation based on conference closure," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 2, pp. 311–322, 2019.

[6] W. Wang, J. Ren, M. Alrashoud, F. Xia, M. Mao, and A. Tolba, "Early-stage reciprocity in sustainable scientific collaboration," *Journal of Informetrics*, vol. 14, no. 3, p. 101041, 2020.

[7] W. Wang, X. Zhao, Z. Gong, Z. Chen, and W. Wei, "An attention-based deep learning framework for trip destination prediction of sharing bike," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2020.

[8] K. Xiangjie, T. Shiqin, G. Haoran et al., "Mobile edge cooperation optimization for wearable internet of things: a network representation-based framework," *IEEE Transactions on Industrial Informatics*, pp. 1–10, 2020.

[9] M. Li and P. Li, "Crowdsourcing in cyber-physical systems: stochastic optimization with strong stability," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 2, pp. 218–231, 2013.

[10] A. Nieto, A. Acien, G. Fernandez, Q. He, C. Xiang, and G. Wei, "Crowdsourcing analysis in 5g IoT: cybersecurity threats and mitigation," *Science of the Total Environment*, vol. 24, no. 3, pp. 881–889, 2019.

[11] M. Strohmeier, M. Smith, M. Schfer, V. Lenders, and I. Martinovic, "Crowdsourcing security for wireless air traffic communications," in *2017 9th International Conference on Cyber Conflict (CyCon)*, Tallinn, Estonia, 2017.

[12] B. M. Booth, K. Mundnich, and S. Narayanan, "Fusing annotations with majority vote triplet embeddings," in *AVEC'18: Proceedings of the 2018 on Audio/Visual Emotion Challenge and Workshop*, pp. 83–89, Seoul Republic of Korea, 2018.

[13] Y. Chen, M. Guizani, Y. Zhang et al., "When traffic flow prediction and wireless big data analytics meet," *IEEE Network*, vol. 33, no. 3, pp. 161–167, 2019.

[14] P. Germain, A. Lacasse, F. Laviolette, M. Marchand, and J. F. Roy, "Risk bounds for the majority vote: from a PAC-Bayesian analysis to a learning algorithm," *The Journal of Machine Learning Research*, vol. 16, no. 1, pp. 787–860, 2015.

[15] A. Hotaling and J. P. Bagrow, "Accurate inference of crowdsourcing properties when using efficient allocation strategies," 2019, https://arxiv.org/abs/1903.03104.

[16] Y. Zheng, G. Li, Y. Li, C. Shan, and R. Cheng, "Truth inference in crowdsourcing: is the problem solved?," *Proceedings of the VLDB Endowment*, vol. 10, no. 5, pp. 541–552, 2017.

[17] R. M. Neal and G. E. Hinton, "A view of the em algorithm that justifies incremental, sparse, and other variants," in *Learning in Graphical Models*, pp. 355–368, Springer, 1998.

[18] X. Wang, C. Li, J. Zhang, Q. Zhang, and J. Hu, "Variational inference-based em for quantized fir system parameter identification," in *2018 IEEE 14th International Conference on Control and Automation (ICCA)*, pp. 636–640, Anchorage, AK, USA, 2018.

[19] S. Hao and X. Yang, "A globally convergent mean-field inference method in dense Markov random fields," in *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, Wuhan, China, 2017.

[20] Q. Liu, J. Peng, and A. T. Ihler, "Variational inference for crowdsourcing," *Advances in neural information processing systems*, vol. 1, pp. 692–700, 2012.

[21] M. D. Hoffman, D. M. Blei, C. Wang, and J. Paisley, "Stochastic variational inference," *Journal of Machine Learning Research*, vol. 14, no. 1, pp. 1303–1347, 2013.

[22] J. Liu, W. C. Tang, Y. Chen, M. Li, and M. Guizani, "A novel crowd-sourcing inference method," in *2019 15th International*

Wireless Communications and Mobile Computing Conference (IWCMC), Tangier, Morocco, Morocco, 2019.

[23] D. Cai, D. T. Nguyen, L. S. Hong, and L. Wynter, "Variational Bayesian inference for crowdsourcing predictions," 2020, https://arxiv.org/abs/2006.00778.

[24] A. Buchholz, F. Wenzel, and S. Mandt, "Quasi-monte carlo variational inference," in *Proceedings of the 35th International Conference on Machine Learning*, Stockholmsmässan, Stockholm Sweden, 2018.

[25] M. Yusoff, M. N. S. B. M. Ikram, and N. Janom, "Task assignment optimization for crowdsourcing using genetic algorithm," *Advanced Science Letters*, vol. 24, no. 11, pp. 8205–8208, 2018.

[26] A. Moayedikia, W. Yeoh, K. L. Ong, and Y. L. Boo, "Improving accuracy and lowering cost in crowdsourcing through an unsupervised expertise estimation approach," *Decision Support Systems*, vol. 122, article 113065, 2019.

[27] J. Shu, X. Liu, X. Jia, K. Yang, and R. H. Deng, "Anonymous privacy-preserving task matching in crowdsourcing," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 3068–3078, 2018.

[28] Y. Xing, L. Wang, Z. Li, and Y. Zhan, "Multi-attribute crowdsourcing task assignment with stability and satisfactory," *IEEE Access*, vol. 7, pp. 133351–133361, 2019.

[29] D. R. Karger, S. Oh, and D. Shah, "Iterative learning for reliable crowdsourcing systems," *Advances in Neural Information Processing Systems*, pp. 1953–1961, 2011.

[30] L. E. Ortiz and Z. Gong, "Correlated equilibria for approximate variational inference in MRFs," 2016, https://arxiv.org/abs/1604.02737.

[31] Y. Yang, X. Cong, K. Long, Y. Luo, W. Xie, and Q. Wan, "MRF model-based joint interrupted SAR imaging and coherent change detection via variational Bayesian inference," *Signal Processing*, vol. 151, pp. 144–154, 2018.

[32] M. Venanzi, J. Guiver, P. Kohli, and N. R. Jennings, "Time-sensitive Bayesian information aggregation for crowdsourcing systems," *Journal of Artificial Intelligence Research*, vol. 56, pp. 517–545, 2016.

[33] Y. Yang, Q. Bai, and Q. Liu, "Modeling random guessing and task difficulty for truth inference in crowdsourcing," in *AAMAS '19: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 2288–2290, Montreal, Canada, 2019.

[34] S. Mandt, J. Mcinerney, F. Abrol, R. Ranganath, and D. Blei, "Variational tempering," *Computer Science*, 2014.

[35] D. Arjo, "Statistical models: theory and practice," *Technometrics*, vol. 48, no. 2, pp. 315–315, 2009.