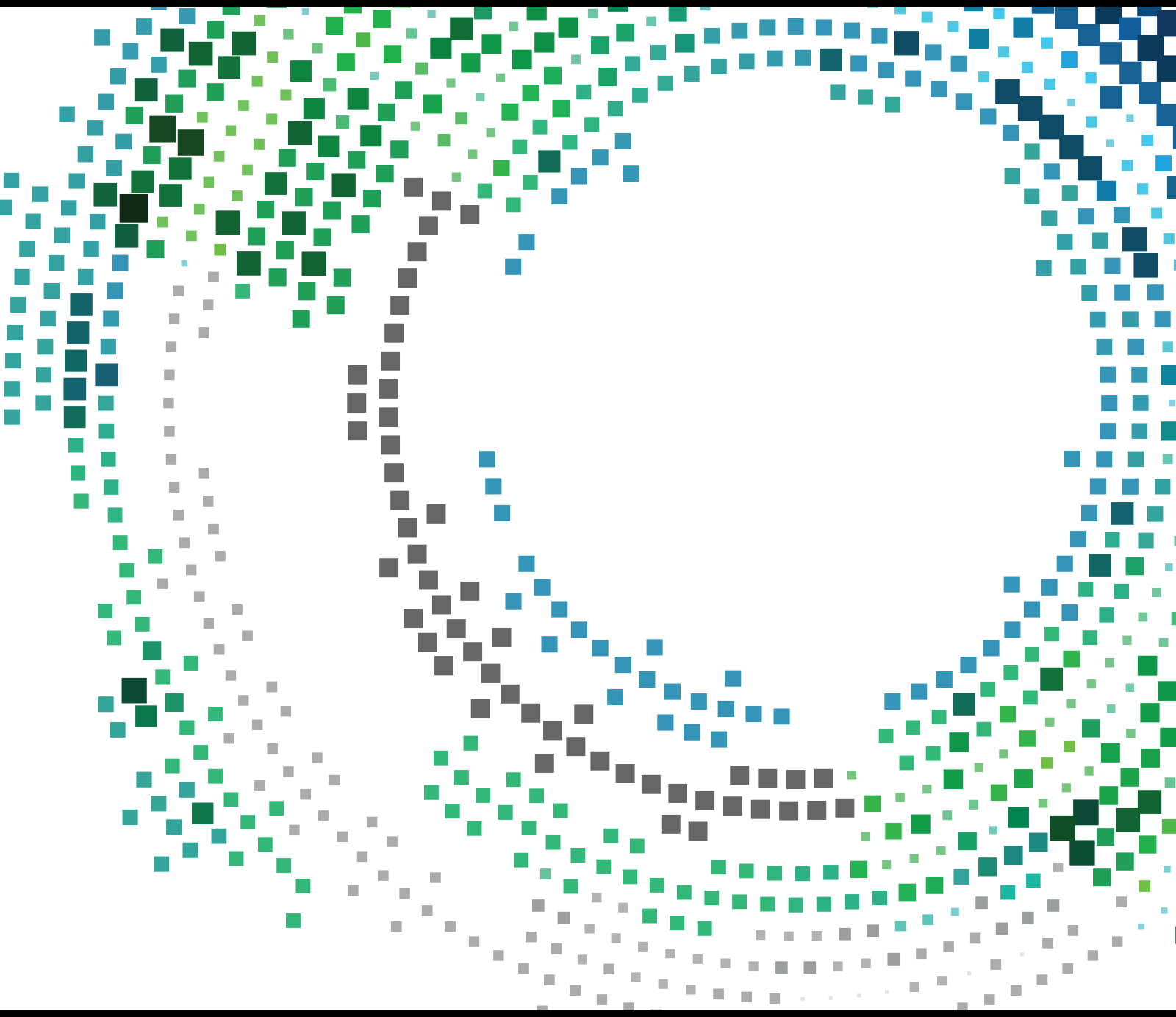
Mobile Information Systems

Artificial Intelligence-Based Traffic Delivery for Mobile Networks

Lead Guest Editor: Jianhui Lv
Guest Editors: Hui Cheng and Qing Li



Artificial Intelligence-Based Traffic Delivery for Mobile Networks

Mobile Information Systems

Artificial Intelligence-Based Traffic Delivery for Mobile Networks

Lead Guest Editor: Jianhui Lv
Guest Editors: Hui Cheng and Qing Li

Contents

Hindawi
Mobile Information Systems
Volume 2022, Article ID 1169035, 13 pages
https://doi.org/10.1155/2022/1169035

Research Article

An Efficient Method for Online Detection of DRDoS Attacks on UDP-Based Services in SDN Using Machine Learning Algorithms

Mitra Akbari Kohnehshahri, Reza Mohammadi, Hatam Abdoli, and Mohammad Nassiri

Computer Department, Engineering Faculty, Bu-Ali Sina University, Hamedan 6517838695, Iran

Correspondence should be addressed to Hatam Abdoli; abdoli@basu.ac.ir

Received 4 February 2022; Revised 5 April 2022; Accepted 6 May 2022; Published 26 May 2022

Academic Editor: Jianhui Lv

With advances in mobile devices and systems and the emergence of new ideas such as cloud computing and big data, as well as the tremendous growth in the number of network users, the need to modify the current network architectures has been very much in the foreground in recent years. One of the promising solutions to overcome these challenges is software-defined networking (SDN). SDN is a unique innovative architecture in which network control and traffic flows are independent of each other and planned directly. The SDN's focused view of networks is more comprehensive than other methods, which is why SDN is more efficient in coping with malicious attacks including amplification attacks. The response to amplification of distributed denial of service (DDoS) attacks is larger than the request. In an amplification attack, the attacker fakes the victim's address as the source address and the responses are forwarded to the victim instead of the attacker. This is why these attacks are more difficult to discover in traditional networks, while the focused method of SDN can contribute to the detection of such attacks. There are different methods for detecting these attacks, one of which is to use machine learning (ML) algorithms. In line with this, the present paper is aimed at the detection of distributed reflection denial of service (DRDoS) attacks using ML algorithms. Simulation was performed by the use of ML algorithms, and the findings suggest a significant improvement in the detection of DRDoS attacks in comparison with previous methods.

1. Introduction

Growing innovation in network applications and reduction in the costs of network operators has resulted in the idea of software-defined networking (SDN). In this model, the network becomes more intelligent and controllable, which facilitates the innovation and management of the network. The main difference with conventional networks is the separation of the control plane from the network devices. Separating these two planes, a centralized control becomes possible, which offers a comprehensive view of the network. Such a view of SDN enables network administrators to control and manage a large number of network devices, network topology, security policies, and routing automatically and dynamically by means of high-level languages [1]. The advantages of separating data and control units in SDN are as follows.

Centralized control allows for integrated management and control by a central controller. Central control is a logical concept, and the control unit can be implemented in a distributed manner [2, 3]. Network devices are managed and controlled in an integrated and centralized manner, which can increase the productivity of network resources [4].

Dynamicity of the configuration helps to easily add new features to the network infrastructure without updating all network devices [2, 3].

Due to the integrated network management in SDN, the network performance is optimized. By using information about the current status of the existing resources, we can utilize methods that make optimum use of network resources in order to improve network efficiency [5].

Workload is balanced. The network controller can be set in a way that, based on the current status of the network, it

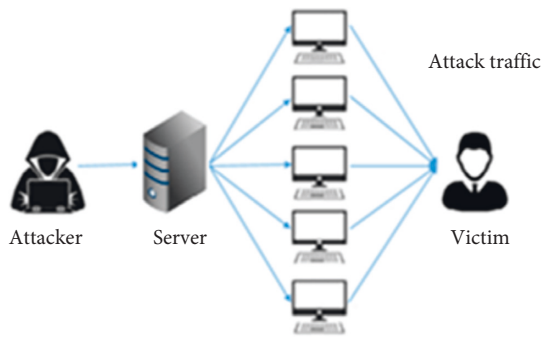


Figure 1: A DDoS attack.

makes the data unit forward the traffic in a balanced manner to prevent congestion on the network and make optimum use of the resources [2, 3].

The required parameters are provided. The users of a network may need certain parameters of quality of service based on the traffic they produce. Thus, to reach an acceptable level of productivity among users, we can develop a program that performs this operation on the network [6].

Error tolerance is increased. By monitoring the network equipment, the controller can detect errors and determine alternative routes if needed. Furthermore, by calculating and announcing alternative routes, the traffic can be instantly forwarded to these routes in the event of error occurrence [2, 3].

Another advantage is that the complexity is decreased. As SDNs are designed at the level of software and do not depend on hardware manufacturers, innovative ideas can be easily implemented in the network [4]. With the increasing growth in networks, security has turned into a major challenge for organizations, governments, and important persons. One type of security threat is denial of service (DoS) attacks which are operated through a single source. The aim is to block the access of authorized users to a certain network resource or make a victim unavailable. Thus, the server can no longer respond to the authorized users. Performing successful DoS attacks on today's powerful systems is not possible through a single system. Moreover, an attack from multiple sources is far more difficult to trace than an attack from a single system. In distributed denial of service (DDoS) attacks, the attacker floods the target with millions of requests per second, which renders the host incapable of offering services to the authorized users. The attackers usually exploit the existing vulnerabilities to attack the target. When the favorable conditions for the attack exist, the attacker can perform a large integrated attack on a target. Ordinary defense mechanisms can easily neutralize single-source attacks. One of the aims of the research in the field of detecting DDoS intrusions has been to devise a comprehensive defense mechanism against such attacks.

DDoS attacks are one of the main security challenges which constitute the largest part of all security threats. Figure 1 illustrates the general structure of a DDoS attack. A DDoS attack attempts to disrupt the network services for various purposes. It uses public services as reflectors in order to increase network traffic [7, 8]. As mentioned regarding distributed reflection denial of service (DRDoS) attacks, the attackers use amplification attacks to intensify the effects of the attack. Therefore, the notion of amplification factor (AF) was introduced to measure the effect of DRDoS attacks. AF is divided into two types: packet amplification factor (PAF) and bandwidth amplification factor (BAF) which are defined by equations (1) and (2), respectively.

$$\mathrm{BAF} = \frac{\mathrm{Len\,(udp\ payload)\ reflector\ to\ victim}}{\mathrm{Len\,(udp\ payload)attacker\ to\ reflector}}, \tag{1}$$

$$\mathrm{PAF} = \frac{\mathrm{Number\ of\ packet\ \ reflector\ to\ victim}}{\mathrm{Number\ of\ packet\ attacker\ to\ reflector}}. \tag{2}$$

BAF is the volume in bytes of the data sent from the amplifier to the victim divided by the volume into bytes of the data sent from the attacker to the amplifier. PAF is the number of packets sent from the amplifier to the attacker divided by the number of packets sent from the attacker to the amplifier. Based on the protocol used, amplification DDoS attacks can be classified as TCP-based or UDP-based attacks. The AF of some of these attacks is listed in Table 1 [9].

As can be seen in Table 1, there are several protocols that cause high values of BAF in UDP-based attacks. For this reason, UDP is the main protocol of DRDoS attacks. As shown in Figure 2, the response produced by these attacks has a greater size than the request. This type of attack fakes the source IP and makes it difficult to filter troublesome packets [7]. The attacker uses UDP-based protocols to disrupt the networks and servers related to network services. In this type of attack, a server which is used as a reflector produces a large amount of data in response to the requests. This server hides the attacker's identity from the victim [10, 11].

These attacks can be detected using SDN technology which adopts a comprehensive, integrated method to networks.

Machine learning (ML) techniques are widely used in intrusion detection systems. Thus, these techniques are effectively used in the detection of DDoS attacks in SDN. In general, ML techniques distinguish between normal and destructive traffic flows according to certain features of the traffic [12].

Table 1: The amplification factor for UDP-based protocol [10].

Protocol	BAF	PAF	Scenario	Description	Port (s)
SNMP v2	6.3	1.00	GetBulk request	Monitoring network-attached devices	161
NTP	556.9	3.84	Request client statistics	Time synchronization	123
DNS	54.6	2.08	ANY lookup at author	Domain name resolution	53
NetBIOS	3.8	1.00	Name resolution	Name service protocol of NetBIOS API	137
SSDP	30.8	9.92	SEARCH request	Discovery of UPnP-enabled hosts	1900
CharGen	358.8	1.00	Character generation request	Legacy character generation protocol	19
QOTD	140.3	1.00	Quote request	Legacy "quote-of-the-day" protocol	17
BitTorrent	3.8	1.58	File search	BitTorrent's Kademlia DHT impl.	Any
Kad	16.3	1.00	Peer list exchange	eMule's Kademlia DHT impl.	Any
Quake 3	63.9	1.01	Server info exchange	Games using the Quake 3 engine	27960
Steam	5.5	1.12	Server info exchange	Games using the steam protocol	27015
ZAv2	36.0	1.02	Peer list and cmd exchange	P2P-based rootkit	164XY
Sality	37.3	1.00	URL list exchange	P2P-based malware dropper	Any
Gameover	45.4	5.39	Peer and proxy exchange	P2P-based banking Trojan	Any

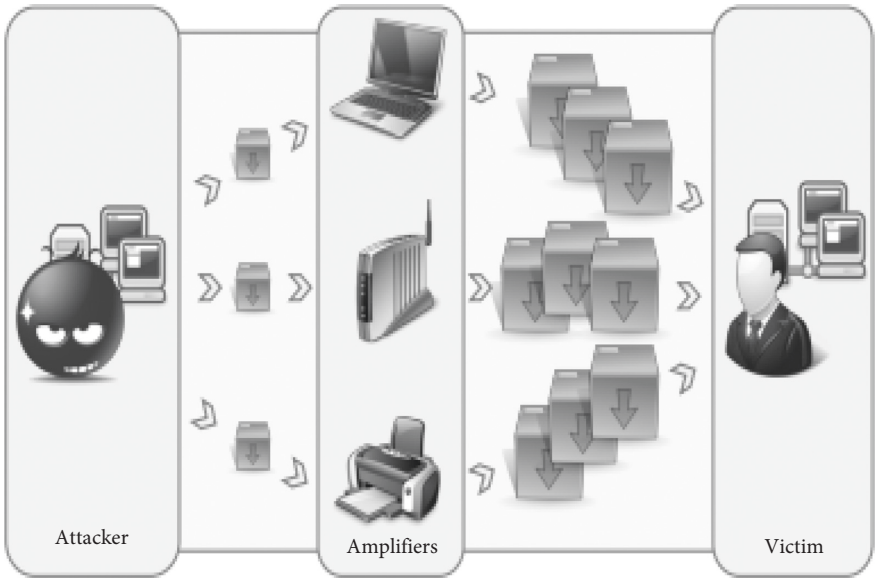


Figure 2: A DRDoS attack [10].

ML methods use techniques for detecting anomalies in a network which may be based on models, statistical and mathematical computations, unsupervised ML, or supervised ML. In fact, any system designed for the detection of network anomalies does this task by collecting the traffic and extracting some kind of information from it. The ML method tries to make a distinction between normal and abnormal traffic patterns [13]. Characteristic of ML techniques is that data are easily available and progress can be seen quite early. They are also useful for solving problems in the functionality and management of the network. The ML techniques for solving fundamental problems in the network include traffic prediction, routing and classification, density control, resource and error management, and network security [14]. As a whole, ML can improve the network's performance through experience. The aim of ML is to design computer systems that learn by experience and are able to adapt to the surrounding environment [15]. The basic procedures of all ML systems are similar. First, the algorithm is provided with training data. The algorithm is actually responsible for learning and looking for different patterns in the data. After finding the patterns, the algorithm devises a model that can be stored in the memory. Afterward, the system can use the models to predict the behavior [16]. As shown in Figure 3, ML algorithms can be classified as supervised learning, unsupervised learning, and reinforcement learning.

To this end, the present paper is aimed at online detection of amplification DDoS attacks in an SDN using ML algorithms.

The paper is organized as follows. In the first section, we shall explain some basic concepts including SDN, attacks, DRDoS attacks, and ML algorithms. The second section reviews the previous works regarding the attacks performed

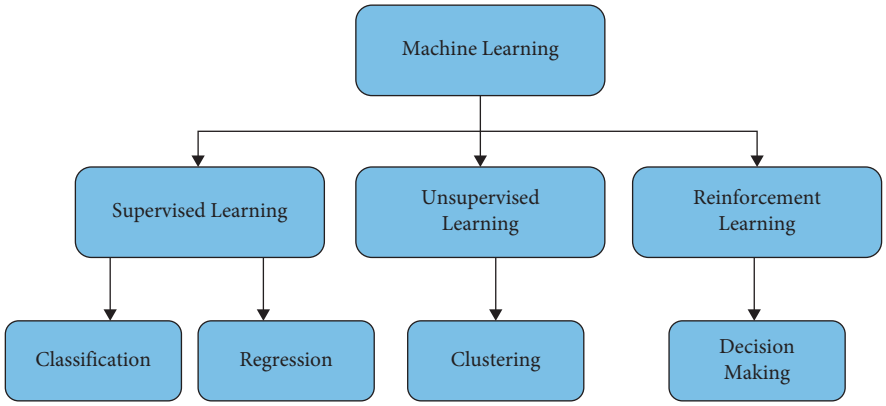


Figure 3: Typology of machine learning.

on SDN. Section 3 introduces the proposed method for online detection of amplification DDoS attacks using ML algorithms. Finally, Section 4 describes the implementation of the proposed method and compares its model with the other methods. The final section makes some conclusions.

2. Related Works

The emergence of SDN has caused novel ideas in this field. After the introduction of SDN, researchers began to investigate the capabilities of these networks. Detection of attacks was one of the key topics that received much attention.

As mentioned in Section 1, DDoS attacks are among the major challenges and vulnerabilities that this young network architecture is facing. Amplification attacks are one of the most serious and frequent DDoS attacks. To obtain an in-depth knowledge of the problem, we reviewed the previously proposed method for defending against DDoS attacks, particularly amplification attacks. Previous works in this field have studied some types of attack through the lens of ML algorithms. We will compare them based on whether they were conducted in the context of traditional networks or SDN.

Santos et al. [13] evaluated four different ML algorithms with the aim of examining the precision and speed of classification of attacks and normal traffic in the vulnerable part of SDNs including controller, flow table, and the bandwidth between switch and controller. The algorithms used in this study include support vector machine (SVM), neural network (NN), decision tree (DTree), and random forest (RF). The emulator used was Mininet, and the controller was POX. In these emulations, the RF algorithm was the most precise ML algorithm in terms of their focus on the vulnerable points of the SDN, followed with a small difference by the DTree algorithm. The shortest processing time belongs to the DTree algorithm. Rahman et al. [17] compared four ML algorithms, namely, J48, RF, SVM, and K-nearest neighbor (K-NN). J48 had the performance in terms of training time and testing time.

The paper used hping3 tool in Python to produce normal traffic and simulate ICMP and TCP flood DDoS attacks. Chen et al. [18] studied the detection of amplification DDoS attacks using ML algorithms in SDN. This paper examined the precision of detection of the SVM algorithm in different time intervals.

Gao et al. [9] conducted research into the detection of amplification DDoS attacks in traditional networks. This study focused on DNS protocols. It sought to detect amplification DDoS attacks using PAF and BAF values. Filho et al. [19] addressed the online detection of DoS and DDoS attacks. They evaluated different datasets and finally proposed CIC-DoS, CICIDS2017, and CSE–CIC–IDS2018 customized datasets. The datasets of different attacks were generated using several tools, i.e., hping3, GoldenEye, hulk, and slowhttptest. This study obtained different precision values with the ML algorithms including RF, DTree, logistic regression (LR), stochastic gradient descent (SGD), perceptron, and AdaBoost with different numbers of features.

Another study by Nanda et al. [20] addressed four ML algorithms, namely, BayesNet, C4.5, NaiveBayes, and DTree. The average prediction precision for BayesNet was found to be 91.68. The aim of this paper was to evaluate the precision and speed of processing. Irom Meite et al. [21] sought to detect amplification DNS attacks in traditional networks using four ML algorithms including DTree, multilayer perceptron (MLP), Naïve Bayes, and SVM. The detection precision of the DTree algorithm was greater than the other algorithms.

3. The Proposed Method

As can be seen in Table 2, the mentioned methods do not provide all the requirements for efficient online detection of amplification DDoS attacks in SDN. For this reason, we decided to incorporate all these items into our proposed method. The importance of amplification DDoS attacks was described above. Our methodology is aimed at online detection of amplification DNS attacks. For implementation, we prefer SDN due to their integrated perspective which consists of data plane and control plane. In the data unit, we

Table 2: A summary of the previous studies.

Source	Online	Offline	SDN-based	DNS amplification
[13]	✗	✓	✓	✗
[17]	✗	✓	✓	✗
[18]	✗	✓	✓	✓
[9]	✗	✓	✗	✓
[19]	✓	✗	✗	✗
[20]	✗	✓	✓	✗
[21]	✗	✓	✗	✓
Proposed method	✓	✗	✓	✓

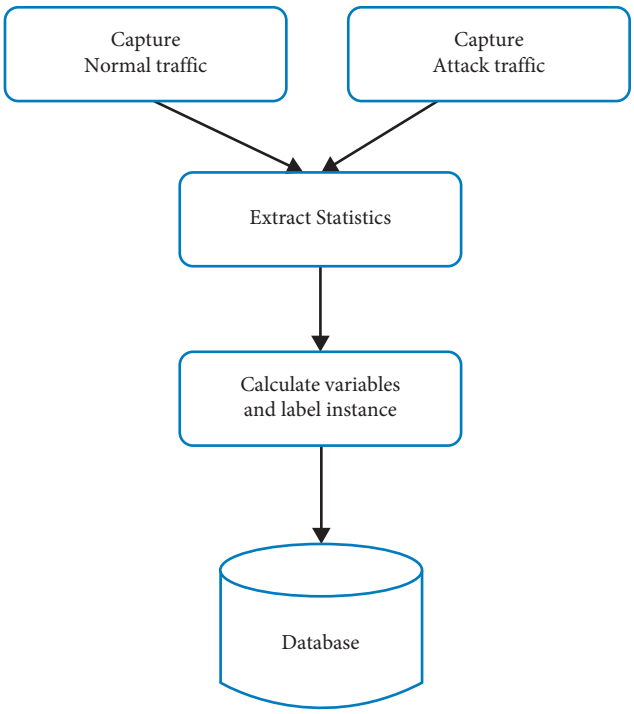


Figure 4: Combination of normal and malicious DNS traffic.

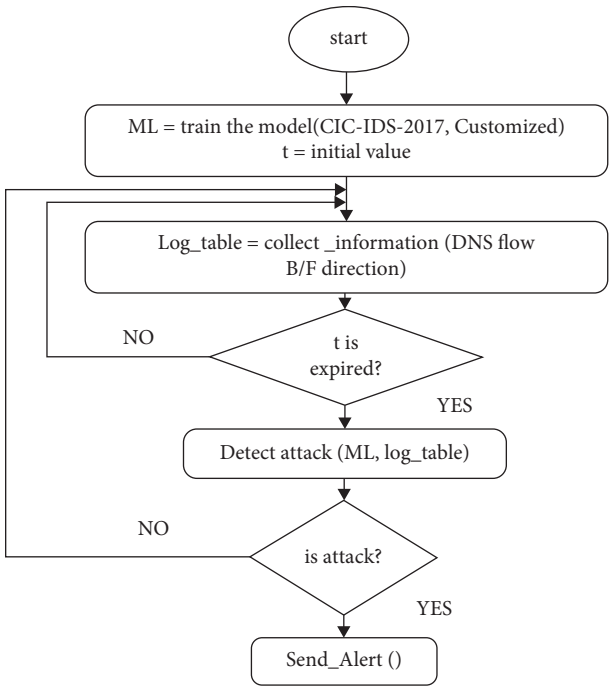


Figure 5: Flowchart of the proposed method.

assume a topology for generating online traffic. The network requires a controller which must be managed by the network on the data plane. On the control plane of an SDN, statistical information about the network links is collected by switches and transmitted to the controller.

In normal DNS traffic conditions, the volume of responses is more than the volume of requests, which is opposed to the attacking conditions where the volume of requests multiplies. Next, the appropriate ML algorithms are selected. The algorithms are trained based on the existing dataset, and then, the model is developed. Finally, online traffic is produced and the controller collects the information about the traffic in intervals of 20 seconds. The ML algorithms will make use of the developed model to distinguish malicious traffic from normal traffic.

As mentioned earlier, the aim of this paper is to detect amplification DNS attacks with the help of the controller's integrated perspective as well as ML algorithms which are modeled based on several attributes that have been extracted from the dataset according to their importance. Our proposed method makes use of five ML algorithms including DTree, RF, SVM, gradient boosting classifier (GBC), and AdaBoost. In this study, the proposed dataset for the detection of DDoS attacks was CICIDS2017 [22] for normal DNS traffic. For malicious traffic, amplification DNS attacks were produced using Scapy tool [23]. CIS-IDS2017 consists of several sets of different types of attack traffic and normal traffic. As can be seen in Figure 4, due to the existence of normal DNS traffic in this dataset, we only extracted the normal traffic from this dataset and combined it with the generated malicious traffic.

The majority of amplification DNS attacks can be detected through the volume and number of response packets. These two attributes were extracted using different algorithms. Figure 5 shows how the proposed method works. First, the ML algorithms are trained in an offline manner using CIS-IDS-2017 as well as the generated malicious dataset. In the determined timespan, statistical information about the DNS response and request flows is collected. When the time ends, attacks can be detected via the trained

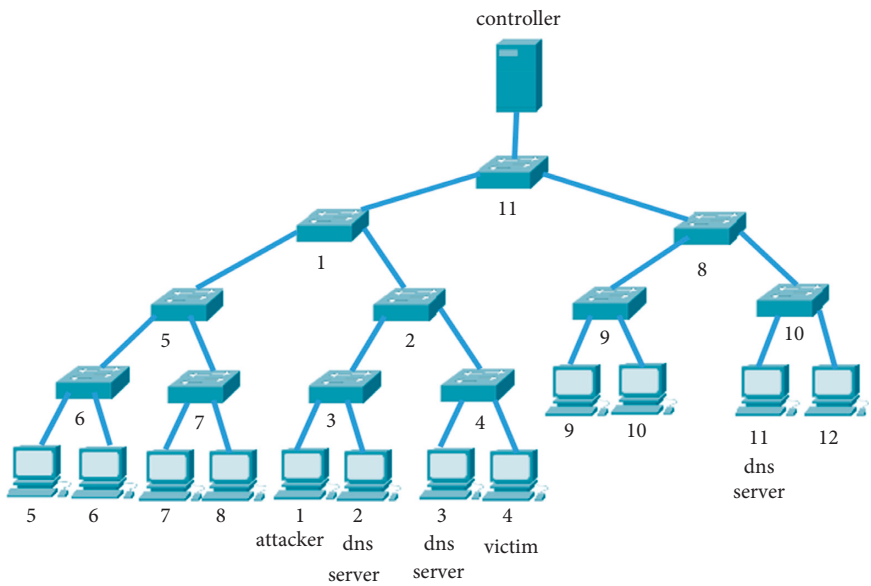


Figure 6: Topology of the online attack traffic.

Table 3: Simulation settings.

Environment	SDN
Emulator	Mininet
Controller	Ryu
ML algorithms	AdaBoost, DTree, RF, SVM, GBC
Number of attributes selected from the dataset	12
Number of switches in the topology	11
Traffic type	Normal DNS traffic, malicious DNS traffic, and iperf traffic
Intervals of preparing reports on the network	20 seconds
Type of attack detection	Online (simultaneously with traffic generation)

ML algorithms and the statistical information about the flows. On detecting an attack, an alert is declared and the system returns to the information collection step.

4. Performance Evaluation

We implemented the control unit using Ryu controller [24], which is a tool written in Python and is well known among academic researchers. The data unit was simulated using Mininet [25].

As shown in Figure 6, the first step in designing the simulation scenario is to consider a system composed of 11 switches and 12 hosts. To investigate the performance of the proposed method in different conditions, we consider three different modes for the number of attackers and victims in the topology. With these different conditions, we can do more experiments and examine more results. In the first scenario, the topology consists of an attacker and one victim, in the second scenario the topology consists of two attackers and two victims, and in the last scenario the topology consists of three attackers and two victims. In all three scenarios, the number of DNS servers is three. The traffic generated in the topology includes amplification DNS attack, normal DNS traffic, and iperf traffic.

Table 4: Selected features in the proposed method. The average importance of each attribute in the proposed method as obtained by the learning algorithms is presented in Figure 7.

Number	Features
1	Total requested packets
2	Total response packets
3	Total volume of requested packets
4	Total volume of response packets
5	Maximum volume of response packet
6	Minimum volume of response packet
7	Average volume of response packets
8	Standard deviation of the volume of response packets
9	Minimum packet volume
10	Maximum packet volume
11	Average packet volume
12	Standard deviation of the volume of packets

Table 3 provides a summary of the implementation conditions of the proposed method. In the first scenario, Host 1 is the attacker and Hosts 2, 3, and 11 are DNS servers. Host 4 is the victim.

As aforementioned, on the control plane of an SDN, statistical information about the network links is collected by switches and transmitted to the controller. In the proposed

method, this information includes source IP address, destination IP address, source port address, destination port address, the number of packets in a flow, and the volume of packets in a flow. As mentioned above, one of the most important criteria for the detection of amplification DDoS attacks is difference in the size and number of packets in request and response flows. To allow for a correct detection, therefore, we have separated the request and response flows of normal and malicious traffic.

To create an amplification DDoS attack, Scapy is used for creating an amplification DNS attack.

After creating the dataset, we reduce the number of attributes down to 12 by examining the existing attributes and evaluating their importance through the ML algorithms. The final attributes are listed in Table 4.

As shown in Figure 8, the results were evaluated using an ROC curve. This graph was used to specify which ML algorithms had provided a correct classification for the dataset of the proposed method. The horizontal axis of the curve represents false-positive rate (FPR), and the vertical axis shows true-positive rate (TPR). The algorithm with the largest area under the curve has the best classification capability.

As can be seen in Figure 8, in the proposed method with the conditions of the first scenario, the SVM and RF algorithms had a better performance than the other algorithms, while the decision tree algorithm was found to have the weakest classification capability. We shall now compare our results with the two papers discussed in the review of the literature section. The first paper is [18]. The implementation method of this paper was offline, and it was implemented in an SDN. It was implemented solely by a support vector algorithm, and its results are referred to as Chen in the graphs. The second study is [9], which was implemented offline in traditional networks and whose results are specified as Gao in the graphs. Both papers are aimed at the detection of amplification DDoS attacks. The first criterion for comparing the three methods is accuracy. Figure 9 compares the accuracy level of the three methods. The horizontal axis shows the algorithms, and the vertical axis shows the accuracy level ranging from 0 to 1. Accuracy can be calculated via equation (3). True positive (TP) denotes the traffic which is an attack and has been detected correctly; true negative (TN) denotes traffic which is normal and has been detected correctly; false positive (FP) denotes normal traffic which has been wrongly specified as malicious; false negative (FN) denotes malicious traffic which has been wrongly specified as normal.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{3}$$

As can be seen from the graph, the proposed method with the conditions of the first scenario has the highest value for the GBC and SVM algorithms. Overall, it is more accurate than Chen and Gao methods in all algorithms. The next graph compares recall rates. As in the accuracy graph, the horizontal axis represents the ML algorithms, and the vertical axis shows the recall rate. Recall rate can be calculated as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{4}$$

As can be seen in Figure 10, the recall rate of the proposed approach with the conditions of the first scenario and the Chen method is the highest for the GBC, RF, DTree, and SVM algorithms, whereas the Yuxuan method has the lowest recall rate for the GBC algorithm. The next criterion for comparison is precision. It refers to the quantitative ratio of the items correctly classified by the ML algorithms to the total number of classified items (whether correctly or incorrectly) by the algorithm. It is calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{5}$$

As in the previous graphs, the horizontal axis represents the ML algorithms and the vertical axis shows precision. Figure 11 compares the precision level of the three methods.

As can be seen in Figure 11, the highest precision in the proposed method with the conditions of the first scenario belongs to the GBC and SVM algorithms and the highest precision in the Chen method belongs to the GBC and RF algorithm. The lowest precision belongs to the DTree algorithm in the Yuxuan method. The next comparison

Table 5: FPR values.

Algorithm	Proposed method	Chen	Gao
GBC	0	0.0625	0
RF	0.02	0.05	0
DTree	0.06	0.08	0.12
SVM	0	0.02	0.05
AdaBoost	0.08	0.06	0.02

Table 6: TNR values (first scenario).

Algorithm	Proposed method	Chen	Gao
GBC	1	0.937	1
RF	0.977	0.944	1
DTree	0.937	0.914	0.872
SVM	1	0.979	0.942
AdaBoost	0.916	0.936	0.979

Table 7: FNR values (first scenario).

Algorithm	Proposed method	Chen	Gao
GBC	0	0	0.03
RF	0	0	0.16
DTree	0	0	0.06
SVM	0	0	0
AdaBoost	0	0.06	0.16

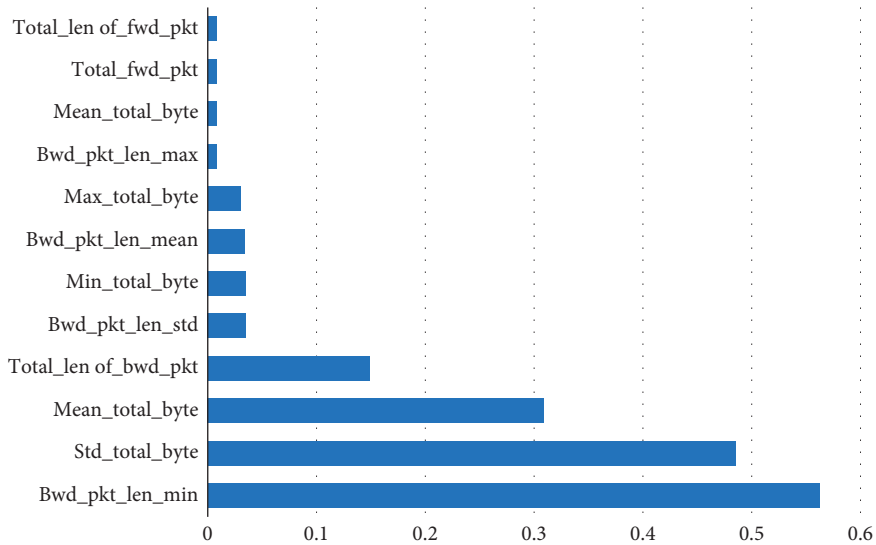


Figure 7: The importance of the selected features as obtained by the ML algorithms.

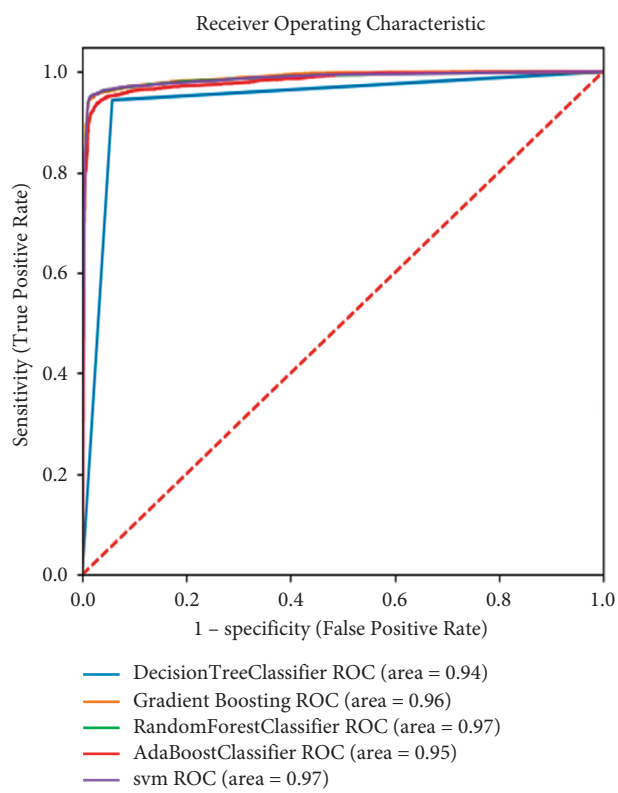


Figure 8: ROC curve of different algorithms.

criterion is F-score. On this graph, the horizontal axis lists the ML algorithms and the vertical axis shows the obtained F-scores. Figure 12 compares the F-scores of the three methods, which can be calculated via equation (6). This equation combines equations (4) and (5).

$$\text{F} - \text{score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{6}$$

As shown in Figure 12, all ML algorithms in the proposed method with the conditions of the first scenario have gained better F-scores than the other two methods. False-positive rate (FPR) is obtained by dividing the number of normal items that were detected as malicious by the total number of normal items that were detected as malicious and normal items that were detected correctly as normal. Equation (7) calculates the value of FPR.

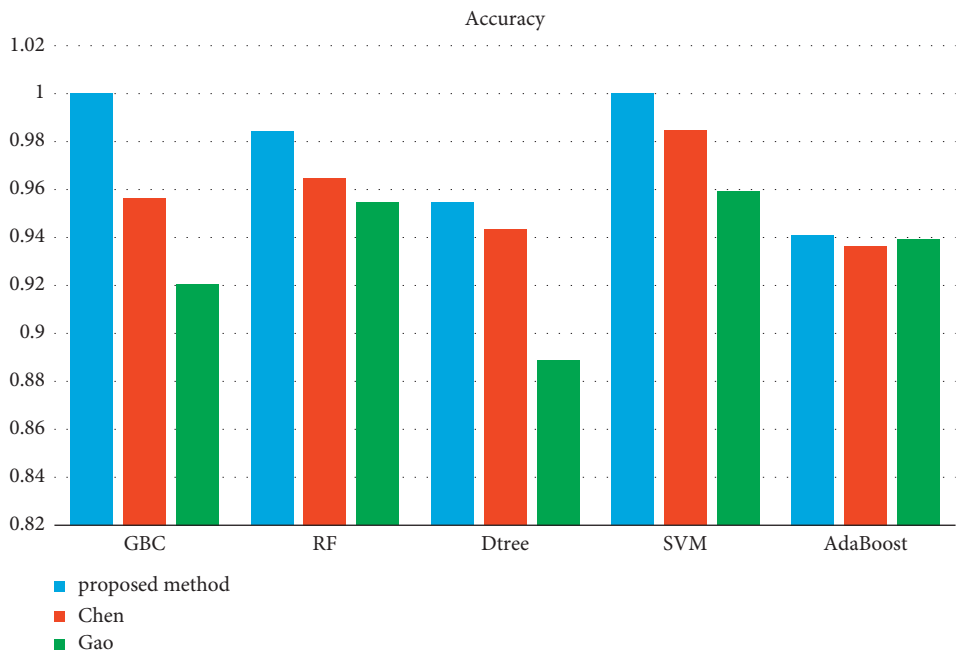


Figure 9: Accuracy for different ML algorithms (first scenario).

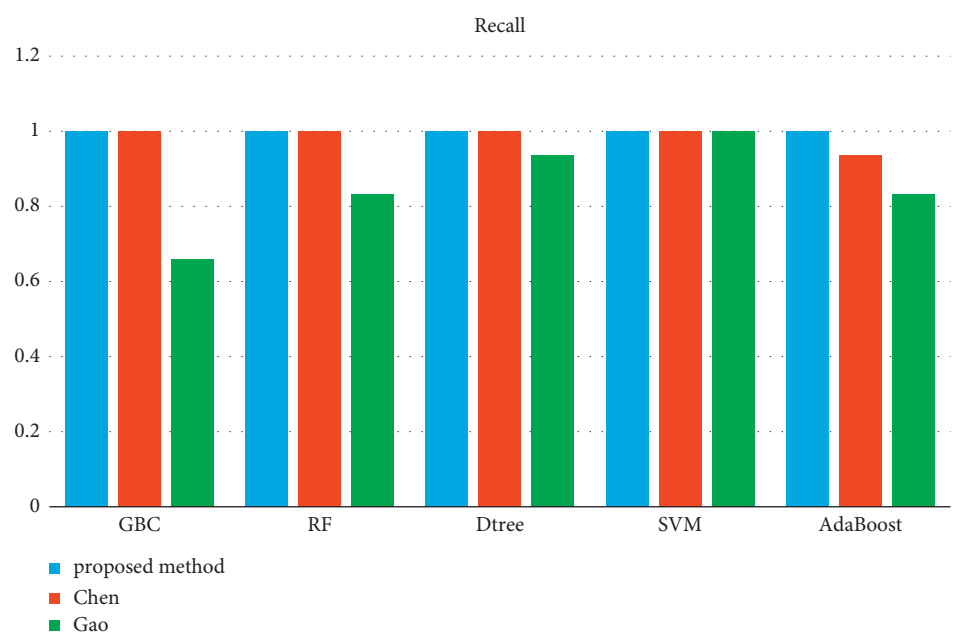


Figure 10: Recall for different ML algorithms (first scenario).

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}. \tag{7}$$

As shown in Table 5, the FPR of the proposed method with the conditions of the first scenario is less than that of the other methods. This value is smallest for the GBC and SVM algorithms in the proposed method. True-negative rate (TNR), which is also known as sensitivity, is obtained by dividing the number of normal items that were detected correctly by the total number of normal items that were

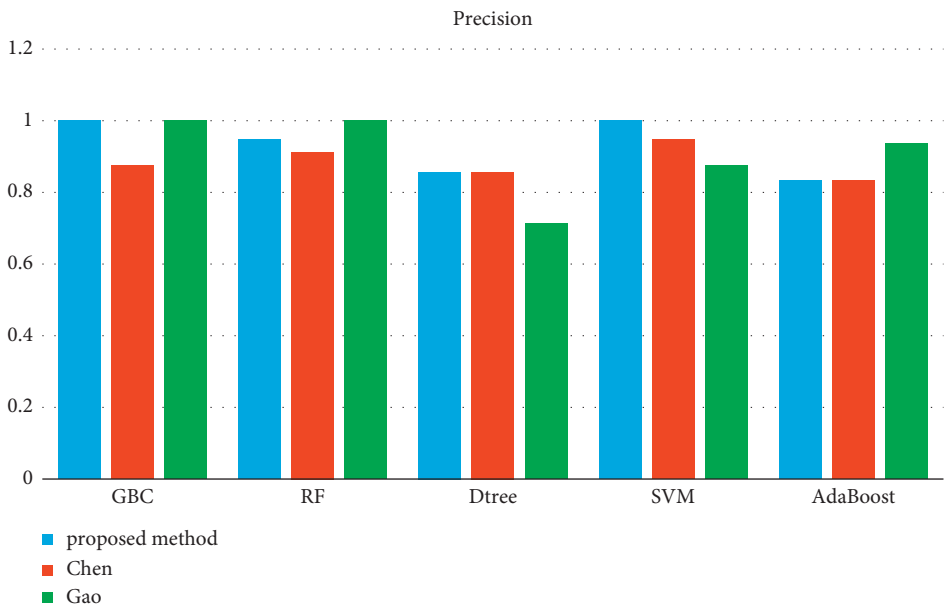


Figure 11: Precision for different ML algorithms (first scenario).

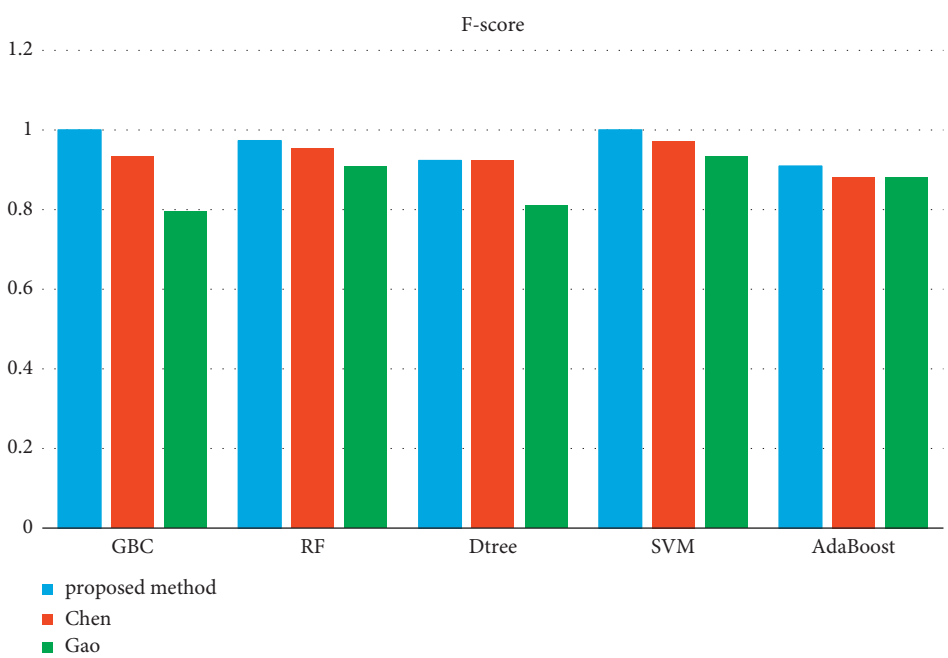


Figure 12: F-score for different ML algorithms (first scenario).

detected correctly and the normal items that were detected as malicious. The following equation is used to calculate TNR as follows:

$$\mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{FP} + \mathrm{TN}}. \tag{8}$$

As it can be seen in Table 6, the TNR of most of the ML algorithms in the proposed method with the conditions of the first scenario is greater than in the other methods. False-negative rate (FNR) refers to the ratio of the number of items that were wrongly detected as normal to the total number of items that were wrongly detected as normal and malicious items that were detected as normal. This value can be calculated by

$$\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{FN} + \mathrm{TP}}. \tag{9}$$

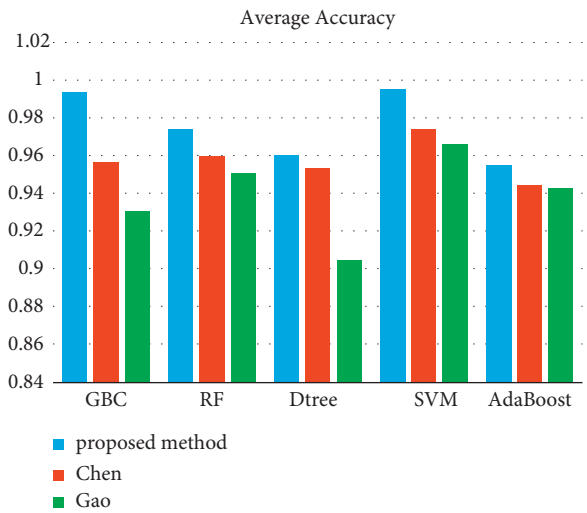


Figure 13: Average accuracy of three scenarios.

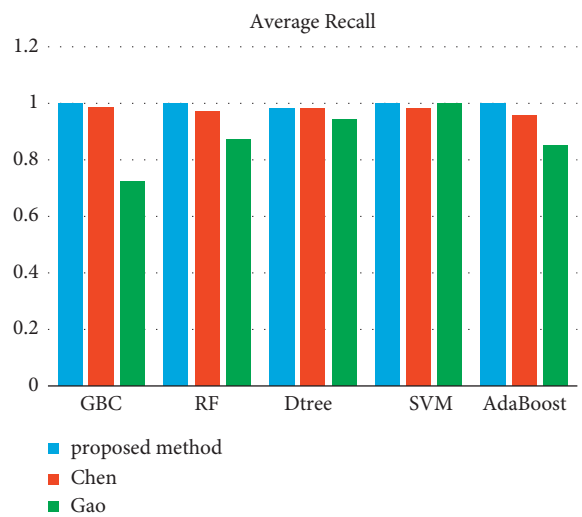


Figure 14: Average recall of three scenarios.

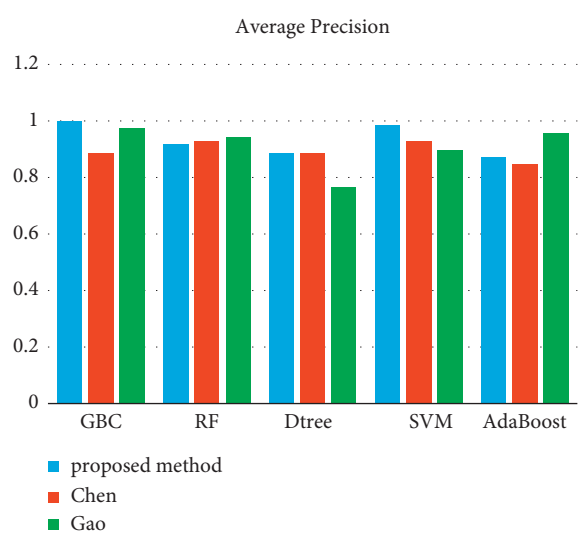


Figure 15: Average F-score of three scenarios.

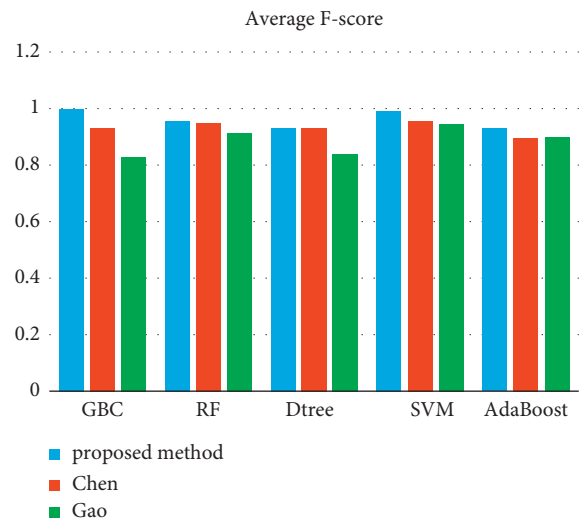


Figure 16: Average F-score of three scenarios.

Table 8: Average of FPR values.

Algorithm	Proposed method	Chen	Gao
GBC	0.008	0.055	0.007
RF	0.053	0.044	0.02
DTree	0.047	0.057	0.104
SVM	0.006	0.028	0.044
AdaBoost	0.063	0.057	0.014

Table 9: Average of TNR values.

Algorithm	Proposed method	Chen	Gao
GBC	0.991	0.944	0.992
RF	0.963	0.953	0.98
DTree	0.951	0.938	0.12
SVM	0.993	0.971	0.952
AdaBoost	0.950	0.940	0.984

Table 10: Average of FNR values.

Algorithm	Proposed method	Chen	Gao
GBC	0	0.015	0.17
RF	0	0.028	0.125
DTree	0.018	0.017	0.076
SVM	0	0.017	0.04
AdaBoost	0	0.04	0.014

As shown in Table 7, all ML algorithms used in the proposed method with the conditions of the first scenario have the lowest FNR.

We also tested the criteria obtained in the first scenario for the second and third scenarios. The results of the second and the third scenarios were close to the first scenario test. The average case results of all three scenarios are displayed in the form of graphs and tables.

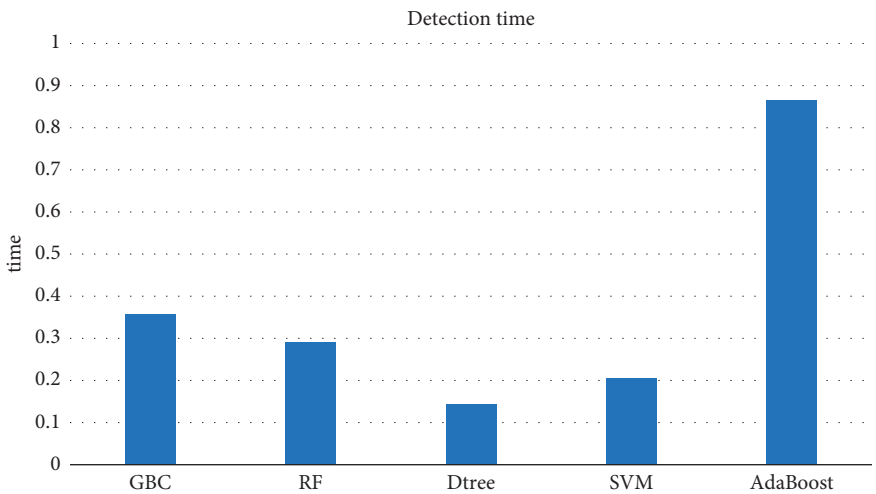


Figure 17: Speed of detection for different ML algorithms.

According to the graphs, the averages obtained in Figures 13–16 are close to the values of the first scenario. In the arrangements made, the values obtained for the first scenario, the second scenario, and the third scenario are approximately equal. According to the mean graphs, it can be concluded that the proposed method has worked better in all three scenarios comparing to the Chen and Gao methods. We also obtained the average of FPR, TPR, and TNR criteria in all three scenarios which can be seen in Tables 8–10.

Considering the values of Tables 8–10, it can be concluded that the proposed method outperforms the Chen and Gao methods.

The speed of processing refers to the classification time required by each ML algorithm to distinguish malicious attacks from normal traffic. On this graph, the horizontal axis shows the ML algorithms and the vertical axis shows the processing time required by each of the algorithms. In fact, the graph in Figure 13 shows the average processing time required for the detection of attacks.

As can be seen in Figure 17, the DTree algorithm has a higher processing speed than the other algorithms. AdaBoost was found to be the slowest algorithm in distinguishing malicious attacks from normal traffic.

5. Conclusion

In this paper, a new ML method has proposed to detect DRDOS attacks in SDN-based networks. The proposed method is implemented on the controller and leverages the capabilities of SDN architecture to gather useful information about DNS traffic flows. Then, it makes use of ML algorithms to detect the presence of DNS amplification attack. To train the ML algorithms, we have selected the most important features from CIC-IDS-2017 dataset. The performance of the proposed method compared to other state-of-the-art techniques has been conducted in a comprehensive manner. The simulation results confirm that using decision tree algorithm as an ML model in the proposed method achieves the shorter attack detection time, whereas SVM and GBC algorithms outperform other algorithms in terms of accuracy. Furthermore, the results showed that our proposed method is applicable and can be beneficiary in online detection of DNS-based amplification attacks in SDN. In the future work, we aim to use other new machine learning methods such as deep learning, and we can identify other types of DRDoS attacks.

Data Availability

The experimental data consist of public datasets: CIC-IDS-2017 dataset [22] and Scapy data generator [23].

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] Y. Zhang, L. Cui, W. Wang, and Y. Zhang, "A survey on software defined networking with multiple controllers," *Journal of Network and Computer Applications*, vol. 103, pp. 101–118, 2018.

[2] D. Kreutz, F. M. V. Ramos, P. E. Verissimo, C. E. Rothenberg, and S. Azodolmolky, "Software-defined networking: a comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2014.

[3] M. Jammal, T. Singh, A. Shami, R. Asal, and Y. Li, "Software defined networking: state of the art and research challenges," *Computer Networks*, vol. 72, pp. 74–98, 2014.

[4] F. Hu, *Network Innovation through OpenFlow and SDN: Principles and Design*, CRC Press, Florida, USA, 2014.

[5] J. S. H. Miranda, "Fault isolation in software defined networks," MSc thesis, Tecnica Lisboa University, Lisboa, Portugal, 2016.

[6] F. Hu, Q. Hao, and K. Bao, "A survey on software-defined network and openflow: from concept to implementation," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2181–2206, 2014.

[7] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks," *IEEE communications surveys & tutorials*, vol. 15, no. 4, pp. 2046–2069, 2013.

[8] S. Dong, K. Abbas, and R. Jain, "A survey on distributed denial of service (DDoS) attacks in SDN and cloud computing environments," *IEEE Access*, vol. 7, pp. 80813–80828, 2019.

[9] Y. Gao, Y. Feng, J. Kawamoto, and K. Sakurai, "A machine learning based approach for detecting DRDoS attacks and its performance evaluation," in *Proceedings of the 2016 11th Asia Joint Conference on Information Security (AsiaJCIS)*, IEEE, Fukuoka, Japan, Auguest 2016.

[10] C. Rossow, "Amplification hell: revisiting network protocols for DDoS abuse," in *Proceedings of the Network and Distributed System Security Symposium*, NDSS, San Diego, California, USA, February 2014.

[11] F. J. Ryba, M. Orlinski, M. Wahlisch, C. Rossow, and T. C. Schmidt, "Amplification and DRDoS attack defense--a survey and new perspectives," 2015.

[12] N. Z. Bawany, J. A. Shamsi, and K. Salah, "DDoS attack detection and mitigation using SDN: methods, practices, and solutions," *Arabian Journal for Science and Engineering*, vol. 42, no. 2, pp. 425–441, 2017.

[13] R. Santos, D. S. Silva, W. D. E. Santo, and A. L. Ribeiro, "Machine learning algorithms to detect DDoS attacks in SDN," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 16, Article ID e5402, 2020.

[14] P. Amaral, J. Dinis, P. Pinto, L. Bernardo, J. Tavares, and H. S. Mamede, "Machine learning in software defined networks: data collection and traffic classification," in *Proceedings of the 2016 IEEE 24th International Conference on Network Protocols (ICNP)*, November 2016.

[15] J. Xie, F. R. Yu, T. Huang et al., "A survey of machine learning techniques applied to software defined networking (SDN): research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393–430, 2018.

[16] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT press, Cambridge, USA, 2018.

[17] O. Rahman, M. A. G. Quraishi, and C.-H. Lung, "DDoS attacks detection and mitigation in SDN using machine learning," in *Proceedings of the 2019 IEEE World Congress on Services (SERVICES)*, July 2019.

[18] C.-C. Chen, Y. R. Chen, W. C. Lu, S. C. Tsal, and M. C. Yang, "Detecting amplification attacks with software defined networking," in *Proceedings of the 2017 IEEE conference on dependable and secure computing*, Auguest 2017.

[19] F. S. D. Lima Filho, F. A. F. Silveira, A. D. M. Brito Junior, G. V. Solar, and L. F. Silveria, "Smart detection: an online approach for DoS/DDoS attack detection using machine learning," *Security and Communication Networks*, vol. 201915 pages, Article ID 1574749, 2019.

[20] S. Nanda, F. Zafari, C. Decusatis, E. Wedaa, and B. Yang, "Predicting network attack patterns in SDN using machine learning approach," in *Proceedings of the 2016 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, November 2016.

[21] I. L. Meitei, K. J. Singh, and T. De, "Detection of DDoS DNS amplification attack using classification algorithm," in *Proceedings of the International Conference on Informatics and Analytics*, Pondicherry, India, Auguest 2016.

[22] ids-2017@www.unb.ca, 2021, https://www.unb.ca/cic/datasets/ids-2017.html.

[23] Scapy Python Library@www.scapy.net, 2021, https://scapy.net/.

[24] Ryu Software Defined Networking Framework@ryu.readthedocs.io, 2021, https://ryu.readthedocs.io/en/latest/getting_started.html.

[25] Mininet Network Emulator@mininet.org, 2021, http://mininet.org/overview/.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5960976, 15 pages
https://doi.org/10.1155/2022/5960976

Research Article

CloudConsumerism: A Consumer-Centric Ranking Model for Efficient Service Mapping in Cloud

Neeraj,[1] **Neha Garg**,[1] **Nipun R. Navadia**,[2] **Anupam Lakhanpal**,[3] **Indrajeet Gupta**,[1] **Wubshet Ibrahim**,[4] **and Manish Raj**[1]

[1]*School of Computer Science Engineering and Technology, Bennett University, Greater Noida, India*
[2]*Department of Computer Science Engineering and Technology, Dronacharya Group of Institutions, Greater Noida, India*
[3]*Department of Computer Science Engineering and Technology, Galgotias University, Greater Noida, India*
[4]*Department of Mathematics, Ambo University, Ambo, Ethiopia*

Correspondence should be addressed to Wubshet Ibrahim; wubshet.ibrahim@ambou.edu.et

Received 23 February 2022; Revised 25 March 2022; Accepted 5 April 2022; Published 2 May 2022

Academic Editor: Jianhui Lv

In cloud, service providers and consumers are primary stakeholders that maintain a business liaison. Cloud service providers (CSPs) offer the services, and consumer uses the services on a payment basis. From a business perspective, the selection of a service based on mutual evaluation benefits both the CSPs and consumers. This paper presents an efficient CloudConsumerism model where the multicriteria decision-making method (MCDM) method, TOPSIS, is used for evaluating the performance of CSPs and consumers. For performance evaluation of CSPs, the performance attributes defined by Cloud Service Measurement Initiative Consortium (CSMIC) are exploited. For evaluating the consumers, this paper is the first approach towards identifying the behavioral attributes for evaluating the cloud consumers analogous to the business models. A service mapping algorithm is proposed for efficient (less overhead and higher robustness) mapping. Extensive simulation experiments are conducted; the results show that the proposed framework can be used for the online cloud-based platform due to limited overhead and high robustness.

1. Introduction

Cloud computing is a business model [1] where provider-consumer liaison is formed. The consumer goes through the services from the catalog of the multiple cloud service providers (CSPs) and selects the services from the CSP, which can provide higher performance at the lowest rate. The services offered by the CSPs fall in one of three categories, i.e., infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) [2]. Manually selecting a service having multiple attributes offered from different CSPs is very tedious for the consumers. A new entity, broker, is evolved, which selects the services from the different CSPs on behalf of the consumer. The broker organization can be a third party or a part of the participating organizations. Some well-known organizations, namely, AWS Service Broker (https://http://aws.amazon.com/partners/servicebroker/), IBM Multicloud Management Services (https://www.ibm.com/in-en/services/cloud/multicloud), Cloudmore (https://web.cloudmore.com/cloud-broker-cloudmore), Cloud Services Brokerage (https://www.jamcracker.com/cloud-service-brokerage) etc., are playing the role of cloud broker.

The demand for cloud-based services is increasing day by day. COVID-19 disruption has also accelerated the usage of cloud-based services [3]. To fulfill the ever-increasing demand for cloud-based services and generate more revenue, many big IT organizations are playing the role of CSPs, namely, Amazon, Microsoft, Google, IBM, etc. [4]. Due to enormous expansion in the number of CSPs and consumers, their efficient selection while mapping has become complicated. In literature, quality of service (QoS)-based selection of a CSP for a consumer is recommended for the efficient service mapping [5–7]. The selection of a CSP for the

consumer is a unidirectional approach where the selection of the service is based on the CSP's evaluation in the context of the QoS requirements of a consumer. For making the cloud business successful, the bidirectional evaluation of the participating entities (CSPs and consumers) involved in the mapping process is necessary for increasing their satisfaction.

For evaluating the CSPs, Cloud Service Measurement Initiative Consortium (CSMIC) has introduced performance metrics that are combined in the form of Service Measurement Index (SMI) [8]. In contrast, no standard parameters are defined for evaluating the behavior of consumers for the CSPs. Consumer behavior is about the approach of how people buy and use merchandise and services [9]. Consumer behavior helps in identifying whom to target, how to target when to reach them, and what message is to be given to them to reach the target consumers to buy the product. Analyzing consumer behavior helps the CSPs in improving brand equity and boosting sales. For evaluating consumer behavior, some consumer behavior analytics tools such as Mixpanel (https://mixpanel.com/), Google Analytics (https://support.google.com/analytics/answer/7126596?hl=en), and Kissmetrics (https://www.kissmetrics.io/) are exploited. These tools maintain the complete profile of consumers from website usage to social media engagement. In marketing research, RFM (recency, frequency, and monetary) model [10] [11, 12] is also proposed for evaluating consumer behavior.

Consumer evaluation plays an important role to make the business successful. Earlier frameworks are proposed for the evaluation of CSPs and consumers. The TRCSM framework is presented for evaluating the CSPs and consumers during service mapping [7]. The proposed framework uses prepurchase transaction attributes turnover, duration, and transaction for evaluating the cloud consumers. Another framework, MECSM [12], is proposed which uses a standard RFM model for evaluating the consumers parallel to the CSPs. RFM model uses prepurchase behavior attributes, recency, frequency, and monetary attributes for evaluating the consumers. The prepurchase behavior shows their behavior in past transactions. In the cloud, the same service is offered from different geographical regions, and sometimes service is available in some geographical regions, and it is not available in another geographical region. Observe the consumer behavior during the owning of the service (in-purchase behavior) while the change in service, location, price, and quality is also important. To improve the service quality and increase the consumer base, it is also necessary to observe consumer postpurchase behaviors.

Due to the dynamic need of the current IT industry, the demand of IT resources varies frequently [6, 13]. For example, for a large computing task, efficiency is the main deciding factor. For time-critical tasks, distance is a high priority. Further, a server's distance is different from different consumers. Also, a consumer may have multiple demands with different priorities. On the other hand, service providers may also have multiple different criteria for evaluating a consumer, e.g., consumers involved for a long duration, consumers with a high transaction amount, and a consumer with less retention rate. Evaluating an entity based on multiple attributes is a multicriteria decision-making problem (MCDM) [14–17]. The glaringly used MCDM methods in the cloud environment are AHP [18], TOPSIS [19], VIKOR [20], and PROMETHEE II [21]. These MCDM methods can be used for ranking both the provider as well as the consumer. Since cloud computing is an internet-based technology, therefore, the execution time of the MCDM method will greatly impact the service performance. In our previous work [6], the TOPSIS method is suggested in terms of lower execution time and higher robustness.

In this paper, CloudConsumerism model is proposed which evaluates the CSPs by exploiting the QoS attributes defined by CSMIC framework 2.0 [8], and consumer behavior attributes are defined for evaluating the consumers in the cloud environment. The service mapping algorithm is developed for efficient service mapping covering all the mapping scenarios (one to one, one to many, many to one, and many to many). A case study is presented for showing the process of service mapping. The experimental analysis is performed on large-scale synthetic dataset for showing its applicability in online cloud-based service selection.

1.1. Motivation and Contribution. Cloud-based services are geographically distributed to support the service consumer organizations in terms of lower upfront cost and dynamic change of hardware/software needs. The COVID-19 disruption has also accelerated the usage of cloud-based services [3]. According to a Gartner report [22], end-user spending on CSPs is forecast to grow 23.1% in 2021 to total $332.3 billion, up from $270 billion that was in 2020. The extensive adoption of cloud-based services in various emerging domains is imposing new challenges and forcing researchers to think about new strategies. To the best of our knowledge, a few research [7, 12] has been conducted where cloud consumers are analyzed parallel to the CSP. The main aim of this paper is to identify the consumer behavior attributes for evaluating the cloud consumer parallel to the CSPs. After the performance evaluation of both the CSPs and consumers, a ranking-based service mapping algorithm is designed which maps a large-scale CSPs and consumers within the limited time overhead of 10 seconds.

The next section discusses the performance evaluation attributes of cloud consumers. Section 3 presents a system model. In Section 4, a case study is presented to show the process of service mapping and the satisfaction of mapped CSPs, and consumers are also evaluated. The experimental analysis is presented in Section 5. The last section concludes the research paper with some future directions.

2. Consumerism in Cloud

In the cloud, consumers can be an individual or an organization. In this work, small organizations are considered as consumers. Organizational buying behavior is mostly a group process. In an organization, a single person does not typically make a buying decision. The service provider must understand group behavior. This group of people may be

actual buyers, the people who impact directly or indirectly on the buying behavior, and the employees of the organization who are going to use these services. For evaluating the group behavior, the attributes are defined in three categories as presented in Figure 1.

2.1. Prepurchase Behavior. The prepurchase behavior of the consumer demonstrates the previous transactions with the provider organization and the complete history of the consumer buying behavior. The buying behavior of the consumer can be predicted by using the following attributes.

(i) Recency. Recency refers how recently a consumer had made a purchase. It is believed that the more recently a consumer has purchased with the CSP, the more likely he or she will continue to keep the service provider and their services in mind for consequent purchases. Compared with consumers who have not acquired from the CSP for a longer period, the probability of engaging in future transactions with recent consumers is arguably higher. Such information can be used to hark back recent consumers to review the business soon to continue meeting their purchase needs. The tendency of recency attribute is considered as negative (lower value is better)

(ii) Frequency. Frequency refers that within a given time, how many times the business deal has been made. The frequency of a consumer's transactions may be affected by factors such as the type of product, the price point for the purchase, and the need for replenishment or replacement. If the purchase cycle can be predicted, for example, when a customer needs to buy new services, marketing efforts could be directed towards reminding them to visit the business. The tendency of frequency attribute is considered as negative (lower value is better)

(iii) Monetary. Monetary value stems from the lucrativeness of expenditures the consumer makes with the business during their transactions. A natural tendency is to put more prominence on encouraging consumers who spend the most money to continue to do so. While this can produce a better return on investment in marketing and consumer service, it also helps in separating those consumers who have been consistent but have not spent as much with each transaction. The tendency of monetary attribute is considered as positive (higher value is better)

(iv) Duration. Duration shows the time the consumer is engaged with the CSP. According to some theories, a consistent consumer buys the services slowly and returns small profit regularly. The tendency of duration attribute is considered as negative (lower value is better)

(v) Churn rate. It is defined as the percentage rate of consumers who cancel or do not subscribe to a service or consumer cancelation over time, usually on an annual basis. The consumer organization which has a lower churn rate will have a constant or higher demand for services. Also, the probability that a single consumer will cancel during a specific time is less. The tendency of churn rate attribute is considered as negative (lower value is better).

$$churn\ rate = a = \frac{\Delta C_{cancel}}{C \times \Delta t}, \tag{1}$$

where C denotes the number of consumers, Δt is the amount of elapsed time, and ΔC_{cancel} is the number of consumers canceling in time Δt

(vi) Consumer acquisition cost. It is defined as the total marketing cost of acquiring a new consumer. Total marketing expenses contain advertising costs, commissions and bonuses paid, and salaries of marketers and sales managers. The tendency of customer acquisition cost attribute is considered as negative (lower value is better). The consumer acquisition cost is defined as

$$Consumer\ Acquisition\ Cost = \frac{Total\ Marketing\ expenses}{number\ \text{of}\ new\ consumers}. \tag{2}$$

(vii) Retention cost. Consumer retention cost (CRC) is defined as the total cost of retaining a consumer. It is assumed that a 5% increase in retention rate boosts the organization's profitability by 25% to 95%. A strategic organization manager should not only focus on retaining the consumers but also focused on the CRC used for retaining the consumer. The tendency of retention cost attribute is considered as negative (lower value is better)

(viii) Recurring revenue. Recurring revenue is the economy metric of subscription of the service over a fixed time. The consumer organization which has higher recurring revenue will have a higher monetary value. The tendency of recurring revenue attribute is considered as positive (higher value is better). The amount of subscription revenue owed by a consumer over a fixed time, usually measured monthly (MRR), quarterly (QRR), or annually (ARR).

$$\begin{gathered} Recurring\ revenue = RR = \frac{R}{\Delta t}, \\ ARR = 4 \times QRR = 12 \times MRR, \end{gathered} \tag{3}$$

where R denotes the subscription revenue owed during the amount of elapsed time Δt

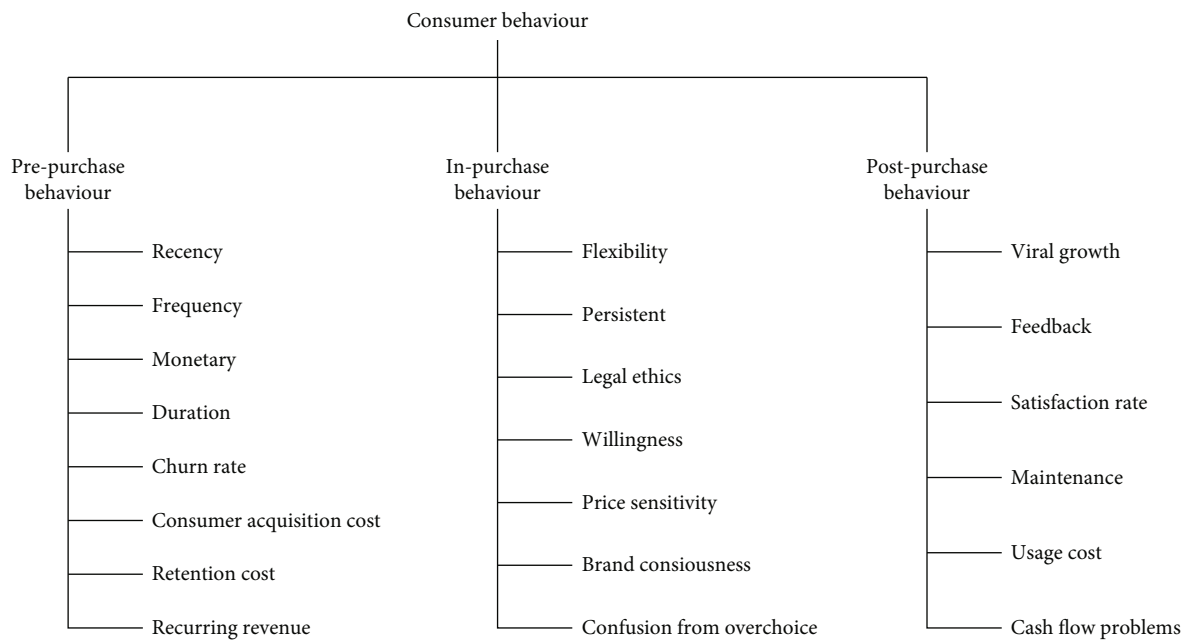


Figure 1: Consumer's behavioral attributes.

2.2. In-Purchase Behavior. In-purchase behavior of the consumer describes the behavior of the consumer during the purchase of services. The behavior of consumers during the purchase can be estimated by the following attributes.

(i) Flexibility. Flexibility is the inherent ability of the consumer to change or adapt or react to a decision-making environment with little change in cloud service cost and performance. The tendency of attribute is considered as positive (higher value is better)

(ii) Persistent. Despite a little difficulty in in-service performance, the consumer is continuing firmly with the existing CSP. The tendency of the attribute is considered as positive (higher value is better)

(iii) Legal ethics. Legal ethics are principles and values which, together with rules of conduct and laws, regulate a profession, such as the legal profession. A consumer, who is aware of the legal ethics of the geographical region from where the services are provisioned, does not blame the CSP regarding privacy issues. The tendency of attribute is considered as positive (higher value is better)

(iv) Willingness. The different consumers have different willingness to purchase a service and to pay for the service. The willingness of the consumer for the modified old services is lesser than the new technology-based services. A consumer who is willing to adopt the changes in technology is considered good for the perspective of the CSP. The tendency of attribute is considered as positive (higher value is better)

(v) Price sensitivity. Many consumers are price-centric if they do not see any deal; they are not going to make a transaction. The consumers who are focused on the performance of the service without the rigid boundary cost are considered good consumers. The tendency of attribute is considered as negative (lower value is better)

(vi) Brand consciousness. Consumers who are purchasing the services from the nationally renowned, expensive, and best-selling brands come in this category. These consumers are also called "price equals quality." Consumers assume that a high price tag is an indicator of a product of higher quality

(vii) Confusion from over choice. Consumers are very quick in deciding on the service selection among the available services, but due to largely available competitors, sometimes consumers get confused. The consumers who are clear about their demand and goals are beneficial for the CSP

2.3. Postpurchase Behavior. Postpurchase behavior describes the behavior of the consumer, such as the way of consumer thinks, feels, and act after provisioning the service.

(i) Viral growth. Viral growth refers to the number of new consumers an existing consumer brings to your services/product in a defined period. These consumers are also called an advocate of the service

(ii) Feedback. Consumer feedback is the information provided by the consumer about the experience of

usage of the service. Positive feedback on the service motivates other consumers to make a purchase

(iii) Satisfaction rate. The repurchase of the services by the consumer is a positive sign of satisfaction rate. The repurchase of the services from the same brand also helps the CSPs in generating more revenue

(iv) Maintenance. Maintenance cost refers to any cost incurred by an organization to keep its services in operational condition. These costs may be for general maintenance like running antivirus software, or they may be fixing some hardware/software problems. A consumer who requires low maintenance of services is beneficial for the service provider

(v) Usage cost. Cloud services have usage-based pricing. Consumers whose usage of the services after provisioning is more is considered beneficial for the service providers

(vi) Cash flow problems. A cash flow problem arises when an organization is unable to pay its debt due to late and slow-paying consumers. A consumer is considered valuable in terms of cash flow if he pays fast

3. CloudConsumerism: A Consumer-Centric Ranking Model for Efficient Service Mapping

This section presents the hypothesis and components such as possible service mapping scenarios, MCDM method TOPSIS used for ranking the CSPs as well as consumers, and Algorithm 1: service mapping algorithm for efficiently mapping the CSPs to the consumer in all the possible scenarios.

3.1. Service Mapping Scenarios. In the brokerage-based service selection, the number of cloud entities that needs to be evaluated is two; therefore, the degree of mapping is two, and the number of possible scenarios is four. The proposed CloudConsumerism model deals efficiently with all possible scenarios.

Case 1. One to one mapping. In this case the single cloud broker provides service that is provisioned by one cloud service provider to a consumer as shown in Figure 2.

Case 2. One to many mapping. In this case a cloud broker provides service that is provisioned by a cloud service provider to many consumers as shown in Figure 3.

Case 3. Many to one mapping. In this case many consumers request a service to the cloud broker, and only one cloud service provider is available for the service provisioning as shown in Figure 4.

Case 4. Many to many mapping. In this case many consumers are requesting a service to the cloud brokers, and many cloud service providers are available for the service provisioning as shown in Figure 5.

3.2. MCDM Method TOPSIS. CloudConsumerism model uses TOPSIS method [23] for evaluating the stakeholders. The method is based on the idea that the selected CSP must be closest to the positive ideal solution and farthest from the negative ideal solution. The steps of TOPSIS methods are the following:

Step 1. Normalize the decision matrix $DM_{\text{norm}} = [nv_{ij}]$, where nv_{ij} for i^{th}CSP for j^{th} attributes are given by

$$nv_{ij} = \frac{v_{ij}}{\sqrt{\sum_{i=1}^{m} v_{ij}^2}}, \tag{4}$$

$$nm = \begin{pmatrix} nv_{11} & \cdots & nv_{1n} \\ \vdots & \ddots & \vdots \\ nv_{m1} & \cdots & nv_{mn} \end{pmatrix}. \tag{5}$$

Step 2. Calculate the weighted normalized decision matrix $WDM_{\text{norm}} = [r_{ij}]$, where r_{ij} for i^{th}CSP for j^{th} attribute is given by

$$\mathrm{r_{ij}} = \mathrm{nv_{ij}} \times \mathrm{w_j}, \tag{6}$$

$$WDM_{\text{norm}} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{pmatrix}. \tag{7}$$

Step 3. Calculate the positive ideal solution A_p and the negative ideal solution A_n.

$$A_p = \left(\min\left(r_{ij} | i = 1, 2, 3, \cdots\cdots\cdots m\right) | j \in J_-, \max \cdot \left(r_{ij} | i = 1, 2, 3, \cdots\cdots\cdots m\right) | j \in J_+\right) \cong \left(r_{pj} | j = 1, 2, 3, \cdots n\right), \tag{8}$$

$$A_n = \left(\max\left(r_{ij} | i = 1, 2, 3, \cdots\cdots\cdots m\right) | j \in J_-, \min \cdot \left(r_{ij} | i = 1, 2, 3, \cdots\cdots\cdots m\right) | j \in J_+\right) \cong \left(r_{nj} | j = 1, 2, 3, \cdots n\right), \tag{9}$$

where

$J_+ = \{j = 1, 2, \cdots..n | j$ is associated to the attribute which is beneficial attribute.

$J_- = \{j = 1, 2, \cdots..n | j$ is associated to the attribute which is cost attribute.

Step 4. Evaluate the separation distance from positive ideal solution S_{ip} and negative ideal solution S_{in}.

$$S_{ip} = \sqrt{\sum_{j=1}^{n} \left(r_{ij} - r_{bj}\right)^2}, i = 1, 2, \cdots\cdots..m, \tag{10}$$

```
Input: Directory of CSPs and Consumers.
Output: Service mapping of CSPs and Consumers.
Begin
  For all the available CSPs set flagCSP ←1.
    IF CSPs ==1 && consumers ==1 THEN
      bind consumer to CSP
      flagCSP ←0.
    END IF
    IF CSPs >1 && consumers ==1 THEN
      get the ranking score of the CSPs using TOPSIS;
      sort the ranking score of CSPs and find top rank CSP;
      map the consumer to CSP;
      flagCSP ← 0;
    END IF
    IF CSPs ==1 && consumers > 1 THEN
      get the ranking score of consumers using TOPSIS;
      sort the ranking score of consumers and find top rank consumers;
      map the consumer to CSP;
      flagCSP ← 0;
    END IF
    IF CSPs >1 && consumers > 1 THEN
      FOR each consumer DO
        get the ranking score of CSPs using TOPSIS Method;
        sort the ranking score of CSPs and find top rank CSP;
      END FOR
      FOR each CSP DO
        get the list of competitive consumers;
      END FOR
      FOR each CSP DO
        IF competitive consumers list is empty THEN
          continue;
        ELSE IF competitive consumers ==1 and flagCSP ==1 THEN
          map the consumer to CSP;
          flagCSP ← 0;
      END IF
      ELSE flagCSP ==1 THEN
        get the score of competitive consumers using TOPSIS;
        rank the consumers and get the top rank consumer;
        map the consumer to CSP;
        flagCSP ← 0;
      END FOR
    END IF
    After completion of the service the flagCSP is set to 1;
End
```

Algorithm 1: Service mapping algorithm.

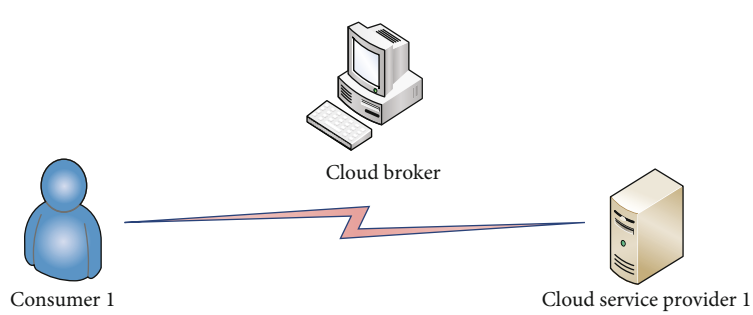


Figure 2: One to one mapping.

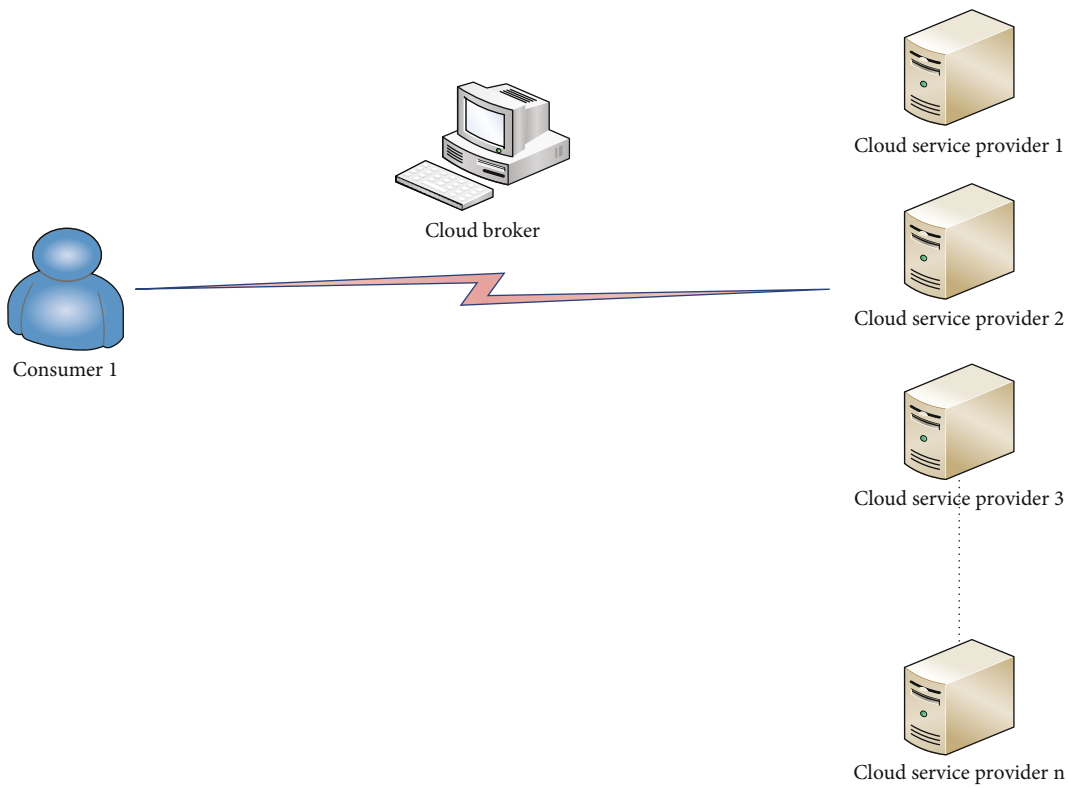


Figure 3: One to many mapping.

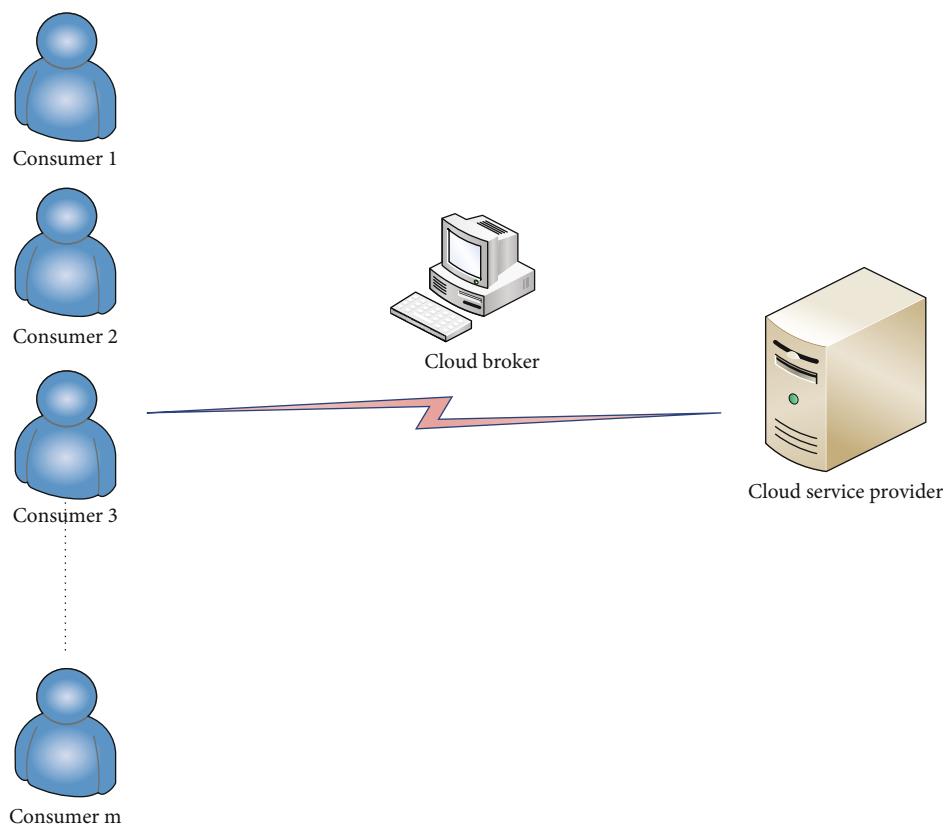


Figure 4: Many to one mapping.

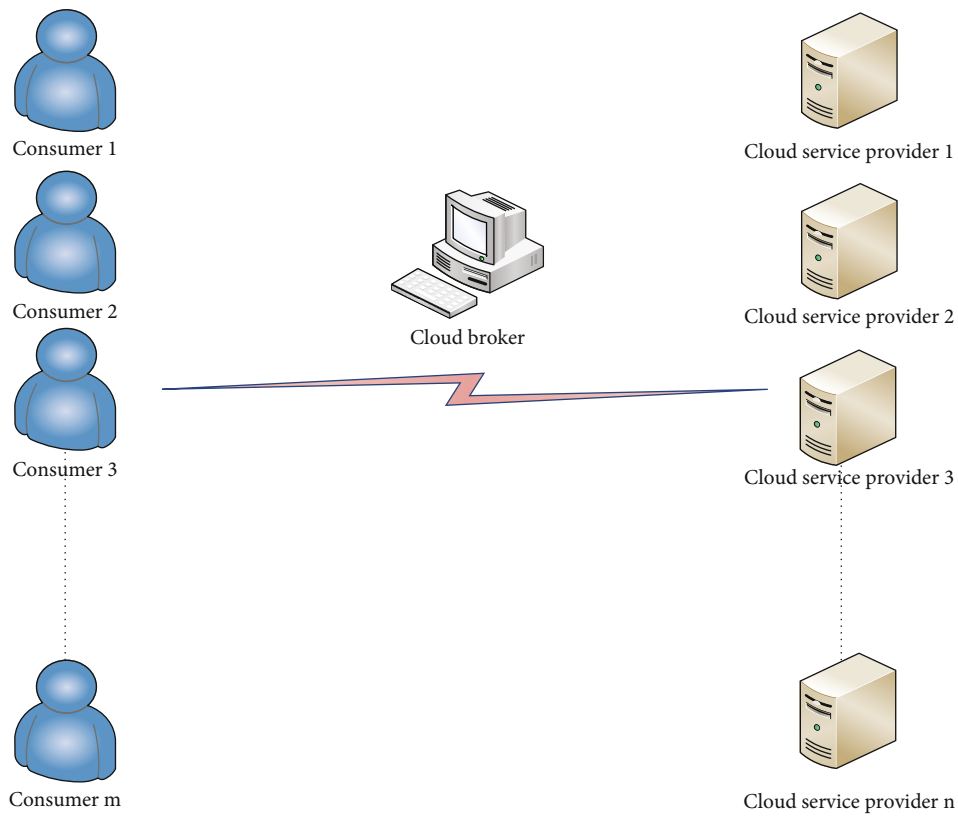


Figure 5: Many to many mapping.

$$S_{\text{in}} = \sqrt{\sum_{j=1}^{n} \left(r_{ij} - r_{nj}\right)^2}, i = 1, 2, \cdots \cdots ..m, \tag{11}$$

where S_{ip} and S_{in} are L2-norm distance of the ith target CSP from the positive and negative ideal solution, respectively.

Step 5. Calculate the relative closeness.

$$C_i = \frac{S_{\text{in}}}{S_{ip} + S_{\text{in}}}, 0 \le C_i \le 1, i = 1, 2, \cdots \cdots \cdots \cdots m.,$$
$$C_i = \begin{cases} 1, \text{if the } \textit{alterative solution has} \text{ the } \textit{best condition}, \\ 0, \text{if the } \textit{alterative solution has} \text{ the } \textit{worst condition}. \end{cases} \tag{12}$$

Step 6. Rank the CSP based on relative closeness.

$$Rank = sort(C_i). \tag{13}$$

3.3. Service Mapping Algorithm. This subsection presents the service mapping algorithm. The Algorithm 1 is used by the proposed CloudConsumerism model during the service mapping process.

4. Case Study

In the presented case study, the offered QoS data is collected from three IaaS public CSPs: Amazon EC2, Rackspace, and Windows Azure. The SMI attributes, namely, accountability, agility, assurance, cost, performance, and security, are considered for estimating the ranking the CSPs. The QoS data offered by CSP_1 (Amazon EC2), CSP_2 (Rackspace), CSP_3 (Microsoft), and relative weights of consumer's requirement is presented in Table 1. The QoS data is collected from the paper [5]. Due to unavailability of data of QoS attributes accountability and security, the values are randomly assigned. QoS attributes comparison rules and scaling are adopted from the paper [24]. For the ranking estimation of consumers, the transaction attributes recency, frequency, monetary, duration, recurring revenue, churn rate, and acquisition cost are considered, and data values are assigned randomly, as presented in Table 2. The preferences of CSPs for selecting the valuable consumers is presented in Table 3.

The TOPSIS method is applied to rank the CSPs and consumers based on the preferences of their counterparts. The ranking of CSPs is shown in Figure 6. In business perspective, every consumer follows a cost-quality trade-off. Thus, cost-quality-based service ranking is considered as the final stage of selection of CSPs. Since the offered cost of the CSP for the services is different, the ranking score of CSPs per unit cost is estimated as presented in Figure 7. From the figure, it is seen that CSP_1 is ranked as the first choice by all the consumers. The CSP_2 is ranked as the

Table 1: Offered QoS of the CSPs.

QoS attributes	Top-level QoS attributes required by the consumers										First level attributes (weights)	Second level attributes (weights)		Service 1 (S1)	Service 2 (S2)	Service 3 (S3)	Value type
	1	2	3	4	5	6	7	8	9	10							
Accountability	0.05	0.25	0.10	0.25	0.05	0.05	0.05	0.05	0.30	0.05	Level 0-10			4	8	4	Numeric
												CPU	0.5	9.6	12.8	8.8	Numeric
Agility	0.10	0.20	0.20	0.10	0.15	0.10	0.10	0.10	0.10	0.30	Capacity (0.6)	Memory	0.3	15	14	15	Numeric
												Disk	0.2	1690	2040	630	Numeric
											Elasticity (0.4)	Time	1	80-120	520-780	20-200	Range
											Availability (0.7)	0.7	0.7	99.95%	99.99%	100%	Numeric
											Service stability (0.2)	Upload time	0.3	13.6	15	21	Numeric
Assurance	0.30	0.10	0.15	0.20	0.20	0.20	0.45	0.20	0.20	0.20		CPU	0.4	17.9	16	23	Numeric
												Memory	0.3	7	12	5	Numeric
											Serviceability (0.1)	Free support	0.7	0	1	1	Boolean
												Type of support	0.3	24/7, diagnostic tools, phone, urgent response	24/7, diagnostic tools, phone, urgent response	24/7, phone, urgent response	Unordered set
												VM cost	0.6	$0.68	$0.96	$0.96	Numeric
Cost	0.30	0.05	0.30	0.10	0.25	0.50	0.05	0.10	0.05	0.10	Ongoing cost (1)	Data	Inbound	10	10	8	Numeric
													Outbound	11	15	18	
												Storage	0.2	12	15	15	Numeric
											Service response time (0.1)	Range	0.5	80-120	520-780	20-200	Range
Performance	0.30	0.30	0.10	0.30	0.30	0.10	0.30	0.30	0.30	0.30		Average value	0.5	100	600	30	Numeric
Security	0.05	0.10	0.15	0.05	0.05	0.05	0.05	0.25	0.05	0.05	Level: 0-10 (1)			4	8	4	Numeric

Table 2: Service transaction of consumers.

	Recency	Frequency	Monetary	Duration	Recurring revenue	Churn rate	Acquisition cost
$Consumer_1$	18	10	520	2	100	0.10	50
$Consumer_2$	4	1	190	5	500	0.15	50
$Consumer_3$	12	6	65	3	200	0.10	45
$Consumer_4$	7	3	2000	4	400	0.25	55
$Consumer_5$	9	4	250	7	700	0.05	35
$Consumer_6$	55	8	726	4	100	0.30	45
$Consumer_7$	17	14	580	3	200	0.20	30
$Consumer_8$	35	6	150	2	500	0.10	20
$Consumer_9$	23	1	325	4	1100	0.40	10
$Consumer_{10}$	3	15	140	2	1200	0.10	40

Table 3: Required transaction preferred by the CSPs.

	Recency	Frequency	Monetary	Duration	Recurring revenue	Churn rate	Acquisition cost
CSP 1	0.30	0.20	0.10	0.10	0.15	0.10	0.05
CSP 2	0.10	0.10	0.25	0.20	0.10	0.15	0.10
CSP 3	0.20	0.20	0.10	0.20	0.10	0.05	0.15

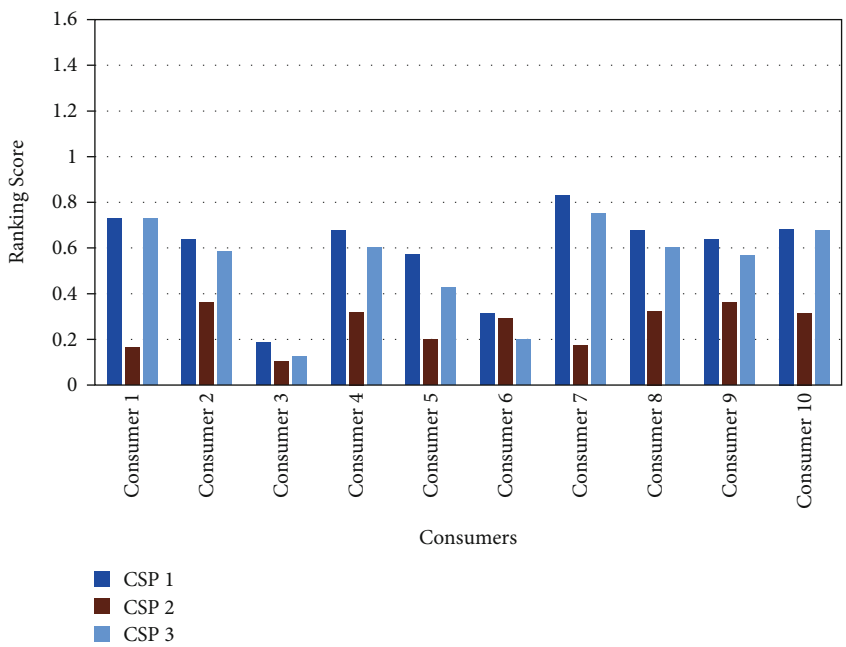


Figure 6: Ranking score of CSPs.

second choice except for $consumer_6$, and CSP_3 is ranked as the last choice by all consumers except $consumer_6$. Thus, for CSP_1, all the consumers are in competition. Therefore, consumers are ranked based on the consumers transaction behavior attribute value and desired consumer weighting by the CSPs. Ranking results of consumers are shown in Figure 8. The duration of the transactions differs from consumer to consumer. The ranking score per unit time is generated for the healthy comparison as shown in Figure 9, which is considered at the final stage of ranking. From the results, it is observed that $consumer_{10}$ is ranked at first choice by the CSP_1 and others also. Considering the mutual choice, the $consumer_{10}$ is mapped to the CSP_1. CSP_2 is preferred at second choice only by $consumer_6$; therefore, $consumer_6$ is directly mapped (no need of ranking the consumer because of no competition) to the CSP_2. CSP_3 is preferred as the second choice by the rest consumers. The rest of consumers are ranked based on the requirement of CSP_3, and it is found that CSP_3 prefers $consumer_1$ as their second choice (first choice is $consumer_{10}$ which is already

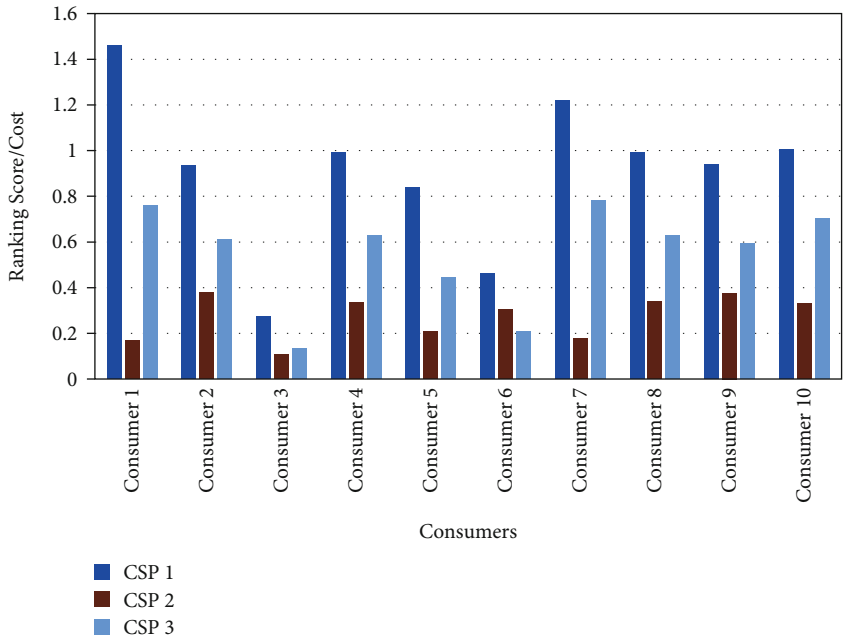


Figure 7: Ranking of CSPs per unit cost.

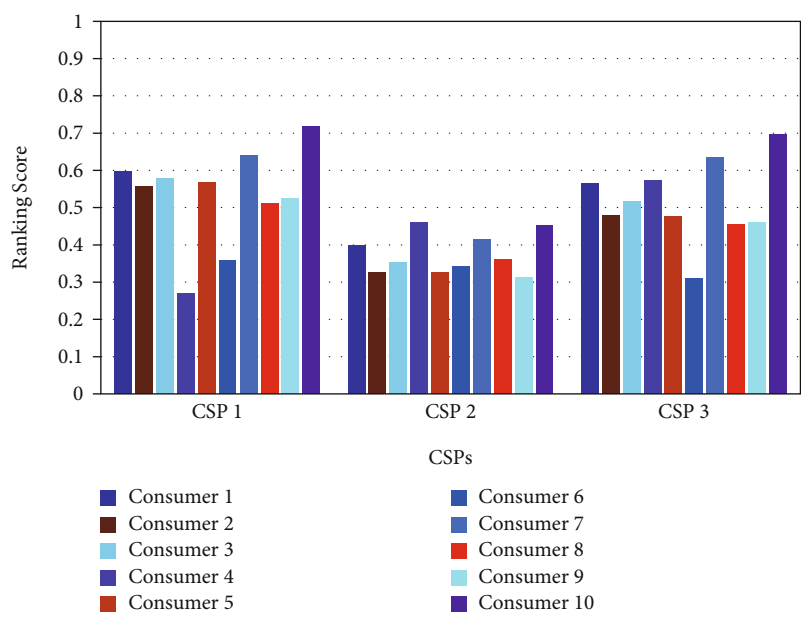


Figure 8: Ranking score of consumers.

mapped), so consumer_1 (second choice) is mapped with the CSP_3. The mapping of consumers and CSPs after considering mutual evaluation is presented in Figure 10.

4.1. Satisfaction of Mapped Consumers and CSPs. After mapping the of consumers and CSPs, the service satisfaction of consumers and CSPs are computed [12]. For computing the satisfaction, the ranking score of the CSPs/consumers is considered. The satisfaction of the CSP/consumer is calculated using the offered and received performance. For the satisfaction of consumers, the ranking score of CSP, which is at the first choice based on the requirement of a consumer, is considered as offered performance for the consumer. The ranking score of actual mapped CSP after considering the ranking of consumers based on the requirement of CSP is considered as received performance and vice versa. The case study results show that consumer_1 is mapped to CSP_3, consumer_6 is mapped to CSP_2, and consumer_{10} is mapped to CSP_1.

The CSP_1 is ranked as the first choice by consumer_1. The ranking score of CSP_1 (1.4577) is considered as the desired/offered performance for the consumer_1, and at the final stage

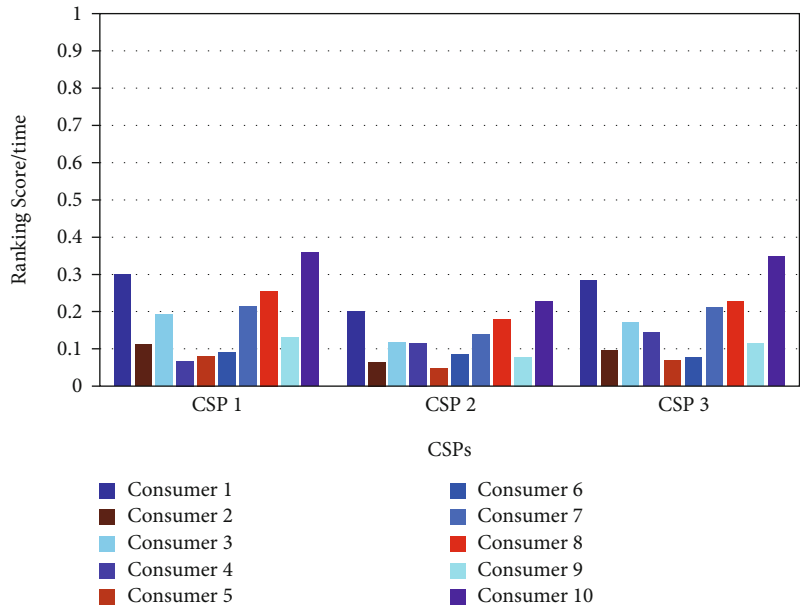


Figure 9: Ranking sore of consumers per unit time or duration.

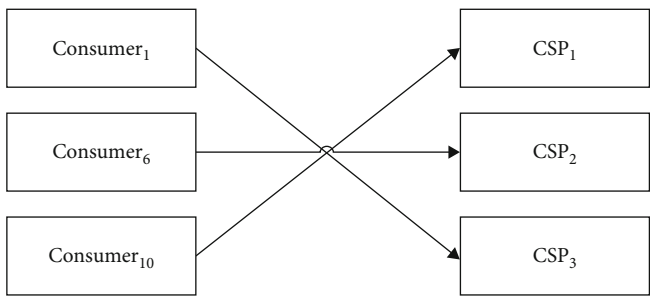


Figure 10: Mapping of consumers and CSPs.

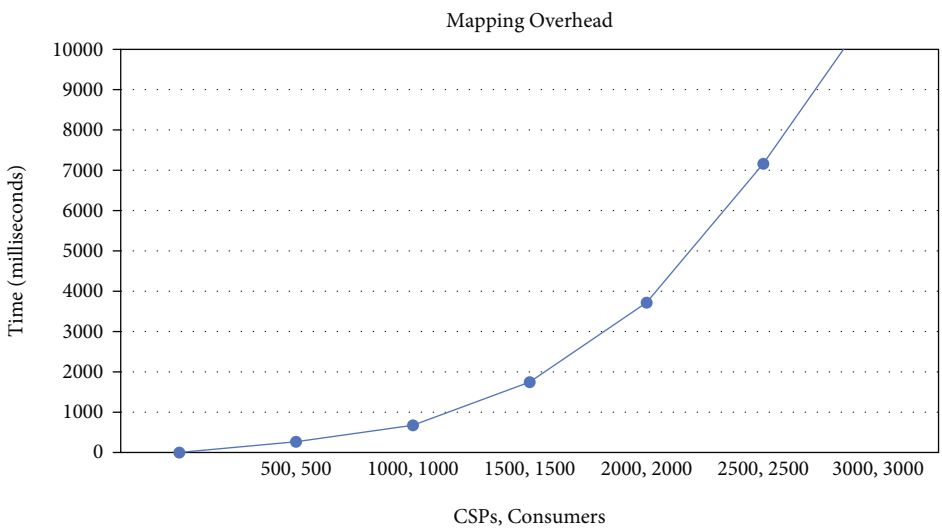


Figure 11: The number of CSPs and consumers is fixed to 2500.

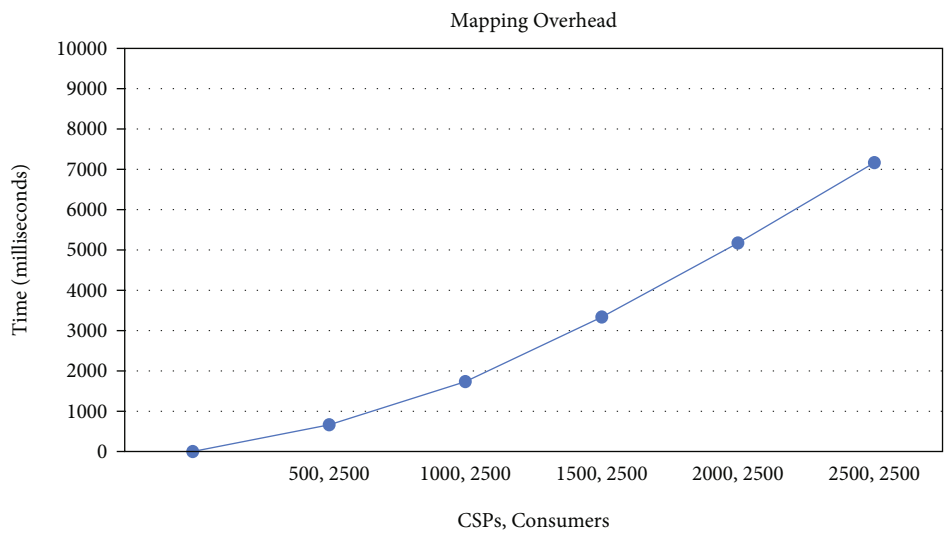


Figure 12: The CSPs are varied from 500 to 2500 and consumers are fixed to 2500.

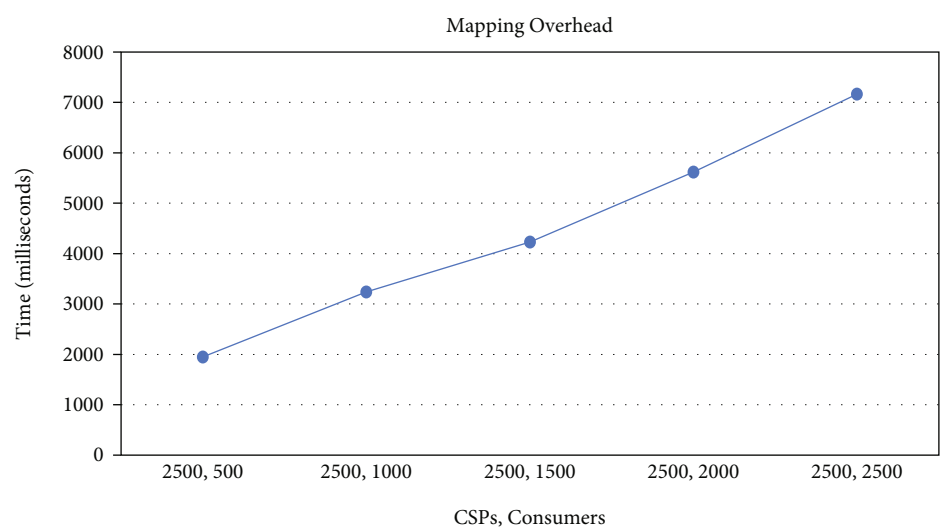


Figure 13: The number of CSPs is varied from 500 to 2500, and the number of consumers is fixed to 2500.

Table 4: Comparative analysis with existing frameworks.

Framework	Number of CSPs	Number of consumers
SMICloud [5]	1000	1
TRCSM [7]	700	700
MECSM [12]	1800	1800
Modified MECSM	2500	2500

of mapping, CSP_3 (0.7592) is mapped to the $consumer_1$, which is considered as actual received performance; therefore, the satisfaction of $consumer_1$ is evaluated as:

$$e^{-\left|\frac{1.4577-0.7592}{0.7592}\right|} = 0.6193 = 61.93\%. \quad (14)$$

The CSP_1 is ranked as the first choice by the $consumer_6$. The ranking score of CSP_1 (0.4605) is considered as the desired/offered performance for the $consumer_6$, and at the final stage of mapping, CSP_6 (0.3040) is mapped to the $consumer_6$, which is considered as received performance for the $consumer_6$; therefore, the satisfaction of $consumer_6$ is evaluated as

$$e^{-\left|\frac{0.3040-0.4605}{0.4605}\right|} = 0.7119 = 71.19\%. \quad (15)$$

The CSP_1 is ranked as the first choice by the $consumer_{10}$. The ranking score of CSP_1 (1.0058) is considered as the desired/offered performance for the $consumer_{10}$, and at the final stage of mapping, CSP_1 (1.0058) is mapped to the $consumer_{10,}$ which is considered as the received performance. For $consumer_{10}$, the offered performance and the received is equal; therefore, the satisfaction of $consumer_{10}$ is 100%.

5. Experimental Results and Discussion

For experimental analysis, the MCDM methods are implemented in JAVA, running on Windows 10 Home Single Language with Intel Core i7-4510U @ 2.00GHz 2.60 GHz having 8 GB RAM. The ranking overhead is evaluated with variations in the number of CSPs and the number of consumers. The number of attributes for CSPs as well as consumers is fixed. The benchmark for ranking overhead is considered 10 seconds. The mapping overhead of three different scenarios (by varying the number of CSPs and consumers) is recorded. In the first scenario, the number of consumers and CSPs are considered equal, i.e., 500 each and the ranking overhead is recorded. The results show that within a limited time overhead of 10 milliseconds, the mapping overhead increases linearly and scales mapping up to 2500 consumers and 2500 CSPs, as shown in Figure 11. In the second scenario, the number of CSPs is fixed to 2500 (considering as a scaling limit as obtained from the first scenario), and the number of consumers is varied from 500 to 2500 with a step size of 500. The results show that the ranking overhead increases linearly, as shown in Figure 12, and within a time overhead of 8 milliseconds, the 2500 consumers and 2500 provides are mapped. In the third scenario, the number of consumers is fixed to 2500, and the number of CSPs is varied from 500 to 2500 with a step size of 500. The ranking overhead increases linearly as the number of CSPs increases, and within 8 milliseconds, 2500 consumers and 2500 CSPs are mapped as shown in Figure 13. For each data value, the results are taken 40 times, and the average execution time is considered the final ranking overhead in each scenario.

6. Conclusion and Future Work

In this paper, we have addressed a consumer-centric quality of service-based service mapping model, CloudConsumerism. In the proposed model, for evaluating the consumers, the behavioral attributes of consumers are defined, and while mapping, the consumers and CSPs both are evaluated. The number of consumers and service providers is scaleup up to 2500 consumers and 2500 providers. We demonstrate the applicability of our work in the online cloud service selection model by increasing the scaleup of the number of consumer and the number of providers within the limited time overhead of 10 seconds. By applying the requirement filtering from both the side, i.e., cloud consumer as well as CSP, it is shown that the valuable consumers are mapped to the high-performance service provider and hence the satisfaction of the cloud consumer and CSP is increased. The comparative analysis with existing frameworks has been presented in Table 4. In the future, the separate broker for each consumer and provider will be involved so that the third-party broker can also be evaluated from both the side consumer as well as provider. For evaluating the broker, the new attribute will also be defined. From Table 4, it is concluded that the modified MECSM is more efficient in terms of overhead and robustness.

Data Availability

The data used to support the findings of this study are currently under embargo, while the research findings are commercialized. Requests for data, 12 months after the publication of this article, will be considered by the corresponding author.

Conflicts of Interest

There is no conflicts of interest of any authors.

References

[1] C. Madhavaiah, "Defining cloud computing in business perspective : a review of research," *Metamorphosis*, vol. 11, no. 2, pp. 50–65, 2012.

[2] T. Leadership and W. Paper, "Defining a framework for cloud adoption: How common ground can help enterprises drive success with cloud computing," *IBM Global Technology Services*, 2010.

[3] Z. R. Alashhab, M. Anbar, M. M. Singh, Y. B. Leau, Z. A. Al-Sai, and S. A. Alhayja'a, "Impact of coronavirus pandemic crisis on technologies and cloud computing applications," *Journal of Electronic Science and Technology*, vol. 19, no. 1, article 100059, 2021.

[4] M. S. Kumar, I. Gupta, S. K. Panda Gupta, and P. K. Jana, "Granularity-based workflow scheduling algorithm for cloud computing," *The Journal of Supercomputing*, vol. 73, no. 12, pp. 5440–5464, 2017.

[5] S. Kumar, S. Versteeg, and R. Buyya, "A framework for ranking of cloud computing services," *Future Generation Computer Systems*, vol. 29, no. 4, pp. 1012–1023, 2013.

[6] M. Neeraj, S. Goraya, and D. Singh, "A comparative analysis of prominently used MCDM methods in cloud environment," *Journal of Supercomputing*, vol. 77, no. 4, pp. 3422–3449, 2020.

[7] N. Yadav and M. S. Goraya, "Two-way ranking based service mapping in cloud environment," *Future Generation Computer Systems*, vol. 81, pp. 53–66, 2018.

[8] "The Cloud Service Measurement Initiative Consortium Service Measurement Index Introducing the Service Measurement Index," 2011, http://www.cloudcommons.com/about-smi.

[9] D. L. Motherbaugh, D. L. Hawkins, and A. Mookerjee, "Consumer behaviour, building marketing," *Strategies*, 2010.

[10] W. Jianping, "Research on VIP customer classification rule base on RFM model," in *MSIE 2011*, pp. 336–338, Harbin, Jan 2011.

[11] J. Wei, S. Lin, and H. Wu, "A review of the application of RFM model," *African Journal of Business Management*, vol. 4, no. 19, pp. 4199–4206, 2010.

[12] M. S. Goraya and D. Singh, "Satisfaction aware QoS-based bidirectional service mapping in cloud environment," *Cluster Computing*, vol. 23, no. 4, pp. 2991–3011, 2020.

[13] H. Shen, G. Liu, S. S. Member, G. Liu, and S. S. Member, "An efficient and trustworthy resource sharing platform for collaborative cloud computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 4, pp. 862–875, 2014.

[14] M. Azadi, M. Izadikhah, F. Ramezani, and F. K. Hussain, "A mixed ideal and anti-ideal DEA model: an application to

evaluate cloud service providers," *IMA Journal of Management Mathematics*, vol. 31, pp. 233–256, 2019.

[15] S. Ramadass, R. Krishankumar, K. S. Ravichandran, H. Liao, S. Kar, and E. Herrera-Viedma, "Evaluation of cloud vendors from probabilistic linguistic information with unknown/partial weight values," *Applied Soft Computing*, vol. 97, article 106801, 2020.

[16] R. Ranjan, K. Siba, and C. Kumar, "Prioritizing the solution of cloud service selection using integrated MCDM methods under fuzzy environment," *The Journal of Supercomputing*, vol. 73, no. 11, pp. 4652–4682, 2017.

[17] L. Sun, J. Ma, Y. Zhang, H. Dong, and F. Khadeer, "Cloud-FuSeR: fuzzy ontology and MCDM based cloud service selection," *Future Generation Computer Systems*, vol. 57, pp. 42–55, 2016.

[18] T. L. Saaty, "The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making," in *Multiple criteria decision analysis*, pp. 363–419, Springer, New York, NY, 2016.

[19] K. Yoon, "A reconciliation among discrete compromise solutions," *Journal of the Operational Research Society*, vol. 38, no. 3, pp. 277–286, 1987.

[20] A. Mardani, E. K. Zavadskas, K. Govindan, A. A. Senin, and A. Jusoh, "VIKOR technique: a systematic review of the state of the art literature on methodologies and applications," *Sustainability*, vol. 8, pp. 37-38, 2016.

[21] J. P. Brans and P. Vincke, "Note—a preference ranking organisation method," *Management Science*, vol. 31, no. 6, pp. 647–656, 1985.

[22] C. Gartner, "STAMFORD, Gartner Forecasts Worldwide Public Cloud End-User Spending to Grow 21% in 2021," 2021.

[23] B. Lee, K. H. Oh, H. J. Park, U. M. Kim, and H. Y. Youn, "Resource reallocation of virtual machine in cloud computing with MCDM algorithm," in *2014 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*, Oct 2014.

[24] V. Xuan, H. Tsuji, and R. Masuda, "A new QoS ontology and its QoS-based ranking algorithm for Web services," *Simulation Modelling Practice and Theory*, vol. 17, no. 8, pp. 1378–1398, 2009.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7826838, 11 pages
https://doi.org/10.1155/2022/7826838

Research Article

Formal Credit-Assisted New Agricultural Business: A Multifactor Analysis Based on BP Neural Network

Lihuan Zhang[1] **and Jing Fan**[2]

[1]*Jilin Agricultural Science and Technology College, Jilin 132101, China*
[2]*Jilin Agricultural University, Changchun 130118, China*

Correspondence should be addressed to Lihuan Zhang; zhanglihuan@jlnku.edu.cn

Received 17 February 2022; Revised 20 March 2022; Accepted 23 March 2022; Published 13 April 2022

Academic Editor: Jianhui Lv

In recent years, the new agricultural business entities have grown rapidly in all parts of the country and have become an effective carrier to achieve the basic stability of China's rural family operation system and moderate-scale agricultural operation. However, the systemic defects of rural financial market and the credit rationing of formal financial institutions make the capital needs of rural areas unsatisfied for a long time. Financial demand determines the direction of supply reform. Therefore, in order to improve rural financial supply and promote the transformation and development of agricultural economy under the current situation, it is necessary to effectively understand the formal credit demand and the impact factors of accessibility of the new agricultural business entities, which represents the direction of China's agricultural development in the future. Based on back propagation (BP) neural network, this paper constructs the farmers' formal credit availability prediction model and studies the formal credit demand and the impact factors of the availability of new agricultural business entities. The experimental results show that the farmers' formal credit availability prediction model has good performance in prediction accuracy, prediction time, and mean squared error.

1. Introduction

With the gradual improvement and development of rural credit infrastructure, credit products, and credit services that meet the needs of agriculture, "rural areas and farmers" are emerging [1]. From the perspective of credit suppliers, credit supply can be mainly divided into formal credit and informal credit [2]. Formal credit generally refers to a type of credit provided by the government and other institutions and subject to legal constraints and unified supervision by monetary authorities, mainly provided by banks and other formal financial institutions. Informal credit refers to the private lending permitted by the government but not limited by banks and some financial regulatory agencies, mainly loans between relatives and friends [3–5]. At present, the development of rural economy in China depends on the loan of farmers besides the financial support of the state, and with the continuous improvement of the whole credit system, formal credit is more and more popular. In the process of social transformation and economic development in China's rural areas, the scale and role of informal credit are constantly weakening. Existing studies show that formal credit accounts for nearly two-thirds of the total, and the long-term trend of the development of China's rural financial market is dominated by formal finance and supplemented by informal finance [6]. To better grasp farmers' credit behavior, it is necessary to understand its determinants.

Compared with traditional agriculture, the subject of new agricultural business is to adhere to the basic agricultural management system based on household contract management. It is a modern agricultural management organization that realizes large-scale, professional, intensive, and market-oriented management through land transfer, mainly including professional large households, family farms, farmers' cooperatives, and agricultural leading enterprises [7, 8]. In recent years, with the acceleration of China's urbanization and industrialization, the large-scale transfer of rural labor force, and the growth and differentiation of

agricultural production and management organizations, the new agricultural business entity has continuously developed and become the backbone of China's agricultural modernization and an important subject of rural economy and rural financial demand [9, 10].

Whether the rural financial system, financial products, and financial services accord with the reality of rural areas depends on whether they can meet the financial needs of rural areas. Therefore, the study of the characteristics of rural financial demand and its satisfaction is the basic premise of the rational arrangement of rural financial system [11]. Due to its characteristics of scale, specialization, marketization, and intensification, the new agricultural business entities have much greater demand and dependence on capital than the traditional household leavers and present significantly different financial demand characteristics from the traditional household leavers. However, due to the adverse impact of China's dual economic structure, China's rural financial market has always been underdeveloped. At present, China's rural financial services are still difficult to adapt to the development of rural economic entities. The financing difficulty of new agricultural business entities is becoming increasingly prominent and has become the core element restricting their development [12]. Therefore, it is very important to study the ability of farmers to get loans in the new agricultural business assisted by formal credit.

Due to farmers' limited risk control ability, weak rural credit foundation, and high nonperforming loan rate, farmers' formal credit has a certain difficulty. At the same time, for financial institutions, how to correctly assess farmers' repayment ability plays an important role in reducing their risk. In the farmer household credit rating, generally through the loan officer, the village committee's subjective opinion or the use of scoring scale to determine the score, although these methods are simple and feasible, they are highly subjective and their execution process is not standardized. Without the help of quantitative mathematical models, it is easy to lead to inaccurate evaluation of farmers' credit status, which cannot fully meet the needs of formal credit risk management. This study attempts to use back propagation (BP) neural network to establish the ability of farmers to get formal credit, so as to identify the credit risk of farmers in formal credit, improve the quality of formal credit, and help the development of new agricultural business.

The contributions of this study are as follows.

(i) After the same characteristic parameters are normalized in data normalization, the normalization among different characteristic parameters is added

(ii) Genetic algorithm is introduced to avoid BP neural network falling into local minimum and can accelerate the convergence speed of the network

(iii) Momentum term is added into weights that can effectively prevent BP neural network from falling into local minimum and shorten learning time

The remainder of the paper is structured as follows: Section 2 reviews the related context. In Section 3, BP neural network is optimized. In Section 4, a multifactor farmers' formal credit availability prediction model based on BP neural network is established and experiments are conducted. Section 5 gives the conclusion of this paper.

2. Literature Review

Relevant scholars have carried out abundant studies on the factors affecting farmers' access to loans, and scholars have done a lot of empirical studies on farmers' loan difficulties and put forward many new ideas to solve the problems. Machine learning has been proved to be able to use a large number of financial data to improve prediction; in [13], the authors used agricultural loan arrears to approximate financial pressure and used logistic regression and several machine learning methods to predict financial pressure. In [14], the authors evaluated whether social capital affected farmers' ability to obtain formal and informal loans. In [15], the authors combined the fuzzy comprehensive model and fuzzy control method to establish the model and indicator system of farmers' credit risk characteristics. Using the data of new agricultural management entities, in [16], the authors investigated the impact of agricultural land management scale on loan availability. In [17], the farmers' credit optimization decision-making model based on common risk guarantee fund and its application were proposed, which had scientific guiding significance and practical application value for farmers' credit decision-making. Starting from the traditional analysis model based on loan approval rejection or default rate, in [18], the authors reexamined the discrimination of a few borrowers in agricultural loans, focusing on the decision-making of loan packaging.

Now, many scholars apply neural network to the study of credit. In [19], peer-to-peer (P2P) lending provided convenient and efficient financing channels for enterprises and individuals, but P2P lending faced credit risk crisis of borrowers in the process of transaction, and the default rate of borrowers was relatively high. The authors used P2P-BP-LM algorithm to build the credit risk assessment model of P2P loan borrowers, so as to establish and improve an accurate and effective credit risk management system of borrowers. In [20], the authors proposed a Deep Genetic Hierarchical Network of Learners (DGHNL) for credit scoring. In [21], the authors proposed a Cost-sensitive Neural Network Ensemble (CS-NNE) model for default prediction. In [22], the authors proposed a deep neural network model for behavioral credit rating. In [23], the authors constructed a personal credit evaluation model and compared the weight adjustment method with BP neural network. In [24], the authors studied the application of radial basis function neural network model combined with optimal segmentation algorithm in personal loan credit rating model of banks or other financial institutions. In [25], the authors used genetic algorithm to adjust and determine the initial weight and threshold of BP neural network to evaluate credit risk. Throughout the literature, there are few neural networks to study formal credit-assisted new agricultural business. Therefore, in this paper, the optimized BP neural network

is used for the prediction of farmers' formal credit availability under new agricultural business.

3. Methods

3.1. Data Normalization. In this study, the actual credit status of farmers is taken as the output neuron of BP neural network, and some indicators of farmers are taken as the input neuron. Because the data types of each information indicator are not unified, there are characters and numbers, and the magnitude of data varies greatly, such data source is easy to cause deviation to the prediction results and affect the accuracy of prediction. Therefore, data normalization is necessary before the prediction of farmers' formal credit availability, which is to convert all data into numbers between [0,1] [26]. Its purpose is to cancel the magnitude difference between data of various dimensions and avoid large network prediction error due to large difference in the number and level of input and output data.

In order to solve the problem of less connection between different feature parameters in common normalization methods, this paper proposes that the row vector is normalized after the column vector of sample space $\Omega_{M\times N}$ is normalized; that is, after the same feature parameter is normalized, the normalization between different feature parameters is added, which is called joint normalization method.

Let the number of BP neural network input nodes be N, select N different characteristic parameters $(X^0, X^1, \cdots, X^{N-1})$ under the same conditions as the input of the network, and select M eigenvalues $(X_0^i, X_1^i, \cdots, X_{M-1}^i$, where $i = 0, 1, \cdots, N-1)$ for each characteristic parameter, which constitutes the sample space $\Omega_{M\times N}$. The joint normalization method is divided into two steps.

3.1.1. Column Vector Normalization. A common normalization method is selected, and the maximum value method is taken as an example for column vector normalization, which is similar to other methods. After normalization,

$$\widehat{X_j^i} = \frac{X_j^i - X_{\min}^i}{X_{\max}^i - X_{\min}^i}, \tag{1}$$

where $i = 0, 1, \cdots, N-1$ and $j = 0, 1, \cdots, M-1$.

3.1.2. Row Vector Normalization. A common normalization method is selected, and the maximum value method is taken as an example for row vector normalization, which is similar to other methods. After normalization,

$$\bar{X} = \frac{\widehat{X^i} - \widehat{X_{\min}^i}}{\widehat{X_{\max}^i} - \widehat{X_{\min}^i}}, \tag{2}$$

where $i = 0, 1, \cdots, N-1$.

To sum up, the data after joint normalization can be used as the input of the network.

3.2. Genetic Algorithm Optimized BP Neural Network. Given the limitations of BP neural network, genetic algorithm has a strong adaptability to the environment and self-learning ability; its highly parallel global search algorithm can overcome its own shortcomings [27]. The combination of these not only help to avoid the BP neural network into a local minimum value but also can accelerate the convergence speed of the network. Meanwhile, it can also improve their ability of learning and dissemination of model generalization ability. We mainly use genetic algorithm in the training process of neural network to adjust the initial connection weight and threshold value of the farmers' formal credit availability, so that the model can have better performance. The optimized model mainly includes the following steps.

Step 1. Train individuals in the population, calculate errors and fitness, and determine fitness values for each individual in the population. Here, individual refers to the factors affecting the availability of farmers' formal credit. The error can be calculated as follows.

$$\mathrm{Err} = \sum_{j=1}^{Q}\sum_{i=1}^{P}\left(a_{ji} - b_{ji}\right)^2, \tag{3}$$

where a_{ji} represents the actual output, b_{ji} represents the ideal output, P represents the number of nodes in the output layer, and Q represents the number of samples. Here, the fitness function is defined as follows.

$$\mathrm{Fit} = \frac{1}{\mathrm{Err}}. \tag{4}$$

Step 2. Population initialization. Parameters such as population size are initialized by binary coding.

Step 3. Evaluate each individual in the population, calculate the individual fitness value according to the global error of the neural network corresponding to each group of weights and thresholds, and output the individual meeting the fitness requirements. For those individuals with low fitness, crossover and mutation are then performed, new chromosomes are generated, and fitness is calculated again and sorted according to fitness value. The higher the sequence, the better the chromosome.

Step 4. According to the sorting results of the previous step to calculate the selection probability corresponding to each chromosome.

According to the sorting results, the selection probability corresponding to each chromosome is calculated according to equation (5). The partial derivative of the error function with respect to the jth element of the nth layer of the neural network is defined as follows.

$$\frac{\partial \mathrm{Err}}{2\partial b_j^n} = b_j - a_j. \tag{5}$$

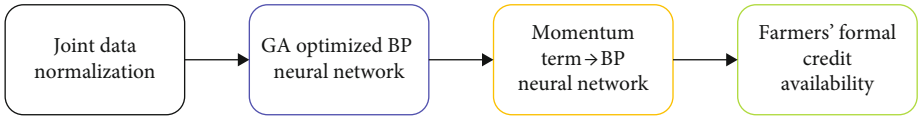


Figure 1: Flowchart of farmers' formal credit availability prediction model.

Table 1: Basic information of farmers.

Age of householder						Gender	
20-30	31-40	41-50	51-60	>60	Average age	Male	Female
15	44	82	69	55	48.6	242	23

Table 2: Relationship between age and formal credit availability.

Formal credit availability	Age ≤ 30	31-40	41-50	51-60	Age > 60
0	7	24	48	39	37
1	8	20	34	30	18
Total	15	44	82	69	55
Formal credit probability	53.30%	45.50%	41.50%	43.50%	32.70%

Table 3: Relationship between education level and formal credit availability.

Formal credit availability	Education level of farmer			
	Elementary school	Middle school	High school	College
0	90	72	9	4
1	18	51	14	7
Total	108	123	23	11
Formal credit probability	16.70%	41.50%	60.90%	63.60%

Step 5. The selection process is performed by using the roulette wheel method. Let c_i represent chromosome number and cp_i represent cumulative probability; that is,

$$\begin{cases} cp_0 = 0, \\ cp_i = \sum eval(c_i) \end{cases}, i = 1, 2, \cdots 50 \tag{6}$$

where eval() function evaluates the string as a valid expression and returns the result and then the random number R is taken from 0 to 1. If $cp_{i-1} \le R \le cp_i$ is satisfied, chromosome i is selected.

Step 6. >In farmers' formal credit availability prediction, the evolutionary process includes two stages: gradual stage and mutation stage. The former has strong intersection, and the latter has strong variation. If the fitness value is large, a smaller value is given to the crossover probability and mutation probability to increase the opportunity left by the individual. On the contrary, a larger value is given to the crossover probability and mutation probability.

Step 7. Go to Step 3, output the corresponding weights and thresholds if the requirements are met, and proceed to the next step. If the requirements are not met, continue the cycle process.

Step 8. The optimal weights and thresholds obtained by the above operations are given to the neural network to start the neural network training process.

Compared with the standard BP neural network, the improved BP neural network by genetic algorithm mainly includes three parts, which are determining the structure of BP neural network, determining the optimal initial weight and bias value by genetic algorithm, and predicting by BP neural network. In the part of the main basis for determining the model structure to solve the problem of input and output of the model parameters, the genetic algorithm is used to determine the initial weights and bias value and to calculate the individual fitness value, through the selection, crossover, and mutation operation to determine the optimal weights and bias of individual corresponding values. Since each individual contains the ownership of the neural network value and threshold value, we only need to assign the corresponding value of the individual to the neural network, and the farmers' formal credit availability prediction of BP

Table 4: Relationship between business type and formal credit availability.

Formal credit availability	Business type				
	Mainly agricultural	Mainly agricultural and engaged in other	Unclear	Mainly nonagricultural	Totally nonagricultural
0	79	50	16	6	5
1	28	38	25	10	6
Total	107	88	41	16	11
Formal credit probability	26.2%	43.2%	61.0%	62.5%	54.6%

Table 5: Relationship between annual household income and formal credit availability.

Formal credit availability	Annual household income					
	<10,000	10000-30,000	30000-50,000	50000-70,000	70000-90,000	>90,000
0	23	63	53	19	10	6
1	2	10	13	22	21	23
Total	25	73	66	41	31	29
Formal credit probability	8.0%	13.7%	20.0%	53.7%	67.7%	79.3%

Table 6: Variables selection.

Type	Variable	Mean	Standard deviation	Minimum	Maximum
Family characteristic	Age of household head	48.6	10.27	25	65
	Education level[a]	1.79	0.78	1	4
	Burden of family[b]	0.46	0.27	0	1
Agricultural production characteristic	Business type	2.02	1.11	1	5
	Land area	6.48	12.56	1	200
	Agricultural insurance[c]	1.96	0.88	1	2
	Agricultural subsidy[d]	1	0.47	1	2
Physical capital characteristic	Fixed assets	2.25	1.12	1	5
	Annual household income	7.81	0.92	3.41	15.16
	Annual household expenditure	5.46	1.02	2.46	9.73
Social capital	Expected loan amount	2.75	1.14	1	5
	Degree of intimacy with relatives and friends[e]	3.74	0.82	1.25	5
	Degree of trust with relatives and friends	3.12	0.81	1	5

[a]Education level: 1 = elementary school; 2 = middle school; 3 = high school; 4 = college. [b]Burden of family: (Number of old men + Number of children)/Total number of households. [c]Agricultural insurance: 1 = buy agricultural insurance; 2 = do not buy agricultural insurance. [d]Agricultural subsidy: 1 = receive subsidy; 2 = without subsidy. [e]Degree of intimacy with relatives and friends: 1, 2, 3, 4, and 5 indicate that the degree increases in turn.

neural network is mainly based on the final structure of the network determined in the previous two steps to predict the new data.

3.3. Improvement of Standard BP Neural Network. Adding momentum term into weights can effectively prevent BP neural network from falling into local minimum and shorten learning time. Since the gradient of the error surface of the standard BP neural network is not uniform, the step size of the standard algorithm is unchanged; if the small step size is adopted, the convergence speed of the error surface in the flat region is slow [28]. Nevertheless, when the step size is large, it is easy to produce oscillation in farmers' formal credit availability prediction. While the BP neural network algorithm with momentum term is improved to solve this problem, consider the weight modification of standard BP neural network, which is summarized as follows.

$$\begin{cases} w_{kl}'{}'(t_0+1) = w_{kl}'{}'(t_0) + \eta \sum_{p1=1}^{p} \mu_{kl}^{p1} x''k^{p1}, \\ w_{jk}'(t_0+1) = w_{jk}'(t_0) + \eta \sum_{p1=1}^{p} \mu_{jk}^{p1} x'k^{p1}, \\ w_{ij}(t_0+1) = w_{ij}(t_0) + \eta \sum_{p1=1}^{p} \mu_{ij}^{p1} x'k^{p1}, \end{cases} \tag{7}$$

where x is input signal of the BP neural network, η is

Table 7: Farmers' formal credit availability prediction indicators.

Target layer	Primary indicator	Secondary indicator
Farmers' formal credit availability	Family characteristic	Education level
		Burden of family
		Business type
	Agricultural production characteristic	Agricultural insurance
		Agricultural subsidy
	Physical capital characteristic	Fixed assets
		Annual household income
		Annual household expenditure
	Social capital	Degree of intimacy with relatives and friends
		Degree of trust with relatives and friends

Table 8: Weight of indicators.

Primary indicator	Weights	Secondary indicator	Weights
Family characteristic	0.197	Education level	0.451
		Burden of family	0.556
Agricultural production characteristic	0.256	Business type	0.368
		Agricultural insurance	0.108
		Agricultural subsidy	0.099
Physical capital characteristic	0.314	Fixed assets	0.292
		Annual household income	0.404
		Annual household expenditure	0.387
Social capital	0.231	Degree of intimacy with relatives and friends	0.392
		Degree of trust with relatives and friends	0.366

learning rate (step size), $\mu_{gh}^{p1}(gh = ij, jk, kl)$ is the error of each layer, $p1$ is the $p1$th sample, and t_0 is the t_0th training. The damping term δ is added into the equation (7), and $\Delta w_{sq}(t_0)$ is the modification direction of the weight at the previous moment.

$$\begin{cases} w_{kl}'{}'(t_0+1) = w_{kl}'{}'(t_0) + \eta \sum_{p1=1}^{p} \delta_{kl}^{p1} x''k^{p1} + \delta \Delta w_{kl}'{}'(t_0), \\ w_{jk}'(t_0+1) = w_{jk}'(t_0) + \eta \sum_{p1=1}^{p} \delta_{jk}^{p1} x'k^{p1} + \delta \Delta w_{jk}'(t_0), \\ w_{ij}(t_0+1) = w_{ij}(t_0) + \eta \sum_{p1=1}^{p} \delta_{ij}^{p1} x'k^{p1} + \delta \Delta w_{ij}(t_0). \end{cases} \tag{8}$$

Consider the change of two adjacent training errors, which is summarized as follows.

$$\begin{aligned} \Delta E &= E_{\text{tot}}(t_0) - E_{\text{tot}}(t_0 - 1), \\ FCA &= \frac{1}{2} \sum_{p1=1}^{p} \sum_{l=1}^{m} \left(e_i^{p1} - y_l^{p1} \right)^2, \end{aligned} \tag{9}$$

Table 9: Parameter settings.

Parameter	Setting
Number of neurons in input layer	14
Number of neurons in output layer	3
Number of hidden layer	1
Number of hidden layer neurons	12
Transfer function between input layer and hidden layer	logsig
Transfer function between hidden layer and output layer	logsig
Training function	traingdx
Error training target	0.0001
Training times	5000
Learning rate	0.2
Damping term	0.6

where E_{tot} is the error of the sample population and FCA is the farmers' formal credit availability.

In equation (9), p is the total number of sample signals, m is the number of output signals, e_i^{p1} is the expected value of input, and y is the calculated value of input.

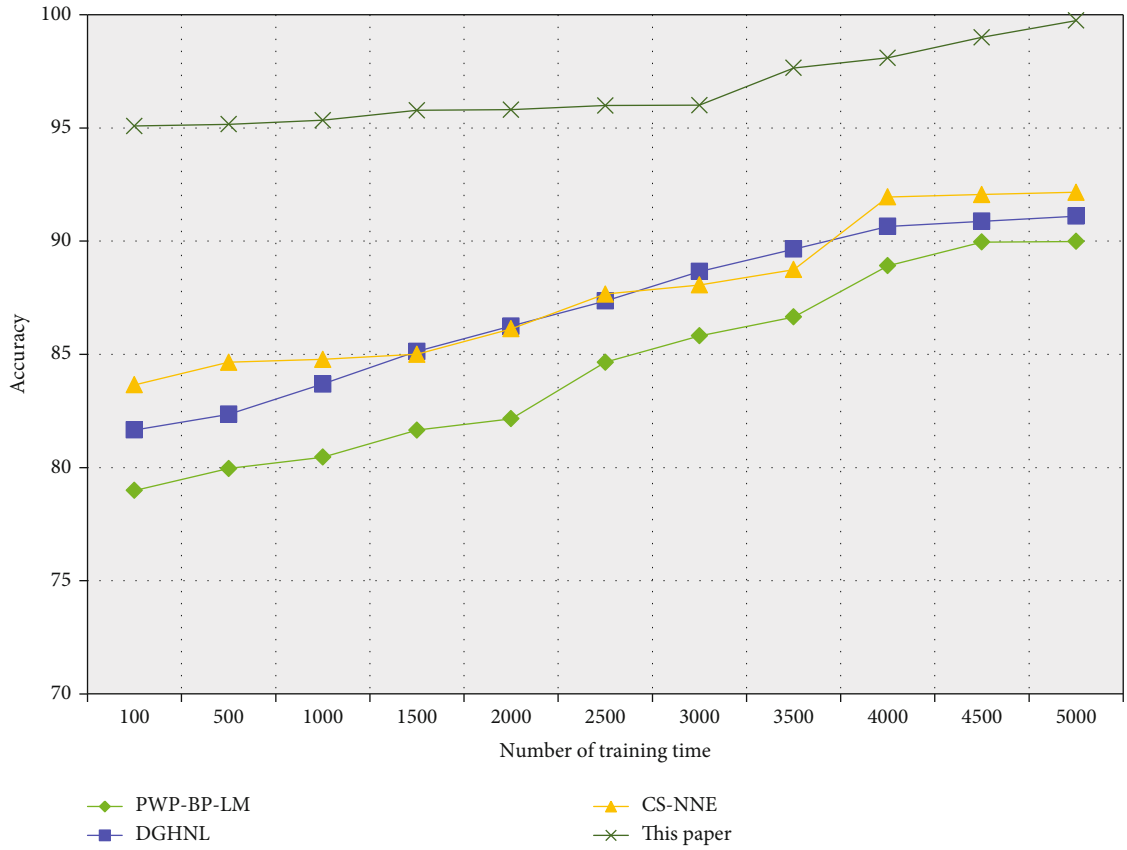


Figure 2: Performance comparison of accuracy.

Given this, the flowchart of farmers' formal credit availability prediction model is shown in Figure 1.

4. Experiments

4.1. Data Source. The analysis of farmers' formal credit availability factors is to establish the indicator system of formal credit availability. The indicator factors obtained based on objective data are scientific. The sample data of this study came from the questionnaire survey of A town and B town, Jiutai District, Changchun City, Jilin Province, in November 2021. Household interviews and questionnaires were conducted in selected villages. A total of 392 questionnaires were issued and 348 were recovered. 265 samples from 25 to 65 years old were selected according to their age. The basic information of farmers is shown in Table 1.

According to the survey data, the surveyed farmers ranged in age from 41 to 60 years old, with an average age of 48.6 years old. There were 242 male householders and 23 female householders. When surveyed whether they had a loan in the past two years, 106 people chose "yes" and 159 people chose "no." Among the most important reasons for not applying, 46 people chose "high interest rate, not cost-effective," 120 people chose "no collateral, no guarantee," and 15 people chose "troublesome procedures, attached many conditions." The other 85 people did not make a choice.

From the overall questionnaire, most of the farmers have loan needs, and the actual people cannot meet all their needs. In the evaluation of financial institutions, most of the farmers are not satisfied with the loan policies of banks. However, to the extent of understanding the loan policies, the farmers are basically half-aware. It can be seen that the farmers have no initiative in financial loans and their own level of knowledge needs to be strengthened.

4.2. Sample Characteristics and Loan Availability Descriptive Analysis. According to the analysis of sample data obtained from the survey, age is negatively correlated with the availability and age of farmers' loans, and the availability exceeds 44% between 30 and 40 years old. Under 30 years old, due to the educational level and the nature of work, the availability is relatively high, while over 60 years old has a strong risk aversion and low availability of loans. For farmers over 70 years old, the bank generally does not accept loan applications. Under 25 years old and over 70 years old are not considered in this sample.

There is no gender difference between farmers in terms of family. At the same time, according to the actual survey data of Jiutai District, there is not much difference between male and female farmers. From the perspective of culture and education, people with a high degree of education will generally have a greater social contact, have advantages in new agricultural business awareness and ability, are more likely to understand and grasp better financial policies, and have better planning for loan investment. This paper makes a statistical analysis of the relationship between age, education level, business type, and annual household income, as shown in Tables 2–5.

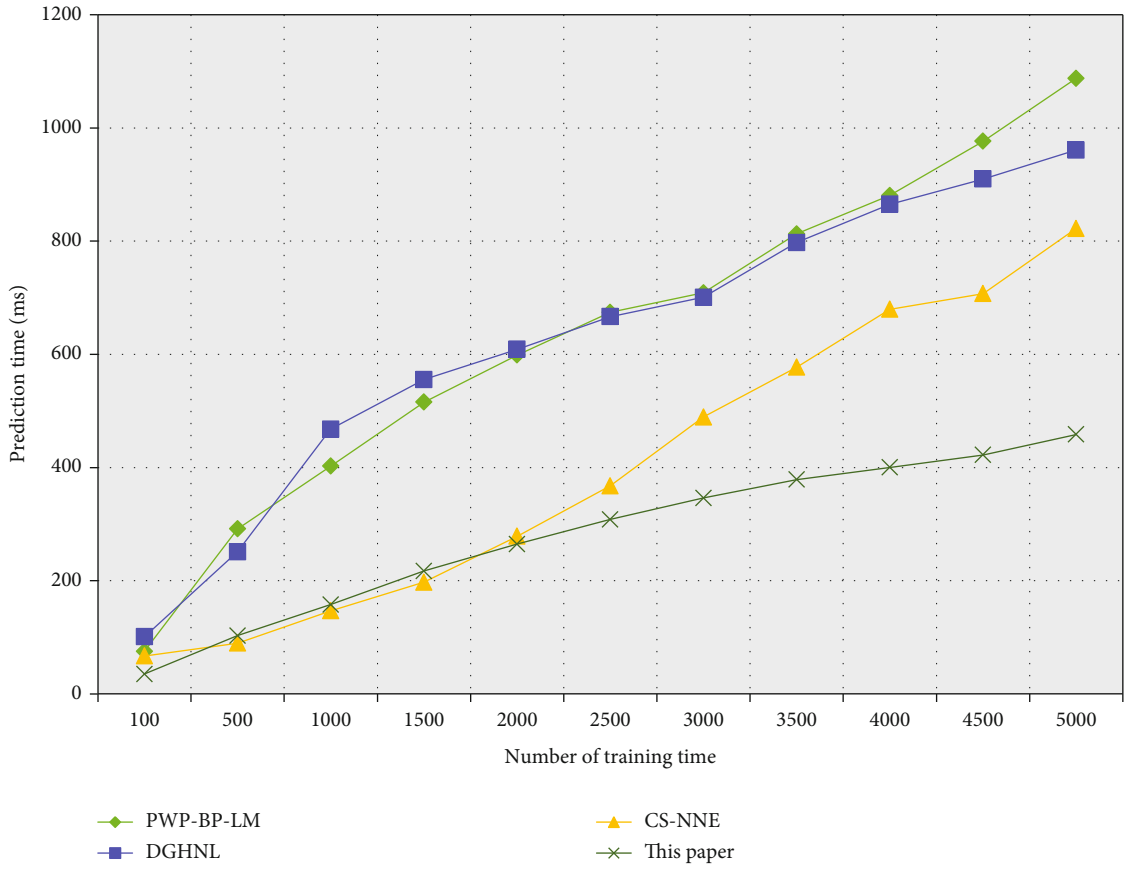


Figure 3: Performance comparison of prediction time.

4.3. Variables Selection. This paper measures farmers' ability to get formal credit from four aspects: family characteristics, agricultural production characteristics, physical capital characteristics, and social capital, as shown in Table 6.

4.4. Impact Indicators of Farmers' Formal Credit Availability. Through the collation and analysis of data, the significant factors affecting the availability of farmers' formal credit in Jiutai District are obtained, which are used as evaluation indicators for the ability of obtaining formal credit. The following formal credit availability prediction indicator system is established, as shown in Table 7.

Equation (7) is used to calculate the weight of each indicator, as shown in Table 8.

On the basis of analyzing the impact factors of farmers' obtaining formal credit, a comprehensive evaluation system is established by selecting several indicators to classify the availability of formal credit and obtain the ability of obtaining formal credit. Ultimately, farmers' formal credit availability can be predicted through equation (8).

4.5. BP Neural Network Training Parameters Setting. Before BP neural network training, the setting of parameters has a great influence on the accuracy and stability of neural network.

4.5.1. Training Times. The training times have direct influence on the accuracy and generalization ability of BP neural network. If the training times are too large or too small, there will be problems. If the training times are too large, it will easily cause overlearning, and if the training times are too small, it will not reach the target error set. To find the best training times, it can be achieved by the change of error. In this paper, the maximum training time is set as 5000 times; when the training time is exceeded, the learning will be terminated even if the network is not within the expected error.

4.5.2. Training Target Error. In the process of BP neural network training, the expected error is obtained by comparing the final network output value with the expected output value. Any prediction model wants the prediction accuracy to get 100%, but it cannot be satisfied under realistic conditions. When the error is reduced to a minimum, it costs longer network training time at least, but the prediction accuracy and prediction time must be weighed in the construction and application of the model. Therefore, from the practical point of view, it is feasible to allow the existence of certain error. According to the actual needs, the expected error determined in this paper is 0.0001.

4.5.3. Transfer Function. The transfer functions of BP neural network mainly include purelin function, tansig function, and logsig function. S-type function converges faster than linear function, and the nonlinear mapping capability of BP neural network is reflected by S-type function. As the

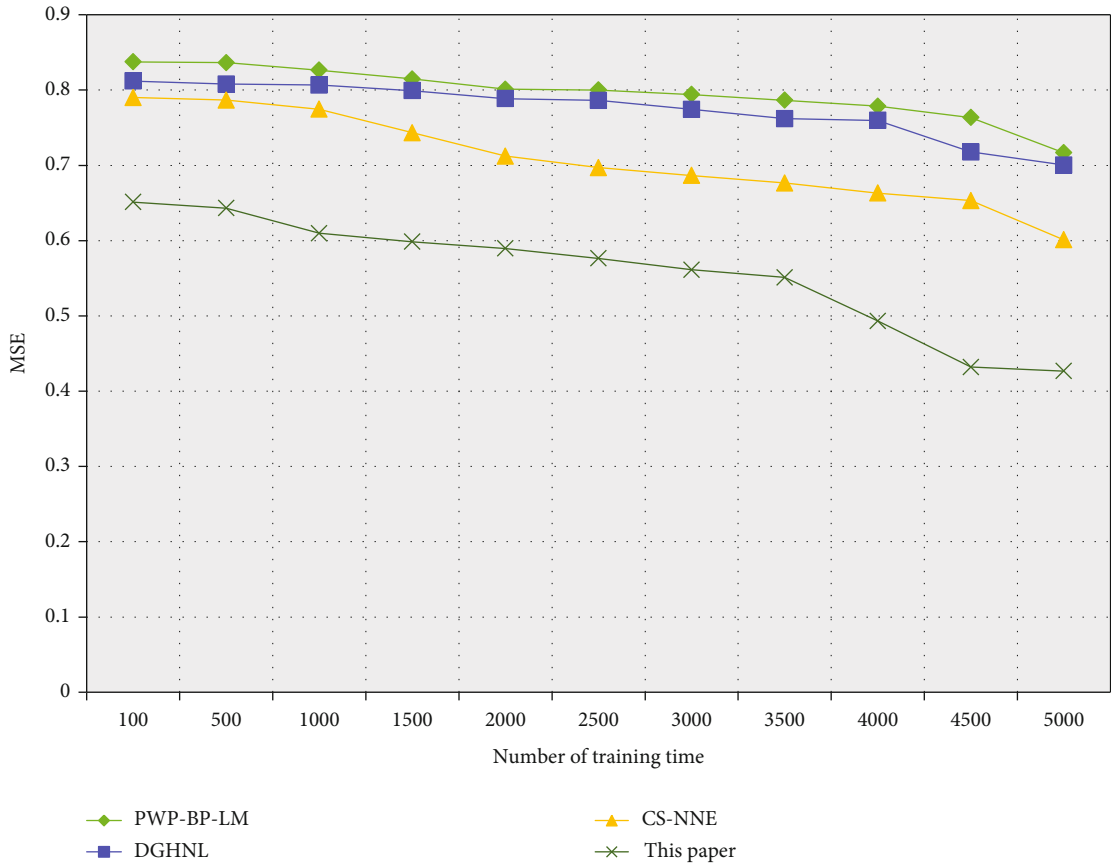


Figure 4: Performance comparison of MSE.

input data is normalized within the range of (0,1), logsig is selected as the transfer function of the hidden layer and output layer in this paper.

4.5.4. Training Function. The commonly used training functions of BP neural network include gradient descent method, L-M momentum method, fast BP algorithm, and traingdx optimization method. The main difference of these algorithms is the difference in learning efficiency. Traingdx optimization method is an optimization method that combines gradient descent optimization with adaptive learning rate method. This method has fast convergence speed, good learning effect, and high precision, so this paper selects the traingdx optimization method as the training function.

4.5.5. Learning Rate. Learning rate refers to the weight adjustment range of each training network layer, which affects the convergence speed of the network. Since the adaptive learning rate adjustment strategy is adopted in this paper, the learning rate can be adjusted according to the variation range of the error curve. After repeated tests, the learning rate is finally selected as 0.2.

Therefore, parameter settings of the BP neural network model are shown in Table 9.

4.6. Comparison Analysis. To verify the effectiveness of the proposed farmers' formal credit availability prediction model, P2P-BP-LM [19], DGHNL [20], and CS-NNE [21] were used for comparison in terms of accuracy, prediction time, and mean squared error (MSE). MSE is the average sum of squares of deviations from the true value.

As can be seen from Figure 2, the farmers' formal credit availability prediction model proposed in this paper is of high accuracy. Compared with other three baselines, the predicted accuracy has been more than 90%. This is because through descriptive analysis and variable selection of research data, the algorithm proposed in this paper makes a significant analysis by empirical method and obtains the main factors affecting the formal credit of farmers. By analyzing the factors of farmers themselves, financial institutions, and external factors, this paper classifies the indicators of farmers into four aspects: family characteristics, agricultural production characteristics, physical capital characteristics, and social capital to measure the ability of farmers to obtain formal credit. As indicated in Figure 3, with the increasing of training times, the prediction time of the model in this paper increases, but the range is small and tends to be stable. In contrast, the prediction time of PWP-BP-LM and DGHNL algorithms show an obvious upward trend. The prediction time of CS-NNE algorithm has a good performance in the first 2000 training sessions, but after 2500 training sessions, the prediction time increases exponentially. This also confirms the validity of the proposed algorithm in prediction time. As shown in Figure 4, with the

increasing number of training times, the MSE of the proposed model has always been decreasing, indicating the good performance of the farmers' formal credit availability model.

5. Conclusion

This paper studies the multifactor analysis based on BP neural network to analyze the formal credit to assist the new agricultural business. At first, since the data types of farmers' relevant information indicators are not unified, there are characters and numbers, and the amount of data changes greatly, and this data source is easy to bias the prediction results and affect the accuracy of prediction, so a joint normalization method is proposed in this paper. Then, genetic algorithm is introduced to avoid BP neural network falling into local minimum and can speed up the convergence speed of the network. Meanwhile, momentum term is added into weights that can effectively prevent BP neural network from falling into local minimum and shorten learning time. Finally, a multifactor farmers' formal credit availability prediction model based on BP neural network is established. The experiments results reveal that the proposed model outperforms baselines with respect to accuracy, prediction time, and MSE.

Nevertheless, this paper only puts forward relevant suggestions to assess farmers' ability to get formal credit from the perspective of meeting farmers' demand for loans but does not analyze how to effectively solve the difficulty of farmers in China to get formal credit from the perspective of fund supply. In addition, this paper is limited by the regional limitations of the sample, and its representativeness is limited. If we can have larger data resources, it will certainly improve the accuracy. Meanwhile, the situation of farmers is a dynamic process, so the application of static method in this paper to predict and study is slightly inadequate. These will be studied in the next step.

Data Availability

All data used to support the findings of the study is included within this paper.

Conflicts of Interest

The authors declare no conflicts of interest in this paper.

Acknowledgments

This work was supported by the Humanities and Social Science Research Project of Education Department of Jilin Province (grant no. JJKH20210429SK).

References

[1] S. Yang, H. Wang, Z. Wang, M. A. Koondhar, L. Ji, and R. Kong, "The nexus between formal credit and e-commerce utilization of entrepreneurial farmers in rural China: a mediation analysis," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 16, no. 4, pp. 900–921, 2021.

[2] S. Chen, E. Luo, L. Alita, X. Han, and F. Nie, "Impacts of formal credit on rural household income: evidence from deprived areas in western China," *Journal of Integrative Agriculture*, vol. 20, no. 4, pp. 927–942, 2021.

[3] J. Dong, W. Liang, Y. Fu, W. Liu, and S. Managi, "Impact of devolved forest tenure reform on formal credit access for households: Evidence from Fujian, China," *Economic Analysis and Policy*, vol. 71, pp. 486–498, 2021.

[4] F. Yang, Y. Jiang, and K. P. Paudel, "Impact of credit constraints from formal financial institutions on rural residents' Health in China," *Healthcare*, vol. 9, no. 1, 2020.

[5] X. Li and X. Huo, "Impacts of land market policies on formal credit accessibility and agricultural net income: evidence from China's apple growers," *Technological Forecasting and Social Change*, vol. 173, 2021.

[6] L. Lin, W. Wang, C. Gan, D. A. Cohen, and Q. T. T. Nguyen, "Rural Credit Constraint and Informal Rural Credit Accessibility in China," *Sustainability*, vol. 11, no. 7, 2019.

[7] L. Cheng, Y. Cui, K. Duan, and W. Zou, "The Influence of New Agricultural Business Entities on Farmers' Employment Decision," *Land*, vol. 11, no. 1, 2022.

[8] L. Cheng, W. Zou, and K. Duan, "The influence of new agricultural business entities on the economic welfare of farmer's families," *Agriculture-Basel*, vol. 11, no. 9, 2021.

[9] Y. Xu, "Research on the improvement of accounting work quality of new agricultural business entities under the background of big data," *Acta Agriculturae Scandinavica Section B-Soil and Plant Science*, vol. 72, no. 1, pp. 440–453, 2022.

[10] Z. Wang, J. Liu, T. Li, J. Chao, and X. Gao, "Factors affecting new agricultural business entities' adoption of sustainable intensification practices in China: evidence from the main apple-producing areas in the Loess Plateau," *Agronomy-Basel*, vol. 11, no. 12, 2021.

[11] W. Dai, "Rural financial information service platform under smart financial environment," *IEEE Access*, vol. 8, pp. 199944–199952, 2020.

[12] S. Zhu, Y. Chen, and W. Wang, "Risk Assessment of Biological Asset Mortgage Loans of China's New Agricultural Business Entities," *Complexity*, vol. 2020, 12 pages, 2020.

[13] J. Chen, A. L. Katchova, and C. Zhou, "Agricultural loan delinquency prediction using machine learning methods," *International Food and Agribusiness Management Review*, vol. 24, no. 5, pp. 797–812, 2021.

[14] H. Sun, V. Hartarska, L. Zhang, and D. Nadolnyak, "The influence of social capital on farm household's borrowing behavior in rural China," *Sustainability*, vol. 10, no. 12, 2018.

[15] H. Wang and J. Wang, "Study of fuzzy-clustering-based unweighted farmers' credit risk estimation and its application," *Journal of Nonlinear and Convex Analysis*, vol. 20, no. 5, pp. 937–948, 2019.

[16] Y. Li, H. Zhang, S. Wu, L. Lei, Y. Yu, and B. Wu, "Does farmland management scale influence credit availability?," *Evidences from Three Provinces of China, Environmental Engineering and Management Journal*, vol. 20, no. 11, pp. 1865–1874, 2021.

[17] S. Pang, M. Wei, and Z. Wen, "Farmers credit optimization decision model and applications based on common risk guarantee fund," *Neural Computing & Applications*, vol. 33, no. 9, pp. 4053–4064, 2021.

[18] C. K. Dhakal, C. L. Escalante, and C. Dodson, "Heterogeneity of farm loan packaging term decisions: a finite mixture

approach," *Applied Economics Letters*, vol. 26, no. 18, pp. 1528–1532, 2019.

[19] Z. Ma, W. Hou, and D. Zhang, "A credit risk assessment model of borrowers in P2P lending based on BP neural network," *Plos One*, vol. 16, no. 8, 2021.

[20] P. Plawiak, M. Abdar, J. Plawiak, V. Makarenkov, and U. R. Acharya, "DGHNL: a new deep genetic hierarchical network of learners for prediction of credit scoring," *Information Sciences*, vol. 516, pp. 401–418, 2020.

[21] W. Yotsawat, P. Wattuya, and A. Srivihok, "A novel method for credit scoring based on cost-sensitive neural network ensemble," *IEEE Access*, vol. 9, pp. 78521–78537, 2021.

[22] A. Mercep, L. Mrcela, M. Birov, and Z. Kostanjcar, "Deep Neural Networks for Behavioral Credit Rating," *Entropy*, vol. 23, no. 1, 2021.

[23] J. Chen and S. Huang, "Evaluation model of green supply chain cooperation credit based on BP neural network," *Neural Computing & Applications*, vol. 33, no. 3, pp. 1007–1015, 2021.

[24] X. Li and Y. Sun, "Application of RBF neural network optimal segmentation algorithm in credit rating," *Neural Computing & Applications*, vol. 33, no. 14, pp. 8227–8235, 2021.

[25] Y. Wu, X. Li, Q. Liu, and G. Tong, "The analysis of credit risks in agricultural supply chain finance assessment model based on genetic algorithm and backpropagation neural network," *Computational Economics*, 2021.

[26] J. Lv, X. Wang, K. Ren, M. Huang, and K. Li, "ACO-inspired information-centric networking routing mechanism," *Computer Networks*, vol. 126, pp. 200–217, 2017.

[27] C. Ding, L. Chen, and B. Zhong, "Exploration of intelligent computing based on improved hybrid genetic algorithm," *Cluster Computing-The Journal of Networks Software Tools and Applications*, vol. 22, no. S4, pp. 9037–9045, 2019.

[28] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Transactions on Cybernetics*, vol. PP, pp. 1–13, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8587896, 7 pages
https://doi.org/10.1155/2022/8587896

Research Article

Integrating Cross-lingual Ontologies through Co-Evolutionary Algorithm

Lili Huang[1] **and Leong Ko**[2]

[1]*International College Fujian Agriculture and Forestry University, No. 15 Shangxiadian Road, Changshan District, Fuzhou, Fujian, 350002, China*
[2]*School of Languages and Cultures University of Queensland Bldg. 32 Gordon Greenwood Building, The University of QLD, Brisbane, Queensland, 4072, Australia*

Correspondence should be addressed to Lili Huang; lilihuang@vip.163.com

Received 13 January 2022; Accepted 25 February 2022; Published 5 April 2022

Academic Editor: Jianhui Lv

To support the collaborations among intelligent applications, it is necessary to integrate various ontologies which are developed and maintained by different organizations. One of the challenges is that different ontologies in the same application domain might use different languages to describe the same concept, which yields the cross-lingual heterogeneity problem, i.e., how to map two identical entities in different languages. To address this problem, this work proposes a problem-specific co-evolutionary algorithm (CoEA)-based matching technique. In particular, we first propose a parallel aggregating framework to aggregate different SMs and then construct a continuous optimization model for defining the problem of cross-lingual ontology integration. To better trade off the algorithm's exploitation and exploration, we use two competitive subpopulations to, respectively, execute the exploitation and exploration. The experiment utilizes OAEI's multifarm track for testing purpose, and the experimental results show that CoEA is able to effectively integrate various cross-lingual ontologies.

1. Introduction

With the rapid development of Semantic Web [1, 2], various ontologies have been developed in diverse domains to annotate the data. To implement the intelligent applications' collaboration, it is necessary to integrate various ontologies which are developed and maintained by different organizations. One of the challenges is that different ontologies in the same application domain might use different languages to describe the same concept, which yields the cross-lingual heterogeneity problem. To address this issue, we need to map two identical entities in different languages, which is the so-called cross-lingual ontology matching [3]. When matching two ontologies, it is important to use the similarity measure (SM) to distinguish the heterogeneous entities [4], and usually, different SMs should be aggregated to make their advantages and disadvantages complement each other, which is of help to improve the confidence of the result. Currently, the parallel framework owns such merits as its flexibility of tuning various weights and the relative independence on different matchers in terms of their executing processes. Before aggregating SMs, their corresponding similarity matrices should be calculated independently first. The similarity matrix's row and column are, respectively, two ontology's entities, and its elements are the corresponding entities' similarity value. Then, these matrices are aggregated into one matrix by using the aggregating weights. Finally, a threshold is used to filter the correspondences with low similarity values.

Since it is a complex optimizing task of optimizing a cross-lingual ontology alignment, metaheuristics approaches, such as evolutionary algorithm (EA) [5],

become the popular methods of determining the high-quality alignments. Being inspired by the success of metaheuristics approaches in the cross-lingual ontology matching domain, this work proposes a co-evolutionary algorithm (CoEA) to determine the cross-lingual ontology alignment. CoEA uses the compact mechanism and co-evolutionary mechanism to overcome two shortcomings of classic EA, i.e., high computational cost and premature convergence. In particular, we make the following contributions: a parallel aggregating framework of matching cross-lingual ontologies is presented; a continuous optimization model is built to define the cross-lingual ontology matching problem; a problem-specific CoEA is proposed to optimize the cross-lingual ontology alignment's quality, which uses two competitive subpopulations to trade off algorithm's exploitation and exploration.

The rest of this study is organized as follows. After defining the cross-lingual ontology matching problem (Section 2), three kinds of SMs are introduced (Section 3), then CoEA is presented (Section 4), and the experimental results are shown (Section 5). Finally, the conclusion is drawn (Section 6).

2. Cross-Lingual Ontology Integrating Problem

An ontology consists of concepts, properties, and axioms, and an ontology alignment is a mapping set between two heterogeneous ontologies. A mapping is a 3-tuple (c_1, c_2, $simValue$), where c_1 and c_2 are, respectively, two ontologies' entities, and $simValue \in\in [0, 1]$ is their similarity [6, 7]. Given an alignment A and a reference alignment A_{ref}, A's quality can be measured with f-measure [8]:

$$\begin{aligned} \text{recall}(A) &= \frac{\left|A \cap A_{ref}\right|}{\left|A_{ref}\right|}, \\ \text{precision}(A) &= \frac{\left|A \cap A_{ref}\right|}{|A|}, \\ f-\text{measure}(A) &= \frac{2 \times \text{recall}(A) \times \text{precision}(A)}{\text{recall}(A) + \text{precision}(A)}, \end{aligned} \tag{1}$$

where | | is the cardinality of a particular set. The cross-lingual ontology integrating problem is modelled as a continuous optimization problem. In particular, its objective is to maximize the f-measure of the cross-lingual ontology alignment, and its decision variable $= \{x_1, x_2 \ldots\}^T$, where $x_i \in [0, 1]$ is the ith aggregating weight, and their sum is equal to 1.

3. Similarity Measure

Generally, there are three broad categories of SM, i.e., syntactic SM, linguistic SM, and taxonomy SM [9–11]. Syntactic SM calculates two strings' similarity by measuring their edit distance. Levenshtein distance [12] is one of the popular syntactic SMs, which is defined as follows:

$$sim_{\text{Levenshtein}}(s_1, s_2) = \max\left(0, \frac{\min\left(|s_1|, |s_2|\right) - \text{d}(s_1, s_2)}{\min\left(|s_1|, |s_2|\right)}\right), \tag{2}$$

where is the character numbers of a string and $d()$ is two strings' edit distance.

Linguistic SM measures two words' similarity with the electronic dictionary, such as Wordnet [13,14]. Given two entities' label label_1 and label_2, their linguistic similarity value is defined as follows:

$$sim_{\text{Levenshtein}}(\text{label}_1, \text{label}_2) = \begin{cases} 1, & \text{if label}_1 \text{ and label}_2 \text{ are synonymous,} \\ 0.5, & \text{if label}_1 \text{ and label}_2 \text{ are hyponymous or hypernymous,} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Taxonomy SM uses two concepts c_1 and c_2's context to determine their similarity, which is defined as follows [15, 16]:

$$sim_{SF}(c_1, c_2) = \frac{sim_{\text{Levenshtein}}(\text{super}_1, \text{super}_2) + \sum sim_{\text{Levenshtein}}\left(sub_i, sub_j\right)}{2}, \tag{4}$$

where $super_1$ and $super_2$ are, respectively, c_1 and c_2's super classes and sub_i and sub_j are, respectively, their ith and jth subclasses.

In this study, by referring to Chen et al.'s work [17], we first use the Babelnet Translate [1] to translate the terminologies in various languages into English and then use the above three SMs to calculate their similarity values. All the similarity measures are executed in parallel to determine their corresponding similarity matrices, which store the similarity values of all correspondences. After that, CoEA is used to determine the aggregating weight for the similarity matrices in this parallel framework.

4. Co-Evolutionary Algorithm

4.1. Compact Encoding. CoEA uses the binary coding mechanism [18], which is of help to reduce the algorithm's computational complexity. Given a set of cut points $C' = \{c_1, c_2, \ldots, c_n\}$, we first sort it in ascending order as $C = \{c_1, c_2, \ldots, c_n\}$, and then, we can get the corresponding weight set through the following equation:

$$w_k = \begin{cases} c_1, & k = 1 \\ c_k - c_{k-1}, & 1 < k < p \\ 1 - c_{p-1}, & k = p \end{cases}. \quad (5)$$

After that, we use n cutting points to obtain $n+1$ aggregating weights.

This work uses two probability vectors (PVs) to describe the gene distribution on two competitive subpopulations, i.e., PV_{better} and PV_{worse}. In particular, PV_{better} describes the subpopulation whose elite solution owns higher fitness value, while PV_{worse} is lower. Each PV's elements are the real number in [0,1], which represents their corresponding gene bit's probability of being 1. With a PV, we are able to generate various solutions with the similar gene distributions.

4.2. Exploration and Exploitation. We apply different strategies on PV_{better} and PV_{worse}'s corresponding subpopulations. The former mainly focuses on the exploitation operator, while the latter uses the exploration strategy. For the sake of clarity, we show these two strategies in Algorithms 1 and 2.

Here, we introduce the exponential crossover operator (EC) [19] to implement CoEA's exploration and exploitation operators. Comparing with traditional crossover operator, EC generates the offspring by inheriting a complete sequential genes from two parents, which is more exploitative. Given two solutions, EC randomly copies a certain number of sequential bits' values from the first one to the second one. With respect to the exploration operator, we use EC to mix a newly generated solution $\text{solution}^{\text{new}}$ and the elite pollen $\text{solution}^{\text{elite}_{\text{better}}}$, while in the exploitation operator, we first mix two newly generated solutions $solution^p$ and $solution^q$ to obtain the mediate individual; then, we mix it with $solution^{new}$, which approximates the evolutionary operator of differential evolution algorithm (DE) [20] to ensure the algorithm's exploration.

```
solution^new = PV^better. generateSolution();
int num = round(ran(0, 1) × solution.length);
int index = 0;
for int i = 0; i < solution^new.length; i++ do
    if index + 1 > num then
        break;
    end if
if index + 1 > solution.length then
        index = 0;
    end if
    solution_i^new = solution_i^elite_better;
end for
```

Algorithm 1: Exploration.

```
solution^new = PV^worse.generateSolution();
solution^p = PV^worse.generateSolution();
solution^q = PV^worse.generateSolution();
int num = round(ran(0, 1) × solution.length);
int index = 0;
    for int i = 0; i < solution^p.length; i++ do
        if index + 1 > num then
        break;
        end if
        solution_i^p = solution_i^q;
end for
int num = round(ran(0, 1) × solution.length);
int index = 0;
    for int i = 0; i < solution^new.length; i++ do
    if index + 1 > num then
        break;
    end if
    if index + 1 > solution.length then
        index = 0;
        end if
    solution_i^new = solution_i^p;
end for
```

Algorithm 2: Exploitation.

4.3. Pseudocode of Co-Evolutionary Algorithm. Given the maximum generation *maxGen*, Algorithm 3 shows CoEA's pseudocode.

CoEA first initializes all the elements of two PVs as 0.5 and then uses them to initialize two elite solutions $\text{solution}^{\text{elite}_{\text{better}}}$ and $\text{solution}^{\text{elite}_{\text{worse}}}$. In each generation, CoEA, respectively, updates PV_{better} and PV_{worse} through exploration and exploitation strategies. At the end of each generation, if $\text{solution}^{\text{elite}_{\text{worse}}}$'s fitness value is higher than that of $\text{solution}^{\text{elite}_{\text{better}}}$, we will switch them and two PVs. If two PVs' Hamming distance [21] is smaller than 0.5, we will re-initialize it by setting all the elements as 0.5, which ensures the population's diversity. Finally, when reaching the maximum

```
********* Initialization **********
Initialize PV_better and PV_worse by setting all their elements as 0.5;
solution^elite_better = PV^better. generateSolution();
generateSolution();
int index = 0;
    ********** Evolution **********
    int gen = 0;
    while gen < max Gen do
        ********* Update PV^better **********
    solution^new = exploration;
    solution^winner = compete(solution^new, solution^elite_better);
        if solution^winner == solution^new then
        solution^elite_better = solution^new;
    end if
    for int i = 0; i < PV^better.length; i++ do
     if PV_i^better == 1 then
    PV_i^better == PV_i^better + 1/PV_i^better.length;
    else
        PV_i^better == PV_i^better − 1/PV_i^better.length;
    end if
    end for
        ********* Update PV^worse **********
        solution^new = exploitation();
     solution^winner = compete(solution^new, solution^elite_worse);
if solution^winner == solution^new then
        solution^elite_worse = solution^new;
    end if
        for int i = 0; i < PV^worse.length; i++ do
    if PV_i^better == 1 then
    PV_i^better == PV_i^better + 1/PV_i^better.length;
        else
    PV_i^better == PV_i^better − 1/PV_i^better.length;
    end if
end for
        ********* Competition *********
        if solution^elite_worse is better than solution^elite_better then
    switch solution^elite_worse and solution^elite_better;
    switch PV^worse and PV^better;
    end if
    if HammingDist(PV^better, PV^worse) < 0.5 then
    Initialize PV^worse by setting all the elements as 0.5;
        end if
        gen++;
    end while
return solution^elite_better;
```

Algorithm 3: Co-evolutionary algorithm.

iteration number $maxT = 3000$, the algorithm terminates and returns $\text{solution}^{\text{elite}_{\text{better}}}$. When the quality of alignment is 1.00, the algorithm can be terminated in advance. However, this situation is barely met in the experiment because it is difficult for the similarity measure to distinguish all the heterogeneous entity pairs. Therefore, this work executes the algorithm until it reaches the maximum generation.

5. Experiment

5.1. Experimental Configuration. In the experiment, Ontology Alignment Evaluation Initiative (OAEI)'s Multifarm track [2], which includes 45 ontology pairs in different languages, is used to test CoEA's performance, and Table 1 gives a brief descriptions on the testing cases.

Table 1: Descriptions on the ontologies in the testing cases.

Testing case	Description
ar-cn	Arabic (ar) vs. Chinese (cn)
ar-cz	Arabic (ar) vs. Chinese (cn)
ar-de	Arabic (ar) vs. German (de)
ar-en	Arabic (ar) vs. English
ar-es	Arabic (ar) vs. Spanish (es)
ar-fr	Arabic (ar) vs. French (fr)
ar-nl	Arabic (ar) vs. Dutch (nl)
ar-pt	Arabic (ar) vs. Dutch (nl)
ar-ru	Arabic (ar) vs. Russian (ru)
cn-nl	Chinese (cn) vs. Dutch (nl)
cn-de	Chinese (cn) vs. German (de)
cn-ru	Chinese (cn) vs. Russian (ru)
cn-fr	Chinese (cn) vs. French (fr)
cn-cz	Chinese (cn) vs. Czech (cz)
cn-pt	Chinese (cn) vs. Portuguese (pt)
cn-es	Chinese (cn) vs. Spanish (es)
cn-en	Chinese (cn) vs. English (en)
cz-de	Czech (cz) vs. German (de)
cz-en	Czech (cz) vs. English (en)
cz-es	Czech (cz) vs. Spanish (sp)
cz-fr	Czech (cz) vs. French (fr)
cz-nl	Czech (cz) vs. Dutch (nl)
cz-pt	Czech (cz) vs. Portuguese (pt)
cz-ru	Czech (cz) vs. Russian (ru)
de-ru	German (de) vs. Russian (ru)
de-es	German (de) vs. Spanish (es)
de-pt	German (de) vs. Portuguese (pt)
de-nl	German (de) vs. Dutch (nl)
de-fr	German (de) vs. French (fr)
de-en	German (de) vs. English (en)
en-es	English (en) vs. Spanish (es)
en-fr	English (en) vs. French (fr)
en-nl	English (en) vs. Dutch (nl)
en-pt	English (en) vs. Portuguese (pt)
en-ru	English (en) vs. Russian (ru)
es-ru	Spanish (es) vs. Russian (ru)
es-pt	Spanish (es) vs. Portuguese (pt)
es-nl	Spanish (es) vs. Dutch (nl)
es-fr	Spanish (es) vs. French (fr)
fr-nl	French (fr) vs. Dutch (nl)
fr-pt	French (fr) vs. Portuguese (pt)
fr-ru	French (fr) vs. Russian (ru)
nl-pt	Dutch (nl) vs. Portuguese (pt)
nl-ru	Dutch (nl) vs. Russian (ru)
pt-ru	Portuguese (pt) vs. Russian (ru)

5.2. Experimental Results. In the experiment, we compare CoEA with EA- and DE-based cross-lingual ontology matching techniques, whose results demonstrated in the tables are the mean values of 30 independent runs. In particular, Table 2 compares CoEA with CEA and CDE in terms of f-measure, and Table 3 carries out T-test [22] among three cross-lingual ontology matching techniques.

Table 2: Comparisons on the alignment's quality in terms of f-measure. The symbols f and std, respectively, stand for mean f-measure and standard deviation.

Testing case	CEA f (std)	HCEA f (std)	CoEA f (std)
ar-cn	0.28 (0.02)	0.28 (0.02)	0.35 (0.01)
ar-cz	0.37 (0.02)	0.37 (0.02)	0.41 (0.01)
ar-de	0.35 (0.01)	0.35 (0.01)	0.43 (0.02)
ar-en	0.36 (0.02)	0.36 (0.01)	0.45 (0.01)
ar-es	0.40 (0.02)	0.40 (0.03)	0.48 (0.01)
ar-fr	0.36 (0.02)	0.36 (0.01)	0.42 (0.02)
ar-nl	0.35 (0.02)	0.35 (0.02)	0.44 (0.01)
ar-pt	0.38 (0.03)	0.38 (0.01)	0.44 (0.01)
ar-ru	0.33 (0.03)	0.33 (0.01)	0.35 (0.02)
cn-cz	0.31 (0.02)	0.31 (0.02)	0.35 (0.01)
cn-de	0.36 (0.02)	0.33 (0.02)	0.41 (0.01)
cn-en	0.30 (0.02)	0.30 (0.01)	0.35 (0.02)
cn-es	0.37 (0.01)	0.37 (0.01)	0.42 (0.01)
cn-fr	0.35 (0.03)	0.35 (0.01)	0.42 (0.01)
cn-nl	0.41 (0.01)	0.41 (0.03)	0.49 (0.02)
cn-pt	0.36 (0.01)	0.36 (0.01)	0.46 (0.01)
cn-ru	0.27 (0.02)	0.27 (0.03)	0.39 (0.01)
cz-de	0.42 (0.02)	0.42 (0.01)	0.45 (0.02)
cz-en	0.44 (0.01)	0.44 (0.03)	0.51 (0.01)
cz-es	0.42 (0.03)	0.42 (0.01)	0.48 (0.01)
cz-fr	0.40 (0.01)	0.45 (0.01)	0.47 (0.01)
cz-nl	0.45 (0.03)	0.42 (0.01)	0.49 (0.02)
cz-pt	0.52 (0.02)	0.44 (0.02)	0.51 (0.01)
cz-ru	0.44 (0.01)	0.42 (0.01)	0.48 (0.01)
de-en	0.42 (0.01)	0.46 (0.01)	0.53 (0.01)
de-es	0.46 (0.02)	0.46 (0.01)	0.55 (0.01)
de-fr	0.46 (0.01)	0.43 (0.03)	0.50 (0.02)
de-nl	0.43 (0.01)	0.41 (0.01)	0.46 (0.01)
de-pt	0.45 (0.03)	0.42 (0.01)	0.53 (0.01)
de-ru	0.42 (0.01)	0.44 (0.02)	0.53 (0.02)
en-es	0.44 (0.02)	0.42 (0.01)	0.49 (0.02)
en-fr	0.42 (0.01)	0.42 (0.01)	0.53 (0.01)
en-nl	0.42 (0.01)	0.45 (0.01)	0.50 (0.01)
en-pt	0.45 (0.01)	0.41 (0.01)	0.51 (0.01)
en-ru	0.45 (0.03)	0.52 (0.02)	0.55(0.01)
es-fr	0.45 (0.02)	0.48 (0.01)	0.55 (0.02)
es-nl	0.48 (0.02)	0.52 (0.02)	0.55 (0.01)
es-pt	0.52 (0.01)	0.52 (0.01)	0.58 (0.01)
es-ru	0.50 (0.03)	0.50 (0.01)	0.63 (0.01)
fr-nl	0.46 (0.01)	0.46 (0.03)	0.52 (0.02)
fr-pt	0.49 (0.01)	0.49 (0.01)	0.58 (0.01)
fr-ru	0.45 (0.01)	0.45 (0.01)	0.55 (0.03)
nl-pt	0.52 (0.02)	0.52 (0.02)	0.58 (0.01)
nl-ru	0.48 (0.01)	0.48 (0.01)	0.58 (0.02)
pt-ru	0.47 (0.03)	0.47 (0.01)	0.51 (0.01)
average	0.41 (0.02)	0.41 (0.01)	0.48 (0.01)

As can be seen from Table 2, CoEA's results are much better than CEA- and HCEA-based ontology matching techniques. CoEA makes use of two different evolving strategies, which is of help to ensures its robustness. From Table 3, we can see that CoEA outperforms other two cross-lingual ontology matching techniques on 5% significant level.

Table 3: T-Test on the alignment's quality.

Testing case	CEA t-value (p value)	HCEA t-value (p value)
ar-cn	93.91 (0.003389)	93.91 (0.003389)
ar-cz	53.66 (0.005931)	53.66 (0.005931)
ar-de	107.33 (0.002966)	107.33 (0.002966)
ar-en	120.74 (0.002636)	190.91 (0.001667)
ar-es	107.33 (0.002966)	75.89 (0.004194)
ar-fr	63.63 (0.005001)	80.49 (0.003954)
ar-nl	120.74 (0.002636)	120.74 (0.002636)
ar-pt	56.92 (0.005592)	127.27 (0.002501)
ar-ru	16.64 (0.019105)	26.83 (0.011858)
cn-cz	53.66 (0.005931)	53.66 (0.005931)
cn-de	67.08 (0.004745)	107.33 (0.002966)
cn-en	53.03 (0.006001)	67.08 (0.004745)
cn-es	106.06 (0.003001)	106.06 (0.003001)
cn-fr	66.40 (0.004793)	148.49 (0.002144)
cn-nl	107.33 (0.002966)	66.56 (0.004782)
cn-pt	212.13 (0.001501)	212.13 (0.001501)
cn-ru	160.99 (0.001977)	113.84 (0.002796)
cz-de	31.81 (0.01)	40.24 (0.007907)
cz-en	148.49(0.002144)	66.40 (0.004793)
cz-es	56.92 (0.005592)	127.27 (0.002501)
cz-fr	148.49 (0.002144)	42.42 (0.007501)
cz-nl	33.28 (0.009561)	93.91 (0.003389)
cz-pt	80.49(0.003954)	93.91 (0.003389)
cz-ru	84.85 (0.003751)	127.27 (0.002501)
de-en	233.34 (0.001364)	148.49 (0.002144)
de-es	120.74 (0.002636)	190.91 (0.001667)
de-fr	53.66(0.005931)	58.24 (0.005465)
de-nl	63.63 (0.005001)	106.06 (0.003001)
de-pt	75.89 (0.004194)	233.34 (0.001364)
de-ru	147.58 (0.002157)	95.45 (0.003334)
en-es	53.03 (0.006001)	93.91 (0.003389)
en-fr	233.34 (0.001364)	233.34 (0.001364)
en-nl	169.70 (0.001876)	106.06 (0.003001)
en-pt	127.27 (0.002501)	212.13 (0.001501)
en-ru	94.86 (0.003355)	40.24 (0.007907)
es-fr	31.81 (0.01)	93.91 (0.003389)
es-nl	93.91 (0.003389)	40.24 (0.007907)
es-pt	127.27 (0.002501)	127.27 (0.002501)
es-ru	123.32 (0.002581)	275.77 (0.001154)
fr-nl	80.49 (0.003954)	49.92 (0.006375)
fr-pt	190.91 (0.001667)	190.91 (0.001667)
fr-ru	94.86 (0.003355)	94.86 (0.003355)
nl-pt	80.49 (0.003954)	80.49 (0.003954)
nl-ru	134.16 (0.002373)	134.16 (0.002373)
pt-ru	37.94 (0.008386)	84.85 (0.003751)

6. Conclusion

To aggregate the ontologies in different languages, this work proposes a CoEA-based cross-lingual ontology matching technique. We first propose a parallel framework of aggregating different SMs and then construct a continuous optimization model to define the problem of cross-lingual ontology integration. To solve this problem, for better trading off the algorithm's exploitation and exploration, we propose a CoEA with two competitive subpopulations. The experiment utilizes OAEI's multifarm track for testing, and the experimental results show that CoEA is able to effectively match cross-lingual ontologies.

In the future, we are interested in further improving the solution's quality by introducing the specific domain knowledge base to distinguish the heterogeneous entity pairs. Moreover, when the scale of ontology becomes large, the searching space of algorithm will definitely grows, which is a challenge for CoEA's efficiency. A feasible approach to face this challenge would be the semantic pruning technique or the ontology partitioning techniques [23], which can be used to improve the efficiency of CoEA when solving large-scale cross-lingual ontologies. [24].

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Education Department of Fujian Province (no. JAS160192).

References

[1] P. Shvaiko and J. Euzenat, "Ontology matching: state of the art and future challenges," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 1, pp. 158–176, 2013.

[2] I. Osman, S. Ben Yahia, and G. Diallo, "Ontology integration: approaches and challenging issues," *Information Fusion*, vol. 71, pp. 38–63, 2021.

[3] M. Abu Helou, M. Palmonari, and M. Jarrar, "Effectiveness of automatic translations for cross-lingual ontology mapping," *Journal of Artificial Intelligence Research*, vol. 55, pp. 165–208, 2016.

[4] X. Xue, J. Chen, and X. Yao, "Efficient user involvement in semiautomatic ontology matching," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 2, pp. 214–224, 2021.

[5] X. Xue, J. Lu, and J. Chen, "Using NSGA-III for optimising biomedical ontology alignment," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 3, pp. 135–141, 2019.

[6] G. Acampora, V. Loia, and A. Vitiello, "Enhancing ontology alignment through a memetic aggregation of similarity measures," *Information Sciences*, vol. 250, pp. 1–20, 2013.

[7] X. Xue and J. Chen, "Optimizing ontology alignment through hybrid population-based incremental learning algorithm," *Memetic Computing*, vol. 11, no. 2, pp. 209–217, 2019.

[8] C. J. Van Rijsbergen, "Foundation OF evaluation," *Journal of Documentation*, vol. 30, no. 4, pp. 365–373, 1974.

[9] S. Mani and S. Annadurai, "Explicit link discovery scheme optimized with ontology mapping using improved machine learning approach," *Studies in Informatics and Control*, vol. 30, no. 1, pp. 67–75, 2021.

[10] X. Xue and Y. Wang, "Optimizing ontology alignments through a memetic algorithm using both MatchFmeasure and unanimous improvement ratio," *Artificial Intelligence*, vol. 223, pp. 65–81, 2015.

[11] X. Xue and Q. Huang, "Generative adversarial learning for optimizing ontology alignment," *Expert Systems*, vol. 2022, pp. 1–12, Article ID 12936, 2022.

[12] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals," *Soviet Physics Doklady*, vol. 10, no. 8, pp. 707–710, 1966.
[13] G. A. Miller, "WordNet," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.
[14] E. Geller, M. Gajek, A. Reibach, and Z. Łapa, "Applicability of wordnet architecture in lexical borrowing studies," *International Journal of Lexicography*, vol. 34, no. 1, pp. 92–111, 2021.
[15] X. Xue and J.-S. Pan, "A segment-based approach for large-scale ontology matching," *Knowledge and Information Systems*, vol. 52, no. 2, pp. 467–484, 2017.
[16] M. AlMousa, R. Benlamri, and R. Khoury, "Exploiting non-taxonomic relations for measuring semantic similarity and relatedness in WordNet," *Knowledge-Based Systems*, vol. 212, pp. 1–27, 2021.
[17] J. Chen, X. Xue, Y. Huang, and X. Zhang, "Interactive cross-lingual ontology matching," *IEEE Access*, vol. 7, pp. 79095–79102, 2019.
[18] Y. Xue, H. Zhu, J. Liang J, and A. stowik, "Adaptive crossover operator based multi-objective binary genetic algorithm for feature selection in classification," *Knowledge-Based Systems*, vol. 227, pp. 1–17, 2021.
[19] S.-Z. Zhao and P. N. Suganthan, "Empirical investigations into the exponential crossover of differential evolutions," *Swarm and Evolutionary Computation*, vol. 9, pp. 27–36, 2013.
[20] J. Liu and J. Lampinen, "A fuzzy adaptive differential evolution algorithm," *Soft Computing*, vol. 9, no. 6, pp. 448–462, 2005.
[21] A. Bookstein, V. A. Kulyukin, and T. Raita, "Generalized hamming distance," *Information Retrieval*, vol. 5, no. 4, pp. 353–375, 2002.
[22] T. K. Kim, "T test as a parametric statistic," *Korean journal of anesthesiology*, vol. 68, no. 6, pp. 540–546, 2015.
[23] X. Xue and J. Zhang, "Matching large-scale biomedical ontologies with central concept based partitioning algorithm and adaptive compact evolutionary algorithm," *Applied Soft Computing*, vol. 106, pp. 1–11, 2021.
[24] P. Wang, Y. Zhou, and B. Xu, "Matching Large Ontologies Based on Reduction Anchors," in *Proceedings of the Twenty-Second International Joint Conference On Artificial Intelligence*, pp. 1–6, Barcelona Catalonia Spain, July2011.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2455782, 9 pages
https://doi.org/10.1155/2022/2455782

Research Article

Metaverse-Empowered Music and Dance: Experience of Emotion and Scene Unification

Fengfei Wang

Xinxiang University, No. 191, Jinsui Road, Xinxiang City, 453003 Henan, China

Correspondence should be addressed to Fengfei Wang; wangfenglei8909@xxu.edu.cn

Received 25 February 2022; Accepted 22 March 2022; Published 4 April 2022

Academic Editor: Jianhui Lv

Music and dance are closely related and symbiotic. On the one hand, dance often requires music accompaniment. On the other hand, dance can enrich the melody and style of music. The emergence of the metaverse has taken the experience of music and dance to a new level. This paper studies the three-dimensional situational experience of music and dance in Virtual Reality (VR) empowered by metaverse to feel the beauty of situational integration. After the spherical video is projected onto a two-dimensional plane to form a panoramic video, the two-dimensional panoramic video needs to be converted into a spherical video for users to watch. Therefore, it is more reasonable to take spherical video distortion as the distortion measure of panoramic video coding. In this paper, spherical video distortion is taken as the measurement standard of video quality, and the panoramic video coding technology is optimized. Furthermore, the corresponding weights are introduced to change the distortion ratio of different interpolation regions in the calculation process of rate distortion cost, and a rate distortion optimization technology based on spherical distortion measurement is proposed. The equal weight feature of spherical pixel is realized on two-dimensional plane, which improves the coding efficiency of panoramic video. Experimental results show that compared with the three benchmarks, the proposed algorithm can achieve 1.6157% bit saving on average and achieve a good Quality of Experience (QoE) when other processes are the same.

1. Introduction

Music contains rich ideological connotations, which can be reflected through vivid musical artistic images [1]. In order to highlight and embody the thoughts and emotions of dance works, it is necessary to analyze and express the musical elements. Music is an art form of expressing rich emotions with sound, rhythm, and melody, while dance is an art form of expressing the emotional connotation of music with human body and movement [2–4]. Therefore, music and dance are closely related. For example, many ethnic dances in China have their own unique styles, and dance music also contains different cultural customs. Therefore, when appreciating the related works, it is necessary to understand not only the skills of dance movements but also the national cultural connotation of the dance, so as to correctly understand the emotional connotation and style rhythm contained in the folk dance music, and improve the appreciation of dance music.

During the dance performance, the music starts first and then the dance performance catches the eye. Therefore, music is the soul of dance art and an important prerequisite to lead people to appreciate dance [5]. Music can touch the deep emotions of human being through beautiful melodies and different rhythms and harmony forms and realize the ideological resonance between the audience and the connotation of dance. Hence, the use of music melody in dance is an important basis for the full expression of dance art. By feeling the music, it can stimulate people's deep desire of emotional expression, so as to realize the real emotional expression of dance with the integration of body and mind, and exercise people's ability of dance expression [6–8]. In addition, music can also be used as a basis for dance, so that people can show the beauty of their body movements to a greater extent in music. In dance, people show the corresponding dance movements by mastering the rhythm and tone of music and grasp the emotional expression of dance through in-depth appreciation of music, so that the dance

shown by people can be more closely connected with music, and at the same time, it can also make people's dance with emotion and richer connotation [9, 10]. Such training for a long time can make people through the mastery of music to innovate dance movements and improve their own dance performance.

Music is an important support of dance art and an important medium to help dance present rich and three-dimensional expressive effects. At present, the experience of dance is no longer the training of body movements and skills, but how to guide people to fully show the thoughts and feelings contained in dance with the help of body movements. Music is used to stimulate artistic emotions so that people can complete dance movements through emotional stimulation while dancing, and the dance performance is richer and more three-dimensional [11]. But when importing music, we should pay attention to whether the emotion of music is expressed clearly enough and also pay attention to the feedback, so as to help people better feel music and show the emotion of dance.

The metaverse is a virtual world created by humans that parallels but does not simply copy the real world [12]. As an emerging concept, it is gradually formed on the basis of Virtual Reality (VR), augmented reality, blockchain, cloud computing, 5G, digital twin, and other technologies [13]. Currently, most of the technologies and products that embody the metaverse are primarily focused on digital entertainment, treating the metaverse as an immersive gaming world [14]. Most people tend to focus on the entertainment properties of the metaverse rather than the potential of edutainment. As we all know, metaverse is the natural expansion of virtual environment. VR determines the manifestation of the metaverse, and virtual human-computer interaction closely related to VR is the symbol of life in the metaverse [15]. It can be said that without VR, there would be no metaverse.

This paper studies the three-dimensional situational experience of music and dance in VR empowered by metaverse to feel the beauty of situational integration. With the rapid development of multimedia technology, VR has become more and more popular due to its immersive experience. The video information in the mainstream VR is spherical video information. Spherical video requires high definition due to user experience. If uncompressed, large amount of data will bring great challenges to common hard disk storage and network transmission. Therefore, it is very important to improve the coding efficiency of spherical video [16]. Due to the limitations of existing video coding standards, spherical video needs to be projected to a two-dimensional plane to form panoramic video before codec operation. At the receiving end, the two-dimensional panoramic video is converted into a spherical video for viewing by the user. Therefore, it is more reasonable to use spherical video distortion as a distortion measure for panoramic video coding. Based on this, this paper takes the spherical video distortion as the measurement standard of video quality to optimize the panoramic video coding technology. Since panorama video is encoded by an existing encoder, rate distortion optimization (RDO) is still used in the mode selection process.

The contribution of this paper is that the corresponding weights are introduced for different projection format designs to change the proportion of distortion in different areas of interpolation in the rate distortion cost calculation process, and a rate distortion optimization technique based on spherical distortion measurement is proposed. The feature of equal weights of spherical pixels is realized in two-dimensional plane, and the coding efficiency of panoramic video is improved.

The remainder of this paper is organized as follows. Section 2 reviews related work. In Section 3, spherical distortion measurement-based ROD is presented. Experimental results are presented in Section 4. Section 5 concludes this paper.

2. Related Work

The rapid development of computer information technology and the rapid update of modern educational technology means make computer technology play an increasingly important role in music and dance. In [17], the authors designed and implemented the dance movement of the basic design idea of the automatic generation algorithm based on genetic algorithm. In [18], the authors proposed to reproduce the scene of closing eyes and listening to music in a computer vision system. In [19], the authors proposed a generative adversarial network-based cross modal association framework, which associated dance movements with music to create the required dance sequence according to the input music. In [20], the authors discussed the examples of electronic dance music from three different angles. In [21], the authors proposed a framework for generating a series of three-dimensional human dance postures for a given music. In [22], the principle and implementation of a digital audio workstation plug-in for chord sequence generation was described. In [23], a novel probability autoregression architecture was proposed to simulate future posture distribution using a standardized flow based on previous posture and music background. In [24], a graph convolution networks based automatic dance generation method were proposed.

As a new video mode, the impact of VR panorama video is beyond doubt. At present, great progress has been made in various technologies of panoramic video, especially in coding. In [25], a new motion model based on spherical coordinate transformation to explicitly solve the deformation problem was proposed to reduce the decoding time. In [26], a new octagonal mapping scheme was proposed to reduce the oversampling area and arrange the points into an octagon. In [27], the authors presented and classified the latest advances in projection methods, video quality evaluation indexes, and transmission optimization methods for video coding.

RDO is a key technology in video coding system. The traditional RDO system measures the distortion of reconstructed video from the perspective of signal processing but does not fully consider the characteristics of human vision and video call. The purpose of making full use of human visual characteristics is to maximize the visual quality of video in the area of concern under the condition of limited

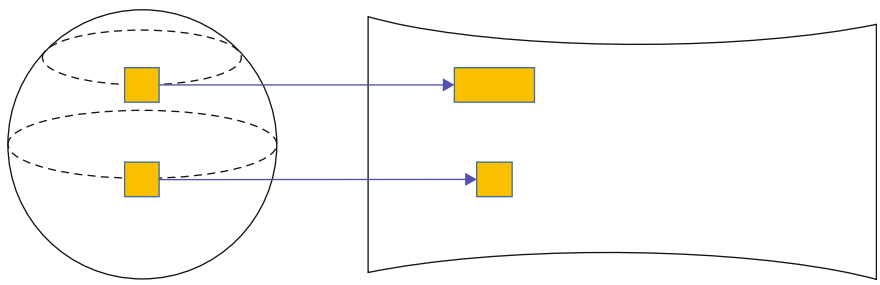


Figure 1: Diagram of spherical video and ERP format video.

bandwidth. Therefore, it is of great significance to study rate distortion model based on human visual perception for video coding. In [28], the authors proposed a method to generate encoded video stream. This method introduced the mathematical theory of decoding-energy-rate-distortion optimization (DERDO), which required less decoding energy than traditional encoded video stream. In [29], the authors proposed a new Lagrange multiplier determination model, which is a key part of rate distortion optimization (LM-RDO). In [30], the authors proposed an accurate coding tree unit level distortion structural similarity and distortion mean squared error (D-SSIM-D-MSE) model to obtain better video coding quality.

3. Spherical Distortion Measurement-Based RDO

Panoramic video coding still adopts the traditional block-based hybrid coding technology, so combined with the theory of encoder's block mode, prediction, rate distortion, quantization, and other operations, we can find ways to optimize panoramic video of different formats. Video distortion is as small as possible to achieve a sufficiently clear video content, and the rate of the video will be high. However, video compression coding hopes to find a way between distortion and bit rate to make the distortion small and the bit rate does not exceed the maximum allowable bit rate. Due to the constraints and contradictions between the two, RDO technology arises at the historic moment. For all lossy compression systems, RDO throughout the video coding system is a very important technology to balance the relationship between distortion and bit rate.

After the completion of the encoding and decoding of panoramic video, the spherical video is actually an output for users to watch. Therefore, for spherical video, regardless of the user viewing window, each pixel in the spherical video is equally important, that is, all pixels on the sphere are equally weighted. However, corresponding to different projection formats, the importance of pixels will change in the process of projection, which will affect the selection process of rate distortion, and further affect the final codec result, so that the video quality loses a certain degree of accuracy under the measurement of spherical video distortion.

In this paper, RDO in coding is improved according to different features of different projection formats of panoramic video. Taking Equirectangular Projection (ERP) as an example, a RDO model of panoramic video based on spherical distortion measurement is proposed.

In panoramic video, ERP format is taken as an example, the video content is stretched at different latitudes due to different interpolation operations in the transformation process. The pixel stretching is the smallest or even nonexistent on the equatorial path. From the equator to the two poles, the horizontal stretching of the pixels becomes more and more serious, and the stretching of the two poles is the largest.

The north and south pole contents of ERP format generally differ greatly in the encoding process due to the interpolation operation of stretching. The larger the stretch, the flatter the area is, resulting in a larger optimal block size after the rate distortion optimization process. For the same mode, selection near the equator obviously has a greater impact on the final video quality than the north and south pole options.

Theoretically, distortion in the coding process should be equally important for content near the equator on the sphere and near high latitudes, as shown in Figure 1. However, when converted to ERP format, the coding process calculates the rate distortion cost by applying the same weight to the distortion of the corresponding content of two blocks in ERP format. After back-projection to spherical surface, the distortion importance of the content in the actual high latitude area becomes greater.

Therefore, this paper adopts the method of introducing weight to distortion to reduce the proportion of distortion d when calculating the cost of rate distortion and to increase the proportion of distortion near the equator, so as to further improve the coding performance.

As for the panoramic video projection format except ERP format, although the severe stretching in ERP format is alleviated, there is still a certain degree of pixel bending. Therefore, when calculating the rate distortion cost, the corresponding weight is given to the distortion. In this paper, Cube Map Projection (CMP), Compact Octahedron Projection (COHP), and Segmented Sphere Projection (SSP) projection formats are used for study.

The rate distortion optimization model proposed in this section assigns different weights to the distortion when calculating the rate distortion cost of coding tree unit (CTU) at different positions of each projection format.

$$d'_k = w_k \times d_k, \tag{1}$$

where w_k and d_k are the weighted factor and distortion of the kth CTU, respectively. Thus, the rate distortion cost function of each CTU can be expressed as follows.

$$f'_{(RDC)} = d'_k + \alpha r_k = w_k \times d_k + \alpha r_k, \tag{2}$$

where α is the slope, and r_k is the bitrate of the kth CTU.

Assuming that a video frame is divided into N CTU during encoding, then its rate distortion cost function is defined as follows.

$$f'(RDC) = \sum_{k=0}^{N-1}\left(f'_{(RDC)}\right) = \sum_{k=0}^{N-1}\left(d'_k + \alpha r_k\right) = \sum_{k=0}^{N-1}(w_k \times d_k + \alpha r_k), \tag{3}$$

where w_k is the weight of distortion d at the calculation of rate distortion cost of the kth CTU, which is obtained by calculating the average weight of all pixels of the current CTU, that is

$$w_k = \frac{1}{s\mathrm{CTU}^2}\sum_{m=0}^{s\mathrm{CTU}-1}\sum_{n=0}^{sCTU-1} wp_{m,n}, \tag{4}$$

where sCTU is the size of CTU, and $wp_{m,n}$ is the pixel weight of the mth row and nth column of the video frame, which is obtained according to the conversion algorithm of spherical video to each projection format.

3.1. Weight for ERP Format. For ERP format, $wp_{m,n}$ is the scale factor from the ERP region to the spherical region, that is, the weight, which is only related to the height of the current CTU row. Let W be the width of a frame of video, H be its height, and $cWidth$ be the height of the current pixel. Assuming that the radius of the sphere is R, the latitude of the current pixel is θ, and the radius of cross-section with angle θ is r.

According to the projection algorithm in ERP format, its weight can be obtained as follows.

$$wp_{m,n} = \frac{2\pi r}{2\pi R} = \frac{r}{R} = \cos\theta. \tag{5}$$

At the same time, θ can be calculated as follows.

$$\theta = \frac{W-1-2n}{2} \times \frac{\pi}{W}. \tag{6}$$

Since sine function is even function, which can be obtained as follows.

$$wp_{m,n} = \cos\left(\frac{2n-W+1}{2} \times \frac{\pi}{W}\right). \tag{7}$$

As can be seen from the Equations (5) to (7), at the two ends of the ERP format, that is, the north and south pixels of the spherical video, have little weight, but the middle pixel row of the ERP format has the largest weight.

3.2. Weight for CMP Format. The CMP format is composed of six square faces, and the pixel weight of the corresponding position of each square face is the same, so only one face should be considered. Let each square face width and height be S.

From the microscopic point of view, the projection algorithm is the change from $\mathrm{d}m\mathrm{d}n$ to $\mathrm{d}\varphi\mathrm{d}\theta$, and the relationship between them can be described as $\mathrm{d}m\mathrm{d}n = J(\varphi,\theta)\mathrm{d}\varphi\mathrm{d}\theta$, where $J(\varphi,\theta)$ is the Jacobian determinant.

$$J(\varphi,\theta) = \frac{\partial(m,n)}{\partial(\varphi,\theta)} = \begin{vmatrix} \frac{\partial m}{\partial\varphi} & \frac{\partial m}{\partial\theta} \\ \frac{\partial n}{\partial\varphi} & \frac{\partial n}{\partial\theta} \end{vmatrix}. \tag{8}$$

At this point, according to the projection relationship between the sphere and a plane in CMP, we can conclude that

$$m = \tan\varphi, \tag{9}$$

$$n = \frac{\tan\theta}{\cos\varphi}. \tag{10}$$

Since the area projected on the square surface is equal to $dmdn$, while the corresponding area on the sphere is actually $\cos\theta|d\varphi d\theta|$, so the weight of a point (m,n) on the surface of CMP format can be obtained as follows.

$$wp_{m,n} = \frac{\cos\theta}{|J(\varphi,\theta)|} = \frac{\cos\theta}{\cos^2\theta\cos^3\varphi} = (\cos\varphi\cos\theta)^3 = \left(1+\frac{d^2}{R^2}\right)^{-3/2}, \tag{11}$$

where R is the radius of the circle tangent to the cube, and $R = S/2$ by this time, and the distance d between the points on the surface (m,n) and the center of the surface is defined as follows.

$$d^2 = \left(m+\frac{1-S}{2}\right)^2 + \left(n+\frac{1-S}{2}\right)^2. \tag{12}$$

So the weight of each pixel position is defined as follows.

$$wp_{m,n} = \left(1+\frac{(m+1-S/2)^2+(n+1-S/2)^2}{(S/2)^2}\right)^{-3/2}. \tag{13}$$

3.3. Weight for COHP Format. COHP format is composed of triangular faces, and the weight of the corresponding position of each triangular face is the same. Therefore, it is the same as CMP format to calculate one surface weight. Suppose the side length of each triangular face is T, and the weight of COHP is defined as follows.

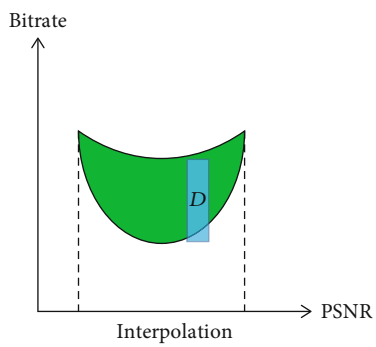


Figure 2: Schematic diagram of $\Delta Bitrate$.

Table 1: Results for ERP format with weight w_k.

Test sequences	$\Delta Bitrate$			
	DERDO	LM-RDO	D-SSIM-D-MSE	Ours
Basketball	-2.3608%	-2.2318%	-2.2441%	-2.4912%
Carphone	-4.1459%	-4.1765%	-4.1789%	-5.0126%
Foreman	-6.0678%	-6.4791%	-6.5110%	-6.8817%
Gaslamp	-0.4624%	-0.4972%	-0.4338%	-0.5241%
Highway	-0.6324%	-0.6223%	-0.6765%	-0.6004%
KiteFlite	-0.1580%	-1.5002%	-1.3900%	-0.1326%
Average	-2.3046%	-2.5845%	-2.5724%	-2.6071%

Table 2: Results for ERP format with weight w'_k.

Test sequences	$\Delta Bitrate$			
	DERDO	LM-RDO	D-SSIM-D-MSE	Ours
Basketball	-2.6775%	-2.7609%	-2.5573%	-2.8699%
Carphone	-4.2016%	-4.4195%	-4.0142%	-4.1277%
Foreman	-5.7500%	-6.5608%	-6.1569%	-6.8102%
Gaslamp	-0.4325%	-0.5496%	-0.5874%	-0.6636%
Highway	-0.5998%	-0.6338%	-0.8012%	-0.7890%
KiteFlite	-1.7631%	-1.6584%	-1.7433%	-1.8867%
Average	-2.5708%	-2.7638%	-2.6434%	-2.8579%

Table 3: Results for SSP format with weight w_k.

Test sequences	$\Delta Bitrate$			
	DERDO	LM-RDO	D-SSIM-D-MSE	Ours
Basketball	-0.0284%	-0.0456%	-0.0897%	-0.7446%
Carphone	-0.1918%	0.1189%	0.3012%	-0.4058%
Foreman	0.5462%	-0.2165%	0.0218%	-0.6915%
Gaslamp	0.0610%	0.0299%	0.0347%	-0.0705%
Highway	-0.1209%	-0.0448%	0.0459%	-0.1202%
KiteFlite	0.2008%	0.3267%	0.4167%	0.4468%
Average	0.0778%	0.0281%	0.7306%	-0.2643%

Table 4: Results for SSP format with weight w'_k.

Test sequences	$\Delta Bitrate$			
	DERDO	LM-RDO	D-SSIM-D-MSE	Ours
Basketball	-0.0454%	-0.0398%	-0.1919%	-0.6989%
Carphone	-0.0741%	-0.0574%	-0.1215%	-0.4365%
Foreman	-0.1563%	0.1550%	0.1384%	-0.5898%
Gaslamp	-0.5156%	0.0012%	0.0301%	-0.0574%
Highway	-0.0754%	0.0899%	0.1257%	-0.1002%
KiteFlite	-0.1887%	0.3418%	0.3352%	-0.3878%
Average	-0.1759%	0.0818%	0.0527%	-0.3784%

$$R = \frac{T}{\sqrt{6}},$$
$$d^2 = \left(m + \frac{1-T}{2}\right)^2 + \left(n + \frac{3-2\sqrt{3}T}{6}\right)^2. \tag{14}$$

The weight of a point (m, n) on each triangular surface is defined as follows.

$$wp_{m,n} = \left(1 + \frac{(m+1-T/2)^2 + \left(n+3-2\sqrt{3}T/6\right)^2}{\left(T/\sqrt{6}\right)^2}\right)^{-3/2}. \tag{15}$$

Since COHP format contains operations such as rotation, segmentation, and combination of triangles, the weight of the whole frame can be obtained by rotation and segmentation and combination of the triangular face with the weight obtained.

3.4. Weight for SSP Format. The SSP format contains invalid fields and the weight is 0. Except invalid regions, the weight of other regions is defined as follows.

$$wp_{m,n} = \begin{cases} \dfrac{2\cos\theta}{\pi - 2|\theta|}, \theta\epsilon\left(\dfrac{\pi}{4}, \dfrac{\pi}{2}\right] \cup \left[-\dfrac{\pi}{2}, -\dfrac{\pi}{4}\right) \\ \dfrac{2\sqrt{2}}{\pi}, \theta\epsilon\left[-\dfrac{\pi}{2}, -\dfrac{\pi}{4}\right] \end{cases}. \tag{16}$$

In this paper, the weight w_k multiplied by calculating distortion for each CTU of each projection format is modified, and the expected value of the weight is changed to 1 by dividing each weight by the expected value. In this paper, the distortion of rate distortion calculation process is directly modified without the operation of adjusting Lagrange multiplier. If the distortion becomes smaller or too large, the bit rate and video quality cannot be well balanced, so the deviation degree of data should be moderate. After a lot of experiments, the variance of the weight is controlled within 0.15 in order to achieve a good balance between bits and video quality without modifying the Lagrange multiplier.

Table 5: Results for CMP format with weight w_k.

Test sequences	$\Delta Bitrate$			
	DERDO	LM-RDO	D-SSIM-D-MSE	Ours
Basketball	0.4248%	0.5164%	0.3979%	0.2306%
Carphone	-0.1589%	-0.1236%	-1.1265%	-1.1189%
Foreman	1.4259%	1.1866%	1.2876%	1.0075%
Gaslamp	0.1898%	0.0900%	0.1267%	0.1048%
Highway	-0.2046%	-0.1923%	-0.2687%	-0.6789%
KiteFlite	0.3744%	0.5551%	0.4367%	0.1007%
Average	0.3419%	0.3387%	0.1423%	-0.0590%

Table 6: Results for CMP format with weight w_k'.

Test sequences	$\Delta Bitrate$			
	DERDO	LM-RDO	D-SSIM-D-MSE	Ours
Basketball	-0.3596%	-0.5215%	-0.4662%	-0.8745%
Carphone	-2.2712%	-1.9846%	-0.2370%	-2.6574%
Foreman	-2.1095%	-2.1005%	-2.2689%	-2.2461%
Gaslamp	-0.0667%	-1.2708%	-1.2227%	-1.3201%
Highway	-1.0909%	-1.0909%	-0.9016%	-1.2107%
KiteFlite	-1.2110%	-1.1234%	-0.9647%	-1.2680%
Average	-1.1848%	-1.3486%	-1.0102%	-1.5961%

Table 7: Results for COHP format with weight w_k.

Test sequences	$\Delta Bitrate$			
	DERDO	LM-RDO	D-SSIM-D-MSE	Ours
Basketball	-1.1394%	-1.1795%	-1.1516%	-1.0678%
Carphone	-1.4500%	-1.2461%	-1.5438%	-2.3684%
Foreman	-0.2268%	-0.3364%	-0.2323%	-0.2941%
Gaslamp	-0.2963%	-0.1919%	-0.3067%	-0.1945%
Highway	0.0634%	0.0361%	0.1891%	-0.1662%
KiteFlite	-0.0091%	-0.0502%	-0.0846%	-0.1108%
Average	-0.5097%	-0.4947%	-0.5217%	-0.7003%

Table 8: Results for COHP format with weight w_k'.

Test sequences	$\Delta Bitrate$			
	DERDO	LM-RDO	D-SSIM-D-MSE	Ours
Basketball	-1.5080%	-1.5398%	-1.5237%	-2.5256%
Carphone	-2.1367%	-2.1397%	-2.0083%	-2.3765%
Foreman	-1.6632%	-1.7665%	-1.6266%	-1.6349%
Gaslamp	-1.0130%	-0.9840%	-1.0487%	-1.0897%
Highway	-1.2464%	-1.2647%	-1.3167%	-1.3136%
KiteFlite	-0.7266%	-0.7787%	-0.8718%	-0.8415%
Average	-1.3823%	-1.4122%	-1.3993%	-1.6303%

Table 9: Comparison of w_k and w_k' with four projection formats.

Projection format	w_k	w_k'	Comp
ERP	-2.6071%	-2.8579%	-0.2508%
SSP	-0.2643%	-0.3784%	-0.1141%
CMP	-0.0590%	-1.5961%	-1.5371%
COHP	-0.7003%	-1.6303%	-0.9300%
Average	-0.9077%	-1.6157%	-0.7080%

$$w_k' = \frac{w_k}{E} \times \sqrt{\frac{0.15}{d}}. \tag{17}$$

To verify the above reasoning, all format weights w_k and w_k' are encoded in the experiment, so as to draw relevant conclusions.

4. Experiment Results and Performance Analysis

4.1. Performance Evaluation Metrics. $\Delta Bitrate$ is used to measure coding performance in this paper, and the calculation method is shown in Figure 2. Different quantization parameters were used for video sequences, and several sets of corresponding Peak Signal to Noise Ratio (PSNR) and bit rate data were obtained after encoding with the proposed method and three benchmarks, respectively. Their curves were plotted in the coordinate system in the figure, respectively. The integral area difference D and the maximum interpolation over PSNR of the two curves is the final $\Delta Bit\ rate$ value. A positive $\Delta Bitrate$ indicates an increase in bitrate, while a negative $\Delta Bitrate$ indicates a saving in bitrate and improved performance.

4.2. Experimental Parameter Settings. This paper uses the 360lib-HM16.14 as a test platform to integrate the improved RDO technology into 360lib-HM16.14 code for performance testing, where 360Lib-HM16.14 is a 360-degree video tool provided by joint video experts team and integrated with HM or JEM for 360-degree video encoding and decoding. To better verify the quality of panoramic video, spherical video distortion is used to measure the quality. The configuration in coding adopts the given configuration [31] in the universal test environment CTC of panoramic video, and the quantization parameters are 22, 27, 32, and 37, respectively. The experiment is based on the RandomAccess (RA) structure suggested by high efficiency video coding general test conditions and the RandomAccessMain profile (RA-main). The test sequences are selected as Basketball, Carphone, Foreman, Gaslamp, Highway, and KiteFlite.

4.3. Performance Analysis. This section compares the performance of the improved distortion algorithm. Three the-state-of-the-art algorithms are selected for comparison, which are DERDO [28], LM-RDO [29], and D-SSIM-D-MSE [30]. We compare the projection format of the four algorithms with the weighted w_k and the modified w_k' in the calculation of rate distortion in 360lib-HM16.14.

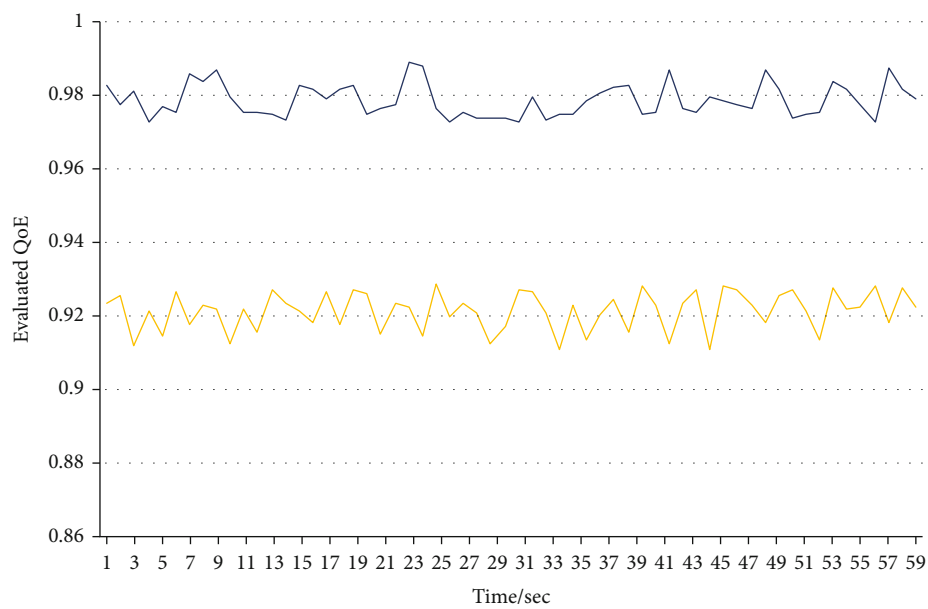


Figure 3: Evaluated QoE with w_k and w_k'.

When weighted w_k is added into ERP format, the results of the proposed algorithm compared with the three benchmarks are shown in Table 1. After adding the optimized weight w_k', the results of the algorithm proposed in this paper compared with the three benchmarks are shown in Table 2.

From the above data, it can find that the introduction of weight w_k in the rate distortion cost calculation of ERP format can improve the coding performance by 2.6071%, while the introduction of optimized weight w_k' can improve the coding performance by 2.8579%.

When weighted w_k is added into SSP format, the results of the proposed algorithm compared with the three benchmarks are shown in Table 3. After adding the optimized weight w_k', the results of the algorithm proposed in this paper compared with the three benchmarks are shown in Table 4.

From the above data, it can find that the introduction of weight w_k in the rate distortion cost calculation of ERP format can improve the coding performance by 0.2643%, while the introduction of optimized weight w_k' can improve the coding performance by 0.3784%.

Summarizing the results of encoding ERP and SSP formats with weighted w_k and w_k' in the process of rate distortion, we can see that the performance improvement is not significant because the values of w_k' and w_k are close to each other.

When weighted w_k is added into CMP format, the results of the proposed algorithm compared with the three benchmarks are shown in Table 5. After adding the optimized weight w_k', the results of the algorithm proposed in this paper compared with the three benchmarks are shown in Table 6.

From the above data, it can find that the introduction of weight w_k in the rate distortion cost calculation of CMP format can improve the coding performance by 0.0590%, while the introduction of optimized weight w_k' can improve the coding performance by 1.5961%, showing a large performance improvement.

When weighted w_k is added into COHP format, the results of the proposed algorithm compared with the three benchmarks are shown in Table 7. After adding the optimized weight w_k', the results of the algorithm proposed in this paper compared with the three benchmarks are shown in Table 8.

From the above data, it can find that the introduction of weight w_k in the rate distortion cost calculation of COHP format can improve the coding performance by 0.7003%, while the introduction of optimized weight w_k' can improve the coding performance by 1.6303%, showing a large performance improvement.

It can be seen from the results that weight Wi and Wi' are used to code CMP format and COHP format in the process of rate distortion. As the modified weight w_k' changes greatly compared with the original w_k, the modified weight improves the coding performance of these two formats especially significantly. Compared with the original weights, the modified weight can achieve a better balance between the bit rate and video quality.

Table 9 compares the bitrate of the four algorithms in the two weight results of the four projection formats and compares the performance of the two algorithms. Comp is obtained by subtracting the performance change of w_k' from the performance change of w_k. If the value is negative, the performance of w_k' is better than w_k.

As can be seen from the above results, for the four projection formats mentioned in this paper, w_k' can improve the coding performance to a certain extent compared with w_k, up to 0.7080% on average. However, the performance of CMP and COHP formats is significantly improved by

1.5371% and 0.9300%, respectively, compared with w_k'. Therefore, when calculating the rate distortion cost of different projection formats, the weight w_k' of the corresponding format for the distortion at different positions of the image improves the coding performance, which can reach 1.6157% on average.

Furthermore, this paper constructs the metaverse-empowered experience of music and dance, so the Quality of Experience (QoE) is evaluated. As can be seen from Figure 3, the projection format with w_k' (blue line) in VR music and dance scenario experience always has a good QoE, which is always higher than the projection format with w_k (orange line), indicating the effectiveness of the RDO model based on spherical distortion measurement proposed in this paper.

5. Conclusions

In this paper, a rate distortion optimization technique for panoramic video based on spherical distortion measurement is proposed for metaverse-empowered music and dance. In high efficiency video coding, rate distortion optimization plays a decisive role in mode selection. However, for different projection formats of panoramic video, the weight of the blocks with equal weight on the sphere will change after the projection algorithm corresponds to different formats. Therefore, this paper optimizes rate distortion technology according to different features of different projection formats and introduces corresponding weights in the process of distortion calculation to restore the properties of equal weights of video content on the spherical surface. Compared with the original distortion algorithm, the encoding performance of the proposed panoramic visual frequency distortion optimization model based on spherical distortion measure is improved by 1.6157% on average.

With the gradual development of VR, people have higher expectations on the quality of panoramic video, and video coding becomes more and more important for panoramic video. How to encode and transmit panoramic video is still a problem that cannot be ignored. In this paper, the operation of panoramic video coding optimization still has some limitations. The optimization of rate distortion cannot only add the corresponding weight to the distortion but also modify the Lagrange multiplier to find a better model. Interframe prediction can also be further optimized and improved.

Data Availability

All data used to support the findings of the study is included within the article.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] T. C. Chin, E. Coutinho, K. R. Scherer, and N. S. Rickard, "MUSEBAQ: a modular tool for music research to assess musicianship," *MUSICAL CAPACITY, MUSIC PREFERENCES, AND MOTIVATIONS FOR MUSIC USE, MUSIC PERCEPTION*, vol. 35, no. 3, pp. 376–399, 2018.

[2] Y. Qi, Y. Liu, and Q. Sun, "Music-driven dance generation," *IEEE ACCESS*, vol. 7, pp. 166540–166550, 2019.

[3] N. F. Bernardi, A. Bellemare-Pepin, and I. Peretz, "Dancing to "groovy" music enhances the experience of flow," *Annals of the New York Academy of Sciences*, vol. 1423, no. 1, pp. 415–426, 2018.

[4] R. T. Solberg and A. R. Jensenius, "Group behaviour and interpersonal synchronization to electronic dance music," *Musicae Scientiae*, vol. 23, no. 1, pp. 111–134, 2019.

[5] Z. Wang, "Music Choreography Algorithm Based on Feature Matching and Fragment Segmentation," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9274098, 9 pages, 2022.

[6] J. Geng, "Personalized analysis and recommendation of aesthetic evaluation index of dance music based on intelligent algorithm," *Complexity*, vol. 2021, 15 pages, 2021.

[7] P. Yao, "Key frame extraction method of music and dance video based on multicore learning feature fusion," *Scientific Programming*, vol. 2022, 8 pages, 2022.

[8] F. A. Raposo, D. M. de Matos, and R. Ribeiro, "Assessing kinetic meaning of music and dance via deep cross-modal retrieval," *Neural Computing & Applications*, vol. 33, no. 21, pp. 14481–14493, 2021.

[9] K. Chen, Z. Tan, J. Lei et al., "ChoreoMaster," *ACM Transactions on Graphics*, vol. 40, no. 4, pp. 1–13, 2021.

[10] R. T. Solberg and N. Dibben, "Peak experiences with electronic dance music: subjective experiences," *PHYSIOLOGICAL RESPONSES, AND MUSICAL CHARACTERISTICS OF THE BREAK ROUTINE, MUSIC PERCEPTION*, vol. 36, no. 4, pp. 371–389, 2019.

[11] C. Howlin, S. Vicary, and G. Orgs, "Audiovisual aesthetics of sound and movement in contemporary dance," *Empirical Studies of the Arts*, vol. 38, no. 2, pp. 191–211, 2018.

[12] S. M. Park, Y. G. Kim, and A. Metaverse, "A metaverse: taxonomy, components, applications, and open challenges," *Components, Applications, and Open Challenges, IEEE ACCESS*, vol. 10, pp. 4209–4251, 2022.

[13] M. S. Smith and R. Brooks, "Meta offers nothing new to the metaverse," *IEEE Spectrum*, vol. 59, no. 1, pp. 22–22, 2022.

[14] A. Jovanovic and A. Milosavljevic, "VoRtex Metaverse platform for gamified collaborative learning," *Electronics*, vol. 11, no. 3, p. 317, 2022.

[15] S. Wu, X. Chen, J. Fu, and Z. Chen, "Efficient VR video representation and quality assessment," *Journal of Visual Communication and Image Representation*, vol. 57, pp. 107–117, 2018.

[16] J. Lv, X. Wang, K. Ren, M. Huang, and K. Li, "ACO-inspired information-centric networking routing mechanism," *Computer Networks*, vol. 126, pp. 200–217, 2017.

[17] Y. He and Q. Zhang, "Automatic generation algorithm analysis of dance movements based on music-action association," *CLUSTER COMPUTING-THE JOURNAL OF NETWORKS SOFTWARE TOOLS AND APPLICATIONS*, vol. 22, no. S2, pp. 3553–S3561, 2019.

[18] X. Guo, Y. Zhao, and J. Li, "DanceIt: music-inspired dancing video synthesis," *IEEE Transactions on Image Processing*, vol. 30, pp. 5559–5572, 2021.

[19] G. Sun, Y. Wong, Z. Cheng, M. S. Kankanhalli, W. Geng, and X. Li, "DeepDance: music-to-dance motion choreography with adversarial learning," *IEEE Transactions on Multimedia*, vol. 23, pp. 497–509, 2021.

[20] B. Burger and P. Toiviainen, "Embodiment in electronic dance music: effects of musical content and structure on body movement," *Musicae Scientiae*, vol. 24, no. 2, pp. 186–205, 2020.

[21] H. Ahn, J. Kim, K. Kim, and S. Oh, "Generative autoregressive networks for 3D dancing move synthesis from music," *IEEE ROBOTICS AND AUTOMATION LETTERS*, vol. 5, no. 2, pp. 3501–3508, 2020.

[22] D. Conklin, M. Gasser, and S. Oertl, "Creative chord sequence generation for electronic dance music," *APPLIED SCIENCES-BASEL*, vol. 8, no. 9, p. 1704, 2018.

[23] G. Valle-Perez, G. Henter, J. Beskow, A. Holzapfel, P. Y. Oudeyer, and S. Alexanderson, "Transflower," *ACM Transactions on Graphics*, vol. 40, no. 6, pp. 1–14, 2021.

[24] J. P. Ferreira, T. M. Coutinho, T. L. Gomes et al., "Learning to dance: a graph convolutional adversarial network to generate realistic dance motions from audio," *COMPUTERS & GRAPHICS-UK*, vol. 94, pp. 11–21, 2021.

[25] Y. Wang, D. Liu, S. Ma, F. Wu, and W. Gao, "Spherical coordinates transform-based motion model for panoramic video coding," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 98–109, 2019.

[26] C. Wu, H. Zhao, and X. Shang, "Octagonal mapping scheme for panoramic video encoding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 9, pp. 2402–2406, 2018.

[27] Z. Chen, M. Li, and Y. Zhang, "Recent advances in omnidirectional video coding for virtual reality: projection and evaluation," *Signal Processing*, vol. 146, pp. 66–78, 2018.

[28] C. Herglotz, A. Heindel, and A. Kaup, "Decoding-energy-rate-distortion optimization for video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 1, pp. 171–182, 2019.

[29] F. Zhang and D. R. Bull, "Rate-distortion optimization using adaptive Lagrange multipliers," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 10, pp. 3121–3131, 2019.

[30] Y. Li and X. Mou, "Joint optimization for SSIM-based CTU-level bit allocation and rate distortion optimization," *IEEE Transactions on Broadcasting*, vol. 67, no. 2, pp. 500–511, 2021.

[31] Y. Liu, L. Yang, M. Xu, and Z. Wang, "Rate control schemes for panoramic video coding," *Journal of Visual Communication and Image Representation*, vol. 53, pp. 76–85, 2018.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2010189, 8 pages
https://doi.org/10.1155/2022/2010189

Research Article

Transmission Line Planning Based on Artificial Intelligence in Smart Cities

Dong Yang, Tao He, Bolin Du, Shou Wang, Zhong Zhang, and Suxin Zhang

Ma'anshan Electric Power Supply Company, State Grid Anhui Electric Power Co., Ltd, Ma'anshan 243000, China

Correspondence should be addressed to Tao He; 18655589207@163.com

Received 25 January 2022; Revised 9 March 2022; Accepted 19 March 2022; Published 31 March 2022

Academic Editor: Jianhui Lv

With the development of smart cities, it is particularly important to realize scientific and reasonable planning of transmission lines as the backbone of the power transmission system. Through the geographic information system (GIS) platform, this paper obtains the geographic information of the planning area. Then, through the typical design cost of the 110 kV transmission line subvolume of the typical cost of the power transmission and transformation project of the State Grid Corporation and the local documents of the smart city planning area, this paper obtains the estimated cost of each grid after gridding in the smart city area. After that, the research uses the ant colony algorithm to search for the optimal transmission line route in the smart city. Finally, the achievements of primary economic and technical indexes are evaluated in transmission line planning of smart city.

1. Introduction

Judging from the development trend of modern society, the scale of future cities will become larger and larger, and the urban population will increase accordingly. At present, the rapid development of cities in my country has led to the increasingly prominent problem of "urban disease" in some areas. The construction of smart cities can effectively solve the development problems of modern cities. A smart city is to analyze, detect, and integrate various key data of the city's core system through communication means and communication technologies so as to quickly and intelligently respond to the problems and needs in the city. Its essence is to realize the intelligent management and operation of the city, use the new generation of information technology to integrate analysis, make the city more suitable for human life and living, and realize the harmonious coexistence of the city and the people.

Geographical information system (GIS) is a computer system used to collect, store, manage, analyze, and display geographic spatial data [1]. Nowadays, GIS is ubiquitous and deeply integrated with other industries. It is widely used in various aspects of social life such as land, cities, resources and environment, public security, emergency response, logistics, electricity, and pipeline networks. In recent years, with the surging wave of new generation technologies such as big data, smart cities, cloud computing, Internet of Things, and artificial intelligence (AI), the development of GIS has ushered in new opportunities and challenges. Among them, AI has developed rapidly and has been widely used in many fields such as face recognition, machine learning, language and image understanding, intelligent control, intelligent search, and automatic planning.

AI was first proposed by American scholars in related fields in a conference held at Dartmouth University in 1956. It has never given a strict definition. This article is more in favor of the AI concept defined by Tsinghua University professor Shi Chunyi and others. AI is summarized as the study of how to make computers imitate the thinking activities such as reasoning, learning, thinking, and planning that the human brain can perform to solve complex problems that can only be handled by human experts [2]. The general algorithms implemented by AI include artificial neural network, genetic algorithm, immune algorithm, ant colony algorithm, and particle swarm optimization algorithm [3].

Ozsari et al. [4] proposed to integrate traditional GIS software with intelligent statistical methods and

mathematical models to achieve location decision-making and strategic planning in the 1990s. The intelligence of their GIS was only reflected in statistical methods and mathematical models. "Intelligence" is still more limited and narrow. Korus et al. [5] had proposed in the book "Artificial Intelligence in Geography" that the heuristic search, neural network, and evolutionary algorithm included in AI methods had their importance and application potential for geography. Li Xia, a Chinese scholar, summed up the concept of intelligent GIS by summarizing the application of AI in GIS combined with previous research. This concept brought AI methods to the spatial data constructed by GIS. In order to solve data extraction, simulation, optimization, and decision-making assistance in complex spatial problems, Li Xia built an automated and intelligent spatial analysis method with self-learning. This concept made up for the goals that could not be achieved by modeling based on traditional mathematical methods [6]. In the past three decades, with the deepening of AI research, the application of AI in the field of GIS had been triggered. Liwei et al. [7] designed a multidimensional unified computing engine based on genetic algorithm to process different types of geographic data and data of different dimensions; Wang et al. [8] used artificial neural network and GIS to simulate soil erosion rate and spatial change; Donoho et al. [9] solved the multiobjective urban land space optimization allocation by constructing a model based on spatial genetic algorithm and obtained urban optimal allocation of space for construction-like land; Fyhn et al. [10] combined ant colony algorithm and GIS for the location selection of logistics distribution center; Yu et al. [11] applied simulated annealing algorithm, genetic algorithm, and particle swarm optimization algorithm, respectively, to the spatial optimization problem of multiobjective land allocation. By introducing AI into the field of spatial information processing, the ability of GIS spatial analysis has been improved, and a model for intelligent processing of spatial data has been built, which has increased the practicability, intelligence, and scientificity of spatial data analysis and processing, and has brought research in various fields of science. There are new ideas, but the combination of AI and GIS in thematic fields is not yet mature, and a mature method system has not yet been formed, and in-depth research is needed.

Path optimization is also an important research content in the application fields involved in AI algorithms. The purpose of path optimization is to find the optimal path. The "optimal" here is not simply the shortest path distance, but it can be a measure of economic consumption, cost, time consumption, efficiency, etc. Ant colony optimization (ACO) is a kind of swarm intelligence algorithm. It was proposed by Italian scholar Dorigo and others in the early 1990s by simulating the behavior of ant colony path optimization in nature [12]. The ACO algorithm is widely used to solve similar traveling salesman (TSP) problems. This algorithm uses a distributed, positive feedback, and parallel computer system. It is easy to combine with other methods and has strong robustness. The optimization problem has great potential [13]. Jiao et al. [14] proposed a robot path planning method in two-dimensional space based on ACO and added artificial ants to improve the flexibility of optimization by considering user requirements and environmental conditions when making decisions. Liu et al. [15] proposed a pheromone distribution and pheromone update strategy based on ACO. It used time-varying characteristics such as inflection point optimization and local optimal expansion path to design inflection point parameters and an overall evaluation method. This method was used for path planning of mobile robots in complex environments, and it was verified that the ant colony expansion path optimization algorithm was better than the traditional ACO and the improved algorithm. The algorithm had better search performance, stronger search ability, and shorter path, avoided falling into local optimum, and realized the search of the optimal path of mobile robot. Obayashi and Masuyama [16] used cellular automata to build a robot's path planning environment model. On this basis, ACO was used to optimize path planning. Experiments proved the feasibility of the method. ACO has the advantages of distributed, positive feedback, and parallel computing in path optimization, but it also has shortcomings such as too long search time and falling into local optimum.

In addition to the overall planning of power grid, the optimization of transmission cable path also affects whether the operation can continue to be stable. Power transmission cables are an important part of the power grid due to their power transmission tasks [17]. Transmission cable design and planning is a huge project with high construction cost, long period, and strict conditions. It includes electrical design, basic design, path selection, insulation coordination design, grounding design, and various types of tower designs. In all tasks, the focus is to select the cable path, because the subsequent design work is directly affected by the path selection, which in turn has an important impact on the input cost and overall effectiveness [18]. The route of the transmission cable is selected to find the best route that meets the requirements of cost-effectiveness, convenient construction, stable operation, and compliance with national policies [19]. At present, a large number of scholars have conducted in-depth research on the optimization of cable selection. With the advancement of geographic information technology, intelligent route selection has become a hot topic in discussion [20]. In this paper, the AI route selection is used to study the route optimization of transmission cable, the main economic and technical indexes are established to evaluate the optimization results, and traditional manual route selection [21] and the Dijkstra algorithm are used for comparison.

2. Principles of AI Path Optimization Methods in Smart Cities

2.1. Process of Path Optimization Method. For route selection design of AI circuit, besides quantifying geographical elements that can influence route selection, comprehensive action index of each element should also be calculated [14, 22–23]. Therefore, the following optimization of AI transmission cable path can only be performed by discovering

the appropriate method to get the geographical information model of the smart city.

The following are detailed steps to use AI to optimize the transmission cable path in a smart city.

(1) Obtain the final cost value matrix through the BP neural network algorithm and turn the geographical elements that affect the transmission cable into a measurable value, and then calculate the overall effect value of all the elements, which is used to perform the AI transmission cable path. Planning calculations

(2) Mark the main cells that constitute the map route of the smart city planning area as the marker points and connect the marker points in sequence from the starting point to the end point to obtain the route plan

(3) The ant colony pheromone updating mechanism is used to solve the path problem and bypass the obstacles in an appropriate way to obtain the best solution for cable transmission path in smart city

2.2. Geographic Information Model of Smart City Planning Area

2.2.1. Classification of Map Cells. In the smart city planning path, both natural and social factors will have an impact on the selection of the transmission cable path. Some of these factors are very restrictive, making it impossible to set up line towers in this area, nor to cross over in this area; but there are also areas that are not too restrictive, and paths can be designed for this area. According to the restrictive strengths of the geographic information factors in the above-mentioned planning area, the cells within the path research range can be divided into three categories: A, B, and C. Category A means that the line can pass through and can erect towers. Category B means that lines can pass through but the erection of towers is prohibited. Category C means that lines can neither pass through nor erect towers. Described from geographic information factors, category A is mainly high-altitude ice-covered areas of Xuefeng Mountain and Nanling Mountain, densely vegetated areas, densely housed areas, Yunma Forest Farm, Nandong Forestry Farm, etc. Category B is mainly rivers, lakes, geological hazards, etc., including 110KV Nanru Line, 110KV Yangbei Line, 35KV Chang'an Line, S219 Provincial Trunk Road, X088 County Trunk Road, X091 County Trunk Road and Township Highway. Category C is mainly for county planning areas, microwave communication towers, etc.

2.2.2. Estimation of Cell Cost Value. There are many factors that affect the selection of transmission cables, such as geology, landforms, hydrology, meteorology, pollution, vegetation coverage, power facilities, military facilities, and transportation facilities. It is difficult to comprehensively consider all factors. The traditional evaluation method is to construct a comparison matrix of the degree of influence between different factors through the analytic hierarchy process and then process the cost of the grid evaluated by the experts through the analytic hierarchy matrix to obtain the final cost value matrix [24]. The disadvantage is that the comparison of the degree of influence of the experts on the pairwise influences on the path selection of the transmission line is obtained through experience, which is highly subjective, and the result may not be accurate; when the amount of data is large, the number of grids is very large, and the evaluation will take a lot of time and energy.

Table 1: Comprehensive cost of unit length for different terrain conditions.

Terrain	Wire model	Maximum wind speed ($m \cdot s^{-1}$), maximum icing (mm)	Comprehensive cost (ten thousand yuan)
The ground	2 × LGJ240/30	25, 5 30, 10	49.49 51.04
River net bathing	2 × LGJ240/30	25, 5 30, 10	72.59 74.54
Hills	2 × LGJ240/30	25, 5 30, 10	51.83 53.19

In this paper, the BP neural network is used to simulate the actual cost of lines under different geographical conditions in the typical design of smart city planning area, and the BP neural network after fitting is used to predict the cost value of the cells across the transmission line under different geographical conditions, which is used as the reference value of the cell generation value. The planning area of this paper is a 110 kV power transmission line somewhere in China, which belongs to the third meteorological region in China. The maximum wind speed is 25 m/s and the maximum ice cover is 5 mm. Refer to the typical cost of power transmission and transformation project of State Grid Corporation of China: the comprehensive cost of line unit length under different terrain conditions in the volume of 110 kV transmission lines in 2006 [25] is shown in Table 1.

Considering that the influence of various factors on the cost value is nonlinear, a three-layer BP neural network function is used to complete the approximation task.

Taking into account that the influence of various factors on the cost value is nonlinear, the three-layer BP neural network function shown in Figure 1 is used to complete the approximation task.

Since the output layer of the network has multiple output nodes, the difference squares of each output node of the output layer need to be summed, so the loss function of each training example is obtained as

$$E\left(\vec{w}\right) = \frac{1}{2} \sum_{k \in \text{outputs}} \left(s_k - o_k\right)^2. \tag{1}$$

Among them, s_k is the expected output of the k neuron, o_k is the actual output of the k neuron, and w is the weight. Adjust the input weight vector in the output node according to the loss function, and then adjust the weight layer by layer

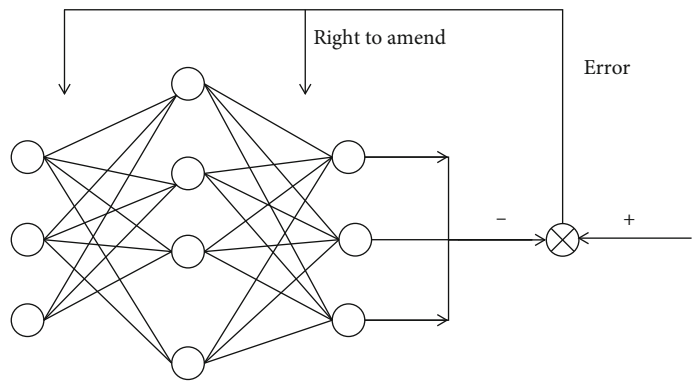


Figure 1: Schematic diagram of 3-layer BP neural network algorithm.

from back to front. Perform the following operations for each training of the metacell cost:

(a) According to the input of the example, calculate from front to back to get the output of each unit of the output layer. Then, the error term of each unit of each layer is calculated backward from the output layer

(b) For each neuron k of the output layer, calculate its error term:

$$\delta_k = o_k(1 - o_k)(s_k - o_k) \tag{2}$$

(c) For each hidden unit h in the network, calculate its error term:

$$\delta_h = o_h(1 - o_h) \sum_{k \in \text{outputs}} w_{kh}\delta_h \tag{3}$$

(d) Update each weight:

$$w_{ji} = w_{ji} + \eta\delta_j u_{ji}, \tag{4}$$

$w_{ji} = \eta\delta_j u_{ji}$ is called the weight update rule, u_{ji} is the input from node i to node j, w_{ji} represents the corresponding weight, and outputs represents the set of cell nodes in the output layer

Enter the samples in Table 1 to obtain the middle layer neural network. After processing the cells of the smart city planning area, the geographic conditions of each cell are obtained, and the cell cost value matrix is acquired by fitting the BP neural network to the geographic conditions of the cells. Additional special crossing and land acquisition costs are given to buffer cells located in areas such as roads, transmission facilities, and residential houses, as shown in Table 2, which ultimately forms a cost substitution value matrix for all cells in the smart city planning area.

Table 2: Special cost obtained by local file of planning region.

Expense item	Standard fee (ten thousand yuan·km^{-1})
Requisition and clean-up fees for construction sites	5.0
Over 110 kV	1.5
Across 35 kV	0.9
Across 10 kV	0.4
Cross the grade highway	2.5
Highway	5.0
Charges for live-over housing measures	0.5

2.2.3. Use GIS to Integrate Map Data of Smart City Planning Areas. Vector data and raster data models are two types of GIS data models [26]. Most of the geographic information in maps used at present are vector data. The grid size can be set according to the environmental parameters of GIS, and the corresponding grid data can be obtained from the original vector geographic information after transformation. Then, we merge and rank relevant geographic information and use BP neural network model to get cell cost value, so as to record coordinates of different grids.

The attribute information of the cell includes the position coordinate field $R(x, y)$, the cell type (A, B, C), and the cell cost value c. The coordinate position of a cell represents the spatial position of the unit on a smart city planning map. Cell type indicates that smart city planning maps can form uncrossed line cells. The cell cost value is the cost value of the line passing over a cell.

2.3. ACO Transmission Cable Path Search Model. Cellular automata (CA) is used as the route selection model for transmission cables, and the search model is acquired by linking points in a straight line [27]. Mark main cells of the path on the map as marking points, connect one after the other from the starting point, and at the end of the search get a transmission cable path. The cells in smart city planning areas have three states: the initial state "0", which indicates that the cell is not threaded through; if a cell is clearly marked as a cell at a time, the state must be changed to "1"; if a cell is on the route, then the status changes to "2,"

and finally the planned route is clarified according to "1" and "2" cells.

The pheromone mechanism of ant colony algorithm is usually used to settle the path searching issue, which helps to avoid obstacles and search for the path to obtain the optimal result. The rules of ACO search path are composed of avoidance rules, movement rules, and pheromone rules. According to these three rules, each ant finds the next moment position from the neighbor cells by roulette.

(a) Search cell area analysis: set at a certain moment, where the ants are located e

(b) Ant colony movement rule: set the neighbor grid to be selected, and then, the probability of moving from to is normalized as

$$p(\mathrm{x}, y) = a_{xy}^{\alpha} b_{xy}^{\beta} m_{xy}^{\gamma}, \tag{5}$$

$$p_t(\mathrm{x}, y) = \frac{p(\mathrm{x}, y)}{\sum_{t=1}^{N} p(\mathrm{x}, \mathrm{y})}. \tag{6}$$

Among them, $p(\mathrm{x}, y)$ is the transformation probability of the cell, and a, b, and c are the information, cost, and direction control coefficients, respectively. Its control force is α, β, and γ, respectively. N is the number of optional neighbor cells.

When the transformation probability $p_t(\mathrm{x}, y)$ is specified for all the neighbor cells, the roulette wheel is applied, and the next transformation cell is randomly chosen from a neighbor cell in a certain degree. The determination process is as follows: generate a random number r in the interval [0, 1]; if r satisfies formula (7), then t is the next grid position.

$$\begin{cases} 0 \le r \le \sum_{v=1}^{t} P_{e \longrightarrow v}(x, y), & t = 1, \\ \sum_{v=1}^{t-1} P_{e \longrightarrow v}(x, y) \le r \le \sum_{v=1}^{t} P_{e \longrightarrow v}(x, y), & t \ge 2 \end{cases} \tag{7}$$

(c) Pheromone update: pheromone concentration is one of the key heuristic factors for ACO path search. After each search obtains a feasible path, the pheromone concentration of the path cell needs to be updated. To avoid the excessive concentration of pheromone accumulated in multiple search results resulting in premature path maturity, the pheromone volatilization coefficient needs to be introduced. The final update formula of cellular pheromone concentration is as follows:

$$m'_{xy} = (1 - \sigma)\left(m_{xy} + \Delta m\right), \tag{8}$$

where m is the control coefficient of pheromone in cell transformation probability, m_{xy} is the pheromone concentration before cellular renewal, Δm is the increase in cellular pheromone concentration, $\Delta m = \lambda / z$, z is the total cost of the path, and λ is the correction coefficient

3. Optimization Process of Transmission Cable Path Based on AI

The high-voltage power transmission cable passes over many places. To obtain high accuracy in searching, the partitioned method is used in the process of processing the covered places. The important points that the line must pass through can be correctly marked on the map between the beginning and end points. The integrated final path results are obtained by automatic searching of each regional path.

(1) Regional geographic information planned for smart city is divided according to grid form, each of which is a cell. The cells can be divided into three types according to the geographic information constraints of grids. The first kind can be connected to tower cell constructing. The second kind can be passed through the line, which cannot construct the tower cell. The third kind cannot pass the line and cannot establish the tower cell. Among them, the network area is county planning area, high altitude repeated ice area, densely populated housing area, rivers and lakes, etc., and the linear shape is power line

(2) Obtain the final value matrix through the BP neural network algorithm, and then, perform the initial setting of the pheromone value of each cell

(3) The cell positions of the start and end points as well as the distance from which the tower and pole are determined

(4) Neighbor cells need to meet some requirements, and the cell of the line traversing is not the cell of the line that has been erected

(5) The transformation cells perform the rules by selecting any of the adjacent cells according to the transformation rate. A cell with a low cost value, a high pheromone concentration, and a small line turning angle has a high conversion probability and low converse conversion probability

(6) Steps (4) and (5) are repeated until the search reaches the end of the cell. Link all the converted cells to produce a transfer path and then calculate the cell cost value

(7) There is only one case where the cost value is lower than the accumulated cell cost value from the previous search. Repeat step (4) to step (7) until the number of calculations is made

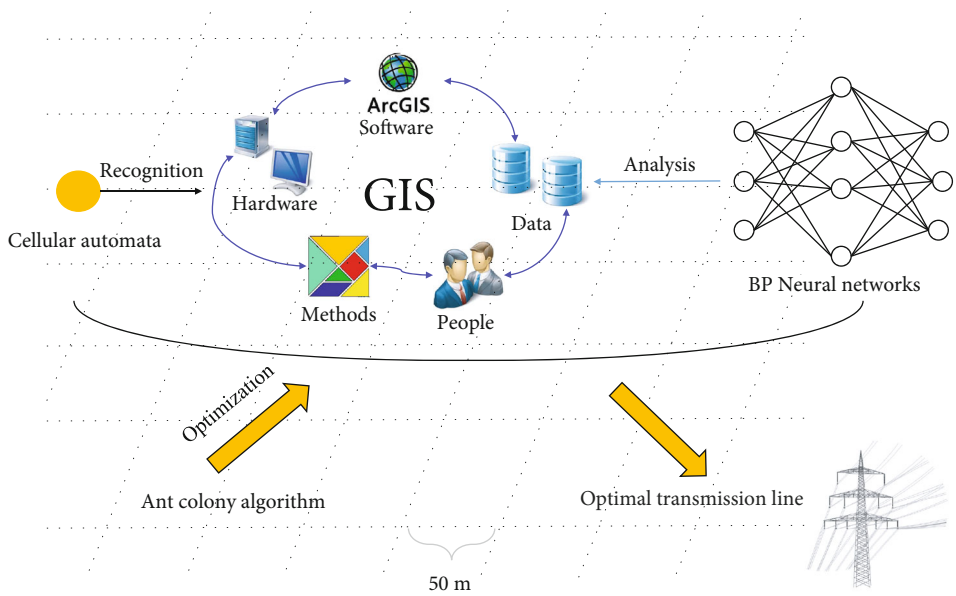


Figure 2: Scenario for transmission line planning.

(8) Update cell pheromone values on scheduled routing, and repeat step (4) to step (8) until a specified number of calculations is made. Record the final solution cell and path

4. Results

4.1. Experimental Data. To verify the feasibility of the model and method, this study used drone aerial photography technology, combined with the transmission line routing requirements, and produced a GIS map (a vector diagram with a resolution of 5 m) of a smart city planning area. The following applies the path search model of the transmission line established in this article to conduct a specific path search test analysis for this area. Considering the accuracy requirements of path selection and the data volume (the smaller the grid size, the larger the data volume), the grid size is set to 50 m × 50 m, and the ArcGIS tool is used to generate a grid format GIS map. At the same time, the elevation data of the original vector GIS map is converted into slope data. Combine terrain information and geographic images to set noncrossable areas (such as airports and protected areas; the cost value of the grid is marked as -9999) and cross-over areas and their buffers (such as highways, rivers, and transmission lines) are 50 m.

4.2. Comparative Analysis. The scenario for transmission line planning is shown in Figure 2. In smart city transmission cable route selection, AI route optimization is used simultaneously with Dijkstra algorithm's route search and manual route selection. Through manual route selection, combined with detailed fund collection and field investigation, the recommended route scheme is finally determined. The comparison between the AI path of this research and the path search and manual routing results of Dijkstra algorithm is shown in Table 3. As indicated in Table 3, the general trend of AI path is basically the same as that of manual route selection, but there is little difference. The evaluation indicators obtained by using this article based on the AI transmission line path search processing mechanism are better than the results obtained by the Dijkstra algorithm. In particular, the number of corners and the number of poles of the line path are significantly reduced, which has a lower cost value.

Table 3: Comparison of different search results.

Parameter	This research	Dijkstra algorithm	Manual route selection
Path length (km)	14.19	14.96	14.56
Number of corners	15	18	20
Number of poles and towers	58	60	63
Cost (yuan)	1580.08	1755.46	1900.56

After analysis, the difference between this algorithm and manual line selection is mainly due to the following reasons.

(1) GIS information lacks elevation information of crossline, so it can be crossed at any point of default crossline in intelligent algorithm. Since the intersection point is actually selected in manual route selection, there are some differences in the selection of intersection point, leading the differences in the path

(2) There is a difference between the route selection process of GIS with that of AI system, which leads to the lack of timely update time of the new route information, such as the lack of AI in the process of construction optimization

(3) Owing to insufficient depth of AI research, the requirements for the mechanical construction of

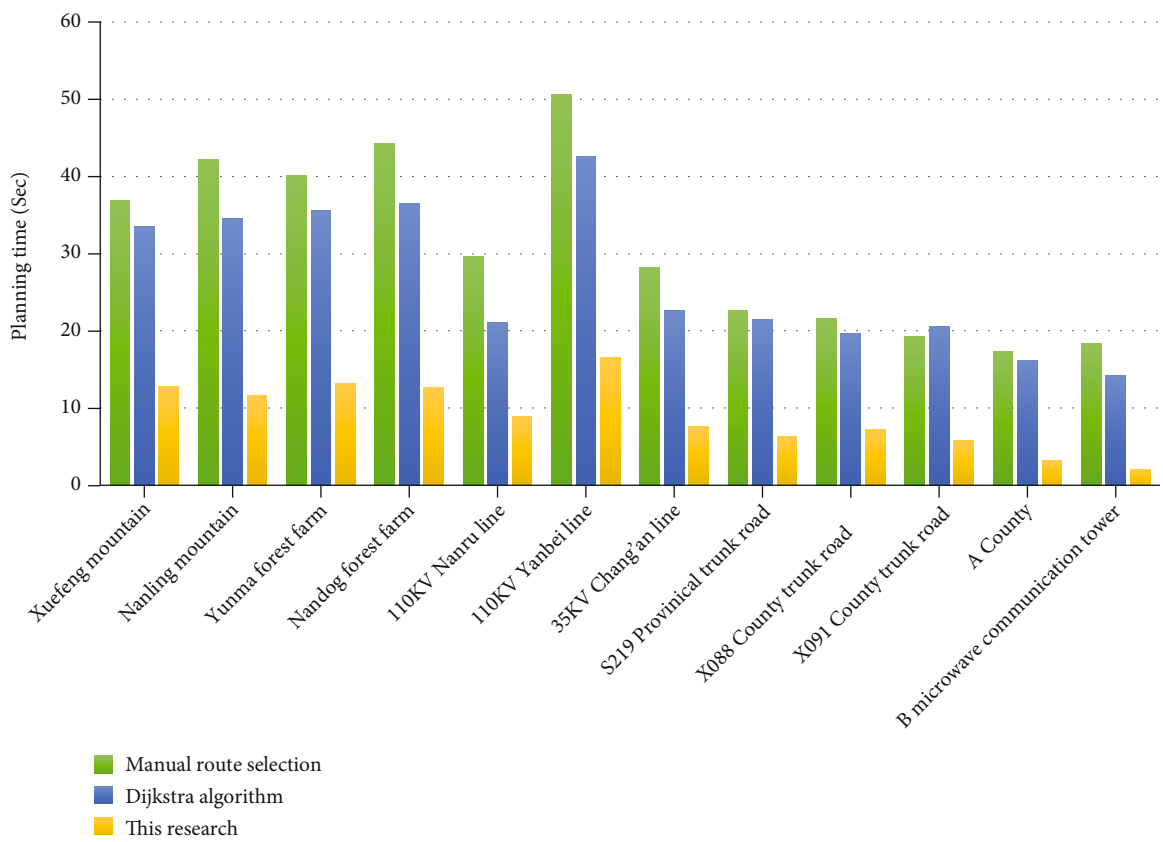


Figure 3: Line planning time in each simulation scenario.

the project cannot list the corresponding constraints. However, where routes are manually selected to guarantee an optimal route, the requirements for mechanistic construction are adjusted accordingly, so there are some differences in route directions

In addition, by simulating the scenarios of the three types of areas mentioned in Section 2.2.1, as indicated in Figure 3, the line planning time of the algorithm studied in this paper is always the lowest. Overall, line planning time is highest for area A due to the complexity of the geographical environment, while the planning time for areas B and C is shorter. However, the 110 kV Yangbei line is a special one, mainly because geological disasters often occur there. In conclusion, the algorithm proposed in this paper has a good performance in line planning time, indicating its effectiveness and pertinence.

5. Conclusions

In this paper, ArcGIS is used as the geographic information collection and analysis platform for the selection of transmission cable routes in the construction of smart cities, and an automatic search model for transmission cable routes based on AI is established. The CA model can accurately recognize geographical information; then, we use BP neural network to analyze the cost that can avoid and selectively traverse the limitations. After rigorous comparison, it can be found that some special factors cannot be considered by AI; the artificial route selection scheme has a small amount of difference between the result of the route and the AI route. Thus, it can be concluded that, without considering some special factors, AI path optimization can obtain the mathematical optimal path of transmission line. The path planning is instructive for the artificial selection of transmission cable routes in smart cities. It can be used to verify that the path plans achieve better benchmark values. In the future, this research direction can further optimize the ACO and introduce a path search direction control mechanism, which overcomes the problems of poor search direction and easy path detours of the traditional ACO and improves the search efficiency.

Data Availability

All data used to support the findings of the study is included within this paper.

Conflicts of Interest

The authors declare no conflicts of interest in this paper.

References

[1] Q. Qi, J. Z. Wu, and C. Long, "Multi objective operation optimization of an electrical distribution network with soft open point," *Applied Energy*, vol. 20, no. 8, pp. 734–744, 2017.

[2] K. Y. Zheng, "Research on the mathematical model construction and algorithm of dynamic reactive power optimization about distribution network," *Applied Mechanics & Materials*, vol. 716, no. 7, pp. 1221–1225, 2014.

[3] J. Lv, X. Wang, K. Ren, M. Huang, and K. Li, "ACO-inspired information-centric networking routing mechanism," *Computer Networks*, vol. 126, pp. 200–217, 2017.

[4] S. Ozsari, H. UGUZ, and H. HAKLI, "Implementation of meta-heuristic optimization algorithms for interview problem in land consolidation: a case study in Konya/Turkey," *Land Use Policy*, vol. 108, p. 105511, 2021.

[5] K. KORUS, M. SALAMAK, and J. Marcin, "Optimization of geometric parameters of arch bridges using visual programming FEM components and genetic algorithm," *Engineering Structures*, vol. 241, p. 112465, 2021.

[6] A. KHALED and K. FARID, "Mobile robot path planning using an improved ant colony optimization," *International Journal of Advanced Robotic Systems*, vol. 15, no. 3, pp. 172988141877467–172988141877467, 2018.

[7] L. Huang, H. Qu, P. Ji, X. Liu, and Z. Fan, "A novel coordinated path planning method using k-degree smoothing for multi-UAVs," *Applied Soft Computing*, vol. 48, pp. 182–192, 2016.

[8] G. G. Wang, H. E. Chu, and S. Mirjalili, "Three-dimensional path planning for UCAV using an improved bat algorithm," *Aerospace Science and Technology*, vol. 49, pp. 231–238, 2016.

[9] D. L. Donoho, Y. Tsaig, I. Drori, and J. L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1094–1121, 2012.

[10] K. Fyhn, M. F. Duarte, and S. H. Jensen, "Compressive parameter estimation for sparse translation-invariant signals using polar interpolation," *IEEE Transaction on Signal Processing*, vol. 63, no. 4, pp. 870–881, 2015.

[11] Y. Cheng, D. Wang, P. Zhou, and T. Zhang, "Model compression and acceleration for deep neural networks: the principles, progress, and challenges," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 126–136, 2018.

[12] F. A. N. Chengli, "Hybrid artificial bee colony algorithm with variable neighborhood search and memory mechanism," *Journal of Systems Engineering and Electronics.*, vol. 29, no. 2, pp. 405–414, 2018.

[13] C. E. F. Caetano, A. B. Lima, J. O. S. Paulino, W. C. Boaventura, and E. N. Cardoso, "A conductor arrangement that overcomes the effective length issue in transmission line grounding," *Electric Power Systems Research*, vol. 159, no. 5, pp. 31–39, 2018.

[14] Z. Jiao, K. Ma, Y. Rong, P. Wang, H. Zhang, and S. Wang, "A path planning method using adaptive polymorphic ant colony algorithm for smart wheelchairs," *Journal of Computational Science*, vol. 25, pp. 50–57, 2018.

[15] J. Liu, J. Yang, H. Liu, X. Tian, and M. Gao, "An improved ant colony algorithm for robot path planning," *Soft Computing*, vol. 21, no. 19, pp. 5829–5839, 2017.

[16] K. Obayashi and S. Masuyama, "Pilot and feasibility study on elderly support services using communicative robots and monitoring sensors integrated with cloud robotics," *Clinical Therapeutics*, vol. 42, no. 2, pp. 364–371.e4, 2020.

[17] C. H. E. N. Gang, W. A. N. G. Yuqi, W. A. N. G. Yifan, J. Liang, L. Zhang, and P. A. N. Guangtang, "Detumbling strategy based on friction control of dual-arm space robot for capturing tumbling target," *Chinese Journal of Aeronautics*, vol. 33, no. 3, pp. 1093–1106, 2020.

[18] T. Zhang, L. Ke, J. Li, J. Li, J. Huang, and Z. Li, "Metaheuristics for the tabu clustered traveling salesman problem," *Computers and Operations Research*, vol. 89, pp. 1–12, 2018.

[19] A. Smith, P. Toth, L. Bam, and J. H. van Vuuren, "A multi-tiered vehicle routing problem with global cross-docking," *Computers & Operations Research*, vol. 137, p. 105526, 2022.

[20] S. L. Gelfand, J. S. Scherer, and H. M. Koncicki, "Kidney supportive care: core curriculum 2020," *American Journal of Kidney Diseases*, vol. 75, no. 5, pp. 793–806, 2020.

[21] R. Cheng, Y. Cheng, D. Chen, and H. Song, "Online quantitative safety monitoring approach for unattended train operation system considering stochastic factors," *Reliability Engineering and System Safety*, vol. 216, p. 107933, 2021.

[22] E. Salazar, C. Henríquez, R. Sliuzas, and J. Qüense, "Evaluating spatial scenarios for sustainable development in Quito, Ecuador," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, p. 141, 2020.

[23] G. Perboli, L. Gobbato, and F. Maggioni, "A progressive hedging method for the multi-path travelling salesman problem with stochastic travel times," *IMA Journal of Management Mathematics*, vol. 28, no. 1, pp. 65–86, 2017.

[24] Q. Luo, H. Wang, Y. Zheng, and J. He, "Research on path planning of mobile robot based on improved ant colony algorithm," *Neural Computing and Applications*, vol. 32, no. 6, pp. 1555–1566, 2020.

[25] A. S. Oshaba, E. S. Ali, and S. M. Abd Elazim, "Speed control of SRM supplied by photovoltaic system via ant colony optimization algorithm," *Neural Computing and Applications*, vol. 28, no. 2, pp. 365–374, 2017.

[26] H. Fatemidokht and M. K. Rafsanjani, "F-Ant: an effective routing protocol for ant colony optimization based on fuzzy logic in vehicular ad hoc networks," *Neural Computing and Applications*, vol. 29, no. 11, pp. 1127–1137, 2018.

[27] K. Long, "Optimization of high voltage transmission line path based on AI," in *IEEE 2nd International Conference on Power, Electronics and Computer Applications*, Shenyang, China, Jan.2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 3735853, 9 pages
https://doi.org/10.1155/2022/3735853

Research Article

Machine Learning Based Distribution of Sports Video Stream Assisting Physical Training

Yanni Wang and Jinhui Li

Qiqihar University, Qiqihar 161006, China

Correspondence should be addressed to Yanni Wang; 01381@qqhru.edu.cn

Received 7 February 2022; Revised 9 March 2022; Accepted 12 March 2022; Published 28 March 2022

Academic Editor: Jianhui Lv

The rapid development of Olympic Games and the intensification of professional sports events make the competition between excellent athletes and teams become more and more fierce, so the demands on athletes' psychology, techniques, and physical strength are becoming higher and higher, especially that the demands on physical strength are more prominent. Therefore, physical training is an important part of sports training and the core link of competitive sports, which is widely valued by coaches and athletes all over the world. At the same time, with the influence and penetration of the rapid development of modern information technology and network big data on competitive sports, the scientific and digital process of athletes' physical training has also been accelerated, and many new ideas, scientific training methods, and advanced sports technology have emerged. During the pandemic, many aggregation activities have been disrupted, as well as physical training. Therefore, this paper improves the quality of physical training through video, so that athletes can do physical training at any time and any place. A cut vertices spanning tree algorithm based on machine learning is proposed to distribute layered multicast streaming and dynamically adjust the number of layers of sports video streaming. The cut vertices spanning tree algorithm is mainly applied to the situation when the network bandwidth resources are relatively scarce. The evaluation results indicate that the proposed method can improve the estimated quality of experience (QoE), packet loss rate, link utilization, and video delay on Mininet simulation platform. Furthermore, it can be seen from the experiments that the proposed method has a good performance on distribution of sports video stream assisting physical training.

1. Introduction

The theoretical research of physical training is developing gradually in China, which has a very important influence on the development of competitive sports and public health [1, 2]. Physical training is the size of cardiopulmonary function support caused by human movement, that is, the ability of cardiopulmonary system, which depends on the strength and scope of human movement system dynamics application [3]. The current physical training is strength training, speed training, aerobic, anaerobic endurance training, flexibility, and coordination training. As a result, physical training is difficult to connect with its supporting disciplines, and it is difficult to combine with biological science, nutrition, medicine, and other high-tech aspects, which also leads to the emergence of many problems such as training difficult to control [4, 5].

The connotation of physical training should be the training of human body to improve the ability of walking, running, jumping, and throwing. Physical training has its systematic characteristics, that is, the integrity and comprehensiveness of training [6–8]. Physical training should combine the characteristics of sports and human body and cross disciplines through multidisciplinary support, and learn from and integrate the excellent achievements and advanced experience in other fields, so as to optimize the bone, joint, muscle, and sports auxiliary system of human body sports system and the large system of sports training in physical training [9, 10].

At present, there are some deficiencies in physical training. (i) In experiential physical training, physical training is a

Table 1: Symbol illustration.

Variable	Description
$G(V, E)$	Undirected connected graph
v	Chile node
u	Father node
n	Node
$\mathrm{ts}[n]$	Timestamp of node n
$\mathrm{ets}[n]$	Earliest timestamp of node n
CV	Cut vertex sets
CV'	Useful cut vertex sets
MT	Multicast tree
n_t	Terminal node
$\mathrm{Path}_{n_t,u}$	Path with higher bandwidth
TN	Terminal node sets
S_N	Node sets
S_L	Link sets
$\mathrm{MT}(S_N, S_L)$	State of MT

part of sports training, and its theoretical basis is inseparable from sports training, so its theoretical content is relatively more empirical. With the development of science and technology and the maturity of human anatomy, physiology, biomechanics, and other subjects, the scientific theory of physical training has become possible, but the deviation of understanding of physical training still leads to the existence of the problem of empirical theory. Especially in China, running is usually considered as the best way for physical training. (ii) For cognitive bias, core strength training, altitude training, and other training theories are being applied to specific sports training. But by definition, its essence is to train the core muscles of the body and its deep muscle capabilities, including strength, stability, balance, and other abilities. With the concept of core strength training, the connotation of this training theory is misinterpreted and its extension is narrowed. (iii) It can be seen from the essence of physical strength that the human body's athletic ability is determined by the structure of various organ systems with different physiological functions. System theory believes that the structure of the system determines the function, and only when the structure is reasonable, the function can be played better.

In the post epidemic era, physical training is usually realized in the form of video. Compared with the traditional mode, video can enable students to perform more effective physical training. However, how to bring good training effect for physical training through video becomes the focus of research. The main contribution of this study is that the spanning tree algorithm of machine learning is improved, and the cut vertices spanning tree algorithm is proposed to realize the distribution of sports video streaming and assist physical training.

The rest of the paper is organized as follows: Section 2 summarizes the related research; in Section 2, the cut vertices spanning tree algorithm is proposed; Section 4 discusses the results; and Section 5 concludes the paper.

2. Related Research

Digital physical training is the inevitable trend of the development of physical training. Digital makes physical training more rational, and the quality of physical training will ultimately depend on scientific digital monitoring. At the same time, digital physical training is also the only way to intelligent physical training.

College students' physical health test is an important part of school physical education and school education evaluation system; in [11], the authors proposed a functional exercise test based on artificial intelligence (AI) to improve college students' physical exercise awareness. To improve the effect of physical training, in [12], the authors combined machine learning to identify physical training characteristics and action prediction, combined with the Internet of things technology to process physical training data, and constructed a machine learning and the Internet of things based on physical education and training system. To solve the problem of human motion recognition, in [13], the authors proposed a recognition system for human motion tracking in physical training. In [14], the authors presented a balanced shunt point field programmable gate array-based method of physical training planning to improve the effectiveness of physical training planning. In [15], the authors proposed a design method of computer multimedia simulation action training model based on fast detection and control of attitude change space, and the computer multimedia simulation method was used to construct human action training and mathematical model of human action.

The development of the Internet technology makes streaming media applications popularized rapidly. However, in the face of ever-expanding users and increasingly diversified content service requirements, the biggest challenge faced by service providers is the high deployment and maintenance cost of distribution system, poor scalability, and other problems, so how to effectively distribute massive video streaming has become a key problem that needs to be solved urgently. In [16], a platform compatible dynamic adaptive streaming over HTTP framework for wireless video sensor networks (WVSNP-DASH) was proposed to flexibly access video from sensors and other miniaturized source nodes. Previous studies had neither addressed appropriate data caching to support vehicle mobility nor provided appropriate seamless forwarding to ensure quality of service and QoE for real-time video streaming services. To solve this problem, in [17], a video packet distribution scheme called Clone was presented to integrate vehicle-to-vehicle communication for video streaming. In [18], a new emergency aware BitTorrent streams mechanism (UR-Aware) was proposed to improve the transmission efficiency of BitTorrent streams on peer-to-peer networks by balancing playback continuity and fragment scarcity distribution. In [19], the authors proposed a new scheme of content centric networking multisource video streaming based on quality of experience (QoE) perception. In [20], the theoretical framework proposed by the authors could analyze the performance of synchronous streaming media, video on demand streaming media, and video downloading. In [21], a new reinforcement learning-based viewport adaptive streaming

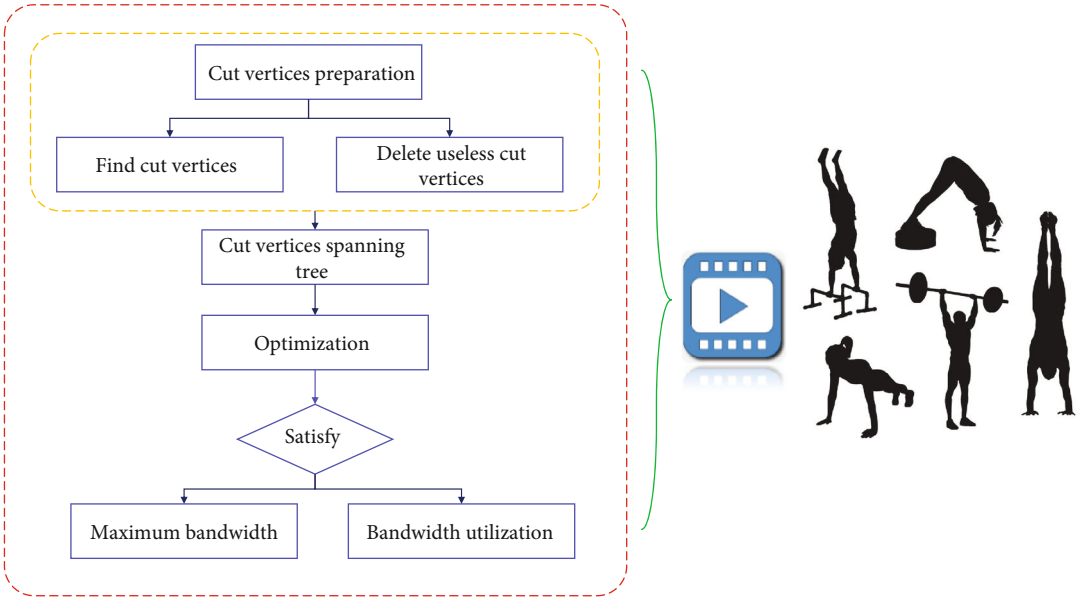


Figure 1: Overall scenario of sports video streaming for physical training.

framework was proposed to optimize 360-degree video streaming from the perspective of viewport prediction, prefetching scheduling, and rate adaptation. In [22], a design is effective in streaming, distribution, and caching multimedia content was proposed. In [23], a new energy efficiency method based on experience quality was proposed to distribute peer-to-peer video in mobile ad hoc networks. In [24], the authors used an online algorithm to jointly and dynamically select the right data center for broadcasters and viewers to save the operating cost of crowdsourced live streaming provider. In [25], the authors proposed a video control plane, which was used to monitor the QoE sent by any content delivery networks (CDN) in the CDN pool and to select the QoE with the best performance when a new video request is received.

3. Proposed Sports Video Streaming Algorithm

Definition 1. Cut vertex. In an undirected connected graph $G(V, E)$, if for a node $v \in V$, after deleting v and all edges connected to v from G, G will become multiple unconnected subgraphs, and then v is called a cut vertex of G.

The distribution of sports video streaming based on cut vertices spanning tree algorithm includes three stages, which are cut vertices preparation, spanning tree, and optimization.

3.1. Cut Vertices Preparation. In the cut vertices stage, two steps are mainly implemented. The first step is to find all the cut vertices in the network and form the cut vertex sets. The second step is to delete the useless cut vertices from the cut vertex sets and form the useful cut vertex sets.

By using the idea of depth first search (DFS) [26], the cut vertices in the network are solved. The idea of finding the cut vertices is roughly as follows. If the node u is reached during DFS, the nodes in the network are divided into two types: one has been DFS, and the other has not been DFS. Then, DFS is performed again on the child node v of node u. Here, u is the father of v, and the visited node is called v's ancestor. If v can return to its ancestors without passing through u, it means that u is not a cut vertex. On the contrary, if v cannot go back to its ancestors, it means that u is the cutting vertex.

$ts[n]$ is used to represent the timestamp of node n during DFS, and $ets[n]$ is used to represent the earliest timestamp that node n can trace back. The timestamp here refers to the number of accessed nodes during traversal. For node u, if there is a child node v, satisfying $\mathrm{ets}[n] \geq \mathrm{ts}[n]$, that is, it cannot be traced back to the ancestor. In this case, if u is not the root node, it can be directly judged that u is the cut vertex. If u is the root node and it has multiple child nodes, it can also be judged as a cut vertex.

After finding all the cut vertices in the network, the cut vertex sets, CV, are obtained. However, since the focus is to construct the spanning tree from the source node to each terminal node, some network nodes do not have to be connected [27]. Therefore, some useless cut vertices can be removed according to specific requirements, and finally a useful cut vertex sets CV' is obtained.

3.2. Spanning Tree. Prim spanning tree [28] is used for useful cut vertex sets, CV'. First, the link with the smallest bandwidth is accessed. Then, start with the cut vertices at both ends of the link, find the link with the smallest bandwidth connected to it, and add it to the multicast tree, MT, and iterate successively until no cut vertices can access MT. Since MT nodes are cut vertices at present, this is called a full-cut spanning tree. When links are added, cut vertices and terminal nodes that are not connected to the MT are grafted to the MT. The priority of grafts is determined by the sum of the bandwidths of the links that are connected to the paths in the spanning tree. The smaller the sum of bandwidths of the links, the higher the priority is. Since the spanning tree at this stage is dominated by cut vertices but contains common nodes, it is called the initial state cut vertices spanning tree.

```
Input: G(V, E), TN
Output: MT(S_N, S_L)
use finding cut vertex sets method to get cut vertex sets CV;
forcv ∈ CV
   if the subnet connected to cv does not have terminal nodes
      delete cv from CV
end for;
use Prim algorithm for all CV to get a basic MT
while (all v ∈ TN are not all in S_N) do
   for all v ∈ TN − n_t
      find v with the smallest bandwidth value used to connect to the current MT
      add the terminal node and path to MT
   end for;
end while;
for all terminal nodes
   if maximum bandwidth requirements are not met
      find other paths that can increase bandwidth reversely
         if find a path with higher bandwidth
            replace the original path
         end if;
   end if;
end for;
for all links
   if find links that can improve link utilization without reducing the bandwidth requirements of the end nodes
      replace the original path
   end if;
end for;
returnMT;
```

Algorithm 1: Cut vertices spanning tree algorithm

3.3. Optimization. This stage mainly optimizes the initial state cut vertices spanning tree obtained in the spanning tree stage. The optimization stage is divided into two metrics. The first metric is to check whether the maximum bandwidth of the terminal node is satisfied or to the best of its ability. This metric is also first priority if no alternative path is available. The second metric is to check whether the bandwidth utilization of the link can be improved again without reducing the first metric. In this stage, the method of reverse path check is also adopted [29]. All terminal nodes are traversed, checking whether it can be the highest bandwidth requirement in the currentMTsatisfied; if meeting the check, whether the link utilization can be improved, that is, after another link is connected to theMT, the link utilization is improved compared with that of the originalMT. If cannot be improved, to continue, check the next terminal node. If the highest bandwidth demand of the terminal node cannot be met in the current,MT, start from the terminal node,n_t,and reverse search to see if there is a path with higher bandwidthPath$_{n_t,u}$to replace the original path. In this process, the change can only be made under the condition that the transmission quality of other terminal nodes is not reduced. If the reverse path reaches the source node, a path with higher bandwidth Path$_{n_t,u}$ is found, and the original path can be replaced. Otherwise, keep the original state and perform the preceding steps to check the link bandwidth utilization.

Let TN be the terminal node sets. MT(S_N, S_L) is used to represent the state of MT in the process of building a branching node spanning tree, where S_N represents the node sets in MT and S_L represents the link sets in MT.

The variables used in this study are shown in Table 1 (according to the order in the paper).

The overall scenario can be shown as Figure 1.

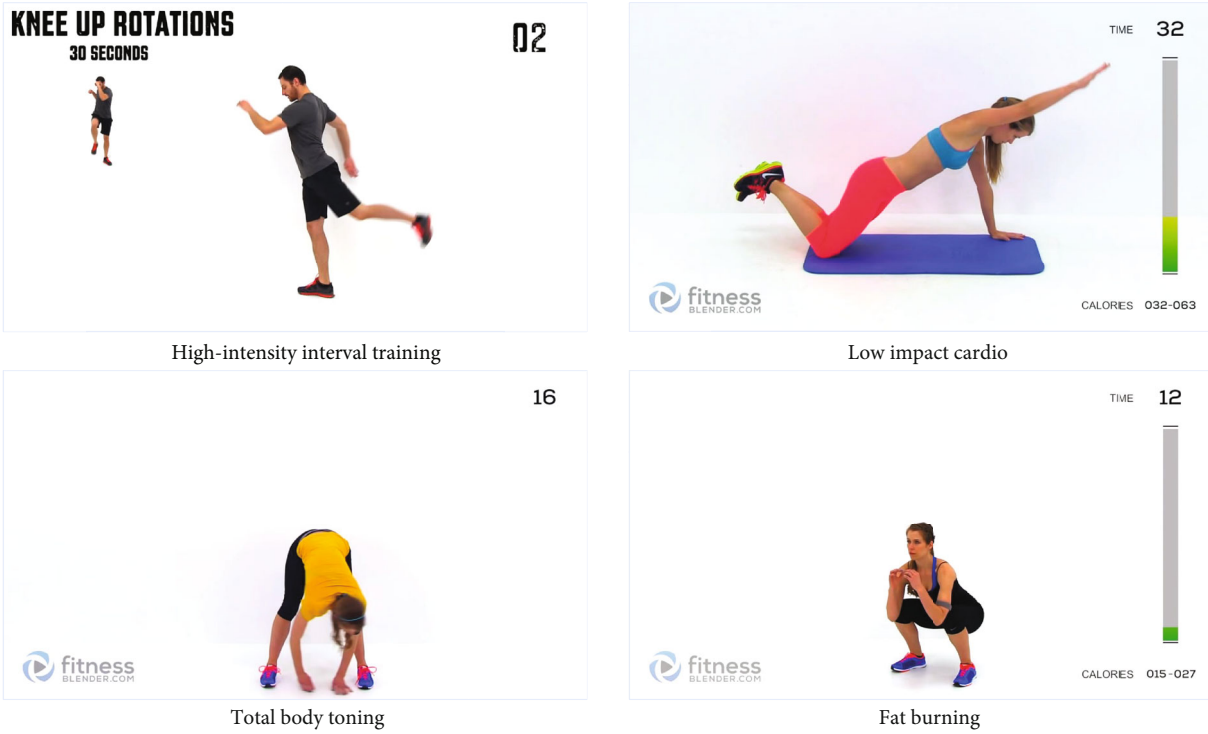
Pseudocode for the cut vertices spanning tree algorithm is shown as follows.

Table 2: Parameter settings.

Parameter	Number
Server	1
Proxy server	2
User	100
User requests	1-5000

4. Evaluation

4.1. Setups. The simulation is driven on Mininet, which is a very useful network simulation tool. It can run a virtual network composed of multiple terminals, multiple switches, and multiple links. The port can be customized for the terminal, and the parameters such as bandwidth and packet transmission latency can be designed for the link. The number of experimental parameters is shown in Table 2.



High-intensity interval training

Low impact cardio

Total body toning

Fat burning

Figure 2: Snapshots of physical training videos.

4.2. Metrics

(1) Estimated quality of experience (QoE)

In physical training video transmission, QoE is the most important metric to evaluate the performance of the algorithm. However, questionnaire survey is the most direct way to obtain QoE. Therefore, the validity of the sports video streaming distribution based on cut vertices spanning tree algorithm proposed in this paper is tested through the obtained estimated QoE.

(2) Packet loss rate

It refers to the amount of data lost in the transmission of a packet as a percentage of the amount of input data. Packet loss rate is a critical metric for video transmission quality. A small packet loss rate indicates that link congestion is rare, while a large packet loss rate indicates that link congestion is frequent.

(3) Link utilization

Link utilization refers to the percentage of throughput in a specific time interval in the link access rate. The higher the efficiency of congestion control algorithm, the better the bandwidth allocation algorithm, and the higher the link utilization.

(4) Video delay

When the layered video arrives at the terminal node, it will be cached and waited and then reordered according to the rule number. If the delay is too large, the layered video will not be sorted in time, and the playback quality of the played video will be very poor.

4.3. Physical Training Video Traces. In this study, four videos were selected from Fitness Blender with over 600 free home workout videos for evaluation. All these video traces were 1920×1080 resolution, and one-minute segments of each video were extracted for the experiment. The video clips of the four videos are shown in Figure 2, which are high-intensity interval training, low impact cardio, total body toning, and fat burning.

4.4. Performance Comparison. To verify the effectiveness of the proposed algorithm, some popular video streaming distribution algorithms were selected for comparison in this study which are WVSNP-DASH [16], Clone [17], and UR-Aware [18].

As can be seen from Figure 3, the estimated QoE of the algorithm proposed in this paper always remains above 9. The QoE of the four algorithms on fat burning video is not very good. Only the QoE of the algorithm proposed in this paper exceeds 9, and the QoE of the other three baselines is below 9, while the QoE of the Clone algorithm does not even exceed 8; this is because the frames per second of fat burning videos is 60. Throughout the QoE of the four algorithms on the four videos, except for the satisfactory QoE of the algorithm proposed in this paper, the QoE of the other three baselines were more or less unsatisfactory with different video contents. Among them, the better is the QoE of WVSNP-DASH algorithm. QoE is a very important indicator in the scenario of improving physical training effects through videos constructed in this paper, and the level of QoE determines users' recognition of this form.

Figure 4 shows that even with the increasing number of user requests, the packet loss rate of the algorithm proposed

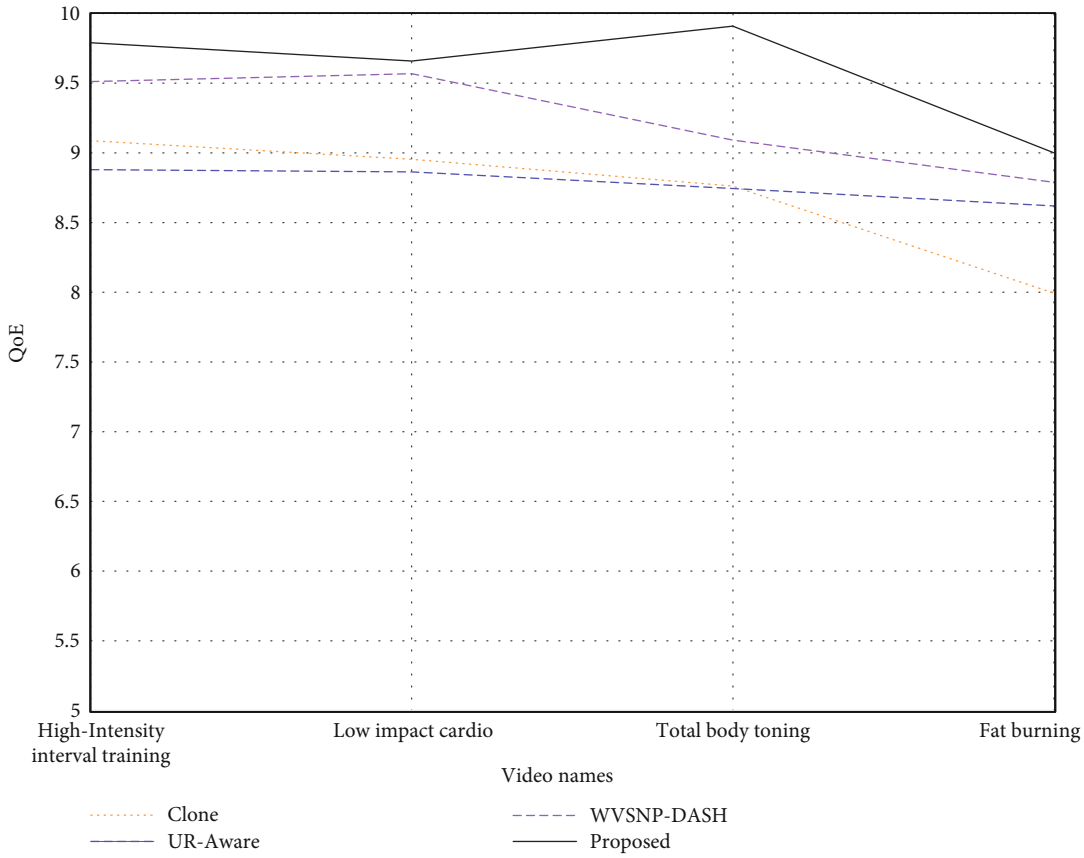


Figure 3: QoE for different videos under four algorithms.

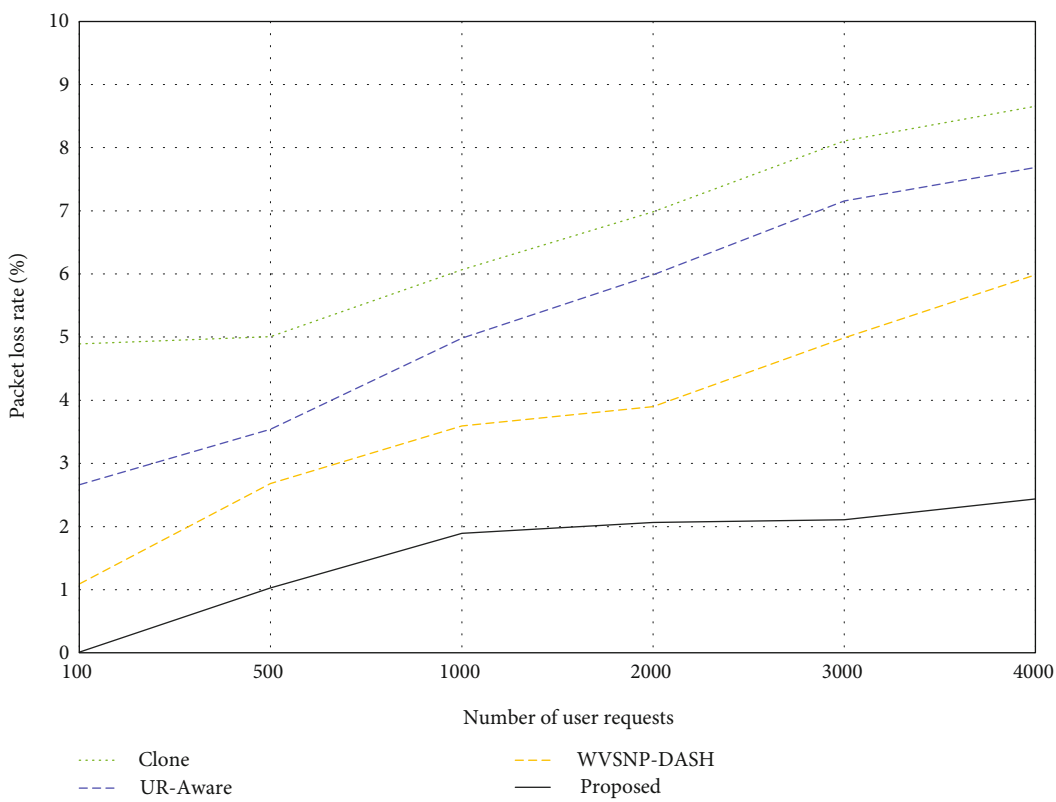


Figure 4: Packet loss rate for different videos under four algorithms.

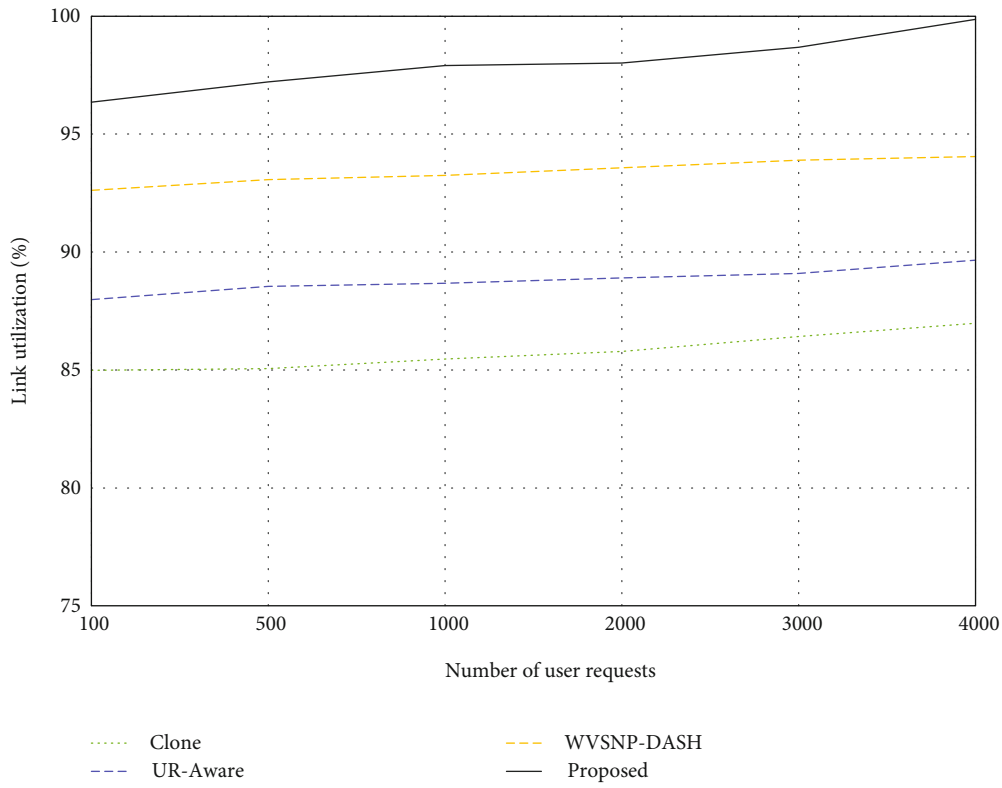


Figure 5: Link utilization of different algorithms under number of user requests.

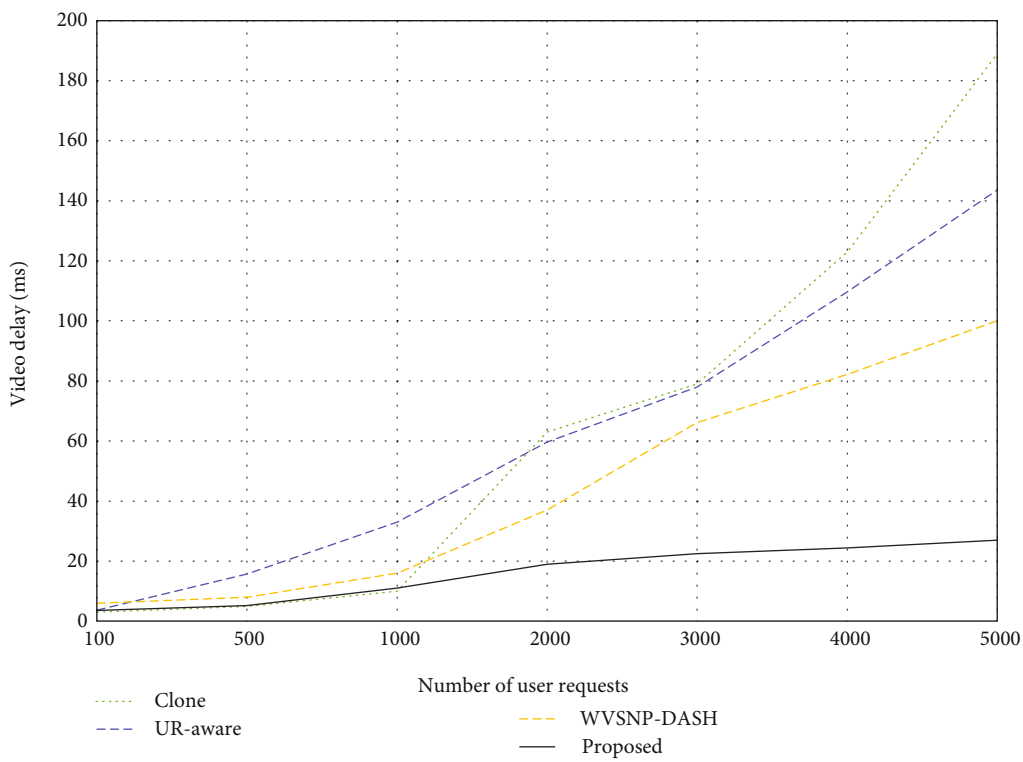


Figure 6: Video delay of different algorithms under number of user requests.

in this paper remains below 3% and tends to be stable. Packet loss rate for the other three baselines have been increasing. However, the packet loss rate of WVSNP-DASH algorithm is acceptable because the WVSNP-DASH framework is based on independent playable video clips, which causes specific naming syntax and can transfer basic metadata, so as to promote flexible search, transmission, distribution, and playback. In the physical training through videos, if the packet loss rate is too high, the video pull stream will be incomplete and the video quality will be reduced to a certain extent, which is not a good experience for users.

As indicated in Figure 5, even with the increasing number of user requests, the link utilization of the algorithm proposed in this paper always remains above 95%, which is a good phenomenon. However, the link utilization of UR-Aware and Clone algorithms is always below 90%. Although the link utilization of WVSNP-DASH algorithm does not perform well, it is always at a very stable level. This is because that WVSNP-DASH framework has low client load and provides significant power saving potential on the source node serving video streams.

It can be seen from Figure 6 that with the increasing number of user requests, the video delay of the three baselines presents an exponential growth trend, while the video delay of the algorithm proposed in this paper is always within 40 ms. But on the whole, the delay of the four algorithms is acceptable. Surprisingly, the video delay of Clone algorithm is lower than that of the proposed algorithm when the number of user requests is less than 1000. Video delay is very important in physical training through videos, and lower video delay ensures the quality of video, thus providing users with a high quality reference for physical training.

5. Conclusions

In this paper, video is used to improve the quality of physical training, and a cut vertices spanning tree algorithm is proposed to distribute layered multicast streaming and dynamically adjust the number of layers of sports video streaming. At first, all the cut vertices in the network are found, and the cut vertex sets are formed, and the useless cut vertices are deleted from the cut vertex sets, and the useful cut vertex sets are formed. Then, Prim algorithm is used to generate the initial state spanning tree of useful cut vertices. Finally, the initial cut point spanning tree that obtained in the spanning tree stage is optimized. The simulation is driven on Mininet, compared with baselines; the proposed algorithm has good performance in terms of estimated QoE, packet loss rate, link utilization, and video delay on physical video streaming distribution.

With the mobile possibilities of fitness training, we will investigate video distribution in mobile edge networks. (i) In mobile edge network, edge nodes have computing capability, which can encode, decode and transcode video content, and this features can be taken into account in video distribution. (ii) The operating environment of edge nodes in video distribution is complex and changeable. If edge nodes need to be trained from every time they receive a new task, the learning speed of edge nodes may not keep pace with the speed of environmental changes or task changes. (iii) Users can take advantage of the computation, storage, and communication capabilities of terminals to participate in video distribution.

Data Availability

All data used to support the findings of the study is included within this paper.

Conflicts of Interest

The authors declare no conflicts of interest in this paper.

Acknowledgments

This work was supported by Homology Study of Olympic Events and Winter Sports of Ethnic Minorities in Heilongjiang Province (Granted No. 145109162).

References

[1] S. Zhao, "Research on scientific sports training of students majoring in physical education," *Revista Brasileira de Medicina do Esporte*, vol. 27, no. 5, pp. 460–463, 2021.

[2] K. Xu, "Whole body physical training in the treatment of chronic low back pain," *Revista Brasileira de Medicina do Esporte*, vol. 27, no. 3, pp. 342–345, 2021.

[3] S. Ketelhut and R. G. Ketelhut, "Type of exercise training and training methods," *Physical Exercise For Human Health*, vol. 1228, pp. 25–43, 2020.

[4] J. Molina-Hernandez and A. Garcia-Leon, "Effects of personal strength training on adolescent physical symptoms," *REVISTA IBEROAMERICANA DE DIAGNOSTICO Y EVALUACION-E AVALIACAO PSICOLOGICA*, vol. 3, no. 60, pp. 105–116, 2021.

[5] M. Xie, "Design of a physical education training system based on an intelligent vision," *Computer Applications in Engineering Education*, vol. 29, no. 3, pp. 590–602, 2021.

[6] J. K. Kim and Y. C. Choi, "The effect of short-term off-season cross-country ski training on body composition, physical fitness, and isokinetic muscle functions of cross-country skiers," *Physical Fitness, And Isokinetic Muscle Functions Of Cross-Country Skiers*, vol. 16, no. 1, pp. 63–74, 2020.

[7] M. M. Mariño, F. J. Grijota, I. Bartolomé, J. Siquier-Coll, V. T. Román, and D. Muñoz, "Influence of physical training on erythrocyte concentrations of iron, phosphorus and magnesium," *Journal of the International Society of Sports Nutrition*, vol. 17, no. 1, 2020.

[8] Y. Zhu, C. B. Sivaparthipan, and V. Vinothraj, "Analysis of physical health with internet of things-based computational narrowband physical health framework," *Technology and Health Care*, vol. 29, no. 6, pp. 1217–1231, 2021.

[9] S. Hermassi, K. Laudner, and R. Schwesig, "The effects of circuit strength training on the development of physical fitness and performance-related variables in handball players," *Journal of Human Kinetics*, vol. 71, no. 1, pp. 191–203, 2020.

[10] A. D. Reubens-Leonidio, T. G. P. de Carvalho, M. B. D. de Cerqueira Antunes, and M. V. G. de Barros, "Educação interprofissional e prática colaborativa na formação em educação

física: reflexões de uma experiência na perspectiva da tutoria," *Saude E Sociedade*, vol. 30, no. 3, 2021.

[11] Q. Guo and B. Li, "Role of AI physical education based on application of functional sports training," *Journal Of Intelligent & Fuzzy Systems*, vol. 40, no. 2, pp. 3337–3345, 2021.

[12] C. Wang and C. Du, "Optimization of physical education and training system based on machine learning and internet of things," *Neural Computing & Applications*, vol. 33, pp. 1–16, 2021.

[13] X. Leng, H. Jiang, and X. Zuo, "Human action tracking design of neural network algorithm based on GA-PSO in physical training," *Cluster Computing*, vol. 22, no. 2, pp. 4149–4155, 2019.

[14] Q. Sun, "Formulation of physical education and training program based on multidimensional education data mining," *Cluster Computing*, vol. 22, no. S2, pp. 5017–S5023, 2019.

[15] J. Dan, Y. Zheng, and J. Hu, "Research on sports training model based on intelligent data aggregation processing in internet of things," *Cluster Computing*, vol. 25, no. 1, pp. 727–734, 2022.

[16] A. Seema, L. Schwoebel, T. Shah, J. Morgan, and M. Reisslein, "WVSNP-DASH: name-based segmented video streaming," *IEEE Transactions on Broadcasting*, vol. 63, no. 3, pp. 346–355, 2015.

[17] Y. Shin, H. Choi, Y. Nam, and E. Lee, "Video packet distribution scheme for multimedia streaming services in VANETs," *Sensors*, vol. 21, no. 21, p. 7368, 2021.

[18] C. Lin and J. Lin, "UR-aware: streaming videos over BitTorrent with balanced playback urgency and rareness distribution," *Peer-To-Peer Networking And Applications*, vol. 9, no. 6, pp. 1114–1125, 2016.

[19] M. N. Sadat, R. Dai, L. Kong, and J. Zhu, "QoE-aware multisource video streaming in content centric networks," *IEEE Transactions on Multimedia*, vol. 22, no. 9, pp. 2321–2330, 2020.

[20] Y. Zhou, L. Chen, M. Jing, Z. Ming, and Y. Xu, "Performance analysis of thunder crystal: a crowdsourcing-based video distribution platform," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 4, pp. 997–1008, 2018.

[21] Z. Jiang, X. Zhang, Y. Xu, Z. Ma, J. Sun, and Y. Zhang, "Reinforcement learning based rate adaptation for 360-degree video streaming," *IEEE Transactions on Broadcasting*, vol. 67, no. 2, pp. 409–423, 2021.

[22] M. H. Bin Waheed, F. Jamil, A. Qayyum et al., "A new efficient architecture for adaptive bit-rate video streaming," *Sustainability*, vol. 13, no. 8, p. 4541, 2021.

[23] M. S. Raheel and R. Raad, "Energy efficient technique for P2P multisource video streaming in mobile ad hoc networks to improve QoE," *Peer-To-Peer Networking And Applications*, vol. 13, no. 1, pp. 219–242, 2020.

[24] C. Dong, W. Wen, T. Xu, and X. Yang, "Joint optimization of data-center selection and video-streaming distribution for crowdsourced live streaming in a geo-distributed cloud platform," *IEEE Transactions on Network and Service Management*, vol. 16, no. 2, pp. 729–742, 2019.

[25] L. De Cicco, S. Mascolo, and V. Palmisano, "QoE-driven resource allocation for massive video distribution," *Ad Hoc Networks*, vol. 89, pp. 170–176, 2019.

[26] N. Enriquez, G. Faraud, and L. Menard, "Limiting shape of the depth first search tree in an Erdos-Renyi graph," *Random Structures & Algorithms*, vol. 56, no. 2, pp. 501–516, 2019.

[27] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Transactions On Cybernetics*, vol. PP, pp. 1–13, 2021.

[28] N. Albin, J. Clemens, D. Hoare, P. Poggi-Corradini, B. Sit, and S. Tymochko, "Fairest edge usage and minimum expected overlap for random spanning trees," *Discrete Mathematics*, vol. 344, no. 5, p. 112282, 2021.

[29] J. Lv, X. Wang, K. Ren, M. Huang, and K. Li, "ACO-inspired information-centric networking routing mechanism," *Computer Networks*, vol. 126, pp. 200–217, 2017.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5444055, 13 pages
https://doi.org/10.1155/2022/5444055

Research Article

A Decomposition and Dominance-Based Multiobjective Artificial Bee Colony Algorithm for Multiple Sequence Alignment

Lei Ye

College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China

Correspondence should be addressed to Lei Ye; hunnan214@163.com

Received 10 January 2022; Accepted 23 February 2022; Published 24 March 2022

Academic Editor: Jianhui Lv

The multiple sequence alignment (MSA) problem is essential in biological research for finding a specific relationship between the biologic sequences and their function. This study proposes a decomposition and dominance-based multiobjective artificial bee colony optimization algorithm for MSA (MOABC/D-MSA). MOABC/D-MSA uses three kinds of searching strategies to obtain a group of nondominated solutions with high quality and diversity of an MSA problem. A decomposition-based employed bee strategy is proposed to search for high-performance solutions of the MSA, while insuring their diversity. A nondominated sorting-based onlooker strategy searches for the solutions near the Pareto front (PF) to guide the subsequent searching. The scout bee strategy facilitates the algorithm to get out of the local optimal. A comparative experiment is implemented on BAliBASE 3.0, a benchmark for MSA algorithms. Experimental results show that the proposed algorithm has competitive performance with state-of-the-art metaheuristic algorithms. Furthermore, nondominated solutions of MOABC/D-MSA have a more uniform distribution in the objective space.

1. Introduction

The multiple sequence alignment (MSA) problem aims to find shared fragments among three or more sequences. In general, the sequences to be aligned are biological sequences, such as nucleotides, amino acids, and proteins. The results of the alignment are a set of sequences that are transformed from the original sequences by inserting several gaps. The sequences in the alignment solution have the same length. From the alignment results, researchers could identify regions that exist the same elements among the sequences. These regions could reveal potential relations among the input sequences, such as the evolutionary relationship among the sequences [1] and the influence of special sites on structure and function.

The MSA problem has been proven to be an NP-complete problem with the sum-of-pairs (SP) metric [2] and NP-hard for most of the existing metrics [3]. Therefore, researchers have paid a lot of attention to developing an effective and efficient approach for MSA. Different kinds of algorithms are proposed to solve the MSA problem.

Literature [4] finds six main groups of different approaches for dealing with biological MSA problems: (1) exact methods, (2) progressive methods, (3) consistency-based methods, (4) iterative methods, (5) evolutionary algorithms, and (6) structure-based methods.

The metaheuristic algorithms, including evolutionary algorithms and swarm intelligence algorithms, have shown competitive performance in optimizing MSA problems. Literature [4] proposed a hybrid multiobjective metaheuristic for MSA, which combines the shuffled frog-leaping optimization algorithm with the fast and accurate Kalgin algorithm. Literature [5] proposes a characteristic-based framework for MSA, which extracts the characteristics of unaligned input sequences and aligns the sequences with the specific configuration according to the characteristics. Ten kinds of characteristics are considered in this study; they are divided into three groups. A multiobjective evolutionary algorithm for MSA (MOMSA) is proposed in reference [2]. MOMSA is a biobjective aligner based on the framework of MOEA/D. The Tchebycheff approach is adopted for subproblem design. Furthermore, this study proposes a tree-

based initialization method and a gap-reinserting-based mutation operator. Literature [6] proposes a quantum-inspired heuristic optimization method for MSA, where a quantum-inspired GVN routine is designed. Literature [1] proposes a multiobjective formulation for MSA for inferring the evolutionary history relating the sequences known as phylogenetic trees. ProbPFP [7] combines the partition function and the hidden Markov model (HMM). The parameters of the HMM are optimized by particle swarm optimization (PSO). Literature [8] proposes an algorithm that employs sparse approximation to reduce the computational cost for the relaxation. Literature [9] proposes a hybrid artificial bee colony (ABC) optimization algorithm for MSA. This algorithm performs a single-point crossover for the employed bee phase. It performs a multiple mutation operator that contains four kinds of mutations operators for the onlooker bee phase. The Kalign2 algorithm is implemented in the scout bee phase to align a segment of the sequences.

Early literature evaluates the alignments by one score function. However, when optimizing the MSA problem, researchers have more than one requirement for aligned sequences. For example, researchers hope to find as many similar fragments as possible. Meanwhile, they want to insert gaps as much as possible. To meet the different requirements of researchers, the MSA problem is designed to contain multiple optimization objectives. Researchers have proposed different definitions of multiobjective MSA problems. Reference [4] adopts the weighted sum-of-pairs function with affine gap penalties (WSP) and the number of conserved total column (TC) score. Reference [5] uses the Q-score (i.e., sum-of-pairs (SP) score) and total column (TC) score as the objective functions. In reference [2], the MSA problem is a biobjective minimization problem, where the first objective function is the scoring function that minimizes the number of gaps and the second objective function is the opposite of the SP function. Four new objectives are proposed in reference [1]as follows: nongap columns for the calculation of entropy, the similarity of columns containing one or more gaps, the similarity of columns containing no gap, and the number of consecutive gaps. Reference [8] develops a relaxed formulation for the MSA problem based on a regression-coding framework.

Since the optimization objectives in a multiobjective MSA are conflicted in most cases, it is hard to optimize all objectives simultaneously. In practice, the optimization algorithm solves the multiobjective problem by working out a set of solutions that reflects the trade-off of the objectives. The relationship between the solutions of the set of solutions is Pareto nondominated.

Metaheuristics have been proven to be effective in optimizing multiobjective problems [10–13]. In this study, a typical metaheuristic algorithm, ABC [14], is adopted to optimize a multiobjective MSA problem. ABC is one of the popular metaheuristic algorithms. It mimics the behaviors of three kinds of bees: employed bees, onlooker bees, and scout bees. The earliest ABC for MSA known in this research is a single-objective algorithm [15]. The motivation for using ABC in this research is that its searching strategy is suitable for solving a complicated multiobjective problem. The ABC uses different kinds of searching strategies; therefore, it can balance the convergence speed and solution quality. Furthermore, the coupling between the three stages of the ABC is low, and the algorithm designer can design targeted strategies for solving MSA problems according to their needs. When optimizing a multiobjective MSA problem, the algorithm needs to handle two tasks: making the solutions converge to the Pareto frontier (PF) of the problem and ensuring the uniform distribution of solutions. The decomposition strategy is adopted in this study. A decomposition-based algorithm could obtain a set of evenly distributed nondominated solutions. It has shown strong performance in approximating the shape of the PF of the multiobjective problem. Furthermore, to improve the effectiveness and efficiency of the algorithm, the ABC needs to be well designed. For the proposed ABC, the employed bees are utilized for making the solutions converge to the PF and be distributed uniformly; the task of the onlooker bees is accelerating the convergence speed of the algorithm; the scout bees aim to prevent the algorithm from falling into local optimums, which is a common phenomenon during the iterations of ABC.

This study proposes a novel ABC algorithm for multiobjective MSA based on decomposition and dominance (MOABC/D-MSA). MOABC/D-MSA uses the decomposition-based multiobjective optimization strategy to ensure the diversity of solutions. Therefore, it can provide MSA users the information about the shape of the PF, which is essential to decision-making. The main contribution of this study is as follows:

(1) A novel ABC algorithm for MSA is proposed. The proposed algorithm considers both the solution quality and diversity. The employed bee stage achieves even distribution of solutions while optimizing the MSA problem. The onlooker bee stage can obtain high-quality solutions based on superior solutions found by employed bees. The scout bee could facilitate the algorithm in avoiding local optimums.

(2) A decomposition-based employed bee searching strategy is employed for optimizing the MSA problem. The proposed algorithm decomposes the multiobjective MSA problem into several scalarized subproblems. The solutions of subproblems are used to construct a nondominated solution set. This strategy allows the algorithm to obtain a group of solutions distributed in the objective space uniformly.

(3) A nondominance sorting-based onlooker bee searching strategy is proposed. This strategy allows the proposed algorithm to improve the quality of alignments by utilizing high-quality solutions that have been searched.

(4) An experimental study on an MSA benchmark is implemented. The experiment compares the proposed algorithm with state-of-the-art evolutionary and metaheuristic algorithms.

The remaining parts of this study are organized as follows: the second section introduces the definition of the multiobjective MSA problem in the proposed study. Section 3 describes the design and implementation of the proposed MOABC/D-MSA. The fourth section compares the MOABC/D-MSA with state-of-the-art metaheuristics on BAliBASE 3.0, a benchmark MSA test suite. The last section summarizes the proposed work and predicts the research direction of metaheuristic algorithms for MSA.

2. Problem Definition

In this study, the multiobjective MSA problem is defined as a three-objective optimization problem.

There are three objectives in the problem: single structure induced evaluation (STRIKE), percentage of totally conserved columns (%TC), and percentage of non-gaps (% nonGap). STRIKE aims to maximize the accuracy of the alignment. Maximizing %TC ensures there are more columns that the residues are exactly the same, that is, more conserved or special regions within an alignment. Maximizing the %nonGap encourages the aligner to reduce the number of gaps in the aligned sequences. The MSA problem is represented by the mathematical form as shown in the following equation:

$$\text{maximize } F(S) = (\text{STRIKE}(S), \%TC(S), \%\text{nonGap}(S)). \tag{1}$$

Strike evaluates the accuracy of an alignment based on structural information of, at least, one sequence of the alignment. This structural information is commonly retrieved from the Protein Data Bank [16].

Using the structural information as a source for amino acid frequencies and contacts, a log-odds contact matrix is estimated by measuring the ratio between the frequency of each possible contact and its expectation, given the background frequency of each single amino acid. Given any pair of amino acids i and j, the score for their contacts can be estimated as follows:

$$M_{ij} = 10 \times \ln\left(\frac{f_{ij}}{f_i f_j}\right), \tag{2}$$

where f_{ij} is the frequency of contacts involving i and j across all observed *residue-residue* contacts and f_i and f_j are the single residue frequencies in the dataset considered.

%TC takes into account the number of columns that are fully aligned with exactly the same compound. TC is defined as shown in the following equation:

$$\%TC(S) = 100 \sum_{l=1}^{L} \frac{\text{totalColumn}(S_l)}{L}, \tag{3}$$

where S_l is the lth column of S, $S_l = s_{il} \forall i = 1, \ldots, k$, and $\text{totalColumn}(S_l)$ is defined as follows:

$$\text{totalColumn}(S_l) = \begin{cases} 1, & \text{if } s_{il} = s_{1l} \forall i = 2, \ldots, k \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

%nonGap measures the number of residues with respect to the number of gaps into the alignment. This objective function is shown in the following equation:

$$\%\text{nonGap}(S) = 100 \sum_{i=1}^{k} \sum_{j=1}^{L} \frac{\text{is NotGap}(s_{ij})}{k \times L}, \tag{5}$$

where s_{ij} represents the symbol in the jth position of the ith sequence in the alignment S. The function *isNotGap* for a specific residue is defined in the following equation:

$$\text{isNotGap}(\text{residue}) = \begin{cases} 1, & \text{if residue} = '' - '' \\ 0, & \text{otherwise} \end{cases}, \tag{6}$$

where the symbol "—" is a placeholder that is used to align subsequences.

3. MOABC/D-MSA

3.1. Representation of Individuals. Since the solution for an MSA problem is a sequence of characters and gaps, this study adopted an encoding strategy that records the positions of gaps for each sequence. The adopted representation uses the format (begin, end) to store the position of a gap or several consecutive gaps.

Figure 1 illustrates the solution representation in this study. The solution s has four regions that exist gap or gaps: the fifth character, the eleventh to the fourteenth characters, the eighteenth character, and the twentieth to twenty-first characters. This solution is represented by the s' in Figure 1.

3.2. Crossover and Mutation Operators. The search behaviors of the bees are based on crossover and mutation operators. Similar to most of the metaheuristics for MSA, this study employs a single-point crossover operator. For two solutions p_1 and p_2 that $\text{len}(p_1) = \text{len}(p_2) = l_p$, the crossover operator generates a random integer d between zero and l_p. Then each sequence of p_1 is divided into two parts after the dth position, assuming that they are k sequences, p_1 is divided to $\{p_{1,1}^1, p_{1,1}^2, \ldots, p_{1,k}^1, p_{1,k}^2\}$. For p_2, each sequence is cut after the last character of the corresponded sequence of p_1 that is not gap.

After the division, a child solution c_1 is constructed by connecting the first part of p_1 and the last part of p_2. The other child solution c_2 is constructed by connecting the last part of p_2 and the last part of p_1. Since all sequences should have the same length, the crossover operator should adjust the new sequences by inserting gaps into them. The length of all sequences of the child solutions is the length of the longest sequence. This length is calculated according to the following equation:

$$l_{\text{new}} = \max\left(\max\left(p_{1,i}^1 + p_{2,i}^2\right)\right), \quad i = 1, \ldots, k. \tag{7}$$

For a sequence whose length is shorter than l_{new}, the crossover operator inserts gaps between the two segments from p_1 and p_2 to make the sequence's length reach l_{new}.

There are three kinds of mutation operators adopted in the proposed algorithm:

s: REDH-PDLIQ----NAK-K--

s' : (5,5),(11,14),(18,18),(20,21)

Figure 1: Illustration of individual representation.

(i) *Shift-closed gaps.* Closed gaps (i.e., a series of sequential gaps) in a sequence are randomly chosen and shifted to another position randomly. At last, gaps columns (i.e., columns that only contain gaps) are removed.

(ii) *Nongap group splitting.* A non-gap group (i.e., a series of characters between two gaps) is selected randomly. Then, this operator split the nongap group into two groups by inserting a gap at a random position. Finally, gaps columns are deleted.

(iii) *Adjacent gap groups merging.* For this operator, two adjacent gaps groups are selected and merged by shifting the nongap group between the two gap groups to the right.

3.3. Algorithm Overview. The proposed algorithm searches for the optimal alignment by simulating the behavior of bees collecting nectar. There are two kinds of bees in the proposed algorithm that performs different kinds of searching behavior: employed bees and onlooker bees. Algorithm 1 shows the framework of the proposed MOABC-MSA. First, the algorithm initializes N random food sources that represent N candidate alignments. Meanwhile, an archive NA for storing nondominated solutions is initialized as NULL. Then the algorithm executes searching procedures of employed bees, onlooker bees, and scout bees. The MOABC-MSA repeats the searching behaviors of two kinds of bees until it meets the stop criterion. Finally, the algorithm outputs nondominated solutions and their corresponding scores. Figure 2 shows the alignment process of the proposed algorithm.

3.4. Food Source Initialization. The initial food sources are created by MUSCLE [17], a nonmetaheuristic method. MUSCLE creates a group of precomputed alignments. This approach can reduce the execution time of the proposed algorithm.

3.5. Decomposition-Based Employed Bee Phase. The employed bees in this study perform a decomposition-based multiobjective optimization, which can obtain solutions with both high quality and diversity. The diversity is essential for the algorithm since it has an important influence on the subsequent optimization. Algorithm 2 shows the execution process of the employed bee stage.

Each employed bee is assigned a food source. The single-point crossover operation is performed. The proposed algorithm calculates the Euclidean distance between the food sources. For each food source, its nearest T food sources are defined as its neighbors. When executing the crossover operator, for the ith bee, the first individual of the crossover operation is the ith food source and the second individual is randomly selected from neighbors of the ith food source.

```
Input: sequences to be aligned;
Output: aligned sequences;
  initialize N random food sources;
  initialize non-dominated set NA;
  evaluate food sources;
  while termination criterion is false do
    employed bees execute search behaviour;
    onlooker bees execute search behaviour;
    for each sub-problem do
      if is not updated for k iterations then
        scout bee execute search behaviour;
      end if
    end for
    update non-dominated set NA;
  end while
  output non-dominated solutions in NA;
```

Algorithm 1: Algorithm framework of MOABC/D-MSA.

After the crossover, two new candidate alignments are generated. Then the proposed algorithm performs shift-closed gaps mutation operator on the two candidate alignments and generates two offspring alignments.

Next, all offspring alignments are evaluated by the objective functions. Then the algorithm should select optimal solutions from original alignments and offspring to update the food sources.

The selection stage adopts the decomposition-based idea. The proposed algorithm generates N weight vectors ω_1 to $\omega_{\mathbf{N}}$ that are uniformly distributed in the objective space. The weight vectors are utilized in building scalarized subproblems. In this study, the penalty-based boundary intersection (PBI) approach [18] is adopted in designing subproblems. The number of subproblems equals the number of weight vectors. For the ith weight vector, its corresponding subproblem is constructed as follows:

$$\text{minimize } g^{pbi}(x \mid \omega_i, z^*) = d_1 + \theta d_2, \text{ subject to } x \in \Omega, \quad (8)$$

where

$$d_1 = \frac{\left\| (\mathbf{z}^* - F(\mathbf{x}))^T \boldsymbol{\omega}_i \right\|}{\|\boldsymbol{\omega}\|} \text{ and } d_2 = \left\| F(\mathbf{x}) - (\mathbf{z}^* - d_1 \boldsymbol{\omega}_i) \right\|, \quad (9)$$

where $\mathbf{z}^*$ is the ideal point that satisfies $\mathbf{z}_i^* = \min_{i=1}^{N} f_i(\cdot)$, and each dimension of $\mathbf{z}^*$ is the minimum value among all candidate solutions.

Since each subproblem is a scalarized function, the algorithm can obtain the unique optimal solution of each subproblem in each iteration. The optimal solutions for the sub-problems are the new food sources.

3.6. Nondominated Sorting-Based Onlooker Phase. The searching behavior of onlooker bees is guided by the searching results of employed bees.

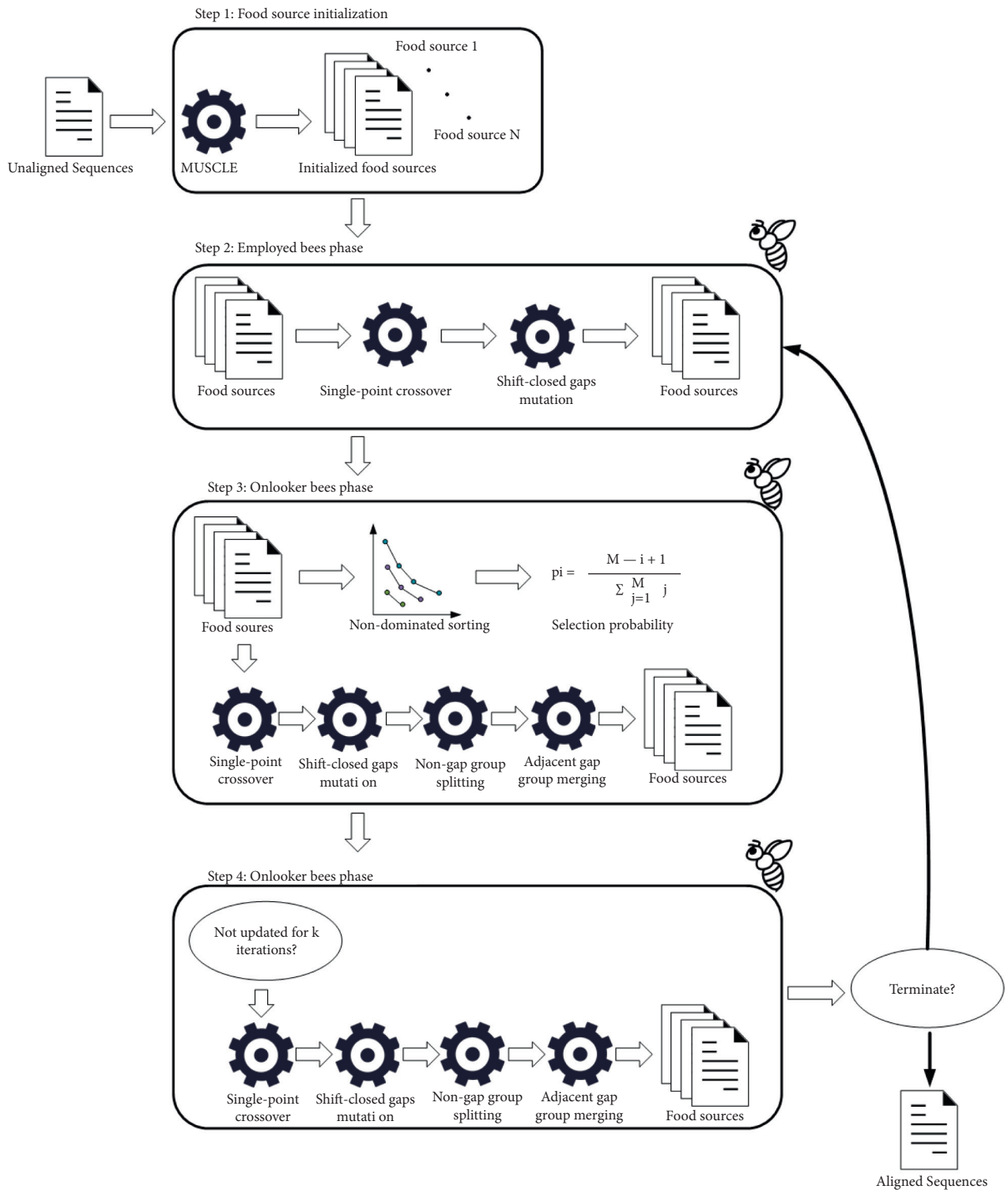


Figure 2: Alignment process of the proposed algorithm.

To guide the food sources converge to the PF, the onlooker bees prefer to exploit high-quality solutions founded by employed bees. Meanwhile, the proposed algorithm should maintain the diversity of food sources. The diversity of food sources is important for fitting the PF. Furthermore, maintaining the diversity of food sources could help the proposed algorithm to avoid being trapped in local optimal.

At this stage, each onlooker selects a food source to search. Since the employed bees have obtained the quality information of food sources, each onlooker bee selects a food

```
Input: N food sources;
Output: N new food sources;
  initialize ω_1 to ω_N;
  calculate neighborhood of each food source;
  while termination criterion is false do
    for each employed bee do
      perform single-point crossover operation;
      perform shift-closed gaps mutation operation;
    end for
    evaluate offspring solutions;
    merge food sources and offspring solutions;
    find optimal solution of each sub-problem;
    update food sources;
  end while
```

Algorithm 2: Searching behavior of employed bees.

source to search based on the quality information. MOABC/D-MSA uses a nondominated ranking-based roulette wheel selection. This method ranks the food sources according to their Pareto dominance relationship. Then, it uses the roulette wheel selection to guarantee that the nondominated food sources are more likely to be selected.

Inspired by the fast nondominated sorting method [19], the proposed algorithm divides the food sources into several ranks. The first rank is nondominated solutions among all food sources. The second rank is nondominated solutions among food sources except the first rank. The rest of the food sources are sorted in the same manner.

When selecting food sources, an onlooker bee first selects a rank. Assuming that there are M ranks among the food sources, the selected probability of the ith rank is calculated according to the following equation:

$$p_i = \frac{M - i + 1}{\sum_{j=1}^{M} j}. \tag{10}$$

Each onlooker bee generates a random number between zero and one. Then, it selects a specific rank according to the roulette wheel method. Finally, the onlooker bee selects a food source randomly from this rank.

After selecting a food source, each onlooker bee performs a single-point crossover operation. Then it performs the shift-closed gaps mutation, nongap group splitting mutation, and adjacent gap groups merging mutation in sequence.

Finally, similar to the employed bee phase, the new solutions are merged with the food sources, and the proposed algorithm updates the food sources by the nondominated and crowded-based sorting selection [20]. The onlooker bee stage is described in Algorithm 3.

3.7. Scout Bee Phase. The scout bees aim to facilitate the algorithm avoiding local optimum. They watch the food sources in every iteration, but they work when the optimal solution of a subproblem is not updated for more than k iterations. The scout bees perform the same crossover operator and mutation operator with the onlooker bees. For the

```
Input: N food sources;
Output: N new food sources;
  perform fast non-dominated sorting;
  calculate selected probability for each food source;
  select a food source for each onlooker by roulette wheel
selection;
  while termination criterion is false do
    for each employed bee do
      perform single-point crossover operation;
      perform shift-closed gaps mutation operation;
      perform non-gap group splitting;
      perform adjacent gap groups merging;
    end for
    evaluate new solutions;
    merge food sources and new solutions;
    perform non-dominated and crowded-based sorting
selection;
    update food sources;
  end while
```

Algorithm 3: Searching behavior of onlooker bees.

crossover operation, the scout bees select a solution from the neighbor of the food source as the mating partner. Finally, the scout bee updates the food sources of the nonupdated subproblems. Algorithm 4 shows the searching behavior of the scout bee.

3.8. Complexity Analysis. According to Algorithms 2–4, the time complexity of the employed bee phase, onlooker bee phase, and scout phase are all $O(n^2)$. Therefore, the time complexity of MOABC/D-MSA is $O(n^2)$. Furthermore, the scout bee phase is not always performed during the optimization process. Meanwhile, MOABC/D-MSA does not require an external archive, and it only maintains a food source set. Therefore, its space complexity is $O(n)$.

4. Experimental Study

This section proposed a comparison study. The performance of the proposed MOABC/D-MSA is compared with metaheuristics on benchmark dataset.

4.1. Benchmark Dataset. Benchmark alignment database (BAliBASE) [21] is a benchmark dataset for evaluating the performance of algorithms for MSA problems. BAliBASE is developed by manually aligning based on 3-D structures of proteins. This study uses the 3.0 version of BAliBASE to test the proposed algorithm. BAliBASE includes thousands of challenging sequences. It is widely accepted in the research of MSA problem. Therefore, using BAliBASE as the benchmark can help researchers to investigate the performance of the proposed algorithm. There are 218 sets of sequences in BAliBASE 3.0, and they are divided into six families as follows: RV11, RV12, RV20, RV30, RV40, and RV50.

This experiment selects twenty-seven test cases from BAliBASE 3.0 to implement the comparative study between

```
Input: sequences to be aligned;
Output: aligned sequences;
  initialize N random food sources;
  initialize non-dominated set NA;
  evaluate food sources;
  while termination criterion is false do
    for each sub-problem do
      if the optimal solution is not updated for more than k iterations then
        perform single-point crossover operation;
        perform shift-closed gaps mutation operation;
        perform non-gap group splitting;
        perform adjacent gap groups merging;
      end if
    end for
    evaluate new solutions;
    merge food sources and new solutions;
    find optimal solution of each sub-problem;
    update food sources;
  end while
```

Algorithm 4: Searching behavior of scout bee.

the proposed MOABC/D-MSA and other genetic and metaheuristic algorithms. The selected instances are as follows: BB11001, BB11005, BB11018, BB11020 in RV11, BB12001, BB12013, BB12022, BB12035, BB12044 in RV12, BB20001, BB20010, BB20022, BB20033, BB20041 in RV20, BB30001, BB30008, BB30015, BB30022 in RV30, BB40001, BB40013, BB40025, BB40038, BB40048 in RV40, BB50001, BB50005, BB50010, and BB50016 in RV50.

4.2. Peer Competitors. This study compares the proposed MOABC-MSA with several state-of-the-art evolutionary or metaheuristic algorithms. The competitor algorithms includes NSGA-II [22], MOEA/D [2], GAPAM [23], MO-SAStrE [24], and HMOABC [25]. The parameters of the peer algorithms are set according to their original literature.

4.3. Static Results. This section compares the statistical results of the algorithms on STRIKE, %TC, and %nonGap. Meanwhile, the executing times of the algorithms are recorded and compared. Each algorithm runs each test cases for 30 times to avoid randomness. The termination criterion is set as 25,000 times of evaluation to guarantee the fairness of the experiment. For MOABC/D-MSA and HMOABC, both the number of employed bees and the number of onlooker bees are set to 20. For the other algorithms, the size of their populations is set to 20.

Tables 1–6 list the best value of STRIKE, %TC, and % nonGap among solutions found by the tested algorithms for each test case. Tables 1–3 list that for instances in RV11, RV12, and RV20, MOABC/D-MSA obtains the optimal STRIKE, %TC, and %nonGap on all test cases. For BB30008, both MOABC/D-MSA and MOEA/D obtain the optimal % TC. For BB30015 and BB50010, MOEA/D obtains the optimal %TC. For BB40048, results of NSGA-II obtain the best %nonGap. For the other test cases in RV30, RV40, and RV50, MOABC/D-MSA outperforms the compared algorithms in all three objectives.

Figure 3 shows the box plots of running time of the tested algorithms. Subfigures (a)–(f) exhibit the running times of the algorithms on BB11001, BB12001, BB20001, BB30001, BB40001, and BB50001, respectively. The box plot is an effective tool in showing the distribution of data. In this experiment, each box in a box plot represents the distribution of running times for the 30 runs of independently repeated experiments of a specific algorithm. There are five horizontal lines in a box, from top to bottom, the lines represent the maximum value, the three-quarter median, the median, the quarter median, and the minimum value. The circles around the box represent an unusual value. The box plots show that the running time of MOABC/D-MSA is stable, in each subfigure, the box of the MOABC/D-MSA is thin. Furthermore, the location of the MOABC/D-MSA's boxes is low, which means that the proposed algorithm consumes less time. Although MOABC/D-MSA takes a little longer time than NSGA-II and MOEA/D, the running time of the proposed algorithm is significantly better than HMOABC, GAPAM, and MO-SAStrE. The boxplots show that the MOABC/D-MSA is an effective algorithm. The efficiency is an essential performance indicator in evaluating the algorithm for MSA. An efficient algorithm can ensure that the algorithm can process a large number of biological information sequences in a short time, which is practical and valuable in practice.

4.4. Hypothesis Results. This study uses the Wilcoxon signed-rank hypothesis test [26] to investigate the difference between the performance of the MOABC/D-MSA and results of the competitors. The objective values of each solution are normalized to a real number between zero and one, and then the normalized solutions are evaluated by the IGD indicator [27]. The IGD indicator works out a scalarized

Table 1: Comparison on RV11.

Test case		MOABC/D-MSA	HMOABC	NSGA-II	MOEA/D	GAPAM	MO-SAStrE
BB11001	STRIKE	**3.25**	3.07	2.94	2.98	2.79	2.88
	%TC	**7.89**	7.48	7.40	7.37	6.84	7.42
	%nonGap	**94.84**	89.46	93.75	92.98	90.57	91.44
BB11005	STRIKE	**3.14**	3.09	2.89	2.94	2.67	2.88
	%TC	**8.04**	6.90	6.59	6.54	6.38	6.52
	%nonGap	**93.68**	87.56	92.64	90.53	88.24	83.20
BB11018	STRIKE	**3.58**	3.07	2.75	2.84	2.39	2.58
	%TC	**8.44**	7.35	7.03	6.89	5.35	5.24
	%nonGap	**94.37**	87.36	92.58	92.46	90.45	91.85
BB11020	STRIKE	**3.38**	3.25	2.93	2.86	2.37	2.25
	%TC	**7.33**	7.28	7.29	7.25	7.32	7.30
	%nonGap	**93.53**	90.45	92.37	92.59	91.43	92.50

Bold numbers indicate the optimum values.

Table 2: Comparison on RV12.

Test case		MOABC/D-MSA	HMOABC	NSGA-II	MOEA/D	GAPAM	MO-SAStrE
BB12001	STRIKE	**2.74**	2.52	2.60	2.58	2.47	2.35
	%TC	**3.79**	3.62	3.74	3.75	3.68	3.71
	%nonGap	**84.40**	80.43	82.13	81.37	78.48	79.63
BB12013	STRIKE	**2.97**	2.73	2.63	2.68	2.74	2.70
	%TC	**3.79**	2.59	3.62	3.55	2.98	2.82
	%nonGap	**84.03**	80.38	83.44	82.56	81.47	78.38
BB12022	STRIKE	**2.81**	2.63	2.72	2.69	2.54	2.60
	%TC	**3.66**	3.47	3.58	3.21	3.46	3.29
	%nonGap	**82.94**	80.65	82.56	81.53	82.08	81.32
BB12035	STRIKE	**2.76**	2.51	2.69	2.65	2.55	2.49
	%TC	**3.73**	3.50	3.71	3.68	3.41	3.59
	%nonGap	**82.30**	79.73	82.21	81.99	80.35	81.02
BB12044	STRIKE	**2.71**	2.46	2.58	2.59	2.32	2.44
	%TC	**3.74**	3.56	3.67	3.66	3.69	3.52
	%nonGap	**82.39**	79.93	81.95	80.61	78.16	77.08

Bold numbers indicate the optimum values.

Table 3: Comparison on RV20.

Test case		MOABC/D-MSA	HMOABC	NSGA-II	MOEA/D	GAPAM	MO-SAStrE
BB20001	STRIKE	**0.69**	0.67	0.54	0.58	0.42	0.51
	%TC	**0.21**	0.09	0.13	0.10	0.14	0.08
	%nonGap	**41.22**	38.85	40.06	39.57	38.30	36.94
BB20010	STRIKE	**0.48**	0.42	0.47	0.41	0.39	0.44
	%TC	**0.27**	0.22	0.25	0.23	0.18	0.20
	%nonGap	**40.18**	36.94	39.82	37.81	38.89	39.06
BB20022	STRIKE	**0.28**	0.24	0.26	0.25	0.25	0.24
	%TC	**0.21**	0.18	0.19	0.17	0.17	0.15
	%nonGap	**39.42**	38.22	38.57	38.26	37.26	37.01
BB20033	STRIKE	**0.56**	0.50	0.55	0.52	0.49	0.54
	%TC	**0.26**	0.19	0.24	0.22	0.17	0.21
	%nonGap	**41.27**	38.48	40.53	36.52	34.83	35.66
BB20041	STRIKE	**0.37**	0.32	0.34	0.35	0.29	0.31
	%TC	**0.20**	0.14	0.18	0.16	0.11	0.13
	%nonGap	**40.28**	36.53	39.39	34.55	32.17	33.68

Bold numbers indicate the optimum values.

Table 4: Comparison on RV30.

Test case		MOABC/D-MSA	HMOABC	NSGA-II	MOEA/D	GAPAM	MO-SAStrE
BB30001	STRIKE	**1.75**	1.63	1.65	1.67	1.54	1.56
	%TC	**0.33**	0.29	0.32	0.30	0.25	0.22
	%nonGap	**50.77**	43.70	49.42	50.04	44.66	42.97
BB30008	STRIKE	**1.85**	1.68	1.74	1.77	1.63	1.66
	%TC	**0.33**	0.27	0.31	**0.33**	0.25	0.22
	%nonGap	**51.40**	42.34	50.06	50.56	49.28	44.30
BB30015	STRIKE	**2.44**	2.18	2.35	2.39	2.22	2.15
	%TC	0.35	0.34	0.35	**0.36**	0.29	0.28
	%nonGap	**49.07**	43.92	48.91	47.40	43.21	40.06
BB30022	STRIKE	**2.06**	1.74	1.93	1.95	1.86	1.88
	%TC	**0.41**	0.34	0.39	0.40	0.36	0.36
	%nonGap	**48.52**	46.03	47.60	47.88	42.52	43.94

Bold numbers indicate the optimum values.

Table 5: Comparison on RV40.

Test case		MOABC/D-MSA	HMOABC	NSGA-II	MOEA/D	GAPAM	MO-SAStrE
BB40001	STRIKE	**3.50**	2.99	3.45	3.47	3.09	3.14
	%TC	**0.42**	0.23	0.36	0.33	0.27	0.29
	%nonGap	**31.13**	27.62	30.75	30.90	28.84	27.56
BB40013	STRIKE	**3.79**	3.22	3.67	3.54	3.48	3.26
	%TC	**0.33**	0.21	0.29	0.27	0.20	0.26
	%nonGap	**31.01**	27.45	29.84	29.98	27.59	25.86
BB40025	STRIKE	**3.59**	3.44	3.53	3.58	3.32	3.17
	%TC	**0.38**	0.26	0.34	0.25	0.27	0.30
	%nonGap	**29.03**	25.34	28.77	27.40	24.59	22.05
BB40038	STRIKE	**3.33**	2.98	3.21	3.25	3.08	3.04
	%TC	**0.36**	0.28	0.35	0.35	0.33	0.29
	%nonGap	**29.88**	26.36	29.24	29.34	28.37	27.75
BB40048	STRIKE	**3.47**	3.05	3.36	3.42	3.06	3.11
	%TC	**0.27**	0.22	0.24	0.25	0.18	0.20
	%nonGap	29.97	26.58	**30.13**	28.44	27.93	26.55

Bold numbers indicate the optimum values.

Table 6: Comparison on RV50.

Test case		MOABC/D-MSA	HMOABC	NSGA-II	MOEA/D	GAPAM	MO-SAStrE
BB50001	STRIKE	**2.11**	1.88	1.97	2.04	1.83	1.65
	%TC	**0.39**	0.27	0.35	0.33	0.28	0.31
	%nonGap	**70.05**	55.93	69.90	67.34	62.98	63.51
BB50005	STRIKE	**2.03**	1.48	1.95	1.88	1.57	1.62
	%TC	**0.35**	0.30	0.34	0.33	0.24	0.25
	%nonGap	**69.01**	62.48	68.45	63.44	61.50	60.24
BB50010	STRIKE	**1.89**	1.56	1.84	1.80	1.61	1.69
	%TC	0.36	0.26	0.36	**0.37**	0.25	0.32
	%nonGap	**63.33**	59.02	62.75	60.45	59.93	57.28
BB50016	STRIKE	**1.89**	1.37	1.88	1.67	1.34	1.42
	%TC	**0.32**	0.19	0.30	0.31	0.24	0.25
	%nonGap	**64.01**	58.99	63.90	61.52	59.31	60.04

Bold numbers indicate the optimum values.

score for each nondominated solution set. Finally, the IGD value of each solution set is utilized in the hypothesis test. Table 7 lists the p values between the results of MOABC/D-MSA and results of the peer algorithms. The significance level is set at 0.05 in this study. If the p value is less than 0.05, it indicates that the result of MOABC/D-MSA is significantly

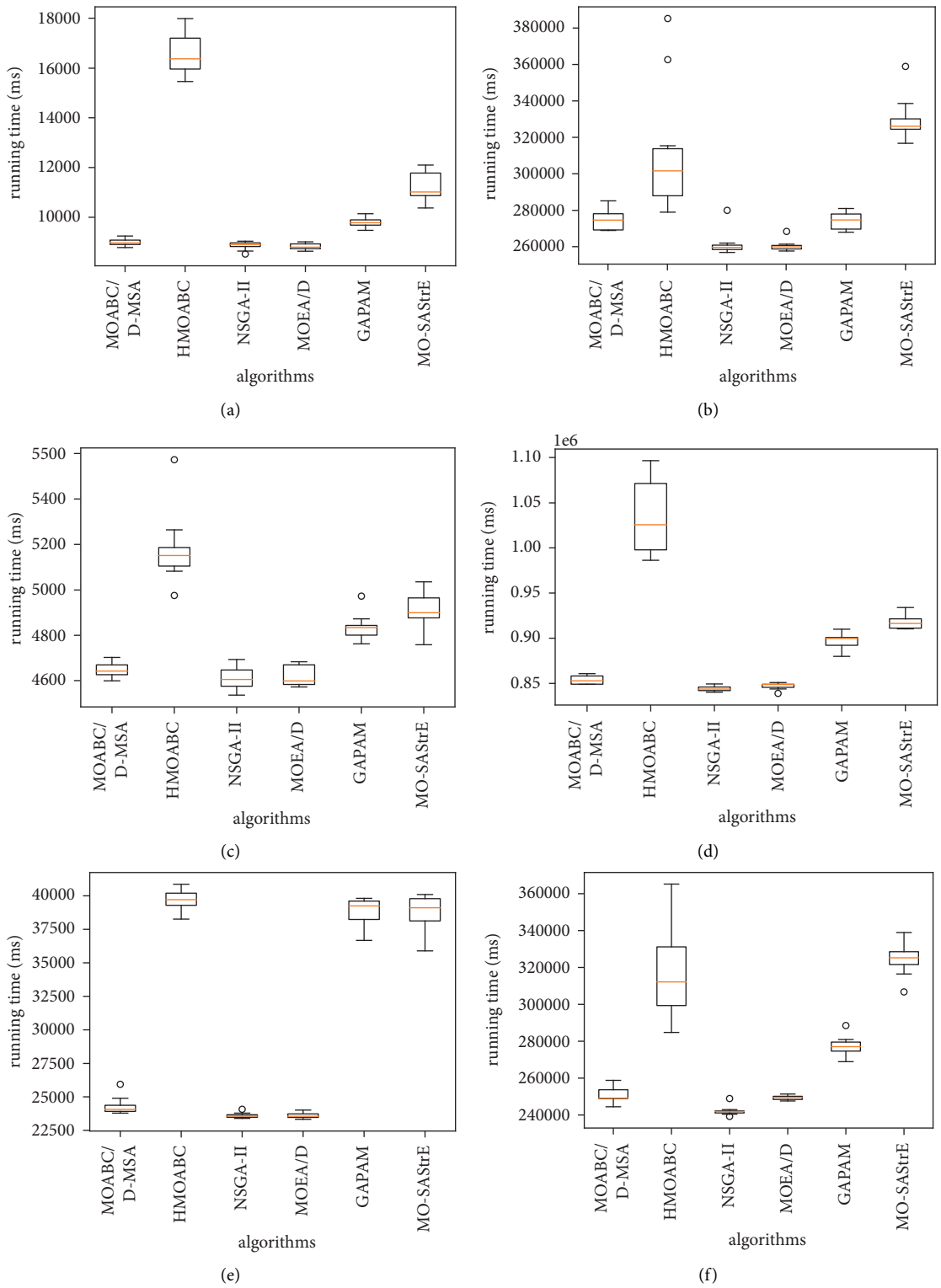


Figure 3: Box plots of running time. (a) BB11001. (b) BB12001. (c) BB20001. (d) BB30001. (e) BB40001. (f) BB50001.

better than the result of the competitor. According to Table 7, except for BB40038, BB50010, and BB50016, MOABC/D-MSA outperforms all its competitors in the test cases. Experimental results show that the results of the proposed algorithm are significantly better on most of the test cases. On BB40038, although the results of MOABC/D-MSA cannot significantly outperform NSGA-II and MOEA/D, according to Table 5, its results outperform the two algorithms in the aspect of %TC, %nonGap, and STRIKE. On BB50010, there is no significant difference between the

Table 7: Comparison on BAliBASE test cases.

Test case	HMOABC	NSGA-II	MOEA/D	GAPAM	MO-SAStrE
BB11001	0.005	0.005	0.005	0.005	0.005
BB11005	0.005	0.005	0.005	0.005	0.005
BB11018	0.005	0.005	0.005	0.004	0.05
BB11020	0.005	0.005	0.005	0.05	0.05
BB12001	0.05	0.005	0.05	0.05	0.05
BB12013	0.005	0.005	0.005	0.005	0.005
BB12022	0.005	0.003	0.004	0.005	0.005
BB12035	0.005	0.005	0.005	0.005	0.005
BB12044	0.005	0.005	0.005	0.005	0.005
BB20001	0.005	0.01	0.05	0.005	0.005
BB20010	0.005	0.05	0.01	0.01	0.01
BB20022	0.005	0.01	0.01	0.005	0.005
BB20033	0.005	0.05	0.01	0.005	0.005
BB20041	0.005	0.05	0.05	0.005	0.005
BB30001	0.005	0.005	0.01	0.005	0.005
BB30008	0.005	0.005	0.005	0.005	0.005
BB30015	0.005	0.01	0.01	0.005	0.005
BB30022	0.005	0.01	0.01	0.005	0.005
BB40001	0.005	0.05	0.05	0.005	0.005
BB40013	0.005	0.01	0.01	0.01	0.01
BB40025	0.01	0.05	0.0 5	0.005	0.01
BB40038	0.005	0.01	0.01	0.005	0.005
BB40048	0.01	0.01	0.10	0.01	0.01
BB50001	0.005	0.005	0.01	0.005	0.005
BB50005	0.005	0.01	0.05	0.005	0.005
BB50010	0.005	0.10	0.10	0.01	0.01
BB50016	0.005	0.05	0.10	0.005	0.01

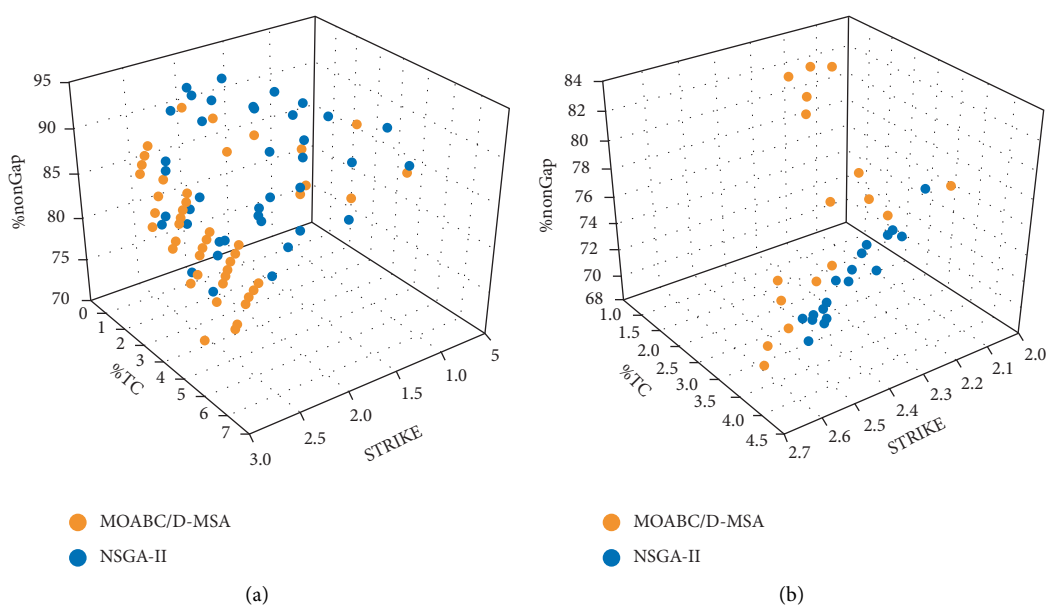


Figure 4: Comparison of the distribution of solutions of MOABC/D-MSA and NSGA-II. (a) BB11001. (b) BB12001.

results of MOABC/D-MSA and the results of NSGA-II, MOEA/D, GAPAM, and MO-SAStrE. On BB50016, although the results of MOABC/D-MSA cannot significantly outperform MOEA/D and MO-SAStrE, according to Table 6, its results outperform the two algorithms in the aspect of %TC, %nonGap, and STRIKE.

4.5. Solution Distribution. The statistical results show that nondominated solutions of MOABC/D-MSA can obtain the optimal results on all three objectives. Since the objectives cannot be optimized simultaneously, the results indicate that the distribution of results of MOABC/D-MSA is more uniform than the results of compared algorithms. Figure 4

illustrates the distribution of nondominated solution sets of MOABC/D-MSA and NSGA-II on BB11001 and BB12001, respectively. The figures show that solutions of MOABC/D-MSA have more comprehensive coverage of the PF. In this case, MOABC/D-MSA can provide solutions that meet different demands of users and researchers.

4.6. Discussion. Experimental results show that MOABC/D-MSA can obtain a better solution on all objectives when optimizing the three-dimensional multiobjective MSA problem. This result means that MOABC/D-MSA not only obtains solutions that are close to the PF of the problems but also obtains uniform solution distribution. In other words, MOABC/D-MSA achieves a good balance between convergence and diversity of solutions.

Meanwhile, MOABC/D-MSA is an efficient algorithm. For the same evaluation times, the running time of MOABC/D-MSA is competitive among the tested algorithms. This performance might be due to the fact that the proposed algorithm does not introduce additional calculations and complex search steps. Efficiency is important for the metaheuristics. In addition, MOABC/D-MSA avoids local optimum and controls the randomness during the research. An efficient algorithm could execute more iterations at the same time; therefore, it is more likely to obtain high-quality solutions.

Since the decomposition-based strategy is adopted in the employed bee phase, MOABC/D-MSA controls the uniform distribution of solutions. This strategy maximizes the difference between understandings while ensuring the quality of the nondominated solutions. In this way, the algorithm can provide more solutions that meet the different demands of algorithm users.

5. Conclusion

This study proposes MOABC/D-MSA, a decomposition-based artificial bee colony optimization algorithm for solving the MSA problem. The searching behavior of the employed bees is based on the PBI method, a decomposition-based strategy. The proposed algorithm considers both the convergence performance and the distribution of the alignments. The employed bees not only make the food sources converge to the PF but also ensure the distribution of the food sources, reflecting the real shape of the PF. The onlooker bees of MOABC/D-MSA accelerate the converging of food sources by utilizing high-quality solutions. MOABC/D-MSA uses the scout bee to get out of the local optimum. This study implements a comparative study on BAliBASE 3.0. The experimental results verify that MOABC/D-MSA has competitive convergence performance. Furthermore, nondominated solutions generated by MOABC/D-MSA show better performance in reflecting the PF the MSA problems. The superior distribution performance provides stronger decision support for biological researchers. For future studies, improving efficiency is a tough task and promising research direction.

Data Availability

There are no applicable datasets.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the Key Project of National Natural Science Foundation of China (U1908212) and the Startup Research Foundation for PhD of Liaoning Province (2020-BS-152).

References

[1] M. A. Nayeem, M. S. Bayzid, A. H. Rahman, R. Shahriyar, and M. S. Rahman, "Multiobjective formulation of multiple sequence alignment for phylogeny inference," *IEEE Transactions on Cybernetics*, 2020.

[2] H. Zhu, Z. He, and Y. Jia, "A novel approach to multiple sequence alignment using multiobjective evolutionary algorithm based on decomposition," *IEEE journal of biomedical and health informatics*, vol. 20, no. 2, pp. 717–727, 2015.

[3] R. K. Ramakrishnan, J. Singh, and M. R. Blanchette, "A reinforcement learning approach for multiple sequence alignment," in *Proceedings of the 2018 IEEE Eighteenth International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 61–66, IEEE, Taichung, Taiwan, October 2018.

[4] Á. Rubio-Largo, M. A. Vega-Rodríguez, and D. L. González-Álvarez, "A hybrid multiobjective memetic metaheuristic for multiple sequence alignment," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 499–514, 2015.

[5] Á. Rubio-Largo, L. Vanneschi, M. Castelli, and M. A. Vega-Rodríguez, "A characteristic-based framework for multiple sequence aligners," *IEEE Transactions on Cybernetics*, vol. 48, no. 1, pp. 41–51, 2016.

[6] K. Giannakis, C. Papalitsas, G. Theocharopoulou, S. Fanarioti, and T. Andronikos, "A quantum-inspired optimization heuristic for the multiple sequence alignment problem in biocomputing," in *Proceedings of the 2019 Tenth International Conference on Information, Intelligence, Systems and Applications (IISA)*, pp. 1–8, IEEE, Patras, Greece, July 2019.

[7] Q. Zhan, N. Wang, S. Jin, R. Tan, Q. Jiang, and Y. Wang, "Probpfp: a multiple sequence alignment algorithm combining partition function and hidden Markov model with particle swarm optimization," in *Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1290–1295, IEEE, Madrid, Spain, December 2018.

[8] P. T. Doan and A. Takasu, "Sparse regression-based multiple sequence alignment," in *Proceedings of the 2019 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1372–1377, IEEE, Shanghai, China, July 2019.

[9] N. Altwaijry, M. Almasoud, A. Almalki, and I. Al-Turaiki, "Multiple sequence alignment using a multiobjective artificial bee colony algorithm," in *Proceedings of the 2020 Third International Conference on Computer Applications & Information Security (ICCAIS)*, pp. 1–6, IEEE, Riyadh, Saudi Arabia, March 2020.

[10] L. Ma, S. Cheng, and Y. Shi, "Enhancing learning efficiency of brain storm optimization via orthogonal learning design,"

IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 11, pp. 6723–6742, 2020.

[11] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Transactions on Cybernetics*, 2021.

[12] L. Ma, X. Wang, X. Wang, L. Wang, Y. Shi, and M. T. Huang, "Truthful combinatorial double auctions for mobile edge computing in industrial internet of things," *IEEE Transactions on Mobile Computing*, 2021.

[13] H. Zhao, C. Zhang, X. Zheng, C. Zhang, and B. Zhang, "A decomposition-based many-objective ant colony optimization algorithm with adaptive solution construction and selection approaches," *Swarm and Evolutionary Computation*, vol. 68, Article ID 100977, 2022.

[14] H. Zhao and C. Zhang, "A decomposition-based many-objective artificial bee colony algorithm with reinforcement learning," *Applied Soft Computing*, vol. 86, Article ID 105879, 2020.

[15] X. Lei, J. Sun, X. Xu, and L. Guo, "Artificial bee colony algorithm for solving multiple sequence alignment," in *Proceedings of the 2010 IEEE fifth international conference on bio-inspired computing: theories and applications (BIC-TA)*, pp. 337–342, IEEE, Changsha, China, September 2010.

[16] S. K. Burley, H. M. Berman, G. J. Kleywegt, J. L. Markley, H. Nakamura, and S. Velankar, "Protein data bank (pdb): the single global macromolecular structure archive," *Methods in Molecular Biology*, pp. 627–641, 2017.

[17] R. C. Edgar, "Muscle: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.

[18] Q. Zhang and H. Li, "Moea/d: a multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.

[19] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[20] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based non-dominated sorting approach, part i: solving problems with box constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2013.

[21] J. D. Thompson, P. Koehl, R. Ripp, and O. Poch, "Balibase 3.0: latest developments of the multiple sequence alignment benchmark," *Proteins*, vol. 61, no. 1, pp. 127–136, 2005.

[22] O. Kaiwartya, S. Prakash, and A. N. Hassan, "Multiple sequence alignment using genetic algorithm and non-dominant sorting genetic algorithm-ii (nsga ii) and variants," *Journal of Bioinformatics and Intelligent Control*, vol. 3, no. 4, pp. 294–299, 2014.

[23] F. Naznin, R. Sarker, and D. Essam, "Progressive alignment method using genetic algorithm for multiple sequence alignment," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 5, pp. 615–631, 2012.

[24] F. M. Ortuño, O. Valenzuela, F. Rojas et al., "Optimizing multiple sequence alignments using a genetic algorithm based on three objectives: structural information, non-gaps percentage and totally conserved columns," *Bioinformatics*, vol. 29, no. 17, pp. 2112–2121, 2013.

[25] Á. Rubio-Largo, M. A. Vega-Rodríguez, and D. L. González-Álvarez, "Hybrid multiobjective artificial bee colony for multiple sequence alignment," *Applied Soft Computing*, vol. 41, pp. 157–168, 2016.

[26] S. M. Taheri and G. Hesamian, "A generalization of the wilcoxon signed-rank test and its applications," *Statistical Papers*, vol. 54, no. 2, pp. 457–470, 2013.

[27] Y. Sun, G. G. Yen, and Z. Yi, "Igd indicator-based evolutionary algorithm for many-objective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 2, pp. 173–187, 2018.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8455164, 8 pages
https://doi.org/10.1155/2022/8455164

Research Article

A Novel Deep Learning-Enabled Physical Education Mechanism

Weiqi Wang[1] and Jianan Jiang[2]

[1]*Fuyang Normal University, Fuyang 236041, China*
[2]*University of Science and Technology LiaoNing, Anshan 114051, China*

Correspondence should be addressed to Jianan Jiang; lnkdjjn@ustl.edu.cn

Received 23 December 2021; Revised 14 January 2022; Accepted 17 January 2022; Published 8 March 2022

Academic Editor: Jianhui Lv

Race walking is one of the key events in the Tokyo Olympic Games, and also one of the strengths of China in athletics events. In recent years, China has made remarkable achievements in various race-walking competitions. However, with the improvement of the performance of race walkers, more and more technical problems have emerged, and the number of fouls due to nonstandard movements has increased significantly. It is a pity that athletes are disqualified for technical fouls in long-distance race-walking competitions. Therefore, it is necessary to introduce scientific training methods to help coaches strictly monitor the training process of athletes and accurately detect their standard degree of action in real-time. This paper mainly proposes a novel mechanism for foul recognition in race walking based on deep learning. Firstly, the image frames in the video are preprocessed by the Yolo algorithm to obtain the athletes' separated images. The U-Net network mixed with the attention mechanism is used to detect the athletes' actions to identify fouls and nonstandard actions, so as to assist the coach to identify the athletes' nonstandard actions in training and adjust them in time. Experiments show that the above method can identify the foul actions and nonstandard actions of multiple athletes in training at the same time quickly, and the recognition accuracy is higher than human eyes. It is more conducive to assist the coach to monitor and standardize the athletes' actions in the long-term training process, so as to reduce the error rate and improve the performance.

1. Introduction

Race walking originated in Britain in the 19th century and was developed on the basis of daily walking. The rules stipulate that the supporting legs must be straight, the two legs move forward alternately, keep uninterrupted contact with the ground, and do not leave the ground at the same time at any time, so as to ensure that there is no "flying" phenomenon, which is also the main difference between race walking and running. It is judged by the time it takes to complete the race. The normal walking speed is about 5 kilometers per hour, and the race walking is much faster. Men's walking race became an official event of the Olympic Games in 1908 and the women's walking race began in the Czech Republic in 1932. Currently, the longest distance in the Olympic event is 50 kilometers for men and 20 kilometers for women. The walking race is also the traditional advantage of the Chinese team in track and field events. From the 1984 Los Angeles Olympic Games to the 2016 Rio Olympic Games, China has won a total of 8 gold medals in track and field events of the Olympic Games, including 5 from walking race. By the end of the 2019 Doha track and field world championships, Chinese women's walking athletes have achieved four championships in five years in the 20 km walking event and made remarkable achievements.

Since race walking competition is usually time-consuming, there are very strict requirements for athletes' physical quality, psychological quality, tactics, and technical actions, especially for long-distance race walking, such as 20 km and 50 km, it is difficult for ordinary people to complete the whole process normally, athletes have to complete the game as soon as possible on the premise of maintaining the characteristic posture, the consumption of athletes in the competition is very large, which is not only testing the endurance of athletes but also testing the technical level of athletes. However, it is inseparable from the

usual hard training to keep the standard actions in such a long competition. With the improvement of the performance of race walkers, the phenomenon of fouls in the competition is also gradually increasing. The influence of fouls on athletes in the competition is very huge. When the race walkers' actions show signs of violating the race-walking technology, they will be given a yellow card warning and the walking referee will give a red card when he watches visible flying or knee bending. When the same athlete receives three red cards, he will be disqualified. At present, in the competition, the referees cannot judge by any equipment. They can only rely on their own eyes to judge whether the athlete violates the rules. According to the literature [1], we know that there will be a certain misjudgment rate if he totally depends on his eyes. The referee may be affected by the factors such as the judgment distance and angle in the process of judgment. Therefore, for the athlete, it is very important to strictly regulate their own actions in daily training, resume the training process timely, adjust non-standard actions, and avoid foul or be a misjudgment.

In the traditional training process, coaches use naked eyes to observe or through video resumes to help athletes correct their actions. However, due to the limitations of human eyes, it is not efficient enough to judge by naked eyes completely and there are many misjudgments and omissions. With the increasing development of artificial intelligence technology, video image capture and wearable device technology are constantly promoting the development of sports behavior recognition. There are many methods to sample and judge athletes' behavior information based on wearable technology [2], but this method is less intuitive than the video capture method, which is not conducive to playback analysis and the recognition accuracy of actions is also not high enough. With the rapid development of the neural network in the field of computer vision, more and more methods to assist pedestrian detection and behavior recognition through image segmentation appear, and u-net is undoubtedly the most commonly used and simplest segmentation model to complete image segmentation. It is simple, efficient, easy to build, and can be trained from small data sets, It can help us quickly identify the athletes in training in the video, but if we want to quickly obtain the most effective information and identify the athletes' foul actions, we need to introduce the attention mechanism. The attention mechanism is being more and more widely used in various fields. Considering that the race-walking competition and training are usually carried out by multiple people at the same time, it is also a problem to be solved to identify whether the actions of multiple athletes are all standardized. Yolo algorithm can well preprocess the image frames and help us solve the problem of separation of human and scene for multi athletes, U-Net network combined with attention mechanism is used to identify the foul actions of all athletes in the video stream containing multiple athletes.

Accordingly, the contributions of this paper are summarized as follows: (i) The video stream of race walking training is obtained by recording video, the image frames in the video stream are preprocessed by the Yolo algorithm to obtain the image of athletes separate with a scene, and the u-net network combined with attention mechanism is used to identify the foul actions and nonstandard actions in the video stream of multiple athletes; (ii) test the effectiveness and accuracy of the method in (i) through experiments to see whether it can effectively detect the foul actions and non-standard actions in a single athlete or multi athletes race walking, and compare the results with the detection results from naked eyes to judge whether the accuracy of foul recognition has been improved, so as to determine whether the method has a positive impact on auxiliary training.

The rest of this paper is organized as follows: The related work is in Section 2. The proposed method is introduced in Section 3. Section 4 shows the implementation method of foul actions recognition. The experimental design and result analysis are introduced in Section 5 and Section 6, respectively. Finally, Section 7 concludes this paper and gives future research directions.

2. Related Work

In paper [1], the author studies the relationship between the observation state of race walking judges and CFF and shows that the CFF of eyes should be regarded as the physiological indexes in choosing the qualified race walking judges. In article [2], the author analyzed the video data recorded of 30 men during the World Cup walking competition and came to the conclusion that step length and stride length are the key areas that must be coordinated in long-distance walking competition.

In recent years, technology plays an important role to help training and judgment in sport; in paper [3], the author proposes the use of a wearable inertial system to derive novel biomechanical indices for the assessment of performance and infringements in race-walking, where the result shows that these indices can be implemented on a wearable inertial system to assist training and judgment in race-walking. Paper [4] shows the preliminary result on the use of a wearable inertial system for the assessment of performances and infringements in race-walking in 2019. Current judging of race walking in international competitions relies on subjective human observation to detect illegal gait, which naturally has inherent problems, the research in [5] aims to determine whether an inertial sensor could improve accuracy based on monitoring every step of seven races walkers in training and competition. In paper [6], Tabori et al. have proposed machine-learning algorithms for automatic detection of infringements (both LOGC and bent knee). Paper [7] presents a new motion analysis protocol for race-walking, through setting up a motion capture system and a force platform to record both kinematic and dynamic aspects of the athletes' action to detect infringement of the rules based on the measure of knee flexion-extension and the loss of ground contact. Paper [8] placed an inertial sensor at L5/S1 of the vertebral column of an Italian national team athlete to acquire timing measurements of the LOGC to validate an inertial system able to detect the loss of ground contact (LOGC) in race-walking in real training conditions, results show that the inertial system can improve the accuracy in detecting the visible LOGC. Paper [9] also aims to

develop an innovative approach based on a wearable inertial system, which enables objective evaluations on the loss of ground contact in race-walking.

Paper [10] proposes an attention mechanism LSTM framework for human action recognition in videos. In [11], a coattention model-based recurrent neural network (CAM-RNN) is proposed, where the CAM is utilized to encode the visual and text features, and the RNN works as the decoder to generate the video caption. Paper [12] introduces a regularized attention mechanism for graph attention networks. Paper [13] proposes an attention mechanism LSTM Framework for Human Action Recognition in Videos.

In the paper [14], the author presents temporal deformable convolutional encoder-decoder networks that fully employ convolutions in both encoder and decoder networks for video captioning. In paper [15], the author proposes a two-stream framework based on combinational deep neural networks to extract both temporal and spatial features by exploring the usage of 3D convolutional networks on both raw RGB frames and motion history images, and tune the weights of different feature channels since the network is trained end-to-end from learning combinational encoding of multiple features to LSTM-based language model. Paper [16] proposes a Multimodal Memory Model (M3) which builds a visual and textual shared memory to model the long-term visual-textual dependency and furthermore guides visual attention on described visual targets to solve visual-textual alignments. In [17], a video summarization technique that uses motion descriptors computed in the compressed domain is described. It can either speed up conventional color-based video summarization techniques or rapidly generate a key-frame-based summary by itself. A key-frames extraction method is proposed with Kekere's Proportionate Error (KPE) codebook generation techniques of vector quantization with ten different codebook sizes and two color-spaces (RGB and KLUV) in [18].

Paper [19] is talking about the usage of the LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages and [20] presents an encoder-decoder model for automatic video captioning based on Yolo Algorithm. Paper [21] proposes an improved YOLOv3-tiny for object detection based on the idea of feature fusion, compared with YOLOv3-tiny, the accuracy of the improved network structure is increased by 6.3%, and the detection speed is 31.8fps in 2019. Paper [22] proposes a faster detection method for real-time object detection based on a convolution neural network model called Single Shot MultiBox Detection (SSD), increase the accuracy in identifying objects. The research of paper [23, 24] is based on real-time multiple object detection through YOLO.

3. The Proposed Method

In recent years, with the continuous development of computer hardware, the computing power required by deep learning has been satisfied, so it has been widely applied in the field of computer vision research, surpassing the results of traditional machine learning. CNN convolutional neural network is a kind of feed-forward neural network with deep structure and convolution computation. As one of the representative algorithms of deep learning, it has achieved good results in large-scale recognition tasks, breaking through traditional classification methods and outperforming recognition naked eyes. However, at present U-Net is the most widely used convolutional neural network architecture in the field of image segmentation. The U-Net architecture combined with attention mechanism can focus limited attention on the effective features of the recognized athletes, thus saving resources rapid access to pertinent information to distinguish whether the athletes' actions are standardized. The overall architecture of this method is shown in Figure 1.

As it can be seen from Figure 1, there are four modules in our pipeline.

In the image acquisition module, the images during race walking competition or training through video are sampled to construct the training and testing dataset. It can be achieved by modern image processing tools, such as the OpenCV.

In the image matting module, the extracted images are preprocessed by YOLO-V3, such that the background can be filtered and only the 2D bounding box of the human is left.

The effective features are provided by the attention based on the U-Net structure, so as to output a three-dimensional probability vector composed of normal, foul, and non-standard action to identify the three kinds of actions.

3.1. The Yolo Algorithm. Object detection is a key task in the field of computer vision, which can be regarded as the combination of image classification and location. Since Ross Girshic K proposed R-CNN in 2013, Fast R-CNN, Faster R-CNN, YOLO, and other algorithms for object detection have been proposed one after another. The full name of Yolo is you only look once, which means that you can identify the category and location of objects (including pedestrians) in the picture only by browsing once. Compared with R-CNN series algorithms, Yolo does not need to generate a large number of candidate boxes first and then use a convolutional neural network to classify and regress the candidate boxes. It takes the whole image as the input of the network, and directly regresses the location and category of the bounding box at the output layer, so as to predict the object category and location in one step, saving time and hardware cost; however, it also can detect from real-time video and has stronger generalization ability.

In terms of implementation method, Yolo∗s convolutional neural network divides the input image into $S \times S$ grid, Then, uses each cell to detect the targets whose center point falls in the grid. Each cell will predict N bounding boxes and confidence score which includes the possibility $P(\text{object})$ of bounding boxes containing targets and the accuracy $A(\text{object})$ of bounding boxes. When the bounding box is a background, $P(\text{object}) = 0$, When the bounding box contains a target, $P(\text{object}) = 1$. Therefore, the confidence score can be defined as C, $C = P(\text{object}) * A(\text{object})$. The size and position of the bounding box are represented by four values: (X, Y, W, H), where (x, y) is the central coordinate of the

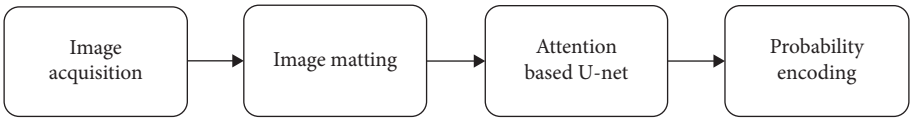


Figure 1: The proposed pipeline for action recognition.

bounding box, and W and H are the width and height of the bounding box. Thus, the predicted value of each bounding box contains five elements: (X, Y, W, H, C), the first four represent the size and position of the bounding box, and the last one is the confidence score. Border category confidence score is actually the conditional probability under each border-box confidence score, which can be expressed as $P(\text{class}|\text{object})$, It represents the possibility that the target object in the bounding box belongs to each category and whether the bounding box matches the target object. In general, the prediction frame of the network is filtered according to the confidence score of the category to identify the target object (including people).

3.2. The U-Net Network. Semantic segmentation is an important branch in the field of image processing and machine vision. It needs to judge the category of each pixel of the image and segment it accurately. At present, it is widely used in the field of automatic matting. U-Net appeared in 2015, which is the most commonly used segmentation model now, it can be regarded as a simplified structure based on the FCN model. It is efficient and simple, can be trained with small data sets, and has higher accuracy than the FCN network. It mainly uses a U-shaped network structure to obtain context information and location information. Figure 2 shows the network model of U-Net [25].

U-Net is actual a U-shaped structure of encoder-decoder. As shown in Figure 2, the left side is the convolution layer for feature extraction, and the right side is the upper sampling layer. The u-net structure contains 4 revolutionary layers and 4 corresponding upsampling layers. Therefore, we can initialize the weight first and then train the model, or use the existing convolution layer mechanism and the trained weight value through the following up sampling layer to train. The feature map obtained from each convolution layer of the U-Net network will be concatenated to the corresponding upper sampling layer so that the feature map of each layer can be effectively used in subsequent calculations, that is, skip connection. Thus, the final feature map contains both high-level features and low-level features, which improves the accuracy of the model.

3.3. Attention Mechanism. Attention mechanism was first applied in computer vision and later developed in the field of NLP. This mechanism focuses limited attention on key information, so as to save resources and quickly obtain the most effective information.

In fact, attention mechanism is a process of filtering out a small amount of important information from a large amount of information, ignoring unimportant information, and better-weighted fusion of information. Focus on its corresponding value according to its weight, the larger the weight, the more aggregation. And the weight indicates the importance of information. Value is its corresponding information, as shown in Figure 3.

The specific calculation process of attention mechanism can be roughly summarized into three stages: (i) calculate the correlation between query and key; (ii) normalize the value obtained from (iii) to obtain the weight coefficient; (iv) assign the weight and summarize the value according to the weighting coefficient.

In stage (i), different functions and calculation mechanisms can be introduced to calculate the correlation by calculating the vector dot product or vector cosine similarity of Query and Key, or by introducing an additional neural network; In stage (ii), a calculation method similar to SoftMax is introduced to convert the values in the first stage. While normalizing, the original calculated values can be sorted into a probability distribution in which the sum of the weights of all elements is 1, highlighting the weights of important elements; In stage (iii), the weight coefficient corresponding to value (i) is calculated in stage (ii) is introduced for weighted summation to obtain the Attention Value for Query.

4. Proposed Method

4.1. Athletes Recognition Based on Yolov3. In this paper, we mainly detect athletes in multi-person race walking competitions or training by pedestrian detection based on yolov3. We use a convolutional neural network to detect athletes in the video. After the video frame is input, it first enters the target detection network based on Yolov3, extracts the features through darknet-53, then performs up sampling and feature fusion, finally performs regression analysis to obtain the prediction frame information, as shown in Figure 4.

As a single-stage detector, Yolo directly classifies and predicts the objects at each position of the feature map without generating candidate regions. In this paper, we use the labelme pedestrian data set, and then through the built Yolo algorithm to generate the model and train, we can accurately identify multiple race walkers and separate the race walkers from the image. Finally, several bounding boxes are output , resize output to $256 * 256$ for actions' recognition of next step.

4.2. Attention Based U-Net Network. Based on the discussion of the above methods, we propose the attention-based U-Net network structure to extract the action state of athletes. For the specific network structure, please refer to Figure 5 below.

According to Figure 5, the network input is 256×256 image frames extracted by YOLOv3, and the output is a

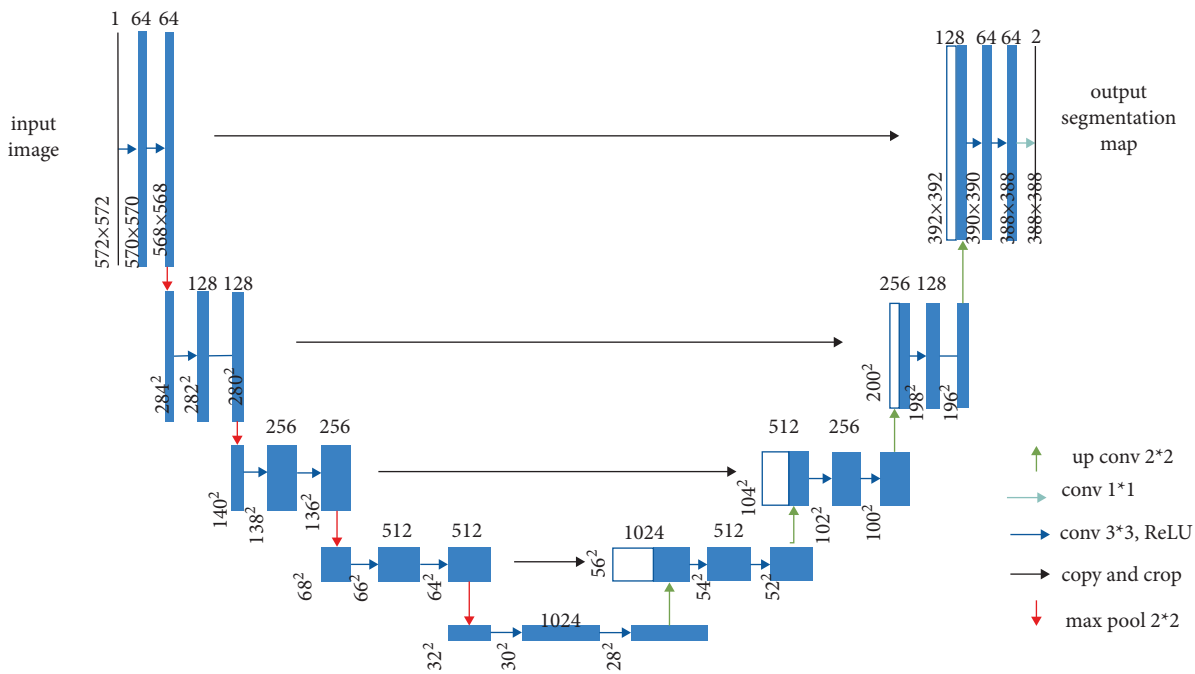


Figure 2: U-Net network model.

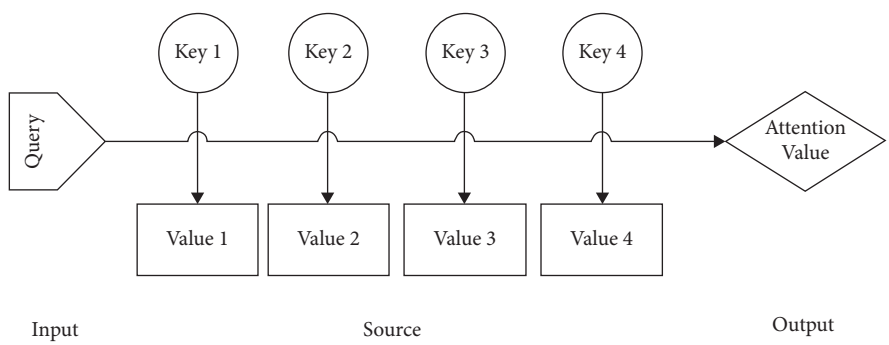


Figure 3: Attention mechanism.

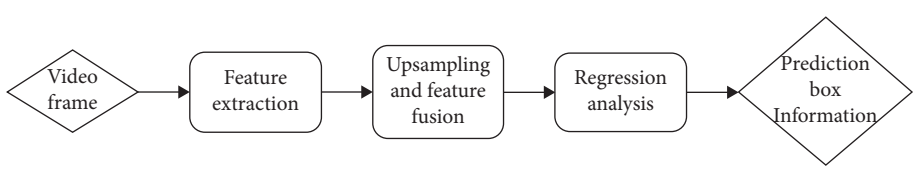


Figure 4: Pedestrian detection process based on Yolov3.

three-dimensional vector X, assuming $X = [X_1, X_2, X_3]\ T$, where X_1 represents the probability of normal action judged from this frame, X_2 represents the probability of nonstandard action judged from this frame, X_3 represents the probability of foul action judged from the frame, and they satisfy $X_1 + X_2 + X_3 = 1$.

The network structure combines the encode part of U-Net architecture and attention mechanism to centrally learn the lower limbs' actions of race-walking athletes and extract the corresponding features. The network front end consists of full convolution and max pooling. Among them, $3 \times 3 \times 16$ represents a 3×3 full convolution network, 16 represents the number of output channels, Relu is used as the activation function, the stripe is 1, zero padding is used, and the number of padding is 1. It can be seen from the figure that the space is down-sampled three times, and the features sampled in the first two times pass through the Attention module. The attention module here adopts CBAM in the document [26–29]. The aggregation layer is used to fuse the features of three scales. In order to ensure consistent resolution, 4x and 2x up samplings are used to align the features [26]. Max pooling channel represents maximum pooling in the channel layer,

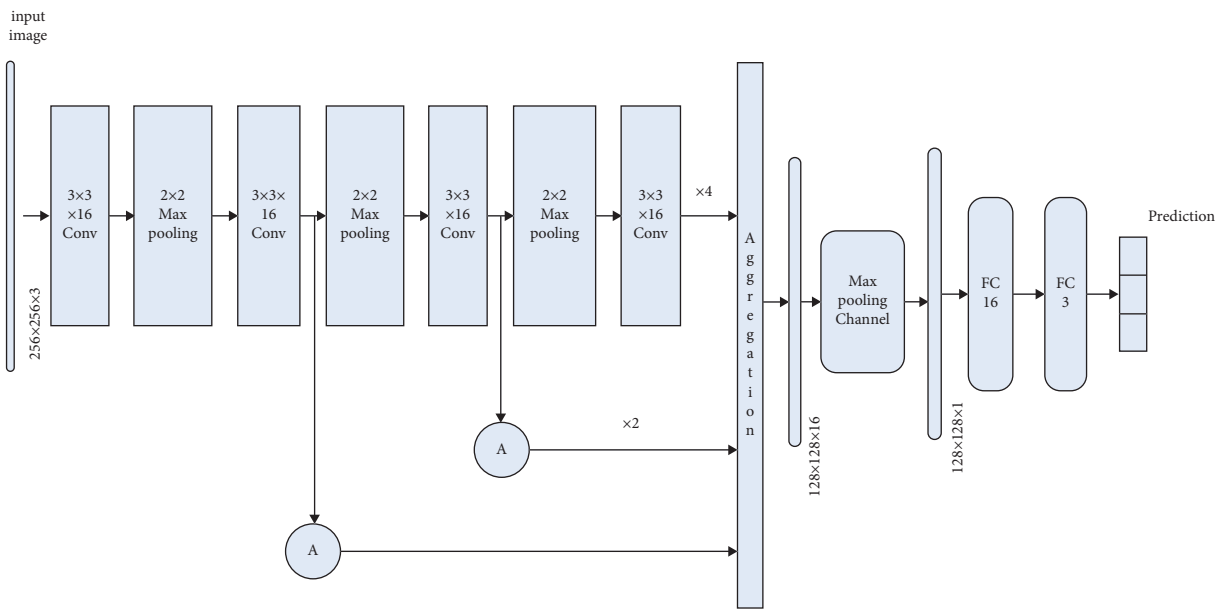


Figure 5: Attention-based U-Net network structure.

which can quickly integrate the information of features. The last two layers FC16 and FC3 represent the whole connection layer of 16 neurons and the whole connection layer of 3 neurons, respectively. Finally, FC3 neurons can obtain a probability value of 0~1 through sigmoid function.

4.3. Training Loss. Each task is associated with a loss function [27]. We select loss function for the task of action classification. Multi classification is actually an extension of two classifications:

$$L = \frac{1}{N}\sum_{i} L_i = -\frac{1}{N}\sum_{i}\sum_{c=1}^{M} w_{ic}\log\left(X_{ic}\right). \quad (1)$$

In formula (1): N is the length of encoding probability, M represents a number of classifications, $w\ ic$ represents Symbolic function (0 or 1), if the actual classification of sample i is equal to C, the value is 1, otherwise, the value is 0; X_{ic} represents prediction probability of observation sample i belonging to classification C.

Now, we use these expressions to calculate the value of the loss function in the above example.

Sample1 loss $= -(0 + 0 \times \log 0.3 + 1 \times \log 0.4) = 0.91$.

Sample2 loss $= -(1 \times \log 0.1 + 0 \times \log 0.2 + 0 \times \log 0.7) = 2.30$.

5. Experimental Design

5.1. Training Model

5.1.1. Experimental Dataset Sources. In this paper, the dataset is obtained by collecting the data from the Race Competition and daily training. Namely, 30% of the 500 groups come from the 20 km race walking competition, and the last comes from the data recorded in the process of athletes' daily training and simulated competition.

5.1.2. Network Training. As mentioned earlier, this paper will use Yolo as the input of the motion detection network, so there is no retraining. In network training, the batch size is 32 and the learning rate is 0.001. Adam optimizer is used to learn 200 epoch.

5.2. Validity Verification. We record the 20 km race walking training process of 5 groups of women and 5 groups of men by video. Six athletes participate in each training, and the race-walking athletes are required to make fouls and use nonstandard actions from time to time. The number of fouls and nonstandard actions in the training of each group of athletes is determined by our Yolov3 algorithm combined with an attention mechanism based on the U-Net model and naked eyes tracking by coaches, respectively.

6. Experimental Results and Analysis

6.1. Experimental Results. The experimental results of action classification of the 20 km race walking training process are shown in Tables 1–3, here group1-group5 are women groups and group6-group10 are men groups.

6.2. Result Analysis. Through the above experimental results, we can observe the following:

(i) Through the new method based on deep learning adopted in this paper we can successfully identify the foul and nonstandard actions of each athlete in the multi person race walking.

(ii) Using the new discrimination method proposed in this paper, more fouls and nonstandard actions are identified in each race, about 8% higher in an average form. This shows that in the process of 20 km race walking, the accuracy rate of identifying fouls and

Table 1: Foul and nonstandard action recognition.

NO.	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
Foul by naked eyes	85	76	62	82	137	62	113	125	107	132
Foul by deep learning	97	86	98	90	190	73	131	200	183	194
Nonstandard by naked eyes	98	102	121	109	198	89	99	201	207	236
Nonstandard by deep learning	118	128	149	121	249	119	128	248	210	271

Table 2: Accurate rate comparison between different methods for Foul action estimation.

NO.	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
GT (Ground truth)	99	88	100	91	191	75	132	203	185	157
Eyes	86.7%	86.3%	88.6%	90.1%	84.6%	82.7%	85.0%	83.9%	88.4%	82.5%
AI	98.9%	97.7%	97.1%	98.9%	98.7%	97.3%	98.5%	97.3%	97.5%	97.5%

Table 3: Accurate rate comparison between different methods for Nonstandard estimation.

NO.	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9	Group10
GT (Ground truth)	120	133	155	123	260	130	140	260	220	275
Eyes	80.8%	64.7%	63.2%	73.2%	190	73.1%	93.6%	76.9%	83.2%	79.2%
AI	98.3%	96.2%	96.1%	98.4%	249	91.2%	91.4%	95.4%	95.5%	98.5%

nonstandard actions by the new method is higher than that by the coach through naked eyes. Hence, it is more effective than human eyes to assist training.

7. Conclusions

In this paper, an attention-based network is proposed to identify the fouls and nonstandard actions of multiple athletes in the process of long-distance race walking, and the effectiveness of this method is verified by experiments. Through the experimental results, we know that the new method has less interference and higher accuracy than human eye recognition.

In the future research, we will continue to focus on improving the accuracy through training more data sets and reduce the complexity of the method so that we can not only use this method to identify actions to assist training, but also apply this method to the real race-walking competition to assist the referee to improve the accuracy of judgment and reduce misjudgment.

Data Availability

The data used to support the findings of the study are included in the paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest in this paper.

Acknowledgments

This work was supported by 2020 Key Research Project of Humanities and Social Sciences in Anhui Universities Granted No. 2020jyxm1989.

References

[1] S. Xu and F. Jiao, "The observation state of race walking judge and CFF," *China Sport Science and Technology*, vol. 35, pp. 39–42, 1999.

[2] B. Hanley, A. Bissas, and A. Drake, "Kinematic characteristics of elite men's 50 km race walking," *European Journal of Sport Science*, vol. 13, no. 3, pp. 272–279, 2013.

[3] T. Caporaso, S. Grazioso, G. D. Gironimo, and A. Lanzotti, "Biomechanical indices represented on radar chart for assessment of performance and infringements in elite race-walkers," *Sports Engineering*, vol. 23, 2020.

[4] T. Caporaso, S. Grazioso, D. Panariello, G. Di Gironimo, and A. Lanzotti, "A wearable inertial device based on biomechanical parameters for sports performance analysis in race-walking: preliminary results," in *Proceedings of the II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0&IoT)*, pp. 259–262, Naples, Italy, June 2019.

[5] J. B. Lee, R. B. Mellifont, B. J. Burkett, and D. A. James, "Detection of illegal race walking: a tool to assist coaching and judging," *Sensors*, vol. 13, p. 16, 2013.

[6] J. Taborri, E. Palermo, and S. Rossi, "Automatic detection of faults in race walking: a comparative analysis of machine-learning algorithms fed with inertial sensor data," *Sensors*, vol. 19, no. 6, 2019.

[7] G. Di Gironimo, T. Caporaso, D. M. Del Giudice, A. Tarallo, and A. Lanzotti, "Development of a new experimental protocol for analysing the race-walking technique based on kinematic and dynamic parameters," *Procedia Engineering*, vol. 147, pp. 741–746, 2016.

[8] G. D. Gironimo, T. Caporaso, G. Amodeo, and D. M. D. Maria, "Outdoor tests for the validation of an inertial system Able to detect illegal steps in race-walking," *Procedia Engineering*, vol. 147, pp. 544–549, 2016.

[9] G. D. Gironimo, T. Caporaso, D. M. D. Giudice, and A. Lanzotti, "Towards a new monitoring system to detect illegal steps in race-walking," *International Journal on Interactive Design and Manufacturing*, vol. 11, no. 2, pp. 1–13, 2016.

[10] C. Yan, Y. Tu, X. Wang, and Y. Zhang, "Corrections to "STAT: spatial-temporal attention mechanism for video captioning"," *IEEE Transactions on Multimedia*, vol. 22, no. 3, 830 pages, 2020.

[11] B. Zhao, Y. Li, and X. Lu, "Co-attention model based RNN for video captioning," *IEEE Transactions on Image Processing*, vol. 28, no. 99, pp. 5552–5565, 2019.

[12] U. S. Shanthamallu, J. J. Thiagarajan, and A. Spanias, "A regularized attention mechanism for graph attention networks," in *Proceedings of the ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 3372–3376, Barcelona, Spain, May 2020.

[13] C. I. Orozco, M. E. Buemi, and J. J. Berlles, "Towards an Attention Mechanism LSTM Framework for Human Action Recognition in Videos," in *Proceedings of the 2020 IEEE Congreso Bienal de Argentina (ARGENCON)*, pp. 105–109, Resistencia, Argentina, December 2020.

[14] J. Chen, Y. Pan, Y. Li, T. Yao, H. Chao, and T. Mei, "Temporal deformable convolutional encoder-decoder networks for video captioning," vol. 33, pp. 6–8, 2019, https://arxiv.org/abs/1905.01077.

[15] C. Zhang and Y. Tian, "Automatic video description generation via LSTM with joint two-stream encoding," in *Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 753–756, IEEE, Cancun, December 2016.

[16] J. Wang, W. Wei, and H. Yan, "M3: multimodal memory modelling for video captioning," in *Proceedings of the 2018 IEEE/CVF conference on computer vision and pattern recognition*, pp. pp.7512–7520, IEEE, Salt Lake City, UT, USA, June 2018.

[17] A. Divakaran, R. Radhakrishnan, and K. A. Peker, "Video summarization using descriptors of motion activity: a motion activity based approach to key-frame extraction from video shots," *Journal of Electronic Imaging*, vol. 10, no. 4, pp. 909–916, 2001.

[18] S. D. Thepade and P. H. Patil, "Novel video keyframe extraction using KPE vector quantization with assorted similarity measures in RGB and LUV color spaces," in *Proceedings of the 2015 International Conference on Industrial Instrumentation and Control (ICIC)*, pp. pp.2512–2520, IEEE, Pune, India, July 2015.

[19] E. Habler and A. Shabtai, "Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages," *Computers & Security*, vol. 78, pp. 155–173, 2017.

[20] A. Shabtai and I. Habler, "Using LSTM encoder-decoder algorithm for detecting anomalous," *Ads-B Messages*, vol. 78, pp. pp1–8, 2019.

[21] H. Gong, H. Li, K. Xu, and Y. Zhang, "Object detection based on improved YOLOv3-tiny," in *Proceedings of the 2019 Chinese Automation Congress (CAC)*, pp. pp.1520–1528, IEEE, Hangzhou, China, February 2020.

[22] S. Kanimozhi, G. Gayathri, and T. Mala, "Multiple Real-time object identification using Single shot Multi-Box detection," in *Proceedings of the 2019 International Conference on Computational Intelligence in Data Science (ICCIDS)*, pp. pp.5512–5522, IEEE, Chennai, India, October 2019.

[23] B. Kumar, R. Punitha, and Mohana, "YOLOv3 and YOLOv4: multiple object detection for surveillance applications," in *Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 1316–1321, Tirunelveli, India, October 2020.

[24] M. Mahendru and S. K. Dubey, "Real time object detection with audio feedback using Yolo vs. Yolo_v3," in *Proceedings of the 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, p. pp.734, Noida, India, March 2021.

[25] P. Li, L. Zhang, J. Qiao, and X. Wang, "A semantic segmentation method based on improved U-net network," in *Proceedings of the 2021 4th international conference on advanced electronic materials, computers and software engineering (AEMCSE)*, pp. 600–603, Changsha, China, August 2021.

[26] S. Woo, J. Park, and J. Y. Lee, "CBAM: convolutional block Attention module," in *Proceedings of the European conference on computer vision*, pp. 541–549, Munich, Germany, July 2018.

[27] V. L. Tran and H. Y. Lin, "3D object detection and 6D pose estimation using RGB-D images and mask R-CNN," in *Proceedings of the 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1890–1898, IEEE, Glasgow, UK, December 2020.

[28] M. Braun, Q. Rao, Y. Wang, and F. Flohr, "Pose-RCNN: joint object detection and pose estimation using 3D object proposals," in *Proceedings of the IEEE International Conference on Intelligent Transportation Systems*, pp. 178–189, IEEE, Rio de Janeiro, Brazil, December 2016.

[29] Z. Zhao, K. Chen, and S. Yamane, "CBAM-Unet++: easier to find the target with the attention module CBAM," in *Proceedings of the 2021 IEEE 10th Global Conference on Consumer Electronics(GCCE)*, pp. 655–657, Kyoto, Japan, December 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8124149, 8 pages
https://doi.org/10.1155/2022/8124149

Research Article

Research on HMM-Based Efficient Stock Price Prediction

Zhi Su and Bo Yi

School of Statistics and Mathematics, Central University of Finance and Economics, Beijing 100081, China

Correspondence should be addressed to Bo Yi; 2018110076@email.cufe.edu.cn

Received 10 January 2022; Revised 11 February 2022; Accepted 15 February 2022; Published 7 March 2022

Academic Editor: Jianhui Lv

Stock market is one of the most important parts of the investment market. Compared with other industries, the stock market not only has a higher rate of return on investment but also has a higher risk, and stock price prediction has always been a close concern of investors. Therefore, the research on stock price prediction methods and how to reduce the error of stock price prediction has become a hot topic for many scholars at home and abroad. In recent years, the development of computer technology such as machine learning and econometric method makes the stock price prediction more reliable. Due to the hidden Markov nature of stock price, this paper proposes a stock price prediction method based on hidden Markov model (HMM). To be specific, since the data of stock price have continuity in time series, it is necessary to extend the discrete HMM to the continuous HMM, and then put forward the up and down trend prediction model based on the continuous HMM. The first-order continuous HMM is extended to the second-order continuous HMM, and the stock price is predicted by combining the prediction method of fluctuation range. As a result, the proposed second-order continuous HMM-based stock price prediction model is simulated on Hang Seng Index (HSI), one of the earliest stock market indexes in Hong Kong. The evaluation results on six months HSI show that the predicted value of the proposed model is very close to the actual value and outperforms three benchmarks in terms of RMSE, MAE, and R^2.

1. Introduction

In recent years, with the increasing number of listed companies, stock has become one of the hot topics in the financial field [1]. On the one hand, the trend of stock price will determine the trend of many economic behaviors to a certain extent, so the stock price prediction is also concerned by more and more financial investors and financial analysts [2, 3]. On the other hand, as the number of investors in the stock market increases year by year, only by accurately analyzing the future trend of stock prices can quickly grasp the market trend and obtain more investment returns. Stock price prediction is the focus of financial research, which is generally considered as a challenging task because of great instability of financial markets [4]. However, in order to obtain profits or understand the nature of the stock market, many market participants or researchers try to use various methods to predict stock prices [5]. The stock market trades with high frequency every day, thus generating a large amount of stock related data [6]. However, as a matter of fact, such huge amount of data are metaphysical for investors. How to extract valuable information from massive stock data with effective methods has become a problem to be solved at present.

However, Brealey's "Fundamentals of Corporate Finance" mentions that the price trend of stocks is unpredictable [7]. At present, the mainstream view of the academic community is that predicting stock price is equivalent to predict the return rate of stock. Return rate of stock is predictable, which is not inconsistent with the efficient markets hypothesis [8, 9]. In fact, any variable with a nonzero correlation can predict returns [10]. The efficient market hypothesis is equivalent to the fact that the stock price has fully reflected all the known information, so any information that affects the price can predict the return rate of the stock.

In the traditional quantitative investment field, the selection of target stocks and the prediction of stock prices are mostly based on the results of long-term stock market experience [11]. The antirisk ability and long-term prediction ability of empirical stock analysis methods are often poor, and it is not easy to spread and promote [12]. In addition, the

analysis speed of traditional method is often slow. Then came the stock analysis methods based on statistics and finance, which is also the beginning of mathematical modeling of stocks, such as autoregressive model [13], stochastic volatility model [14], and Markov model [15]. The prediction and analysis effect of these methods is better than empirical methods. Moreover, due to the use of mathematical modeling, these models are very suitable for computer analysis, which are based on a small amount of input data, and cannot be applied to the current large-scale data scenario.

Thanks to the continuous development of computer technology and the Internet, it is possible for artificial intelligence to enter the field of financial analysis. Since the stock price is an observable time series, the factors that determine the stock price belong to unknown variables [16]. This feature is consistent with hidden Markov model (HMM), which has been applied to stock price prediction by many scholars [17]. HMM is a statistical model that has been used in automatic speech recognition [18], DNA sequence analysis [19], image processing [20], and pattern recognition [21]. The main contribution of this paper is that the second-order continuous HMM-based model is constructed for stock price prediction.

The rest of this paper is organized as follows. Section 2 reviews related work. In Section 3, HMM-based stock price prediction model is presented. Experimental results are presented in Section 4. Section 5 concludes this paper and gives future work.

2. Related Work

Nowadays, stock price prediction has made a lot of achievements and related technologies are becoming more and more mature. Especially since the introduction of artificial intelligence related methods, stock price prediction methods have made great progress. In addition, due to the theoretical development of the financial field, the mathematical models used to describe the stock are also more abundant. At present, the research on stock price prediction can be divided into two directions: statistics-based and nonstatistical. For the stock price prediction method based on statistics, in [22], a news augmented generalized autoregressive conditional heterscedasticity (NA-GARCH) model was proposed to use quantitative news sentiment and its impact on asset price movement as the second information source to predict the fluctuation of asset price return together with asset time series data. In [23], the authors introduced deep learning method into the study of stock market correlation. Based on recurrent deep neural network and GARCH model, a hybrid model was proposed for stock price prediction. In [24], geometric Brownian motion mathematical model was used to predict the future price of stock. In [25], optimized custom moving average most suitable for stock time series smoothing was proposed to smooth stock price series and forecast trend direction. In [26], the authors used the exponential smoothing method to process the initial data, calculate the relevant technical indicators as the characteristics to be selected, and optimize the random forest to predict the stock trend. In [27], the authors proposed a hybrid extreme gradient boosting framework and auto regressive integrated moving average model to predict stock price. In [28], a combined predicting method based on wavelet multiresolution analysis was proposed to predict the stock market more accurately and concisely. In [29], a prediction method of stock price volatility based on time series analysis technology was proposed.

For stock price prediction methods based on nonstatistics in [30], a convolutional neural networks-bi-directional long short-term memory-attention mechanism (CNN-BiLSTM-AM) method was proposed to predict the stock closing price of the next day. In [31], a new method of predicting stock trend based on graph convolution neural network model was proposed, which considered both stock market information and individual stock information. In [32], a model called feature fusion long-term and short-term memory convolution neural network was proposed to predict stock prices. In [33], a convolution neural network model based on deep decomposition machine and attention mechanism was constructed to improve the prediction accuracy of stock price movement. In [34], a stock price predicting method integrating multiple data sources and investor sentiment was proposed. In [35], a prediction method of stock market price trend based on high order HMM was presented. In [36], a multisource heterogeneous data analysis method was constructed to integrate multisource information, namely transaction data, news event data, and investor comments to predict future stock price. In [37], an adaptive hidden Markov abnormal state model (AHMMAS) was proposed to model and detect price manipulation activities.

3. HMM-Based Stock Price Prediction

3.1. HMM Introduction. HMM is usually characterized by the following five elements.

(i) N, number of hidden states in model. Record each state as $S = \{s_1, s_2, \cdots s_t \cdots, s_N\}$, and s_t represents the state at time t.

(ii) M, number of different observations for each hidden state. Record each observation as $O = \{o_1, o_2, \cdots, o_N\}$, and o_t represents the observation at time t.

(iii) $\mathbb{A} = \{a_{ij}\}$, state transition probability matrix, which represents the probability that the state at time $t+1$ is s_i when the known state is s_j at time t.

(iv) $\mathbb{B} = \{b_j(o_k)\}$, observation probability matrix when hidden state is j, which represents the probability that the observed value is o_k when the state is s_j at time t.

(v) $\Pi = \{\pi_i\}$, initial state probability distribution, which represents the probability that the hidden state is s_j at the initial time $t = 1$.

A complete HMM includes state transition probability matrix, observation probability matrix, and initial

probability distribution. For simplicity, a simple notation can be used to represent the complete parameters of HMM.

$$\Lambda = (\mathbb{A}, \mathbb{B}, \Pi). \tag{1}$$

3.2. Second-Order Continuous HMM. In the first-order HMM, the state of the current moment is assumed to be only related to the state of the previous moment and has no relationship with the state of other moments, namely the so-called first-order hidden Markov property. It is obvious that such assumption has certain defects, so it is necessary to establish the second-order HMM to make up for its defects, and the second-order HMM can overcome many shortcomings of the first-order HMM. On this basis, this paper extends the first-order HMM to the second-order HMM for efficient stock price prediction. The closing prices from July 31, 2021 to December 31, 2021, were selected as the research subjects.

The Gaussian mixture model can be used to extend the discrete HMM to the continuous HMM. The HMM also has shortcomings in model assumptions, that is, there are many factors affecting the stock price. The stock price of the day cannot be affected only by the stock price of the previous day, which is very unreasonable for the prediction of the stock price. Therefore, in order to improve the model, the first-order HMM is extended to the second-order HMM, so as to generate the final model, that is, the second-order continuous HMM, which also strengthens the model foundation for the prediction of stock price index.

In order to explain the Baum–Welch algorithm [38] of the second-order continuous HMM, the Baum–Welch algorithm of the first-order continuous HMM can be referred. Therefore, a custom variable [39] $\delta_t(m,n,u)$ is also be used, which represents the probability that the model is in state s_m at time $t-1$, state s_n at time t, and state s_u at time $t+1$ under the condition of given Λ and O, that is

$$\delta_t(m,n,u) = P\left(q_{t-1} = s_m, q_t = s_n, q_{t+1} = s_u | O, \Lambda\right). \tag{2}$$

Define another custom variable $\delta_t(m,n)$, which represents the probability that the model is in state s_m at time $t-1$, state s at time t under the condition of given Λ and O, that is

$$r_t(m,n) = P\left(q_{t-1} = s_m, q_t = s_n | O, \Lambda\right). \tag{3}$$

$$\text{so } r_t(m,n) = P\left(q_{t-1} = s_m, q_t = s_n | O, \Lambda\right)$$
$$= \sum_{k=1}^{n} P\left(q_{t-1} = s_m, q_t = s_n, q_{t+1} = s_u | O, \Lambda\right) = \sum_{k=1}^{n} \delta_t(m,n,u), \tag{4}$$

and we have

$$\begin{aligned} &\hat{a}_{m,t} = \frac{r_2(m,n)}{\sum_{n=1}^{N} r_2(m,n)}, \hat{a}_{m,n,u} = \frac{\sum_{t=2}^{T-1} \delta_t(m,n,u)}{\sum_{t=2}^{T-1} r_t(m,n)}, \\ &\hat{b}_{m,u} = \frac{\sum_{n=1}^{N} r_2(m,n)}{\sum_{n=1}^{N} \tau_t(n)}, \hat{b}_{m,n,u} = \frac{\sum_{t=2}^{T-1} \delta_t(m,n)}{\sum_{t=2}^{T-1} \tau_t(m,n)}, \hat{\Pi}_i = \sum_{n=1}^{N} r_2(m,n), \end{aligned} \tag{5}$$

where $\tau_t(m,n)$ is the probability of the mth mixed component of the observation vector O_t at time t, that is

$$\tau_t(m,n) = \frac{\alpha_t(m,n)\beta_t(m,n)}{\sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_t(m,n)\beta_t(m,n)} \cdot \frac{W_{m,n} N\left(O_t, \varphi_{m,n}, \sum m,n\right)}{\sum_{m=1}^{G} W_{m,n} N\left(O_t, \varphi_{m,n}, \sum m,n\right)}. \tag{6}$$

Because the mixture Gaussian density function is selected, the parameter estimation problem is transformed from the parameter estimation of the observed probability matrix B_2 to the estimation of parameter $W_{m,n}$, $\varphi_{m,n}$, $\sum m,n$, and the parameter estimation is as follows:

$$\begin{cases} \hat{W}_{m,n} = \dfrac{\sum_{t=2}^{T-1} \tau_t(m,n)}{\sum_{t=2}^{T-1}\sum_{m=1}^{M} \tau_t(m,n)}, \\ \hat{\varphi}_{m,n} = \dfrac{\sum_{t=2}^{T-1} \tau_t(m,n) O_t}{\sum_{t=2}^{T-1} \tau_t(m,n)}, \\ \widehat{\sum m,n} = \dfrac{\sum_{t=2}^{T-1} \tau_t(m,n)\left(O_t - \mu_{m,n}\right)\left(O_t - \mu_{m,n}\right)'}{\sum_{t=2}^{T-1} \tau_t(m,n)}. \end{cases} \tag{7}$$

3.3. Stock Price Prediction Model. This section proposes a new prediction method based on the continuous HMM Baum–Welch algorithm combined with K-means clustering and dichotomy to classify data, and the specific steps are as follows:

Step 1. For closing prices series CP_t, $t = 1, 2, \cdots, T$ in one day, let $Y_t = CP_{t+1} - CP_t/CP_t$, $t = 1, 2, \cdots, T$-1. According to the generated observation series Y_t, a moderate initial growing rate α_0 is determined, and $\alpha_0 = \max(Y_t) + \min(Y_t)/2$.

Step 2. Hidden states can be divided into small growing rate g_1 (less than α_0) and large growing rate g_2 (larger than α). When $Y_t < \alpha_0$, hidden state is in small growing state g_1 at time t, that is, $I_t = 1$, else $I_t = 2$. Therefore, the certain initial state probability distribution Π and state transition matrix $\mathbb{A}$ cab be calculated.

Step 3. According to the index series I_t, the closing prices series CP_t is divided into two types of data, and a reasonable number of mixed components M is determined. K-means clustering is performed on both two types of data, and the initial parameter $W_{m,n}$, $\varphi_{m,n}$ and $\sum m,n$ are determined, so the initial continuous HMM is determined.

Step 4. Determine the convergence threshold and carry out multiple iterations according to Baum–Welch algorithm of the continuous HMM until the model enters the convergence saturation state, the output log likelihood tends to be stable, and finally the convergence threshold is determined.

Step 5. Decode the model to obtain the hidden state at time t. If hidden state is the small growing state g_1, then $a_{n+1} = a_n + \min(Y_t)/2$, $\max(Y_t) = a_n$, and repeat Step 2 to Step 5, if hidden state is in the large growing state g_2, then

$a_{n+1} = a_n + \max(Y_t)/2$, $\min(Y_t) = a_n$, and repeat Step 2 to Step 5. This loop continues until the initial state probability distribution Π does not change, and the final growing state α_F is obtained.

Step 6. The predicted data at time $t + 1$ are as follows.

$$CP_{t+1} = (1 + \alpha_F).CP_t. \tag{8}$$

4. Evaluation

4.1. Experimental Setup. For most stocks, because each stock has its own trend, it cannot well measure the local stock market, and Hang Seng Index (HSI) is one of the earliest stock market indexes in Hong Kong, which can better reflect the price movements of major industry sectors of the market. The closing prices from July 1, 2021 to December 31, 2021, were selected as the research subjects. As of December 31, 2021, all information for an index prior to its launch date is back-tested, and back-tested performance reflects hypothetical historical performance. The data should be preprocessed before the prediction. We convert the daily closing price sequence of stocks into the corresponding benefit sequence. Due to the complexity of stock market, normal distribution cannot accurately simulate the probability distribution of observed values. Therefore, the probability of the observed value generated by the HMM model is transformed into the probability of each discrete value by discretizing the stock return rate. The rolling prediction method was used to make 1-month prediction, 3-month prediction, and 6-month prediction. To verify the effectiveness of the algorithm proposed in this paper, NA-GARCH [22], CNN-BiLSTM-AM [30], and AHMMAS [37] were used for the comparison. In order to evaluate the performance of the proposed algorithm on stock price prediction, this paper uses three regression evaluation metrics: root mean square error (RMSE), mean absolute error (MAE), and R^2 to quantify the performance of the model. The three metrics are calculated as follows:

$$\begin{aligned} \mathrm{RMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{N}\left(CP_i - CP_i'\right)^2}, \\ \mathrm{MAE} &= \frac{1}{n}\sum_{i=1}^{N}\left|CP_i - CP_i'\right|, \\ R^2 &= 1 - \frac{\sum_{i=1}^{N}\left(CP_i - CP_i'\right)^2}{\sum_{i=1}^{N}\left|CP_i - \overline{CP_i}\right|}, \end{aligned} \tag{9}$$

where CP_i is the actual value, CP_i' is the predicted value, $\overline{CP_i}$ is the mean value of the actual value, and N is the number of samples. RMSE and MAE are used to measure the deviation between the actual value and the predicted value. The smaller the value is, the closer the predicted value is to the actual value. R^2 is used to measure the fitting degree of model, and the closer it is to 1, the better fitting the model is.

4.2. Experimental Results. As can be seen from Figure 1, the predicted values and actual values of the four algorithms are similar in some places, but there is a big difference in some places, and even the opposite trend appears in some places. On the whole, the volatility is large. The volatility of the proposed algorithm is small, especially on some dates, which is basically consistent with the actual value. In contrast, the other three baselines did not perform well in some aspects, especially when the stock market was volatile, which was quite different from the actual value.

It can be seen from Figure 2 that the overall trend of the broken line is very similar from the comparison of the predicted value with the actual value over the 3 months, except for the opposite fluctuations at certain time, but the overall effect is better. It is still obvious that the curve between the predicted value and the actual value of the algorithm in this paper is the closest. The error between the predicted result and the actual value of the model is small and the fitting degree is high. AHMMAS algorithm and CNN-BiLSTM-AM algorithm have a large error between the predicted value and the true value on HSI, and the fitting degree of the predicted value curve and the actual value curve is low. The error between the predicted value and the actual value of NA-GARCH algorithm on HSI is the largest.

As indicated in Figure 3, in the comparison between the predicted value and the actual value in 6 months, the overall effect is not very ideal. By comparing the prediction accuracy between the predicted value and the actual value of each algorithm, it can be found that the prediction accuracy of NA-GARCH is the lowest, the prediction accuracy of CNN-BiLSTM-AM is higher than NA-GARCH, and the prediction accuracy of AHMMAS is significantly higher than CNN-BiLSTM-AM. Therefore, we can see the efficiency and stability of the proposed algorithm. The Gaussian mixture model can be used to extend the discrete HMM to the continuous HMM. The HMM also has shortcomings in model assumptions, that is, there are many factors affecting the stock price. The stock price of the day cannot be affected only by the stock price of the previous day, which is very unreasonable for the prediction of the stock price. Therefore, in order to improve the model, the first-order HMM is extended to the second-order HMM, so as to generate the final model, that is, the second-order continuous HMM, which also strengthens the model foundation for the prediction of stock price index.

In order to eliminate the result contingency caused by one experiment, 10 experiments were performed on each model, and the average value of multiple results was taken to obtain the final experimental results, as shown in Figures 4–6.

It can be seen from Figures 4–6 that RMSE and MAE obtained by the algorithm in this paper on three data sets are both smaller than those of other models, and R^2 is closer to 1, indicating that compared with other models, the stock price predicted by the algorithm in this paper is closer to the actual value and the model fits better. It is proved that the second-order continuous HMM can effectively improve the performance of stock price prediction. Among them, the prediction error of AHMMAS model is smaller than that of NA-

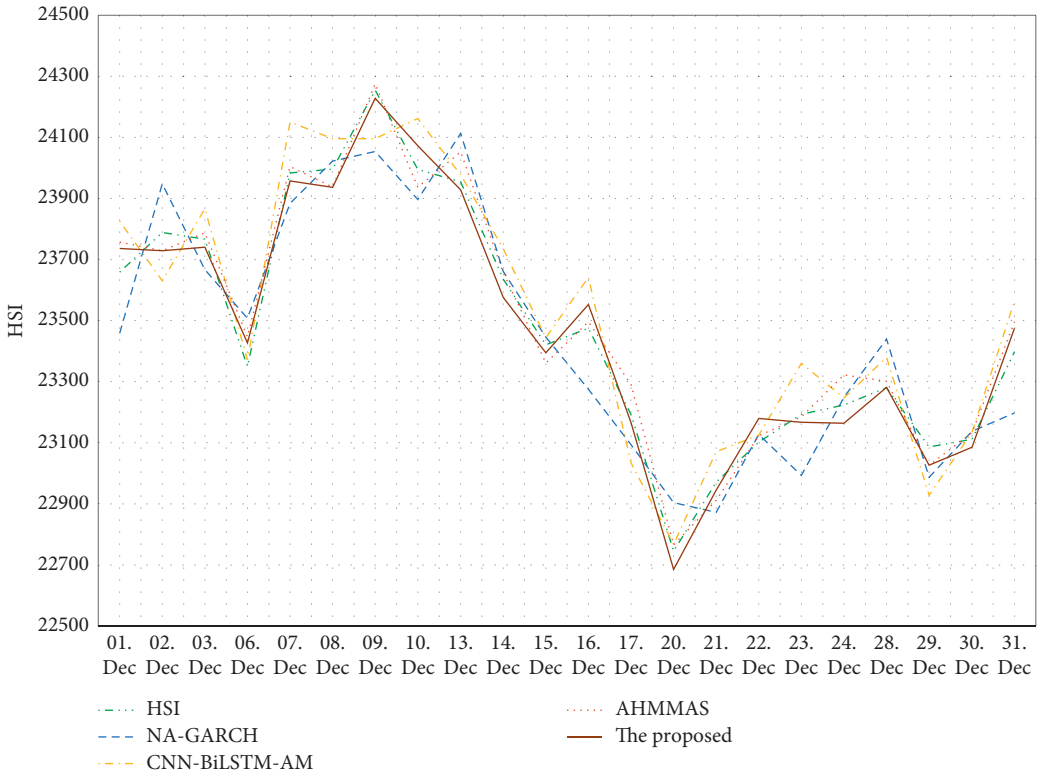


Figure 1: The predicted value and actual value of HSI of each model in 01 Dec to 31 Dec 2021.

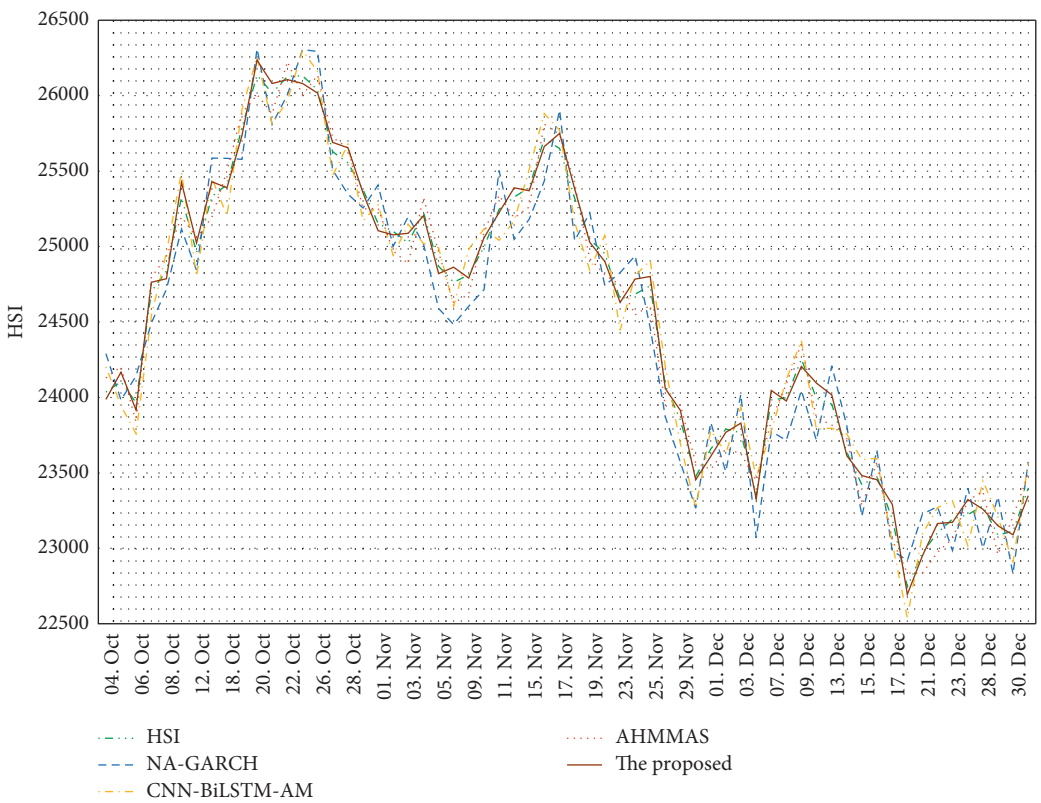


Figure 2: The predicted value and actual value of HSI of each model in 01 Oct to 31 Dec 2021.

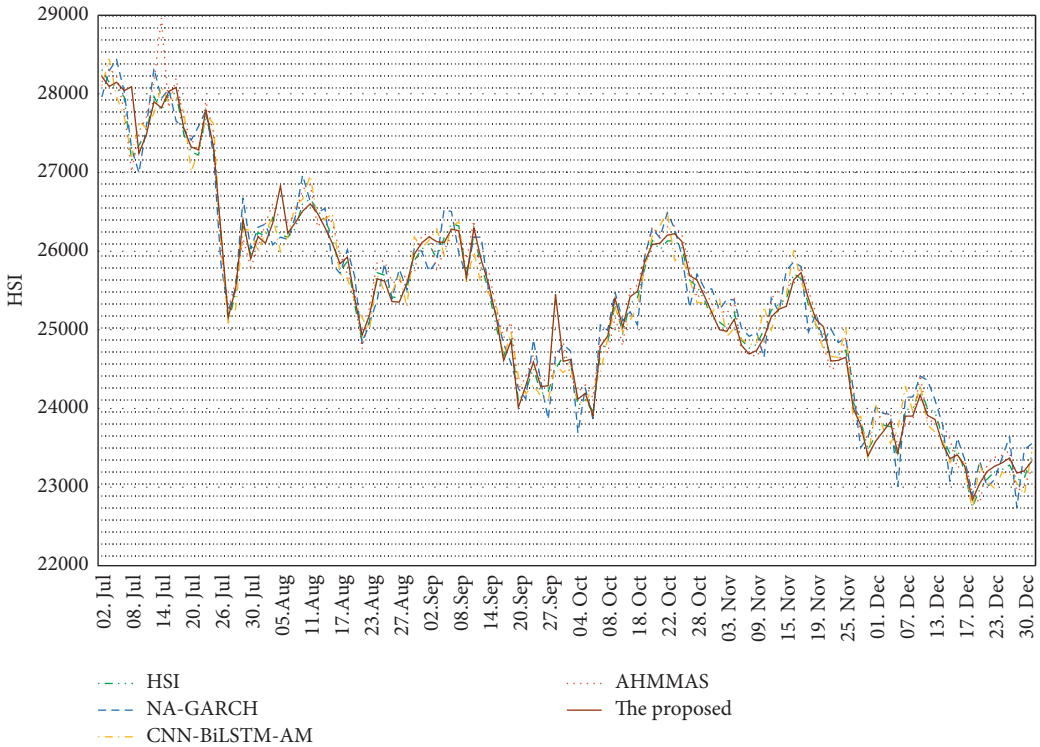


Figure 3: The predicted value and actual value of HSI of each model in 01 Jul to 31 Dec 2021.

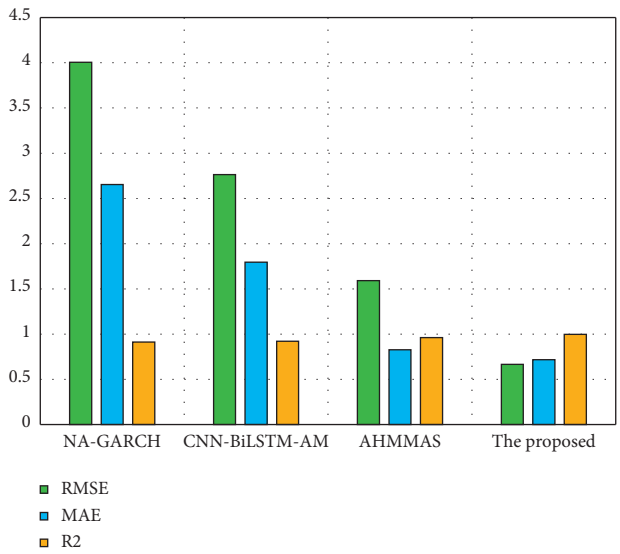


Figure 4: RMSE, MAE, and R^2 of each model from 01 Dec to 31 Dec 2021.

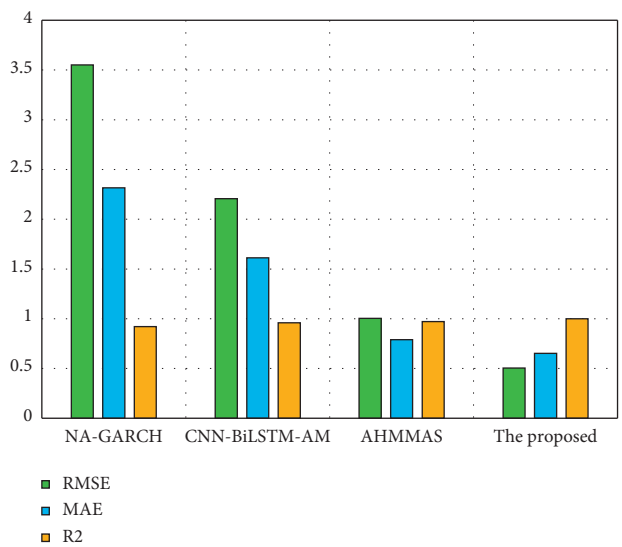


Figure 5: RMSE, MAE, and R^2 of each model from 01 Oct to 31 Dec 2021.

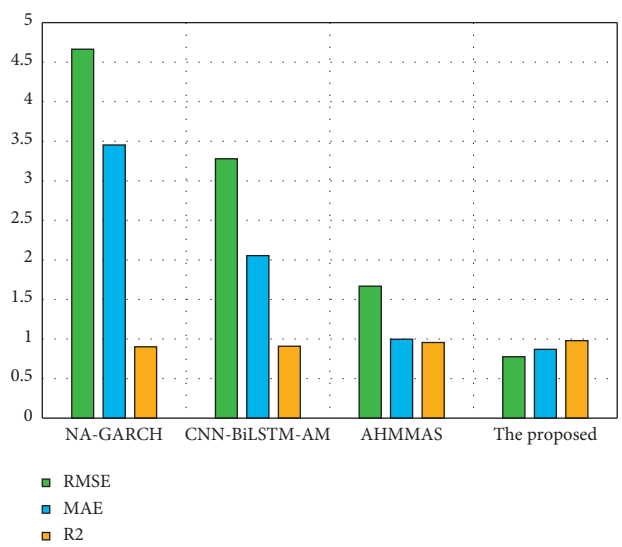


Figure 6: RMSE, MAE, and R^2 of each model from 01 Jul to 31 Dec 2021.

GARCH model and CNN-BilSTM-AM model. This is because AHMMAS extracts more effective features from the original data, so the prediction error is smaller.

5. Conclusion and Future Work

Since the stock price data have a certain Markov property in time series, at first, based on the relevant properties of hidden Markov model, this paper proposes a new efficient stock price prediction method based on HMM. Then, in order to solve the continuity problem of stock price index and the deficiency of model assumption, this paper extends discrete HMM to continuous HMM and extends the first-order continuous HMM to multiorder continuous HMM by deriving Baum–Welch algorithm. The second-order continuous HMM is used to predict the up and down trend of stock price. Finally, combined with K-means clustering algorithm and dichotomy, the stock price is successfully predicted. The evaluation results demonstrate that the proposed model is more efficient than three benchmarks.

In this paper, how to choose the initial value of HMM is not discussed, but the selection of initial value often has a great influence on the model. (1) Whether the initial value obtained according to a specific algorithm can improve the accuracy and efficiency of the model prediction in this paper remains to be further discussed. (2) This paper mainly predicts the market state and price, but it does not explain how to process empirical results and how to make investment. (3) This paper only studies the accuracy of model prediction for HSI. Without targeting individual stocks or other indexes, the adaptability of the model cannot be proved.

Data Availability

All data used to support the findings of the study are included within this paper.

Conflicts of Interest

The authors declare no conflicts of interest in this paper.

References

[1] M. T. I. Khan, S. H. Tan, L. L. Chong, and G. G. G. Goh, "Investment environment, stock market perception and stock investments after stock market crash," *International Journal of Emerging Markets*, 2021.

[2] M. Kim, E. L. Park, and S. Cho, "Stock price prediction through sentiment analysis of corporate disclosures using distributed representation," *Intelligent Data Analysis*, vol. 22, no. 6, pp. 1395–1413, 2018.

[3] J. Schroth, "Managerial compensation and stock price manipulation," *Journal of Accounting Research*, vol. 56, no. 5, pp. 1335–1381, 2018.

[4] G. Sheelapriya and R. Murugesan, "Stock price trend prediction using Bayesian regularised radial basis function network model," *Spanish Journal of Finance and Accounting/Revista Española de Financiación y Contabilidad*, vol. 46, no. 2, pp. 189–211, 2017.

[5] S. Garcia-Vega, X. Zeng, and J. Keane, "Stock returns prediction using kernel adaptive filtering within a stock market interdependence approach," *Expert Systems with Applications*, vol. 160, 2020.

[6] A. Urquhart and R. Hudson, "Efficient or adaptive markets? Evidence from major stock markets using very long run historic data," *International Review of Financial Analysis*, vol. 28, pp. 130–142, 2013.

[7] L. V. Ryan, A. K. Buchholtz, and R. W. Kolb, "New directions in corporate governance and finance: implications for business ethics research," *Business Ethics Quarterly*, vol. 20, no. 4, pp. 673–694, 2010.

[8] G. Cohen, "Investors' exposure rating and stock returns," *The Journal of Behavioral Finance*, vol. 13, no. 3, pp. 143–146, 2012.

[9] P. Arendas and B. Chovancova, "The adaptive markets hypothesis and the BRIC share markets," *Ekonomický časopis*, vol. 63, no. 10, pp. 1003–1018, 2015.

[10] B. Li, L. Chen, and H. Fan, "Non-zero total correlation means non-zero quantum correlation," *Physics Letters A*, vol. 378, no. 18-19, pp. 1249–1253, 2014.

[11] C. Tang, W. Zhu, and X. Yu, "Deep hierarchical strategy model for multi-source driven quantitative investment," *IEEE ACCESS*, vol. 7, pp. 79331–79336, 2019.

[12] A. K. Tiwari, A. B. Dar, N. Bhanja, and R. Gupta, *A Historical Analysis of the US Stock Price Index Using Empirical Mode Decomposition over 1791-2015, ECONOMICS-THE OPEN ACCESS OPEN -ASSESSMENT E-JOURNAL*, vol. 10, 2016.

[13] M. K. Khan, J. Teng, M. I. Khan, and M. F. Khan, "Stock market reaction to macroeconomic variables: an assessment with dynamic autoregressive distributed lag simulations," *International Journal of Finance & Economics*, 2021.

[14] M. Alghalith, "Estimating the stock/portfolio volatility and the volatility of volatility: a new simple method," *Econometric Reviews*, vol. 35, no. 2, pp. 257–262, 2016.

[15] J.-C. Huang, W.-T. Huang, P.-T. Chu, and W.-Y. Lee, "Applying a Markov chain for the stock pricing of a novel forecasting model," *Communications in Statistics - Theory and Methods*, vol. 46, no. 9, pp. 4388–4402, 2017.

[16] A. Harel and G. Harpaz, "Forecasting stock prices," *International Review of Economics & Finance*, vol. 73, pp. 249–256, 2021.

[17] J. Liang, W. Song, and M. Wang, "Stock price prediction based on procedural neural networks," *Advances in Artificial Neural Systems*, vol. 2011, p. 11, Article ID 814769, 2011.

[18] D. Palaz, M. Magimai-Doss, and R. Collobert, "End-to-end acoustic modeling using convolutional neural networks for HMM-based automatic speech recognition," *Speech Communication*, vol. 108, pp. 15–32, 2019.

[19] T. M. Porter and M. Hajibabaei, "Profile hidden Markov model sequence analysis can help remove putative pseudogenes from DNA barcoding and metabarcoding datasets," *BMC Bioinformatics*, vol. 22, no. 1, 2021.

[20] Y. Liu, S. Dhakal, and B. Hao, "Multimedia Image and Video Retrieval Based on an Improved HMM," *MULTIMEDIA SYSTEMS*, 2020.

[21] P. K. H. Gauba, P. P. Roy, and D. P. Dogra, "Coupled HMM-based multi-sensor data fusion for sign language recognition," *Pattern Recognition Letters*, vol. 86, pp. 1–8, 2017.

[22] Z. A. Sadik, P. M. Date, and G. Mitra, "News augmented GARCH(1,1) model for volatility prediction," *IMA Journal of Management Mathematics*, vol. 30, no. 2, pp. 165–185, 2019.

[23] J. Ni and Y. Xu, "Forecasting the Dynamic Correlation of Stock Indices Based on Deep Learning Method," *Computational Economics*, 2021.

[24] W. F. Agustini, I. R. Affianti, and E. R. M. Putri, "Stock price prediction using geometric Brownian motion," *INTERNATIONAL CONFERENCE ON MATHEMATICS: PURE, APPLIED AND COMPUTATION*, vol. 974, 2018.

[25] A. Raudys and Z. Pabarskaite, "Optimising the SMOOTHNESS and accuracy OF moving average for stock price data," *Technological and Economic Development of Economy*, vol. 34, no. 3, 2018.

[26] L. Yin, B. Li, and R. Zhang, "Research on Stock Trend Prediction Method Based on Optimized Random forest," CAAI TRANSACTIONS ON INTELLIGENCE TECHNOLOGY, 2021.

[27] Y. Wang and Y. Guo, "Forecasting method of stock market volatility in time series data based on mixed model of ARIMA and XGBoost," *CHINA COMMUNICATIONS*, vol. 17, no. 3, pp. 205–221, 2020.

[28] S. Luo, J. Huo, and Z. Dai, "Frequency-division combination forecasting of stock market based on wavelet multiresolution analysis," *Discrete Dynamics in Nature and Society*, vol. 2018, Article ID 1259156, 11 pages, 2018.

[29] X. Jiang and L. Zhang, "Stock price fluctuation prediction method based ON time series analysis," *Discrete and Continuous Dynamical Systems - Series S*, vol. 12, no. 4-5, pp. 915–927, 2019.

[30] W. Lu, J. Li, J. Wang, and L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," *Neural Computing & Applications*, vol. 33, no. 10, pp. 4741–4753, 2021.

[31] W. Chen, M. Jiang, W. Zhang, and Z. Chen, "A novel graph convolutional feature based convolutional neural network for stock trend prediction," *Information Sciences*, vol. 556, pp. 67–94, 2021.

[32] T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," *PLoS One*, vol. 14, no. 2, 2019.

[33] X. Zhang, S. Liu, and X. Zheng, "Stock price movement prediction based on a deep factorization machine and the attention mechanism," *Mathematics*, vol. 9, no. 8, 2021.

[34] S. Wu, Y. Liu, Z. Zou, and T. Weng, "S_I_LSTM: stock price prediction based on multiple data sources and sentiment analysis," *Connection Science*, 2021.

[35] M. Zhang, X. Jiang, Z. Fang, Y. Zeng, and K. Xu, "High-order Hidden Markov Model for trend prediction in financial time series," *Physica A: Statistical Mechanics and its Applications*, vol. 517, pp. 1–12, 2019.

[36] L. Chai, H. Xu, Z. Luo, and S. Li, "A multi-source heterogeneous data analytic method for future price fluctuation prediction," *Neurocomputing*, vol. 418, pp. 11–20, 2020.

[37] Y. Cao, Y. Li, S. Coleman, A. Belatreche, and T. M. McGinnity, "Adaptive hidden Markov model with anomaly states for price manipulation detection," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 2, pp. 318–330, 2015.

[38] C. Wang, K. Li, and X. He, "Network risk assessment based on Baum welch algorithm and HMM," *Mobile Networks and Applications*, vol. 26, no. 4, pp. 1630–1637, 2021.

[39] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Transactions on Cybernetics*, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2207252, 7 pages
https://doi.org/10.1155/2022/2207252

Research Article

Using Competitive Binary Particle Swarm Optimization Algorithm for Matching Sensor Ontologies

Lei Xiao,[1] **Junhong Feng**,[1] **Xishuan Niu**,[1] **and Jian-Hong Wang**[2]

[1]*School of Computer Science and Engineering, Yulin Normal University, Yulin 537000, China*
[2]*Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung 411030, Taiwan*

Correspondence should be addressed to Junhong Feng; jgxyfjh@126.com and Xishuan Niu; cookes2000@163.com

Received 14 December 2021; Accepted 26 January 2022; Published 9 February 2022

Academic Editor: Jianhui Lv

Developing sensor ontologies and using them to annotate the sensor data is a feasible way to address the data heterogeneity issue on Internet of Things (IoT). However, the heterogeneity issue exists between different sensor ontologies hampers their communications. Sensor ontology matching aims at finding all the heterogeneous entities in two ontologies, which is a feasible solution for aggregating heterogeneous sensor ontologies. This work investigates swarm intelligence (SI)-based sensor ontology matching techniques and further proposes a competitive binary particle swarm optimization algorithm (CBPSO)-based sensor ontology matching technique. In particular, a guiding matrix (GM) is proposed to ensure the population's diversity and a competitive evolutionary framework is presented. The experiment uses ontology alignment evaluation initiative (OAEI)'s benchmark and three real sensor ontologies to test CBPSO's performance. The experimental results show that the competitive evolutionary framework is able to help CBPSO effectively optimize the alignment's quality, and it significantly outperforms other SIs at 5% significant level.

1. Introduction

Since most Internet of Things (IoT) [1] devices work in diverse environments, the heterogeneity problem of underlying devices makes it difficult to provide a uniform way of representing sensor data. Developing sensor ontologies [2] and using them to annotate the sensor data is a feasible way to address these issues. Currently, various sensor ontologies, such as the CSIRO sensor ontology (CSIRO) [3], semantic sensor network ontology (SSN) [4], and MMI device ontology (MMI) [5], have been widely used in the IoT domain. These sensor ontologies own lots of overlapped information, but there also exists the heterogeneity problem between them since one sensor concept might be represented with different terminologies, granularities, or contexts. Sensor ontology matching technique is able to map the identical ontology entity pairs, which enables the processing, interpretation, and sharing of sensor data that are organized with different ontological schemes [6].

Inspired by the success of metaheuristics-based matching technique in ontology matching domain, this work further proposes a competitive binary particle swarm optimization algorithm (CBPSO) and uses it to address the sensor ontology matching problem. To better trade off the algorithm's searching performance, CBPSO uses guiding matrix (GM) to describe the solutions' distribution and guide newly generated particle's initialization; then, two competitive subpopulations are introduced to, respectively, focus on the exploitation and exploration. The contributions made in this work are as follows: (1) we present the mathematical formula for the sensor ontology matching problem; (2) we propose a CBPSO-based sensor ontology matching technique, which uses the GM to ensure the population's diversity, and two competitive subpopulations to trade off the algorithm's exploitation and exploration; (3) we employ CBPSO on ontology alignment evaluation initiative (OAEI)'s benchmark and three real sensor ontologies' matching tasks; the results reveal that CBPSO is able to determine high-quality sensor alignments.

The rest paper is organized as follows. The sensor ontology matching problem is defined in Section 3; CBPSO is presented in Section 4 in detail; the experimental results are shown in Section 5; finally, Section 6 draws the conclusion.

2. Related Work

Determining high-quality sensor ontology alignment is a challenge since there exist rich semantic relationships among sensor concepts [7]. To face this challenge, in recent years, different researchers have proposed various intelligent matching techniques. FuzzyAlign [8] integrates the semantic sensor web with a fuzzy theory-based method, and it uses evolutionary algorithm (EA) to find the optimal threshold for filtering the final alignment's quality. Differential evolution-based ontology matching (DEOM) [9] first uses the neural network to train the ontology matcher, and then, DE is used to find the correct correspondences. Co-evolutionary algorithm-based ontology matching (cEAOM) [10] uses multiple populations' trade off the algorithm's performance and improves the searching efficiency. Evolutionary tabu search algorithm (ETSA) [11] improves the algorithm's converging speed by combining the tabu search algorithm [12] as the local search strategy with EA. Due to the limitation of f-measure [13], the classic EA might optimize the solution's quality by improving recall (or precision) while sacrificing the other [14]. To overcome this drawback, multiobjective evolutionary algorithm-based ontology matching (MOEAOM) [15] is proposed to simultaneously optimize two objectives. Comparing with single-objective EA, MOEAOM is able to provide the decision makers with more nondominated options.

Besides EA, another category of metaheuristics, i.e., swarm intelligence (SI), is also applied to match the sensor ontologies. Borrowing the idea from cEA, a co-Firefly Algorithm (cFA) [16] is proposed to optimize the sensor ontology alignment's quality. It uses two subpopulations, i.e., the better subpopulation with higher-quality elite, and the worse subpopulation whose elite's fitness value is lower. These two subpopulations, respectively, use the exploitation and exploration, and when the worse subpopulation's elite's quality outperforms the better subpopulation, they adaptively switch the searching strategy. More recently, particle swarm optimization (PSO) [17] is also presented to aggregate different sensor ontology matchers, which borrows the idea from ETSA and introduces the simulated annealing (SA) [18] to execute the local search process.

3. Sensor Ontology Matching Problem

A sensor ontology O is a 3-tuple (C, P, R), where C, P, and R, respectively, represent the sets of concept, data-type property, and relationship between two concepts [19]. Matching sensor ontologies aim at finding the sensor ontology alignment, which is a set of sensor entity correspondences. Each correspondence is a 4-tuple $(e_1, e_2, \text{rel}, \text{conf})$, where e_1 and e_2 are, respectively, two sensor ontologies' entities, rel is two entities' relationship, and conf is a real number in [0, 1] that denotes to what extent the correspondence holds [20]. Often, conf is measured by the similarity value sim Value between two entities. If sim Value = 1, two entities are the same, and if sim Value = 0, they are different.

There are three categories of ontology matcher for measuring the entities' similarity, which are, respectively, based on string, linguistic, and ontology structure [21]. This work uses a hybrid similarity measure that is proposed by Xue and Wang [22], which combines three kinds of ontology matchers to improve the result's confidence. Given an alignment A, the more correspondences it has $(f_1(A) = \text{normlize}(|A|))$ and the higher mean similarity value it owns $(f_2(A) = \sum \text{sim}_i/|A|)$, the better quality it has. On this basis, we use the following metric to evaluate the quality of [23]

$$f(A) = \frac{2 \times f_1(A) \times f_2(A)}{f_1(A) + f_2(A)}. \tag{1}$$

To optimize the alignment's quality, the sensor ontology matching problem is defined as follows:

$$\begin{cases} \max F(M), \\ \text{s.t.} \quad M = \left[m_{i,j}\right]_{|O_1| \times |O_2|}, \\ \quad m_{i,j} \in \{0, 1\}, \end{cases} \tag{2}$$

where M is the matching matrix, $m_{i,j} = 1$ means the ith entity in ontology O_1 is mapped with the jth entity in ontology O_2, and $F(M)$ use the function $f()$ to measure M's corresponding alignment's quality.

4. Competitive Binary Particle Swarm Optimization Algorithm

For trading off the algorithm's exploitation and exploration, CBPSO uses two subpopulations, where the one with better (worse) global elite particle is called the better (worse) subpopulation. In this work, the better subpopulation mainly focuses on the exploitation, and the worse one tends to explore the unknown domain.

4.1. Initialize Particle. CFPA uses the binary encoding mechanism to encode a particle, which is denoted by a 0-1 matrix, whose row and column are, respectively, two entity sets, and its element 1 means two corresponding entities are mapped, while 0 means not. When the scale of entities is huge, this matrix is a sparse one, i.e., most of its elements are 0. Therefore, we need to uniformly assign the elements with value 1 in each subpopulation's matrices on the whole. To this end, we introduce a guiding matrix, which is of help to ensure the diversity of each subpopulation. For the sake of clarity, given a guiding matrix GM and a matching matrix MM [24], we show the pseudocode of initializing a particle in the subpopulation in Algorithm 1.

In Algorithm 1, we initialize a new particle according to GM, whose elements represent the distribution of current subpopulation's particle positions. The larger value of GM's element is, the more likely this dimension has been explored by the existing particles, whose probability of being 1 in newly generated particle should be smaller. In this way, we are able to ensure the diversity of subpopulation to the maximum extent.

```
if (GM is not given)
    set all GM's elements as zero;
end if
for int j = 0 to rand × |O1| × |O2|
    [(m,n), (m',n')] = randomly select two elements from MM;
    if (GM_{m,n} < GM_{m',n'})
        MM_{m,n} = 1;
        GM_{m,n} = GM_{m,n} + 1;
    else
        MM_{m',n'} = 1;
        GM_{m',n'} = GM_{m',n'} + 1;
    end if
end for
return MM
```

Algorithm 1: Initialization.

4.2. Update Particle. In each iteration, particle MM is updated according to the following equations:

$$\mathrm{MM}^{t+1} = \omega_t \mathrm{MM}^t + c_1^t\left(\mathrm{MM}^*_{\mathrm{local}} - \mathrm{MM}^t\right) + c_2^t\left(\mathrm{MM}^*_{\mathrm{global}} - \mathrm{MM}^t\right), \tag{3}$$

$$s\left(\mathrm{MM}^{t+1}_{i,j}\right) = \frac{1}{1 + e^{-\mathrm{MM}^{t+1}_{i,j}}}, \tag{4}$$

$$\mathrm{MM}^{t+1}_{i,j} = \begin{cases} 1, & \text{if rand} \le s\left(\mathrm{MM}^{t+1}_{i,j}\right), \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where $\mathrm{MM}^*_{\mathrm{local}}$ and $\mathrm{MM}^*_{\mathrm{global}}$ are, respectively, the particle's local optima and its subpopulation's global optima and ω_t is the inertia weight in tth generation, which is determined by the following equation:

$$\omega_t = \omega_{\mathrm{start}} + \left(\omega_{\mathrm{start}} - \omega_{\mathrm{end}}\right)\left(1 - \frac{t}{\max T}\right). \tag{6}$$

c_1^t and c_2^t are two learning factors, whose values are, respectively, calculated by the following equations:

$$c_1^t = c_{1,\mathrm{start}} + \left(c_{1,\mathrm{end}} - c_{1,\mathrm{start}}\right)\frac{t}{\max T}, \tag{7}$$

$$c_2^t = c_{2,\mathrm{start}} + \left(c_{2,\mathrm{end}} - c_{2,\mathrm{start}}\right)\frac{t}{\max T}, \tag{8}$$

where $\omega_{\mathrm{start}} = 0.9$, $\omega_{\mathrm{end}=0.4}$, $c_{1,\mathrm{start}} = 2.5$, $c_{1,\mathrm{end}} = 0.5$, $c_{1,\mathrm{start}} = 0.5$, $c_{2,\mathrm{end}} = 2.5$, and $\max T$ is the maximum generation. Here, we use different annealing strategies on ω_t, c_1^t, and c_2^t to trade off the algorithm's exploration and exploitation. All these parameters are configured empirically, which enables the algorithm determine the highest f-measure in average on all testing cases.

4.3. The Pseudocode of Competitive Binary Particle Swarm Optimization Algorithm. The pseudocode of CBPSO is presented in Algorithm 2. The initialization of CBPSO consists of three steps: (1) two subpopulations P_{better} and P_{worse}'s particles are initialized according to Algorithm 1, (2) P_{better} and P_{worse}'s global optimal particles are initialized by selecting the particles with the best objective function values, and (3) each particle's local optima is initialized with newly generated particle. During each iteration, each particle is first updated according to equations (3)–(8); then, their corresponding local optima and each subpopulation's global optima are updated accordingly. After that, P_{better} and P_{worse} are competed through their local optimal particles, and if P_{worse} wins, we will switch P_{better} and P_{worse}. After the competition, two subpopulations' guiding matrices $\mathrm{GM}_{\mathrm{better}}$ and $\mathrm{GM}_{\mathrm{worse}}$ are calculated by summing all the subpopulation's particles' matching matrices together. Finally, we use $\mathrm{GM}_{\mathrm{better}}$ and $\mathrm{GM}_{\mathrm{worse}}$ to, respectively, reinitialize 10% and 50% particles in P_{better} and P_{worse} according to Algorithm 1. When reaching the maximum iteration number $maxT = 3000$, the algorithm terminates and returns $\mathrm{MM}^*_{\mathrm{global,better}}$.

5. Experiment

5.1. Experimental Setup. In this work, ontology alignment evaluation initiative (OAEI)'s benchmark and three real sensor ontologies are used to test CBPSO's performance. In Tables 1–3, we compare CBPSO with five SI-based sensor ontology matching techniques, i.e., differential evolution algorithm (DE) [9], co-evolutionary algorithm (cEA) [10], evolutionary tabu search algorithm (ETSA) [11], co-firefly algorithm (cFA) [16], and simulated annealing particle swarm optimization (SA-PSO) [18] in terms of recall, precision, and f-measure [25], respectively. Their results are the average of thirty independent runs. Table 4 briefly describes OAEI's benchmark and three real sensor ontologies.

5.2. Statistical Experiment. We utilize statistical testing method T-test [26] to compare different competitors' performance in terms of recall, precision, and f-measure, respectively. Tables 1 and 2 show six SI-based sensor ontology matching techniques' mean, recall, precision, and f-measure and the corresponding standard deviation on all the testing cases, and Table 3 presents the t-value on recall, precision, and f-measure.

```
*** Initialization **
Initialize two subpopulations P_better and P_worse;
Initialize P_better and P_worse's global optimas, MM*_global,better and MM*_global,worse;
Initialize each particle's local optima MM*_local;
generation t = 0;
while t < max T
    update each particle according to equations (3)–(8);
    update each particle's local optima MM*_local;
    update MM*_global,better and MM*_global,worse;
    *** Competition **
    [winner, loser] = compete(MM*_global,better, MM*_global,worse);
    if(winner == MM*_global,worse)
        switch (P_better, P_worse);
    end if
    update P_better's guiding matrix GM_better;
    update P_worse's guiding matrix GM_worse;
    *** Trade off Exploitation and Exploration **
    use GM_better to randomly reinitialize 10% particles in P_better;
    use GM_worse to randomly reinitialize 50% particles in P_worse;
    t = t + 1;
end while
return MM*_global,better
```

Algorithm 2: Competitive binary particle swarm optimization algorithm.

Table 1: Comparison among swarm intelligence-based sensor ontology matching techniques in terms of alignment's quality. 1XX, 2XX, and 3XX refer to those starting with number 1, 2, and 3, respectively, and the symbols f, r, and p, respectively, stand for f-measure, recall, and precision.

Testing case	cEA $f(r, p)$	ETSA $f(r, p)$	cFA $f(r, p)$	DE $f(r, p)$	SA-PSO $f(r, p)$	CBPSO $f(r, p)$
1XX	1.00 (1.00, 1.00)	0.81 (0.72, 0.90)	0.95 (0.94, 0.95)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
2XX	0.92 (0.91, 0.93)	0.72 (0.84, 0.61)	0.82 (0.79, 0.86)	0.91 (0.88, 0.94)	0.73 (0.72,0.74)	0.96 (0.96, 0.96)
3XX	0.78 (0.81, 0.74)	0.42 (0.37, 0.48)	0.71 (0.67, 0.77)	0.89 (0.85, 0.94)	0.82 (0.85, 0.79)	0.93 (0.91, 0.95)
MMI-SSN	0.90 (0.86, 0.94)	0.92 (0.90, 0.95)	0.92 (0.90, 0.95)	0.94 (0.92, 0.95)	0.88 (0.90, 0.87)	0.96 (0.94, 0.97)
CSIRO-SSN	0.92(0.89, 0.95)	0.94 (0.94, 0.95)	0.94 (0.94, 0.94)	0.94 (0.94, 0.95)	0.90 (0.88, 0.93)	0.96 (0.95, 0.96)
MMI-CSIRO	0.86 (0.88, 0.91)	0.90 (0.87, 0.94)	0.90 (0.87, 0.94)	0.92 (0.90, 0.93)	0.90 (0.87,0.94)	0.93 (0.94, 0.93)
Average	0.89 (0.88, 0.92)	0.78 (0.77, 0.80)	0.87 (0.85, 0.90)	0.93 (0.91, 0.95)	0.87 (0.87, 0.87)	0.96 (0.95, 0.96)

Table 2: Comparison in terms of standard deviation. 1XX, 2XX, and 3XX refer to those starting with number 1, 2, and 3, respectively, and the symbols f_d, r_d, and p_d, respectively, stand for the standard deviation of f-measure, recall, and precision.

Testing case	cEA $f_d(r_d, p_d)$	ETSA $f_d(r_d, p_d)$	cFA $f_d(r_d, p_d)$	DE $f_d(r_d, p_d)$	SA-PSO $f_d(r_d, p_d)$	CBPSO $f_d(r_d, p_d)$
1XX	0.01 (0.01, 0.01)	0.03 (0.02, 0.02)	0.01 (0.03, 0.02)	0.01 (0.01, 0.01)	0.01 (0.01, 0.01)	0.01 (0.01, 0.01)
2XX	0.01 (0.02, 0.01)	0.04 (0.02, 0.04)	0.02 (0.01, 0.03)	0.02 (0.02, 0.02)	0.03 (0.02, 0.03)	0.01 (0.02, 0.01)
3XX	0.02 (0.01, 0.03)	0.05 (0.02, 0.04)	0.02 (0.01, 0.03)	0.01 (0.02, 0.01)	0.02 (0.02, 0.01)	0.02 (0.01, 0.02)
MMI-SSN	0.01 (0.02, 0.02)	0.03 (0.03, 0.02)	0.03 (0.03, 0.02)	0.02 (0.03, 0.01)	0.04 (0.03, 0.02)	0.01 (0.01, 0.01)
CSIRO-SSN	0.03 (0.01, 0.01)	0.03 (0.02, 0.01)	0.03 (0.02, 0.02)	0.05 (0.03, 0.02)	0.03 (0.02, 0.02)	0.02 (0.01, 0.02)
MMI-CSIRO	0.03 (0.02, 0.01)	0.03 (0.01, 0.01)	0.02 (0.02, 0.03)	0.01 (0.01, 0.01)	0.04 (0.02, 0.04)	0.02 (0.02, 0.01)

In Tables 1 and 2, alignments determined by CBPSO are much better than other sensor ontology matching technique. Due to subpopulations' competitive mechanism, CBPSO is able to better trade off the algorithm's exploitation and exploration, which ensures not only the solution's quality but also the algorithm's stability. In Table 3, the degree of freedom is 61 (two samples' scales are both 33), the significant level is 5%, and except those results that are the same, the p-value of the rest are all smaller than 0.005, which show that CBPSO significantly outperforms other competitors on 5% significant level.

Table 3: T-test's t-value. 1XX, 2XX, and 3XX refer to those starting with number 1, 2, and 3, respectively, and the symbols f_t, r_t, and p_t, respectively stand for the t-value on f-measure, recall, and precision.

Testing case	(cEA, CBPSO) $f_t\,(r_t, p_t)$	(ETSA, CBPSO) $f_t\,(r_t, p_t)$	(cFA, CBPSO) $f_t\,(r_t, p_t)$	(DE, CBPSO) $f_t\,(r_t, p_t)$	(SA-PSO, CBPSO) $f_t\,(r_t, p_t)$
1XX	0 (0, 0)	−180.24 (−375.65, −134.16)	−106.06 (−56.92, −67.08)	0 (0 0)	0 (0 0)
2XX	−84.85 (−53.03, −63.63)	−174.62 (−127.27, −254.66)	−187.82 (−228.07, −94.86)	−67.08 (−84.85, −26.83)	−218.19 (−254.55, −208.71)
3XX	−159.09 (−212.13, −174.73)	−284.11 (−724.4,8, −315.28)	−233.34 (−509.11, −149.76)	−53.66 (−80.49, −13.41)	−116.67 (−80.49, −214.66)
MMI-SSN	−127.27 (−107.33, −40.24)	−37.94 (−37.9,4, −26.832)	−37.94 (−37.94, −26.83)	−26.83 (−18.97, −42.42)	−58.20 (−37.94, −134.16)
CSIRO-SSN	−33.28 (−127.27, −13.41)	−16.64 (−13.41, −13.41)	−16.64 (−13.41, −21.21)	−11.14 (−9.48, −10.60)	−49.92 (−93.91, −31.81)
MMI-CSIRO	0 (0, 0)	−180.24 (−375.65, −134.16)	−106.06 (−56.92, −67.08)	0 (0, 0)	0 (0, 0)

Table 4: Descriptions on the ontologies in the testing cases.

Testing case	Ontology	Scale
OAEI's benchmark	Bibliographic ontology	97 entities
Real sensor ontology	CSIRO sensor ontology (CSIRO)	33,205 entities
	Semantic sensor network ontology (SSN)	32,298 entities
	MMI device ontology (MMI)	24,034 entities

6. Conclusion

Building sensor ontologies and using them to annotate the sensor data is a feasible way to address the IoT data heterogeneity issue. However, before using these ontologies, we need to address the heterogeneity problem among them. Recently, SI-based matching techniques have been emerging as a popular method of matching heterogeneous sensor ontologies, and this work proposes a CBPSO-based sensor ontology matching technique. In particular, CBPSO uses the GM to ensure the diversity of the population, and two competitive subpopulations to trade off the algorithm's exploitation and exploration. The experiment compares CBPSO with five SI-based sensor ontology matching techniques, and the experimental results show that CBPSO's results outperform the competitors.

In the future, we will further improve CBPSO by introducing more subpopulations with different searching strategies. In addition, the competing mechanism can be further improved by referring to binary fish migration optimization algorithm (BFMO) [27]. We are also interested in further improving CBPSO to match biomedical ontologies [28], which is the large-scale matching task. Some efficiency-improving strategies will be used, such as the compact encoding, and the efficient compact evolutionary paradigm will be one of our future studies.

Data Availability

The data used to support this study can be found at http://oaei.ontologymatching.org.

Conflicts of Interest

The authors declare that they have no conflicts of interest in the work.

Acknowledgments

This work was supported in part by the Scientific Research and Technology Development Project of Yulin City (no. 20204012), Natural Science Foundation of Guangxi Province (no. 2021GXNSFAA220076), and Science Research Foundation for High-level talents of Yulin Normal University (no. G2021ZK17).

References

[1] P. Pande and A. Padwalkar, "Internet of things–A future of Internet: a survey," *International Journal of Advance Research in Computer Science and Management Studies*, vol. 2, pp. 354–361, 2014.

[2] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001.

[3] H. Neuhaus and M. Compton, "The semantic sensor network ontology," in *Proceedings of the AGILE Workshop on Challenges in Geospatial Data Harmonisation*, pp. 1–33, Hanover, Germany, June 2009.

[4] C. Rueda, N. Galbraith, R. A. Morris, L. Bermudez, R. Arko, and J. Graybeal, "The MMI device ontology: enabling sensor integration," in *Proceedings of the AGU Fall Meeting Abstracts*, pp. 1–35, San Francisco, CA, USA, December 2010.

[5] M. Compton, P. Barnaghi, L. Bermudez et al., "The SSN ontology of the W3C semantic sensor network incubator group," *Journal of Web Semantics*, vol. 17, pp. 25–32, 2012.

[6] P. Shvaiko and J. Euzenat, "Ontology matching: state of the art and future challenges," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 1, pp. 158–176, 2011.

[7] H. Jiang, Q. Li, Y. Jiang et al., "When machine learning meets congestion control: a survey and comparison," *Computer Networks*, vol. 192, pp. 1–28, 2021.

[8] S. Fernandez, I. Marsa-Maestre, J. Velasco, and B. Alarcos, "Ontology alignment architecture for semantic sensor web integration," *Sensors*, vol. 13, no. 9, Article ID 12581, 2013.

[9] X. Xue and C. Jiang, "Matching sensor ontologies with multi-context similarity measure and parallel compact differential evolution algorithm," *IEEE Sensors Journal*, vol. 21, no. 21, Article ID 24570, 2021.

[10] X. Xue and J.-S. Pan, "A compact Co-evolutionary algorithm for sensor ontology meta-matching," *Knowledge and Information Systems*, vol. 56, no. 2, pp. 335–353, 2018.

[11] X. Xue and J. Chen, "Using compact evolutionary tabu search algorithm for matching sensor ontologies," *Swarm and Evolutionary Computation*, vol. 48, pp. 25–30, 2019.

[12] D. de Werra and A. Hertz, "Tabu search techniques," *Operations-Research-Spektrum*, vol. 11, no. 3, pp. 131–141, 1989.

[13] A. Maratea, A. Petrosino, and M. Manzo, "Adjusted F-measure and kernel scaling for imbalanced data learning," *Information Sciences*, vol. 257, pp. 331–341, 2014.

[14] X. Xue and Y. Wang, "Optimizing ontology alignments through a memetic algorithm using both MatchFmeasure and unanimous improvement ratio," *Artificial Intelligence*, vol. 223, pp. 65–81, 2015.

[15] X. Xue, X. Wu, C. Jiang, G. Mao, and H. Zhu, "Integrating sensor ontologies with global and local alignment extractions," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6625184, 10 pages, 2021.

[16] X. Xue, X. Wu, and J. Chen, "Optimizing biomedical ontology alignment through a compact multiobjective particle swarm optimization algorithm driven by knee solution," *Discrete Dynamics in Nature and Society*, vol. 2020, no. 7, 10 pages, Article ID 4716286, 2020.

[17] H. Zhu, X. Xue, A. Geng, and H. Ren, "Matching sensor ontologies with simulated annealing particle swarm optimization," *Mobile Information Systems*, vol. 2021, Article ID 5510055, 11 pages, 2021.

[18] D. Bertsimas and J. Tsitsiklis, "Simulated annealing," *Statistical Science*, vol. 8, no. 1, pp. 10–15, 1993.

[19] X. Xue and J. Zhang, "Matching large-scale biomedical ontologies with central concept based partitioning algorithm and

adaptive compact evolutionary algorithm," *Applied Soft Computing*, vol. 106, pp. 1–11, 2021.

[20] X. Xue and J. Chen, "Matching biomedical ontologies through compact differential evolution algorithm with compact adaption schemes on control parameters," *Neurocomputing*, vol. 458, pp. 526–534, 2021.

[21] J. Martinez-Gil and J. F. Aldana-Montes, "Evaluation of two heuristic approaches to solve the ontology meta-matching problem," *Knowledge and Information Systems*, vol. 26, no. 2, pp. 225–247, 2011.

[22] X. Xue and Y. Wang, "Using memetic algorithm for instance coreference resolution," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 580–591, 2016.

[23] X. Xue and Q. Huang, "Generative adversarial learning for optimizing ontology alignment," *Expert Systems*, vol. 2022, pp. 1–12, 2022.

[24] X. Xue, P.-W. Tsai, and Y. Zhuang, "Matching biomedical ontologies through adaptive multi-modal multi-objective evolutionary algorithm," *Biology*, vol. 10, pp. 1–16, 2021.

[25] G. Hripcsak and A. Rothschild, "Agreement, the f-measure, and reliability in information retrieval," *Journal of the American Medical Informatics Association*, vol. 12, no. 3, pp. 296–298, 2005.

[26] D. Semenick, "Tests and measurements: the T-test," *National Strength & Conditioning Association Journal*, vol. 12, no. 1, pp. 36-37, 1990.

[27] J. S. Pan, P. Hu, and S. C. Chu, "Binary fish migration optimization for solving unit commitment," *Energy*, vol. 226, pp. 1–11, 2021.

[28] X. Xue, J. Lu, and J. Chen, "Using NSGA-III for optimising biomedical ontology alignment," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 3, pp. 135–141, 2019.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8344814, 10 pages
https://doi.org/10.1155/2022/8344814

Research Article

AI-Based Heterogenous Large-Scale English Translation Strategy

Chuncheng Wang

Tongling University, Tongling 244061, China

Correspondence should be addressed to Chuncheng Wang; chad825@tlu.edu.cn

Received 3 January 2022; Revised 21 January 2022; Accepted 26 January 2022; Published 9 February 2022

Academic Editor: Jianhui Lv

English has become one of the most widely used languages in the world. If there is no good translation mechanism for such a widely used language, it will bring trouble to both study and life. At present, the world's major platforms are committed to the study of English translation strategies. There are translation platforms from different regions and different translation mechanisms. These translation data from different translation platforms have the characteristics of large-scale, multisource, heterogeneity, high dimensions, and poor quality. However, such inconsistent translation data will increase the translation difficulty and translation time. Therefore, it is necessary to improve the quality of translation data to achieve a better translation effect. How to provide a large-scale and efficient translation strategy needs to integrate the translation strategies of various platforms to perform heterogeneous translation data cleaning and fusion based on machine learning. At first, this paper represents the multisource, heterogeneous translation data model as tree-augmented naive Bayes networks (TANs) and naturally captures the relationship between the datasets through the learning of TANs structure and the probability distribution of input attributes and tuples, using data probability value to complete the classification of translation data cleaning. Then, a multisource, heterogeneous translation data fusion model based on recurrent neural network (RNN) is constructed, and RNN is used to control the node data of hidden layer to enhance the fault-tolerant ability in the fusion process and complete the construction of fusion model. Finally, experimental results show that TANs-based translation data cleaning method can effectively improve the cleaning rate with an average improvement of approximately 10% and cleaning time with an average reduce about 5%. In addition, RNN-based multisource translation data fusion method improves the shortcomings of the traditional fusion model and improves the practicability of the fusion model in terms of root mean square error (RMSE), mean absolute percentage error (MAPE), fusion time, and integrity.

1. Introduction

The importance of translation is seldom paid attention to in English learning. However, with the progress and development of society, the demand for English translation talents is increasing, making translation gradually receive attention in teaching [1, 2]. Translation is conducive to exercise students' divergent thinking, promote the comprehensive development of English level, and let students fully understand foreign culture. It is very important to cultivate students' translation ability, which is not only conducive to improve students' comprehensive English quality, but also enhance their ability and level of listening, speaking, reading, and writing. It can also exercise students' thinking ability and truly translate works that can be accepted universally. In the process of translation, students can fully understand foreign knowledge and foreign culture [3–5]. Based on this, various translation platforms emerge one after another. However, there is no large-scale English translation strategy [6]. The so-called large-scale refers to the text translation data from different regions, different platforms, and different translation mechanisms [7]. These data have large-scale, multisource, heterogeneous types and modes, high dimensions, and poor quality. Such data inconsistency will increase the difficulty of translation and the translation time. Therefore, it is necessary to improve the quality of data to achieve better translation effect. But how to provide a large-scale and efficient translation strategy? Data cleaning and data fusion are needed to integrate the translation strategies of each region.

Datasets inevitably exist as redundant data, missing data, uncertain data, and inconsistent data, and these data is called "dirty data" [8, 9]. The purpose of data cleaning is to detect the incorrect and inconsistent data in the English translation dataset and then delete or correct them so as to improve the quality of the translation data. Multisource, heterogeneous translation data cleaning refers to removing noise data and irrelevant data from datasets of different translation platforms, processing omitted data, and removing white noise in blank data fields and knowledge background. In multisource, heterogeneous translation environment, traditional data cleaning methods are generally divided into two ways: (1) Data is fused into the same data source through data integration or data fusion, and data cleaning is carried out during and after data fusion. (2) Through the development of unified data cleaning standards, each data source carries out inaccurate data cleaning at the same time. The latter method is more difficult to clean, and the cleaning efficiency is uncertain, but if the cleaning model is proper, the cleaning efficiency is greatly improved compared to the first method.

The successful application of machine learning (ML) [10] in pattern recognition, information retrieval, data mining, and other fields provides a new solution for data processing, but the traditional statistical ML method cannot be simply used to process multisource, heterogeneous data because traditional statistical ML methods assume that the data to be processed comes from the same feature space and has the same distribution. Classification and clustering are two main problems in ML [11], and the typical solution is to use Bayesian network. Bayesian network [12] is a method to obtain a priori probability from a specific domain by using graphic patterns, and Bayesian network is suitable for processing incomplete data. For processing instances with missing values, all possible attribute values can be accumulated or integrated. In addition, Bayesian network has strong robustness for the overfitting problem of the model. Based on the advantages of Bayesian network in data processing, this paper considers Bayesian network to process uncertain translation data. Considering the characteristics of large-scale, multisource, heterogeneous translation data, and the advantages of Bayesian network in processing inconsistent data, a multilayer reduction model based on Tree-Augmented Naive Bayes Networks (TANs) [13] is constructed, which includes data source reduction, data attribute layer and tuple layer reduction, and duplicated data cleaning.

Accordingly, the main contributions of this study can be summarized as follows. (i) TANs-based translation data cleaning method is proposed. (ii) The RNN-controlled translation data fusion method is proposed.

The rest of this paper is organized as follows. Section 2 reviews related work. In Section 3, TANs-based heterogeneous translation data cleaning method is presented. In Section 4, RNN-controlled heterogeneous translation data fusion is proposed. Experimental results are presented in Section 5. Section 6 concludes this paper.

2. Related Work

In the process of data acquisition, based on the comprehensiveness of data collection and the integrity of relevant data, data collection usually involves multiple data sources, including a variety of databases, file systems and service interfaces, resulting in complex data types and large data scale. Therefore, it is necessary for data cleaning after data acquisition. In [14], a novel data cleaning technique was introduced to remove dirty data effectively, and based on optimization, a new hybrid firefly update enabled rider optimization algorithm was proposed. In [15], a hybrid data cleaning system that integrated bias detection and repair was proposed to process multiple bias types. In [16], a federated data cleaning protocol was presented to realize data cleaning without damaging data privacy in edge intelligence. In [17], a privacy-aware data cleaning as a service model was proposed, which promoted the interaction with the parties requesting data query from the client, as well as the service provider using the data pricing scheme, which calculated the price according to the data sensitivity. In [18], a set of conditional functional dependencies was introduced into the density-based data cleaning algorithm, and the algorithm used the set to repair inconsistent data. Internet of things is a multisource information integration technology, which collects multisource data every day, but most of the data may be irrelevant and redundant. In [19], a deep Q-network-based feature selection method was proposed for multisource data cleaning. In clinical research, the abundance of data resources also brings the challenge of data cleaning; some data cleaning methods had been proposed [20–22].

Multisource, heterogeneous data refers to data with multiple sources and different component characteristics. The fusion of multisource, heterogeneous data can assist researchers to obtain effective information in data. Therefore, multisource, heterogeneous data fusion has become a hot topic in current research. In [23], a multiscale deep coupled neural networks was proposed to fully fuse the fault information. In [24], the author presented a multisource, heterogeneous data fusion method for narrowband Internet of things based on perceptual semantics. In [25], a novel and dynamic opportunistic clustering and data fusion scheme based on self-organizing hesitation fuzzy entropy was presented to overcome energy consumption and the bottleneck of network lifetime. In [26], the authors had made a comprehensive research on the data fusion method based on ML. Data fusion also plays a pivotal role in providing environmental information. In [27], the authors listed data fusion algorithms provided by the general data fusion framework and illustrated the method of the general data fusion framework by an example of three-dimensional reconstruction from two-dimensional images. In [28], the authors analyzed the detection process of massage chair intelligent detection robot and made theoretical research from decision-level fusion and data-level fusion. Data generated by sensors in the Internet of Things always is large-scale, multisource, and heterogeneous, so data fusion is necessary for providing intelligent services [29–32].

Translation data cleaning and fusion are pivotal steps for the following data processing, so this paper pays attention to translation data cleaning and fusion in heterogeneous large-scale English translation.

3. TANs-Based Translation Data Cleaning

The main method to classify multisource, heterogeneous data is to establish and examine the multisource, heterogeneous data network model formed by data relationship. TANs model is a method of knowledge representation, learning, and reasoning on the structure and relationship between data based on probability framework, which can better describe the uncertainty of translation data. Due to the dependence between various attributes in the translation data source, this paper represents the multisource, heterogeneous translation data model as TANs and naturally captures the relationship between the datasets through the learning of TANs structure and the probability distribution of input attributes and tuples.

The basic idea of translation data cleaning based on TANs is as follows: according to different eigenvectors of translation data, the translation data attributes to be cleaned are divided into different classes to form multiple Bayesian network structures. A parent node is made up of several child nodes. The translation data generated by different data objects in Bayesian networks in different Bayesian network structures indicate a data object according to the importance of data attributes and tuples in the network. Therefore, the Bayesian network formed by data objects in one data domain becomes a comparable quantity with the Bayesian network formed by data objects in another data domain. Divide all translation data generated by data objects containing the same task into a dataset. Finally, the associated data objects of different data sources are combined into a dataset to be cleaned using TANs to represent data attributes.

Suppose a dataset $D = \{T_1, T_2, \ldots, T_n, C\}$, C is a class variable and its value ranges are $\{c_1, c_2, \ldots, c_m\}$, m is the total number of classes, $\{t_1, t_2, \ldots, t_n\}$ is the attribute value of $\{T_1, T_2, \ldots, T_n\}$ that represents classification feature, and n is the number of attributes of the classification. TANs classifier is assumed by attribute nodes $\{T_1, T_2, \ldots, T_n\}$. The structure of TANs network composed of attribute nodes $\{t_1, t_2, \ldots, t_n\}$ is a tree, and each attribute variable has no more than one attribute parent node besides the parent class. Each attribute variable has no more than one attribute parent node except the parent class, and a tree is formed between the attribute nodes as the maximum weight span tree.

For probability distribution $P(T_1, T_2, \ldots, T_n, C)$, and we have

$$\arg\max\left\{P(C)\prod_{i=1}^{n}P(t_i)|\delta_{t_i}, C, WS_T\right\}. \tag{1}$$

The classifier for predicting variable C is TANs classifier, where WS_T represents the maximum weight span tree of $T_1, T_2, \ldots, T_n$ under the constraint of class variable C, and δ_{t_i} is the value of the attribute parent $\prod(t_i)$ of t_i in the maximum weight span tree.

The TANs attribute tree can be characterized by the function δ: $\{1, 2, \ldots, n\} \longrightarrow \{0, 1, \ldots, n\}$, where the node $\pi(i) = 0$ is the parent node. There is no sequence $\{i_1, i_2, \ldots, i_k\}$ to make $\delta(i_j) = i_j + 1$, where $i \le j < k$, and $\delta(i_k) = i_1$; that is, no undirected loop can be generated. When $\delta(i) > 0$, $\prod t_i = \{t_{\delta(i)}\}$. When $\delta(i) = 0$, $\prod t_i = \varnothing$. Therefore, the function δ defines the structure of the TANs classifier.

The weights of TANs attributes are constructed by calculating the mutual information of attributes among variables. Mutual information refers to the degree of association between two random variables, that is, the degree of uncertainty weakening of another random variable given a random variable. According to the chain rule of entropy, $H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$. Therefore, the difference $H(X) - H(X|Y) = H(Y) - H(Y|X)$ is denoted as $MI(X, Y)$, and mutual information $MI(X, Y)$ is defined as follows.

$$MI(X, Y) = \sum_{xy} P(x, y)\log\frac{P(x, y)}{P(x)P(y)}, \tag{2}$$

where $P(x, y)$ is the joint distribution of variable (X, Y) and $P(x)$ and $P(y)$ are the marginal distribution, respectively.

The mutual information between attributes is the correlation of attributes, and the correlation values calculated by different class attributes are also different. Considering the addition of TANs class variable attributes, the mutual information equation of a certain classification attribute needs to be redefined, so the mutual information of TANs can be calculated as follows.

$$\begin{aligned} MI_{ij} &= \left(c_i, c_j \mid C\right) \\ &= \sum_{c_i, c_j, c} P\left(c_i, c_j, c\right) D\left(c_i \| c_j\right) \log\frac{P\left(c_i, c_j \mid C\right)}{P\left(c_i \mid C\right) P\left(c_j \mid C\right)}, \end{aligned} \tag{3}$$

where c_i and c_j are attribute variables and C are class variables, and $D(c_i \| c_j)$ is the relative entropy of c_i and c_j.

The construction for TANs is summarized as follows.

Step 1. Obtain the mutual-information values $MI(T_i, T_j)$ of all attribute pairs through equation (3).

Step 2. $MI(T_i, T_j)$ is sorted in descending order, and node pairs are output successively.

Step 3. According to the principle that TANs does not generate a loop, edges are selected in descending order of edge weight until $n - 1$ edges are selected, and a completely undirected graph with mutual-information value as weight is constructed.

Step 4. Select any node in the completely undirected graph as the root of TANs and set the direction of all edges outward from the root node. The process of converting an undirected tree into a directed tree is completed by setting the direction between attribute nodes.

Step 5. Add a class node (i.e., class attribute node) to each node in TANs and the directed edges of the class node pointing to all attribute nodes.

Therefore, TANs-based translation data cleaning steps are as follows.

(1) The translation data attribute dataset TD is sampled.

(2) The importance measurement algorithm [33] is used to reduce the data level of the sampled translation data.

(3) Construct TANs.

(4) The mutual-information value is used to score the TANs.

(5) The Top-k problem is to get the maximum number of k from an array or list. Judge the Top-k mutual-information value $MI(T_i, T_j)$ and the empirical parameter value. If the Top-k mutual-information value $MI(T_i, T_j)$ is less than the empirical parameter value, remove the top-level node and query the score R of the result of removing the top-level node.

(6) The TANs nodes are sorted and output in descending order according to the score R.

4. RNN-Controlled Translation Data Fusion

4.1. Position of Multisource, Heterogeneous Translation Data Nodes. Due to the structural diversity of multisource, heterogeneous translation data, the data structure will produce diverse fusion results. Therefore, before multisource, heterogeneous fusion, nodes in multisource, heterogeneous translation data need to be located. The improved firefly algorithm [34] is used to calculate the distance of the coordinates of heterogeneous anchor nodes. DV-HOP algorithm is used to calculate the hop count of nodes and anchor nodes as follows.

$$t_{k+1/k} = NP_{k|k}N^T + S, \tag{4}$$

where $t_{k+1/k}$ represents the prior estimate value of state at time $K+1$, $P_{k/k}$ represents the posterior estimate matrix at time K, S is the state estimate value, T is the time, and N is the anchor node.

Equation (3) is used to preliminarily lock the region where the heterogeneous translation data is located, collect all sample points in this region, and predict the location of unknown mobile nodes. Suppose that the moving speed of the unknown node meets the interval $[0, v_{\max}]$ and presents an interval uniform distribution, and its position is defined as follows.

$$\text{pos} = \begin{cases} \dfrac{1}{\delta v^2}, & d \le v_{\max}, \\ 0, & d > v_{\max}, \end{cases} \tag{5}$$

where d represents the distance of moving node from time $K-1$ to time K, and v represents the average speed of moving node. Set the communication radius between hops and remove nonconforming nodes, if the nodes are within a communication range and less than the communication radius with anchor nodes. On the contrary, the nodes that do not meet the conditions are filtered, the calculation results of all nodes are integrated, the data fusion mapping relationship of nodes is described, and the fusion model is constructed.

4.2. Analysis of Translation Data Fusion Mapping. Any data fusion process can be viewed as a process of external to internal mapping. When constructing the multisource, heterogeneous translation data model, the node relation is obtained to describe the translation data fusion mapping relation. The quintuple in the fusion model is defined as follows.

$$FM = \{S, M, MS, TS, Rm\}, \tag{6}$$

where S represents the state data in the prefusion space, M represents the measurement space, MS represents the fusion space, TS represents the target space for fusion judgment, and Rm represents the mapping set relationship between different spaces.

Suppose the mapping set has the following triplet relationship.

$$Rm = \{\alpha, \beta, f\}, \tag{7}$$

where α represents the measurement space mapping in the space to be fused, β represents the process of transforming original multisource, heterogeneous translation data into integrated spatial data after mapping processing, and f represents the mapping spatial relationship. The space before the fusion of j multisource, heterogeneous translation data can be expressed as follows.

$$S = \begin{pmatrix} S_{11} & S_{1j} \\ S_{j1} & S_{ij} \end{pmatrix}, \tag{8}$$

where rows represent the target contained in the space before translation data fusion, and columns have multisource, heterogeneous attributes. i represents the maximum number of features of the fusion target, which is zero if the target does not contain the related feature. Assume that the fusion space M at time t is expressed as follows.

$$M_t = \begin{pmatrix} m_{11} & m_{1q} \\ m_{q1} & m_{pq} \end{pmatrix}, \tag{9}$$

where m_{pq} represents the qth heterogeneous translation data obtained by information source p in the fusion model at time t. The maximum value of translation data provided by each information source is u, and the number of heterogeneous translation data sources is v, so the matrix of the fusion space is expressed as follows.

$$MS = \begin{pmatrix} ms_{11} & ms_{1v} \\ ms_{v1} & ms_{uv} \end{pmatrix}. \tag{10}$$

Combining equations (9) and (10), and the mapping relation can be calculated as follows.

$$MS = \sigma(M_t). \tag{11}$$

In the underlying dataset fusion, the translation data corresponding to nodes has been simply preprocessed. At this point, the mapping relation of σ is 1 : 1, and the final space TS is composed of the final result of the fusion model, and the space can be expressed as follows.

$$TS = (d_1, d_1, \ldots, d_n)^T, \tag{12}$$

where d_i is the final fusion degree of fusion target i, and the fusion mapping relation can be expressed as follows.

$$TS = f\left(\sum_{t=0}^{q\Delta t} MS_t\right), \tag{13}$$

where Δt represents the time interval between fused translation data and q represents fusion times of translation data. Finally, the translation data fusion mapping relationship is described by using equation (13). By using the characteristics of the recurrent neural network (RNN), the fusion process of the fusion model is controlled and the multisource heterogeneous translation data fusion model is constructed.

4.3. RNN-Controlled Fusion Process. Before the fusion model is completed, RNN is used to control the multisource heterogeneous translation data fusion process, and the neural network structure is used to control. For the mapping set ts formed after mapping, it is assumed that the output of the input set of the neural network is O_{ts}, where the ith input is O_{ts_i}; that is,

$$W = \sum_{t=0}^{n} O_{ts\_i} O_{ts}. \tag{14}$$

The sigmoid function is a common "S" shape that is often used as the threshold function of neural networks, mapping variables to between 0 and 1, so equation (14) can be regarded as sigmoid function, and then it can be transformed into

$$\begin{aligned} O_{ts\_i} &= \text{sigmoid}(W) \\ &= \frac{1}{1 + e^{-O_{ts}}}. \end{aligned} \tag{15}$$

In order to reduce the bias in the neural network, the bias function [35] is normalized, and the bias function calculates the biases in the whole fusion process.

$$B_{ts} = \frac{1}{3} \sum_{ts} (w_i - O_{ts\_i})^2, \tag{16}$$

where w_i represents the training weight of the RNN, which is used to limit the error function to a minimum. In order to ensure that the RNN can control all hidden layer nodes and enhance fault-tolerant during the whole fusion process, hidden layer node information is selected, and the calculation is as follows.

$$h = \frac{I + O}{2} + \varepsilon, \tag{17}$$

where h represents the number of nodes in the hidden layer, I represents the number of input nodes, O represents the number of output nodes, and ε represents the constant ranges from 0 to 1. Constantly delete and increase the number of nodes, determine the nodes in the control fusion process, realize the neural network control fusion process, and complete the construction of multisource, heterogeneous translation data fusion model based on RNN.

5. Experiment and Results Analysis

5.1. Setup. This paper selected five different types of English corpora, which were Corpus of Contemporary American English (COCA), 20 Newsgroups, Auslan, Reuters-21578, and UNIX_user_data. Aiming at imprecise corpus data and similar repeated corpus data, the effectiveness of TANs-based translation data cleaning method proposed in this paper was verified by using two metrics: cleaning rate and cleaning time. FU-ROA [14], IHCS [15], and FedClean [16] are selected for comparison.

Root mean square error (RMSE), mean absolute percentage error (MAPE), fusion time, and integrity were used for comparison of heterogeneous translation data fusion. In comparison, MDCN [23], NB-IoT [24], and HFECS [25] were used.

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}, \\ \text{MAPE} &= \sqrt{\frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|}. \end{aligned} \tag{18}$$

5.2. Comparison Analysis

5.2.1. Translation Data Cleaning Experiment. Figure 1 shows that the cleaning rate of TANs-based translation data cleaning method proposed in this paper is generally in COCA, 20 Newsgroups, UNIX_user_data, and Reuters-21578 showed significantly higher cleaning rates in four English corpora than in the other three baselines. However, the improvement in Auslan English corpus is not obvious because there is a conditional independence relationship between the sample attributes. Except for Auslan English corpus, the method in this paper maintains a relatively stable cleaning rate, basically maintaining or even exceeding 90%, and can maintain such a high level of classification accuracy, indicating that the method in this paper is feasible. In the heterogeneous large-scale English translation strategy, due to the heterogeneity of translation data on different platforms, the good translation data cleaning rate solves the problem that the correlation of translation data in multisource, heterogeneous environment may lead to overfitting. The core tuple and boundary tuple are retained according to

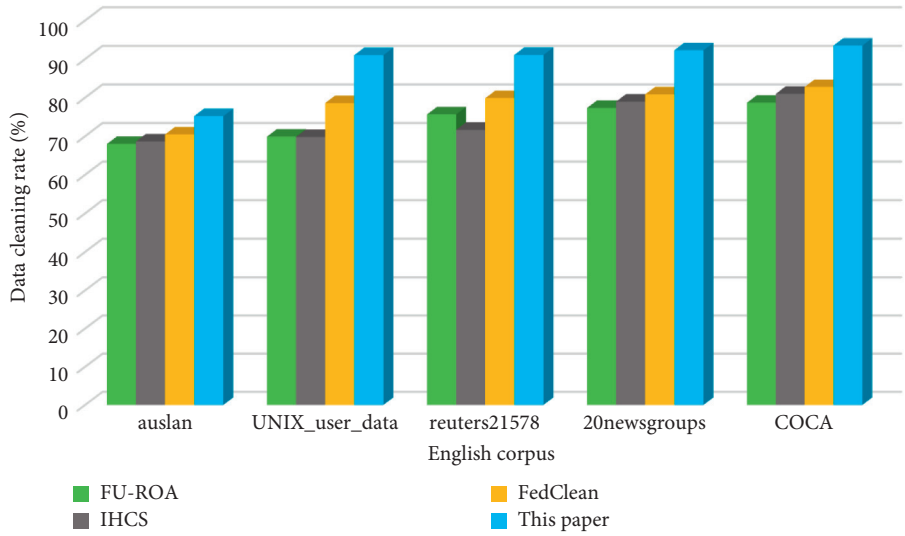


Figure 1: Translation data cleaning rate under different English corpora.

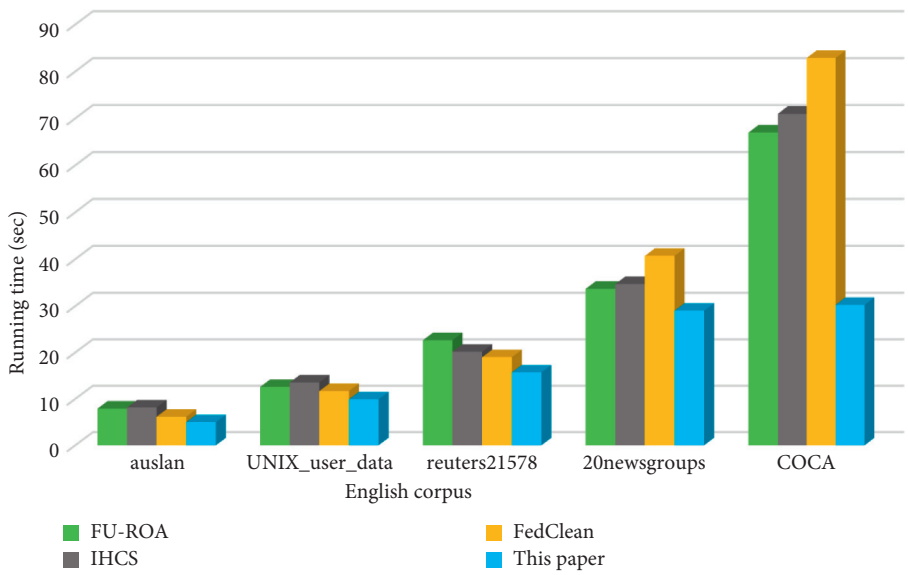


Figure 2: Translation data cleaning time under different English corpora.

the weight to reduce the outlier tuple, which greatly reduces the workload of data cleaning and the difficulty of translation.

As can be seen from Figure 2, for the small-scale English corpora Auslan, UNIX_user_data, Reuters-21578, and 20 Newsgroups, the performance effects of the four algorithms are not different. However, for the English corpus COCA with high dimension and large scale, the time spent by this algorithm is relatively less and the performance improvement is relatively obvious. The main reason is that when the amount of data is large, the improved attribute reduction algorithm can be mapped to a less search space, and the simplest attribute set can be solved by searching only a few simplified elements, which reduces the time consumption of attribute reduction and improves the efficiency of the algorithm, especially for large-scale English corpus. At the same time, it also solves the problem of redundancy and dependence of data features extracted from different interrelated English corpus. In heterogeneous large-scale English translation strategies, less translation data cleaning time corresponds to less translation time, and the translation effect is better.

5.2.2. Translation Data Fusion Experiment. As can be seen from Figure 3, RMSE of the proposed method in this paper is generally low in all English corpora. RMSE of NB-IoT algorithm performed well in Auslan, UNIX_user_data, Reuters-21578, 20 Newsgroups English corpus, better than MDCN and HFECS, but weakened in COCA corpus. The RMSE of MDCN algorithm on COCA corpus even exceeds 10. RMSE actually describes a degree of dispersion, while the

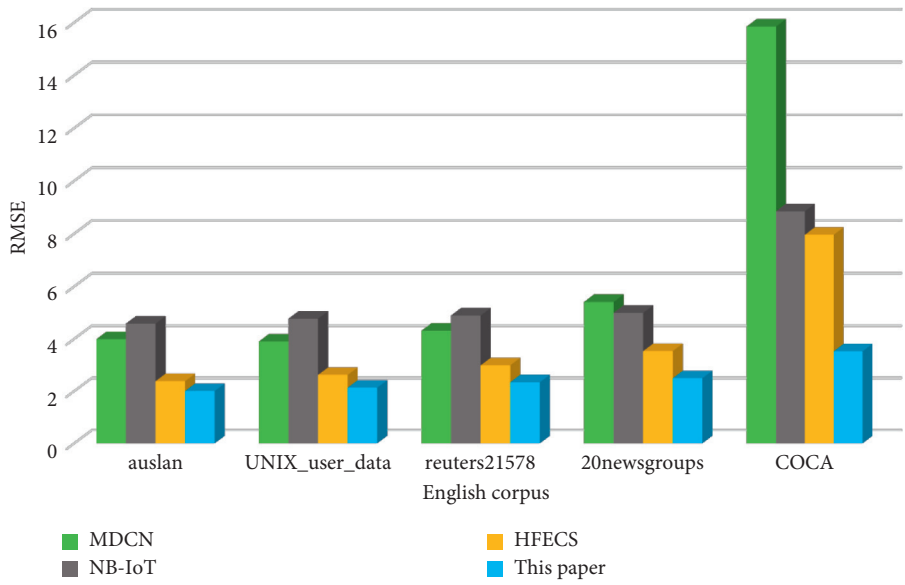


Figure 3: RMSE of translation data fusion under different English corpora.

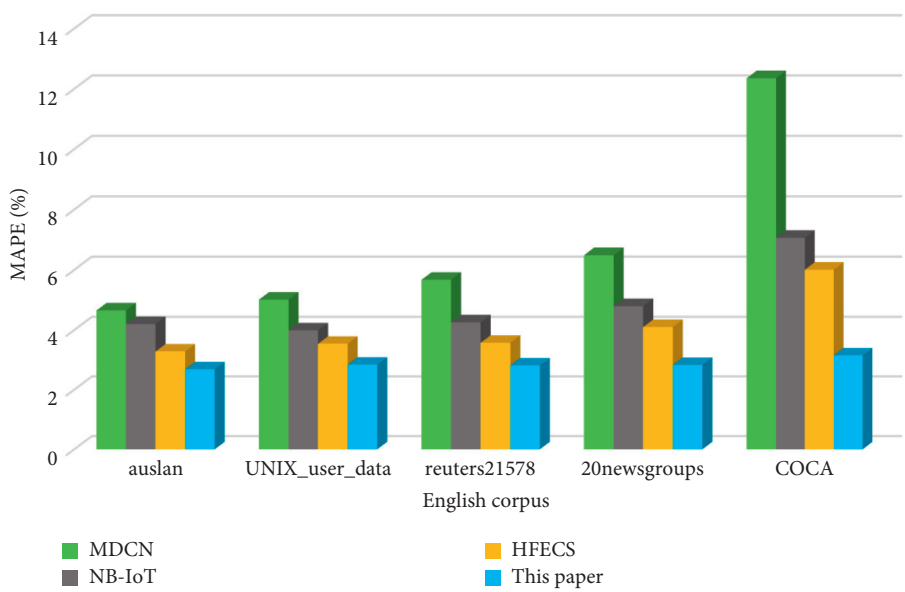


Figure 4: MAPE of translation data fusion under different English corpora.

translation data fusion algorithm proposed in this paper has a low RMSE, which further indicates that the algorithm presented in this paper has a good stability, while the other three baselines have great differences in different English corpora. It is not conducive to translation data fusion of heterogeneous translation platforms. From Figure 4, it can be seen that the MAPE of the translation data fusion algorithm in this paper always remains below 5% in each English corpus, showing a good fitting effect.

As indicated in Figure 5, the translation data fusion time of the four algorithms on Auslan, UNIX_user_data, and Reuters-21578 English corpora is similar. However, in 20 Newsgroups and COCA English corpus, the fusion time of the translation data fusion algorithm in this paper is much lower than the other three baselines. In large-scale, heterogeneous English translation, less heterogeneous data fusion time provides strong support for translation strategies. Figure 6 shows the integrity of the translation data fusion algorithm in this paper is always the best and remains above 80%. A large-scale and efficient translation strategy requires a complete integration of translation strategies in all platforms and the fusion of translation data. However, the algorithm in this paper maintains the integrity of heterogeneous translation data.

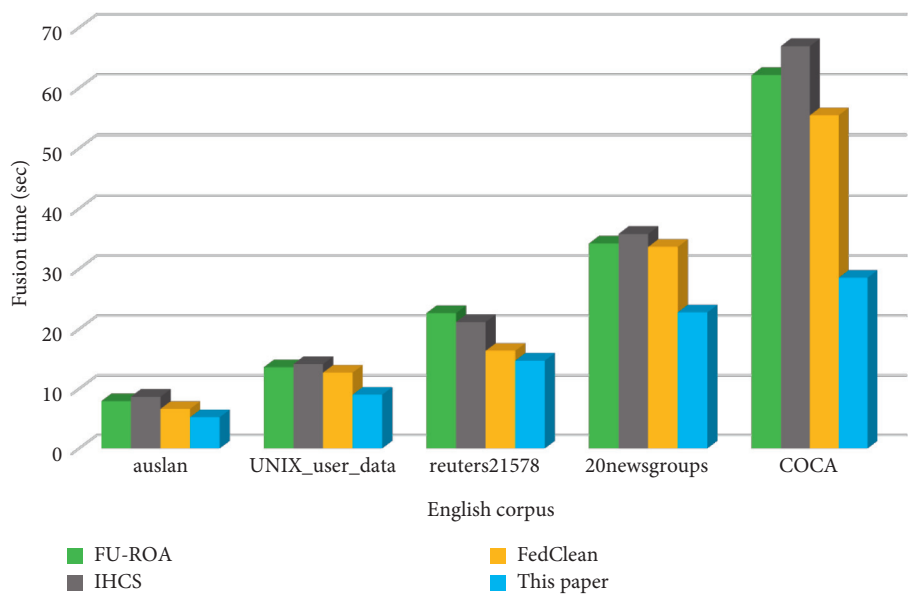


Figure 5: Fusion time of translation data fusion under different English corpora.

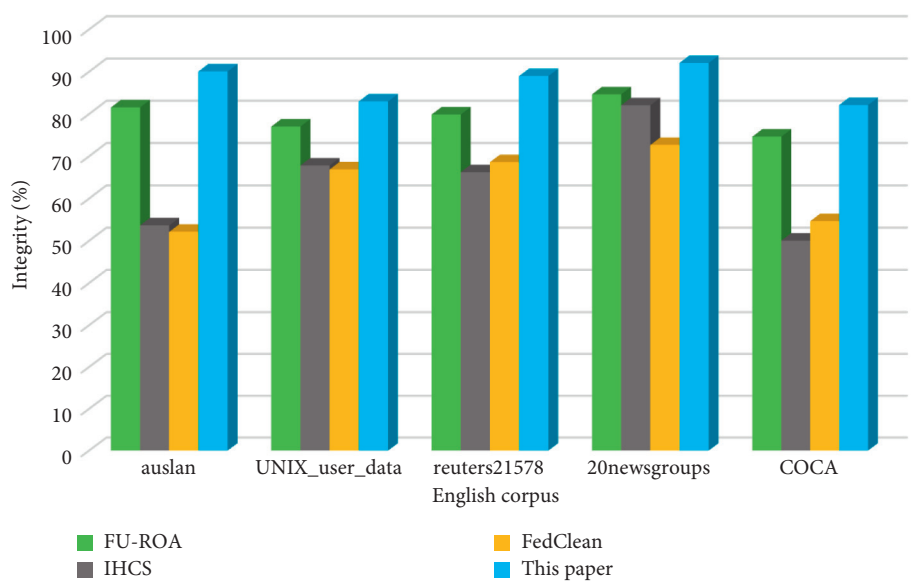


Figure 6: Integrity of translation data fusion under different English corpora.

6. Conclusions

Aiming at the problem of large amount of translation data in multisource, heterogeneous translation environment, TANs-based translation data cleaning is proposed to clean translation data. Translation data attributes and tuples are weighted by data hierarchical reduction, which solves the problem that translation data correlation may lead to overfitting in multisource, heterogeneous environment. Core and boundary tuples are retained according to the weight, and outlier tuples are reduced, which greatly reduces the workload of translation data cleaning. In addition, aiming at the problem that the number of unique elements obtained by the traditional multisource, heterogeneous translation data fusion model is small, resulting in the lack of strong integrity of the final fusion data, a multisource, heterogeneous translation data fusion model based on RNN is constructed. By locating the node data in the multisource, heterogeneous data and analyzing the mapping relationship, the multisource heterogeneous translation data fusion model is realized. Experimental results show that TANs-based strategy can effectively improve the cleaning rate and cleaning time of translation data cleaning compared with baselines. Moreover, the translation data fusion experiment results demonstrate that the RNN-based translation data fusion method outperforms baseline with respect to RMSE, MAPE, fusion time, and integrity.

In the future work, we will deeply study various data types, data storage methods, data association status, and errors in the process of data cleaning. At the same time,

although the translation data fusion model in this paper improves some metrics to a certain extent, there are still some deficiencies. It still needs to be improved in future research to obtain a better fusion model.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] S. Bi, "Intelligent system for English translation using automated knowledge base," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5057–5066, 2020.

[2] S. Wang, "Simulation of English translation text filtering based on machine learning and embedded system," *Microprocessors and Microsystems*, vol. 83, 2021.

[3] Y. Liu and H. Bai, "Teaching research on college English translation in the era of big data," *International Journal of Electrical Engineering Education*, 2021.

[4] Z. Li, "Simulation of English Education Translation Platform Based on Web Remote Embedded Platform and 5G Network," *Microprocessors and Microsystems*, vol. 81, 2021.

[5] L. Shi and X. Wang, "Strategies of cross-cultural eco-education in college English translation teaching," *Ekoloji*, vol. 28, no. 107, pp. 3045–3050, 2019.

[6] M. Kolhar and A. Alameen, "Artificial intelligence based language translation platform," *Intelligent Automation and Soft Computing*, vol. 28, no. 1, 2021.

[7] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Transactions on Cybernetics*, 2021.

[8] J. A. DeSimone and P. D. Harms, "Dirty data: the effects of screening respondents who provide low-quality data in survey research," *Journal of Business and Psychology*, vol. 33, no. 5, pp. 559–577, 2018.

[9] J. Rammelaere and F. Geerts, "Cleaning data with forbidden itemsets," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 8, pp. 1489–1501, 2020.

[10] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021.

[11] R. Choi, A. S. Coyner, J. Kalpathy-Cramer, M. F. Chiang, and J. P. Chambell, "Introduction to machine learning, neural networks, and deep learning," *Translational Vision Science & Technology*, vol. 9, no. 2, 2020.

[12] L. Azzimonti, G. Corani, and M. Zaffalon, "Hierarchical estimation of parameters in Bayesian networks," *Computational Statistics & Data Analysis*, vol. 137, pp. 67–91, 2019.

[13] Y. Long, L. Wang, and M. Sun, "Structure extension of tree-augmented naive Bayes," *Entropy*, vol. 21, no. 8, 2019.

[14] K. Rahul and R. K. Banyal, "Detection and correction of abnormal data with optimized dirty data: a new data cleaning model," *International Journal of Information Technology and Decision Making*, vol. 20, no. 02, pp. 809–841, 2021.

[15] C. Ge, Y. Gao, X. Miao, L. Chen, C. S. Jensen, and Z. Zhu, "IHCS: an integrated hybrid cleaning system," *Proceedings of the Vldb Endowment*, vol. 12, no. 12, pp. 1874–1877, 2019.

[16] L. Ma, Q. Pei, L. Zhou, H. Zho, L. Wang, and Y. Ji, "Federated data cleaning: collaborative and privacy-preserving data cleaning for edge intelligence," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6757–6770, 2021.

[17] Y. Huang, M. Milani, and F. Chiang, "Privacy-aware data cleaning-as-a-service," *Information Systems*, vol. 94, 2020.

[18] S. Al-Janabi and R. Janicki, "Data repair of density-based data cleaning approach using conditional functional dependencies," *Data Technologies and Applications*, 2021.

[19] Q. Wang, Y. Guo, L. Yu, X. Chen, and P. Li, "Deep Q-network-based feature selection for multisourced data cleaning," *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 16153–16164, 2021.

[20] X. Shi, C. Prins, G. Van Pottelbergh, P. Mamouris, B. Veas, and B. D. Moor, "An automated data cleaning method for Electronic Health Records by incorporating clinical knowledge," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, 2021.

[21] M. B. Gesicho, M. C. Were, and A. Babic, "Data cleaning process for HIV-indicator data extracted from DHIS2 national reporting system: a case study of Kenya," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, 2020.

[22] S. Liu, G. Li, S. Jiang et al., "Investigating data cleaning methods to improve performance of brain-computer interfaces based on stereo-electroencephalography," *Frontiers in Neuroscience*, vol. 15, 2021.

[23] J. Tian, D. Han, L. Xiao, and P. Shi, "Multi-scale deep coupling convolutional neural network with heterogeneous sensor data for intelligent fault diagnosis," *Journal of Intelligent and Fuzzy Systems*, vol. 41, no. 1, pp. 2225–2238, 2021.

[24] Y. Liu, "Multi-source heterogeneous data fusion based on perceptual semantics in narrow-band Internet of Things," *Personal and Ubiquitous Computing*, vol. 23, no. 3-4, pp. 413–420, 2019.

[25] J. Anees, H. Zhang, S. Baig, B. G. Lougou, and T. G. Robert Bona, "Hesitant fuzzy entropy-based opportunistic clustering and data fusion algorithm for heterogeneous wireless sensor networks," *Sensors*, vol. 20, no. 3, 2020.

[26] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, "A survey on machine learning for data fusion," *Information Fusion*, vol. 57, pp. 115–129, 2020.

[27] R. Dominguez, M. Post, and A. Fabisch, "Common data fusion framework: an open-source common data fusion framework for space robotics," *International Journal of Advanced Robotic Systems*, vol. 17, no. 2, 2020.

[28] B. He, X. Cao, and Y. Hua, "Data fusion-based sustainable digital twin system of intelligent detection robotics," *Journal of Cleaner Production*, vol. 280, 2021.

[29] I. A. Al-Baltah, A. A. Abd Ghani, G. M. Al-Gomaei, F. H. Abdulrazzak, and A. A. Al Kharusi, "A scalable semantic data fusion framework for heterogeneous sensors data," *Journal of Ambient Intelligence and Humanized Computing*, 2020.

[30] M. Simjanoska, S. Kochev, J. Tanevski, A. Madevska Bogdanova, G. Papa, and T. Eftimove, "Multi-level information fusion for learning a blood pressure predictive model using sensor data," *Information Fusion*, vol. 58, pp. 24–39, 2020.

[31] I. Ullah and H. Y. Youn, "Intelligent data fusion for smart IoT environment: a survey," *Wireless Personal Communications*, vol. 114, no. 1, pp. 409–430, 2020.

[32] K. K. Kumar, E. Ramaraj, and P. Geetha, "Multi-sensor data fusion for an efficient object tracking in internet of things (IoT)," *Applied Nanscience*, 2021.

[33] J. Lv, X. Wang, K. Ren, M. Huang, and K. Li, "ACO-inspired Information-Centric Networking routing mechanism," *Computer Networks*, vol. 126, pp. 200–217, 2017.

[34] A. M. Altabeeb, A. M. Mohsen, L. Abualigah, and A. S. Ghalllab, "Solving capacitated vehicle routing problem using cooperative firefly algorithm," *Applied Soft Computing*, vol. 108, pp. 1–10, 2021.

[35] R. T. Godwin and D. E. Giles, "Analytic bias correction for maximum likelihood estimators when the bias function is non-constant," *Communications in Statistics-Simulation and Computation*, vol. 48, no. 1, pp. 15–26, 2019.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 7334686, 9 pages
https://doi.org/10.1155/2022/7334686

Research Article

Intelligent English Tense Collocation and Evaluation Based on Deep Reinforcement Learning

Yiling Ding[1,2] **and Tianhua Wang**[1]

[1]*Heilongjiang University, Harbin 150080, China*
[2]*Harbin Normal University, Harbin 150080, China*

Correspondence should be addressed to Yiling Ding; dingyiling@hrbnu.edu.cn

Received 4 January 2022; Accepted 17 January 2022; Published 8 February 2022

Academic Editor: Jianhui Lv

The representation of time in sentences is the key problem for tense collocation. Based on the relationship expression among regions in Allen's interval algebra theory, we propose a vector representation method, i.e., relationship vector, and several operations are defined based on the relationship vector for temporal reasoning in this work. This method transforms the original matrix representation into vector representation, which reduces the amount of computation of temporal reasoning. In addition, we propose a temporal classification and collocation method based on deep learning and deep reinforcement learning. This method uses a bidirectional cyclic neural network and a convolutional neural network for text expression and achieves temporal word classification and temporal collocation based on the deep reinforcement learning model. In the experiments, the proposed method obtains the average accuracy of 92.17% in five datasets, i.e., MPQA, CR, MR, Subj, and TREC, which proves its effectiveness in tense collocation.

1. Introduction

Mathematical proof and simulation are the important foundation of modern science and technology, and the development of artificial intelligence (AI) is also based on strong mathematical theories [1]. Symbolization and formulation of human mind on computer have become an important topic in the research of AI. Establishing a unified theoretical system of knowledge-based human intelligence and machine intelligence is the mainstream idea of AI, and the key problem is knowledge representation, knowledge reasoning, and the application of knowledge. In recent years, with the application and development of AI and big data technology, deep learning (DL) technology has made breakthroughs in many fields, such as image recognition, speech recognition, and natural language processing (NLP) [2]. Among them, semantic analysis technology for NLP plays an important role in studying human communication mode. It affects not only the development of AI, but also the mode of human-computer interaction [3]. Using AI semantic analysis technology to analyze the language materials can significantly improve the analysis efficiency of user research results and even dig out important information that is difficult to find by artificial analysis. NLP is a technology that uses computer as a tool to process the information of human unique natural language [4, 5]. Semantic analysis of language materials is the basis of NLP involving linguistics, computational linguistics, machine learning, cognitive language, and other disciplines. The goal of semantic analysis is to realize the automatic semantic analysis of vocabulary, sentences, and chapters and understand the meaning of language text expression by establishing effective computational analysis models. Among them, the semantic analysis of words is the basis, and the commonly used knowledge expression methods include semantic field, semantic network, concept map, and ontology [6].

There are a lot of emotions and ideas in natural language that cannot be described by simple logical relations. It is difficult to achieve because of the complexity of recognizing and understanding natural languages by computer. From the

development history of NLP, relevant research can be divided into two categories: symbolic approaches and stochastic approaches. The symbolic approaches adopt the analysis method based on linguistic definition rules focusing on the reasoning model and the logical language understanding model, which is called rationalism. The stochastic approaches mainly study random and statistical algorithms to identify and calculate large-scale corpus [7]. Although the symbolic approaches can complete the basic analysis of a single sentence, the understanding of complex sentence patterns and even the analysis of the whole paragraph or chapter are not of high quality. The statistical method contained in the stochastic approaches provides a new idea for the research method of natural language understanding, but the early effect is not satisfactory until the emergence of DL technology in recent years. Relying on statistical methods and DL algorithms, machines can learn from big data and obtain the rule of language usage, which is greatly improved in efficiency and accuracy compared with traditional methods. In particular, in the application of Chinese semantic analysis, the effect improvement is particularly prominent. It uses the ability of automatic feature extraction in the DL model to replace the manual feature construction process in traditional machine learning, which greatly improves the effectiveness of the feature model. From the historical evolution of NLP, we can see that semantic analysis technology has been developed greatly with the intervention of machine learning methods represented by DL. Through the study of literature, it is found that many scholars have studied the application methods of intelligent semantic recognition technology, and their methods and analysis steps are also different according to different application purposes. Allen [8] proposes a temporal reasoning system based on continuous interval representation of time information. It can easily represent the temporal relationship between any two-time intervals, which have strong expression ability and are easy to understand. It is widely used in natural language understanding and other fields. However, in the reasoning operation of interval algebraic representation, repeated table searching is required, which is very time-consuming. It is very impractical for the system with high real-time requirements. Therefore, it is proposed to use interval matrix to represent temporal intervals and convert the traditional table searching into matrix operation, which is very fast. Its effect is very excellent when the real-time requirements are relatively quotient. However, the temporal relationship expressed by matrix is not intuitive.

The traditional text representation methods include word bag model and vector space model. The word bag model uses unique heat coding to represent each word as a vector composed of 0 and 1. This method is simple and intuitive, but does not consider word order and semantic information. The vector space model treats text as points in vector space. The term frequency-inverse document frequency (TF-IDF) is commonly used to represent the importance of words and construct text feature representation. However, with the rapid increase in text data on the Internet, traditional text representation methods have some disadvantages, such as high dimension, sparse features, and loss of word order. In this case, many scholars begin to use DL, and there are three types of models for text representation and classification using neural networks [9]. The first type is the serialization model, such as convolutional neural network (CNN), recurrent neural network (RNN), and long-term and short-term memory network (LSTM). They record the text serialization information for training and carry out the text classification task after obtaining the feature representation. The second type is based on attention mechanism. The third category is structured text representation and classification model, such as recursive automatic encoder [10] and short-term and long-term memory network with tree structure. They use the predefined parse tree to divide the text hierarchy and then represent and classify the text. These text representation and classification models based on deep neural networks have been verified on public datasets and have achieved good text representation and classification performance. When constructing text features, traditional text representation methods and classification models based on deep learning often delete some useless words by means of stopping vocabulary or using parsing tree to divide text structure. They cannot delete words irrelevant to classification tasks according to different texts. Parsing trees are also constructed manually, which is time-consuming and laborious. Because reinforcement learning (RL) has the ability of autonomous learning with the rapid improvement of computing performance in recent years, the combination of DL and RL for task decision-making has been widely used. Deep reinforcement learning (DRL) has the advantage of learning knowledge by interacting with the environment [11–14]. Many scholars have applied DRL to text representation and classification tasks and achieved good results. On the basis of DRL, the text representation and classification model can independently learn words related to text classification tasks, divide sentence phrase structures, and make text feature tables. It is worth studying to show and improve the performance of text classification.

To realize the accurate collocation of English sentence tenses, we propose a word-level text representation and classification model brand model based on DRL in this work. This method extracts words and divides phrase structures related to text classification tasks from main learning and optimizes text representation and classification tasks for words and phrases, respectively. Then, the Allen interval algebra is used to represent the temporal relationship, and it is applied to the path consistency algorithm of the network graph composed of a group of temporal intervals and the constraint relationship between them. In the method, the temporal constraint consistency of the subnetwork graph composed of three adjacent nodes is used as the preprocessing technology. Furthermore, the calculation result matrix is transformed into a temporal relationship representation, and one or two eigenvalues in the atomic matrix are used to determine its corresponding temporal relationship. Thus, while improving computational efficiency, we can visually represent and understand the temporal relationship and then make interval algebra obtain better representation and reasoning knowledge in our model. The

experimental results show that the accuracy of tense collocation in English sentences can reach 92.17%.

The rest of this study is organized as follows. Section 2 presents the related work of tense collocation methods. Section 3 presents the detailed design on tense collocation method based on DRL. Experimental results and discussion are reported in Section 4. Finally, the conclusion of this study is in Section 5.

2. Related Work

Focusing on the text representation and classification model of sentences and documents, Zhang et al. [15] proposed a character-level convolutional neural network (char-CNN) model, which attempted to represent and classify text by constructing alphabet and using one-hot coding and CNN. Kim [16] proposed a word-CNN model that achieves word vector representation of static and non-static channel and English text classification. Mikolov et al. [17] proposed a language model using a cyclic neural network to study text representation and classification. Based on the work, Lai et al. [18] introduced a two-way structure to capture context information and retain a wide range of word sequence information combined with cyclic and CNN models to achieve better text representation and classification results. Tang et al. [19] proposed a gate-based cyclic neural network, which analyzed document-level content through a two-way gated recurrent neural network, and used forward and backpropagation to obtain context feature representation to improve the classification effect.

The model of text feature representation and classification using DL has achieved good performance. However, in the task of text representation and classification, CNN, RNN, and other models do not consider the text hierarchy information. For structured DL models, such as tree LSTM, the predefined parsing tree is used for text structured processing, and the parsing tree is constructed by relying on a large amount of manpower and expert experience, which requires a lot of time and energy. Therefore, some scholars began to explore whether there is an interactive way that can make the text representation and classification model extract the vocabulary information related to the text classification task through autonomous learning, divide the text into phrase structures, and improve the accuracy of text feature representation and the performance of text classification. In this demand, DRL has entered the vision of many scholars. RL belongs to the branch of machine learning, and many classical algorithms have been proposed in the early stage, such as Q-learning algorithm and policy gradient algorithm. With the in-depth application of deep neural network, many scholars began to try to combine DL and RL in various fields to solve complex perceptual decision-making tasks. In view of the successful application of DRL in the image field, this method has been introduced in the text field. Yu et al. [20] used the strategic algorithm of DRL in generative adversarial networks (GANs), which provided an idea for solving the task of serializing text.

Temporal reasoning includes the formalization of time concept and provides tools for the representation and reasoning of temporal knowledge. The research of temporal reasoning originated in the late 1950s and early 1960s and initially appeared in the form of temporal logic. In the 1980s, the research on temporal logic and reasoning was very active and had never been interrupted. A logical method is one of the main methods to study temporal reasoning. The research on temporal reasoning is mainly reflected in reliability, completeness, expression ability, decidability, and computational complexity. The research of spatial logic could be traced back to the 1950s, but it had not attracted great interest of researchers in AI, spatial database, image processing, and other fields until nearly two decades, mainly including topological spatial logic, metric spatial logic, directional relationship logic, and spatial logic integrating topology and measurement. In 2002, Bennett and Cohn [21] studied the multimodal system for spatiotemporal reasoning and proposed a multidimensional modal logic for representing spatio-temporal information. In 2006, Kudinoy et al. [22] added a difference operator based on logic S4 in topological space to enhance the expression ability of logic and proved its logical completeness after adding modal operator. The azimuth relation is an important spatial relation in spatial reasoning. The topological relation between spatial objects under topological transformation remains unchanged, but the azimuth relation may change. The logic complexity based on the azimuth relation is low, which is very effective for the formal representation of natural space, and has enough ability to represent meaningful spatial knowledge. Therefore, modal logic based on the azimuth relationship can better represent knowledge in natural space.

Different from the previous work, this study proposed the DRL-based method to realize the accurate collocation of English sentence tenses. The proposed method shows more effective than baselines in tense collocation.

3. The Detailed Methodology

The representation of tenses is an important basis for the accurate collocation of tenses in text generation. In this work, we propose a temporal vector representation, i.e., relationship vector, based on the relationship expression method between regions in Allen's interval algebra theory. Then, we design an intelligent English tense collocation method, in which several operations are defined, including relationship vector and text representation. This section will describe the details of the algorithm.

3.1. Vector Representation of Temporal Relations. The Allen interval algebra defines 13 basic interval relations, and their meanings are shown in Figure 1, where x and y represent continuous temporal intervals, $x, y \in [0, T]$, and T represents sufficiently large time values. Among the 13 relationships, 6 pairs of relationships are mutually inverse, and the other equal relationships are mutually inverse with itself. Each basic interval relation describes the definite relationship between two temporal intervals. To describe the uncertain temporal relationship, it needs to be represented by a group

Relation	Symbol	Inverse	Meaning
x before y	b	bi	
x meets y	m	mi	
x overlaps y	o	oi	
x starts y	s	si	
x during y	d	di	
x finishes y	f	fi	
x equals y	eq	eq	

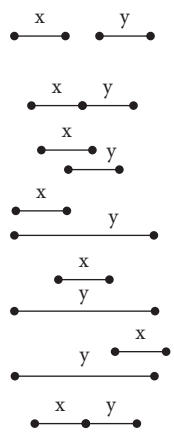


Figure 1: 13 relations of basic interval.

of basic interval relations, which are called general interval relations.

The temporal relationship represented by matrix has some disadvantages, such as less intuition and high computational complexity. To express the temporal relationship and calculation more simply, we consider using the left and right endpoints of the interval to characterize the relationship between the endpoints of the interval. Let $[0, T]$ be a time interval taken from the time axis, and T is large enough, and the interval to be investigated is always its subinterval. $I = [I^-, I^+]$ $(0 \le I^- \le I^+ > T)$ is one of its subintervals, i.e., $I \subseteq [0, T]$. $I^L = [0, I^-]$ is the left outer point interval of I, I^- is the left endpoint of I, $I^i = (I^-, I^+)$ is the inner interval of I, $I^R = [I^-, T]$ is the right outer point interval of I, and I^+ is the right endpoint of I. $(I^L, I^-, I^i, I^+, I^R)$ elements are called interval I. Therefore, any time interval I can be written as follows:

$$I^L \cup I^- \cup I^i \cup I^+ \cup I^R = [0, T]. \tag{1}$$

We set up α and β as two subintervals of interval $[0, T]$ and define $\Delta_{\alpha\beta}^{LL}$, $\Delta_{\alpha\beta}^{RL}$, $\Delta_{\alpha\beta}^{LR}$, and $\Delta_{\alpha\beta}^{RR}$ as follows:

$$\begin{cases} \Delta_{\alpha\beta}^{LL} = \alpha^- - \beta^-, \\ \Delta_{\alpha\beta}^{RL} = \alpha^+ - \beta^-, \\ \Delta_{\alpha\beta}^{LR} = \alpha^- - \beta^+, \\ \Delta_{\alpha\beta}^{RR} = \alpha^+ - \beta^+, \end{cases} \tag{2}$$

where α^-, α^+, β^-, and β^+ are the left and right endpoints of α and β, respectively.

Definition 1. Setting up α and β as two temporal intervals and vector $v = (\Delta_{\alpha\beta}^{LL}, \Delta_{\alpha\beta}^{RL}, \Delta_{\alpha\beta}^{LR}, \Delta_{\alpha\beta}^{RR})$ is called basic relationship vector, in which each element is -1, 0, or 1. If $\Delta_{\alpha\beta}^{LL} > 0$, $\Delta_{\alpha\beta}^{LL} = 1$; if $\Delta_{\alpha\beta}^{LL} = 0$, $\Delta_{\alpha\beta}^{LL} = 0$; and if $\Delta_{\alpha\beta}^{LL} < 0$, $\Delta_{\alpha\beta}^{LL} = -1$. $\Delta_{\alpha\beta}^{RL}$, $\Delta_{\alpha\beta}^{LR}$, and $\Delta_{\alpha\beta}^{RR}$ have similar value rules.

The purpose of Definition 1 is to express the basic relationship between intervals through the relationship between the left and right endpoints of intervals. Each vector obtained according to Definition 1 corresponds to a basic relationship, and thus, the basic interval relation of the Allen interval algebra can be expressed by vector, which is called the basic relationship vector. The 13 basic relationship vectors of the Allen interval algebra are listed as follows:

(1) α before β can be denoted as $V_b = [-1, -1, -1, -1]$
(2) α after β can be denoted as $V_b = [1, 1, 1, 1]$
(3) α meets β can be denoted as $V_b = [-1, 0, -1, -1]$
(4) α met by β can be denoted as $V_b = [1, 1, 0, 1]$
(5) α overlaps β can be denoted as $V_b = [-1, 1, -1, -1]$
(6) α overlapped by β can be denoted as $V_b = [1, 1, -1, 1]$
(7) α finishes β can be denoted as $V_b = [1, 1, -1, 0]$
(8) α finished by β can be denoted as $V_b = [-1, 1, -1, 0]$
(9) α during β can be denoted as $V_b = [1, 1, -1, -1]$
(10) α includes β can be denoted as $V_b = [-1, 1, -1, 1]$
(11) α starts β can be denoted as $V_b = [0, 1, -1, -1]$
(12) α started by β can be denoted as $V_b = [0, 1, -1, 1]$
(13) α equals β can be denoted as $V_b = [0, 1, -1, 0]$

In the vector representation of relationship, the rules of compound operation and transfer operation of relationship are defined as follows.

Definition 2. Compound operation of relationship (V). If two intervals x and y satisfy the relationship set $P = (p_1, p_2, \ldots, p_k)$, $1 \le k \le 13$, and V_j $(j = 1, 2, \ldots, k)$ is the atomic relationship vector corresponding to each atomic relation, then the temporal relationship vector V_p between x and y can be linearly expressed by V_j $(j = 1, 2, \ldots, k)$ as follows:

$$V_p = \sum_{j=1}^{k} V_j = [1, 1, \ldots]_{1\times k} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_k \end{bmatrix}. \tag{3}$$

It can be known from Definition 1 that we should obtain the relationship between the left and right endpoints of α and γ.

Thus, the relationship between the left endpoints of α and γ can be obtained from the left endpoint of β. The relationship between the right endpoint of α and the left endpoint of γ can be obtained from the right endpoint of β. The relationship between the left endpoint of α and the right endpoint of γ can be obtained from the left endpoint of β. The relationship between the left endpoint of α and the right endpoint of γ can be obtained from the right endpoint of β. From point algebra, the relationship between points $\{<, >, =\}$ has the following characteristics:

(1) Transitivity, if the point $a < (>, =)\ b$ and $b < (>, =)\ c$, then $a < (>, =)\ c$
(2) Retention, if the point $a < (>)\ b$ and $b = c$, then $a < (>)\ c$. If $a = b$ and $b < (>)\ c$, then $a < (>)\ c$

(3) Uncertainty, if the point $a < (>)\ b$ and $b > (<)\ c$, then the relationship between a and b is uncertain, which can be denoted as $a\ (<, >, =)\ b$

From the above discussion, the transfer operation of the defined relationship is as follows.

Definition 3. Relation transfer operation (•). Setting time interval α, β, and r_1, $r_2 \in U$, and if $\alpha\{r_1\}\beta$, $\beta\{r_2\}\gamma$, then V_{r1} and V_{r2} are their corresponding relationship vectors, and the relationship vector of α between γ can be obtained by V_{r1} and V_{r2} with operation: $V_{r1r2} = V_{r1} \bullet V_{r2}$, where $V_{ri} | i = 1, 2, 3$, and 4 are the components of the relationship vector, and the value rules are as follows:

$$V^3_{r1r2} = \begin{cases} 1, & \text{if both of } V^3_{r1} \text{ and } V^4_{r2} \text{ are } 0, \\ & \text{or one of them is } 0 \text{ and the other one is } 1, \\ -1, & \text{if both of } V^3_{r1} \text{ and } V^4_{r2} \text{ are } 1, \text{ or one of} \\ & \text{them is } 0 \text{ and the other one is } -1, \\ 0, & \text{if both of } V^3_{r1} \text{ and } V^4_{r2} \text{ are } 0, \\ \{-1, 0, 1\}, & \text{if one of } V^3_{r1} \text{ and } V^4_{r2} \text{ is } 1 \text{ and} \\ & \text{the other one is } -1, \end{cases} \tag{4}$$

$$V^2_{r1r2} = \begin{cases} 1, & \text{if both of } V^4_{r1} \text{ and } V^2_{r2} \text{ are } 0, \text{ or one of} \\ & \text{them is } 0 \text{ and the other one is } 1, \\ -1, & \text{if both of } V^4_{r1} \text{ and } V^2_{r2} \text{ are } -1, \text{ or one of} \\ & \text{them is } 0 \text{ and the other one is } -1, \\ 0, & \text{if both of } V^4_{r1} \text{ and } V^2_{r2} \text{ are } 0, \\ \{-1, 0, 1\}, & \text{if one of } V^4_{r1} \text{ and } V^2_{r2} \text{ is } 1 \text{ and} \\ & \text{the other one is } -1, \end{cases} \tag{5}$$

$$V^1_{r1r2} = \begin{cases} 1, & \text{if both of } V^1_{r1} \text{ and } V^1_{r2} \text{ are } 0, \text{ or one of} \\ & \text{them is } 0 \text{ and the other one is } 1, \\ -1, & \text{if both of } V^1_{r1} \text{ and } V^1_{r2} \text{ are } -1, \text{ or one of} \\ & \text{them is } 0 \text{ and the other one is } -1, \\ 0, & \text{if both of } V^1_{r1} \text{ and } V^1_{r2} \text{ are } 0, \\ \{-1, 0, 1\}, & \text{if one of } V^1_{r1} \text{ and } V^1_{r2} \text{ is } 1 \text{ and} \\ & \text{the other one is } -1, \end{cases} \tag{6}$$

$$V^4_{r1r2} = \begin{cases} 1, & \text{if both of } V^4_{r1} \text{ and } V^4_{r2} \text{ are } 0, \text{ or one of} \\ & \text{them is } 0 \text{ and the other one is } 1, \\ -1, & \text{if both of } V^4_{r1} \text{ and } V^4_{r2} \text{ are } -1, \text{ or one of} \\ & \text{them is } 0 \text{ and the other one is } -1, \\ 0, & \text{if both of } V^4_{r1} \text{ and } V^4_{r2} \text{ are } 0, \\ \{-1, 0, 1\}, & \text{if one of } V^4_{r1} \text{ and } V^4_{r2} \text{ is } 1 \text{ and} \\ & \text{the other one is } -1, \end{cases} \tag{7}$$

The relationship between r_1 and r_2 in transfer operation can be generalized to relational subsets I_1 and I_2.

Definition 4. Transitive operation of relation subset (⊙). Setting time interval α, β, and γ, and relational subsets I_1 and $I_2 \subseteq U$, if $\alpha\{I_1\}\beta$, $\beta\{I_2\}\gamma$, V_{r1}, and V_{r2} are their relationship vectors, then the relationship vector α between γ can be calculated by V_{r1} and V_{r2} with the operation ⊙ as follows:

$$\begin{aligned} V_{I_1 I_2} &= V^T_{I_1} \odot V_{I_2} \\ &= \begin{bmatrix} V_{r1} \\ V_{r2} \\ \vdots \\ V_{rm} \end{bmatrix} \left[V_{R1}\, V_{R2} \ldots V_{RN} \right], \end{aligned} \tag{8}$$

where $r_1, \ldots, r_m \in I_1$, $R_1, \ldots, R_n \in I_2$, and V_{r1} and V_{R1} are their corresponding relationship vectors.

3.2. Word-Level Text Representation with Tense Collocation Based on DRL. Considering the slow training speed of strategic network in DRL, it is necessary to pre-train the original DL model without strategic network to accelerate and stabilize strategic network training and verify its effect on word extraction related to classification tasks in text. Therefore, before designing the overall word-level text representation and classification model based on DRL, we determine a DL model with excellent text representation and classification performance as the basic model, which is selected based on bidirectional recurrent neural network and convolution neural network (BRCNN). This model consists of two parts, including bidirectional cyclic neural network and CNN, in which the bidirectional cyclic neural network is used to extract the forward and reverse word order information of the text, and CNN is used to obtain the feature representation of the text.

3.2.1. Input of the Model. We focus on the text representation and classification of sentences in this model, and the network input is the word vector matrix $(x_1, x_2, \ldots, x_N)^T \in R^{N \times D}$ corresponding to the ith word in the sentence, i.e., $x_i \in R^D$, D represents the dimension of the word vector, and N represents the number of words in the sentence.

3.2.2. Bidirectional Cyclic Neural Network. After the input vector $(x_1, x_2, \ldots, x_N)^T$ passes through the bidirectional cyclic neural network, the forward word order information and backward word order information can be obtained, which are $(\overrightarrow{h}_1, \overrightarrow{h}_2, \ldots, \overrightarrow{h}_N)$ and $(\overleftarrow{h}_1, \overleftarrow{h}_2, \ldots, \overleftarrow{h}_N)$, respectively. Attention mechanism contains two groups of weight factors, i.e., forward weighting factors $(\overrightarrow{\alpha}_1 \cdot \overrightarrow{h}_1, \overrightarrow{\alpha}_2 \cdot \overrightarrow{h}_2, \ldots, \overrightarrow{\alpha}_N \cdot \overrightarrow{h}_N)$ and backward weighting factor $(\overrightarrow{\alpha}_1 \cdot \overrightarrow{h}_1, \overleftarrow{\alpha}_t \cdot \overleftarrow{h}_2, \ldots, \overrightarrow{\alpha}_N \cdot \overrightarrow{h}_N)$, respectively. The word order information vectors output by the cyclic neural network at each forward and backward time are weighted

and cascaded to obtain $\overrightarrow{O}$ and $\overleftarrow{O}$. We cascade $\overrightarrow{O}$ and $\overleftarrow{O}$ to obtain the context vector O_{RNN}, which is shown as follows:

$$O_{RNN} = [\overrightarrow{O}, \overleftarrow{O}], \tag{9}$$

where $O_{RNN} \in R^{N\times 2K}$, in which K is the number of hidden layer neurons.

3.2.3. CNN Model. The input of CNN is the output O_{RNN} of bidirectional cyclic neural network. In BRCNN, the size of convolution kernel f is $S \times 2K$, where S is the window size of convolution kernel corresponding to the number of hidden states O_{RNN} in each convolution. $2K$ is the convolution kernel width, which is consistent with the vector dimension in O_{RNN}. The output O_{RNN} of bidirectional cyclic neural network can obtain the output vector after convolution operation of convolution kernel. The BRCNN model designed in this work uses three different convolution kernels f_1, f_2, and f_3, and their sizes are $S_1 \times 2K$, $S_2 \times 2K$, and $S_3 \times 2K$, respectively. After the output of the bidirectional cyclic neural network, i.e., O_{RNN}, passes through three convolution kernels f_1, f_2, and f_3, $O_1 \in R^{N\_S1+1}$, $O_2 \in R^{N\_S2+1}$, and $O_3 \in R^{N\_S3+1}$ can be obtained as shown in Figure 2.

We assume that the number of convolution kernels in each type is M, and $O_1 \in R^{M\times(N\_S1+1)}$, $O_2 \in R^{M\times(N\_S2+1)}$, and $O_3 \in R^{M\times(N\_S3+1)}$ can be obtained after the convolution operation. The pooling operation is often carried out to reduce the dimension of the features extracted by the convolution kernel and prevent the number of features from being too large and the classification model from being too fitted. In this work, maximum pooling is used to process the output of CNN, i.e., taking the maximum value of the second dimension in $O_1 \in R^{M\times(N\_S1+1)}$, $O_2 \in R^{M\times(N\_S2+1)}$, and $O_3 \in R^{M\times(N\_S3+1)}$ to obtain the vector, which is the result of pooling. As shown in Figure 2, for the convolution kernel f_1, we take the maximum value in the convolution result O_1 to obtain the pooled result of a convolution kernel. In the same way, we calculate the results of f_1 convolution kernels to obtain the m-dimensional vector $O_1' \in R^{M}$, $O_2' \in R^{M}$, and $O_3' \in R^{M}$, which will be used as input for the next phase of the model.

3.2.4. DRL Model. BRCNN model uses the strategic network of DRL to independently extract and classify temporal related words and delete irrelevant words. By predicting the behavior of each word in the sentence, the strategic network determines the deletion or retention behavior of words, i.e., sampling the behavior of each word to achieve the purpose of purifying text information.

In this work, we aim to classify the tenses of a single sentence sample X and produce collocation results. The word vector matrix of X can be expressed as $(x_1, x_2, \ldots, x_N)^T$, where x_t corresponds to the vector representation of the tth word. The status, behavior, and reward of policy network are introduced as follows. We note that the state of the policy network at time t is s_t ($t = 1, 2, \ldots, N$). In the BRCNN model, the state s_t ($t = 1, 2, \ldots, N$) is composed of the word vector input x_t of the tth word and the hidden state information obtained at time t of the bidirectional cyclic neural network. a_t

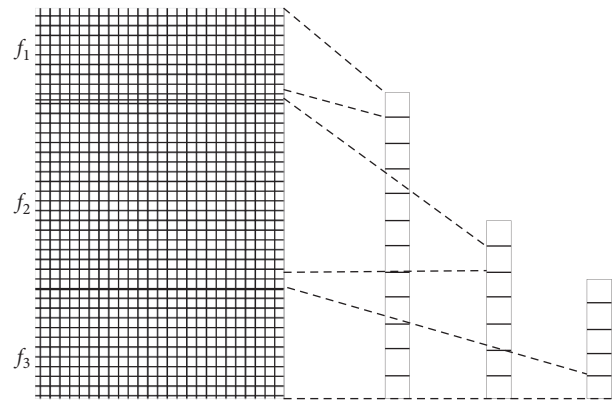


Figure 2: Convolution operation of the convolution layer.

($t = 1, 2, \ldots, N$) denotes the behavior of the policy network at time t. In the BRCNN model, the goal is to extract words related to tenses, and thus, the behavior a_t is set as the deletion or retention of the tth word. We set $a_t \in \{0, 1\}$, where 0 indicates deletion and 1 indicates reservation. During training stage, the strategic network samples the behavior of each word in probability according to the following equation:

$$\pi\left(\frac{a_t}{s_t;\Theta}\right) = \text{Sigmoid}\left(W_{RL}\cdot s_t + b_{RL}\right), \tag{10}$$

where $\pi(a_t|s_t;\Theta)$ denotes the probability of selecting behavior a_t in state s_t. The probability calculation of BRCNN model policy network adopts the sigmoid activation function, which can map each input value to the interval from 0 to 1 as the probability of retaining the current word. Then, we subtract this value from 1 to be the probability of word deletion. $\Theta = \{W_{RL}, b_{RL}\}$ represents the parameters of the BRCNN model policy network. When the policy network is trained and tested, the behavior a_t is selected according to (11), where $t = 1, 2, \ldots, N$.

$$a_t = \arg\max_a \pi\left(\frac{a_t}{s_t;\Theta}\right). \tag{11}$$

The reward value of the strategic network is denoted as R_N. In the BRCNN model, the behavior is sampled by the strategic network, the forward and backward word order information is extracted by the bidirectional cyclic neural network, and the word-level text is represented and classified by the CNN. The classification result label can be obtained to calculate the reward value R_N of the strategic network. $P(y|X)$ represents the probability of dividing text X into category y, where y represents category information and X is classification text.

y_{real} denotes the real label of sample X, and the probability of correct classification is $P(y_{\text{real}}|X)$. To delete as many independent words as possible and ensure the accuracy of classification, the number of words deleted is considered in the reward value function R_N. The number of words deleted by the strategic network is L, and the number of words deleted by X is N, and thus, the definition of reward value function R_N is shown as follows:

$$R_N = \log\left(P\left(y_{\text{real}}|X\right)\right) + \lambda \frac{L}{N}. \quad (12)$$

The reward value function R_N of the policy network combines the probability of correct classification with the word deletion ratio "L/N" for obtaining a higher probability of correct classification when deleting as many words irrelevant to the classification task as possible. λ is a super parameter that balances the two factors of correct classification probability $P(y_{\text{real}}|X)$ and word deletion ratio "L/N."

4. Simulation Experiments

This section compares and analyzes the experimental results of the word-level text representation and tense collocation model based on DRL. In this section, we introduce the content and source of the experimental dataset, and the setting of parameter in the experiment. All the experiments are conducted on 8 GB RAM, Intel Core i7 processor with 3.60 GHz, and Windows 10 operating system. The method is implemented using the MATLAB R2018b Win64.

4.1. Dataset Collection. There are five basic text classification datasets used in this experiment. The number of categories, the average length of text, the number of samples, the number of thesauri, and the number of samples in the test set are shown in Table 1:

(1) MPQA text dataset: MPQA opinion corpus comes from various news articles and is marked by Wiebe et al., which contain opinions and a series of private states, such as emotion, speculation, and belief

(2) CR text dataset: CR is customer comments on 14 Amazon products, which are divided into positive and negative user emotions

(3) MR text dataset: MR, i.e., movie reviews, is about customers' emotional evaluation of a series of films, including positive and negative emotions

(4) Subj text dataset: subjectivity (Subj) dataset is a subjective and objective emotion classification dataset

(5) TREC text dataset: TREC is a dataset for classifying problems

The initialization word vector in this experiment comes from the trained word vector of Google, i.e., GoogleNews-vectors-negative300.bin, which is trained by the CBOW model on the dataset Google News. Google News contains about 100 billion English words, and the CBOW model finally obtains the vector representation of 3 million words or phrases, in which the dimension of each word vector is 300. For the words that do not appear in the 3-million-word list, the model generates 300 dimensional vectors randomly by means of uniform distribution:

Network weight: the network parameters involved in the BRCNN model are initialized with a normal distribution with a mean value of 0 and a standard deviation of 0.1, and the bias terms are initialized to 0.1. The settings come from the optimal simulation based on different combinations.

Hyperparameters: the Adam method is used to train the model in the experiment, and the learning rate is 0.001. L2 regularization parameters with the size of 0.0001 are used in the full-connection layer to prevent overfitting. The data input batch is 64. There are 128 hidden units in the cyclic neural network. The CNN model contains three types of convolution kernels with window sizes of 3, 4, and 5, and 100 convolution kernels are set for each type. λ of reward function R_N in policy network is 0.1. Similarly, the setting method is to make different combinations of simulation and select the optimal simulation.

Table 1: A summary of the datasets.

Dataset	MPQA	CR	MR	Subj	TREC
Number of categories	2	2	2	2	6
Average length	2.9	18.5	19.5	22.4	10.7
Number of samples	10606	3775	10662	10000	5952
Number of thesauri	6246	5340	18765	21323	9592
Number of test samples	450	280	360	680	500

4.2. Results and Analysis. The results of comparative analysis of the words with the largest number of deletions or more typical words in the five datasets are shown in Table 2. The number of occurrences represents the total number of occurrences of a word in the dataset, the number of retentions is the number of words retained by the BRCNN model, and the retentions percentage is the ratio of retained words to the number of occurrences.

As can be seen from Table 2, the BRCNN model can effectively retain tense words in characterization data and make accurate tense collocation based on these words. In the CR dataset, "always" has the highest number of occurrences, but the classification accuracy is lower than "before," which shows that the temporal feature of "before" is stronger. In the MPQA dataset, the retention percentage of "normally" is relatively low, indicating that its temporal feature is weak. In MR and Subj datasets, the number of occurrences and retention percentage of "recently" are relatively high, which indicates that it is a typical word with strong temporal characteristics. In the TREC dataset, the number of occurrences of "yet" is low, but the retention percentage is high. This means that although the word is used less, the temporal feature is obvious.

Table 3 shows the accuracy of BRCNN in the English tense collocation proposed in this study. To objectively evaluate the results of temporal collocation, HSASRL [23], SRML [24], and Usr-mtl [25] are introduced as comparison models. It can be seen from the results in Table 3 that the accuracy of BRCNN is higher than HSASRL and Usr-mtl. Compared with the SRML model, the BRCNN model also obtains the highest accuracy in the MPQA, MR, and TREC databases. The reason for experimental results is that BRCNN optimizes the temporal description of sentences using the Allen interval algebra and establishes the

Table 2: BRCNN reserved word.

Datasets	Words	Number of occurrences	Number of retentions	Retention percentage
MPQA	Always	2228	1982	88.96
MPQA	Before	3419	3324	97.22
MPQA	After	6751	6624	98.11
MPQA	Since	2345	2134	91.00
MPQA	Now	12058	10859	90.05
MPQA	Usually	6843	6029	88.56
CR	Always	79558	70020	88.01
CR	Before	57485	56988	99.14
CR	After	48784	41194	84.44
CR	Since	2934	2756	93.93
CR	Now	3050	2691	88.23
MR	Always	5331	5126	96.15
MR	Before	2569	2366	92.10
MR	After	2302	2103	91.36
MR	Since	5037	4829	95.87
MR	Recently	8975	8690	96.82
MR	Comedy	1843	1657	89.90
Subj	Before	1665	1423	85.47
Subj	After	5641	5564	98.63
Subj	Since	4892	4688	95.83
Subj	Recently	9281	9024	97.23
Subj	Yet	6574	6322	96.117
Subj	Ever	9524	9367	98.35
TREC	Before	9661	9554	98.98
TREC	After	6541	6317	96.58
TREC	Since	5236	4594	87.74
TREC	Recently	8563	8364	97.67
TREC	Yet	2566	2496	97.27
TREC	Ever	9325	8864	95.05

Table 3: Temporal collocation accuracy of comparison models.

Dataset	HSASRL (%)	SRML (%)	Usr-mtl (%)	BRCNN (%)
CR	88.92	92.71	83.26	91.92
MPQA	84.51	89.44	80.86	93.76
MR	85.18	87.49	89.73	92.11
Subj	83.37	94.88	90.26	87.57
TREC	85.56	90.33	88.71	95.51
Avg accuracy	85.51	90.97	86.56	92.17

corresponding relationship between sentences and tenses by combining the high-precision text abstract with the bidirectional cyclic neural network and the convolution neural network. In addition, to further optimize the temporal expression of the text, BRCNN uses the intelligent decision-making ability of RL to effectively retain the tense words and completes the tense collocation with these words.

5. Conclusions

In this work, we propose a vector representation, that is, relational vector based on the relational expression of interval in Allen's interval algebra theory, and define the compound and transfer operations of relations on relational vectors, which provide support for the application of relational vectors in temporal reasoning. In addition, based on this temporal representation method, we propose a BRCNN model for temporal word classification and temporal collocation. The model realizes the text expression based on the bidirectional cyclic neural network and the convolutional neural network, and the classification of temporal words using the DRL model. It carries out the corresponding word collocation according to the calculated tenses. Through the experimental comparison of the five models on the datasets, i.e., MPQA, CR, MR, Subj, and TREC, it is proved that the BRCNN proposed in this work performs well in temporal collocation.

As a novel English tense evaluation mechanism, it has some limitations. For example, this study lacks the large-scale experiments. In addition, this study has no sufficient theory analysis. In future, we will address the above-mentioned limitations. In particular, we plan to make the practical experiments based on the constructed platform.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Heilongjiang Project: Research on Etiquette Discourse Based on Ecological Linguistics, under Grant No. 18YYC263.

References

[1] Y. Wang and W. Xu, "Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud," *Decision Support Systems*, vol. 105, no. 1, pp. 87–95, 2018.

[2] D. S. Kermany, M. Goldbaum, W. Cai et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[3] A. Khalil, M. Jarrah, M. Al-Ayyoub, and Y. Jararweh, "Text detection and script identification in natural scene images using deep learning," *Computers & Electrical Engineering*, vol. 91, no. C, pp. 107043–107058, 2021.

[4] J. Huan, A. A. Sekh, C. Quek, and D. K. Prasad, "Emotionally charged text classification with deep learning and sentiment semantic," *Neural Computing & Applications*, vol. 1, no. 1, pp. 1–11, 2021.

[5] N. Sun and C. Du, "News text classification method and simulation based on the hybrid deep learning model," *Complexity*, vol. 2021, no. 1, 11 pages, Article ID 8064579, 2021.

[6] H. Klaudel, M. Koutny, Z. Duan, and B. Moszkowski, "From box algebra to interval temporal logic," *Fundamenta Informaticae*, vol. 167, no. 4, pp. 323–354, 2019.

[7] F. Grandi, F. Mandreoli, R. Martoglia, and W. Penzo, "Unleashing the power of querying streaming data in a temporal database world: a relational algebra approach," *Information Systems*, vol. 103, no. 13, Article ID 101872, 2021.

[8] J. F. Allen, "Maintaining knowledge about temporal intervals," *Communications of the ACM*, vol. 26, no. 11, pp. 832–843, 1983.

[9] S. Huang, X. Li, Z. Zhang, F. Wu, and J. Han, "User-ranking video summarization with multi-stage spatio–temporal representation," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 2654–2664, 2018.

[10] E. Santos, "Cost-based temporal reasoning," *Information Sciences*, vol. 482, no. 1, pp. 392–418, 2019.

[11] A. Leeuwenberg and M.-F. Moens, "A survey on temporal reasoning for temporal information extraction from text," *Journal of Artificial Intelligence Research*, vol. 66, pp. 341–380, 2019.

[12] F. Ding, G. Ma, Z. Chen, J. Gao, and P. Li, "Averaged soft actor-critic for deep reinforcement learning," *Complexity*, vol. 2021, Article ID 6658724, 6 pages, 2021.

[13] X.-l. Chen, L. Cao, C.-x. Li, Z.-X. Xu, and J. Lai, "Ensemble network architecture for deep reinforcement learning," *Mathematical Problems in Engineering*, vol. 2018, Article ID 2129393, 6 pages, 2018.

[14] J. Liu and L. Feng, "Diversity evolutionary policy deep reinforcement learning," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5300189, 11 pages, 2021.

[15] X. Zhang, J. Zhao, and Y. Lecun, "Character-level convolutional networks for text classification," *Neural Information Processing Sys*, vol. 1, no. 1, p. 1, 2015.

[16] Y. Kim, "Convolutional neural networks for sentence classification," in *Proceedings of the Conference on Empirical methods in natural language processing*, pp. 1746–1751, Doha, Qatar, October 2014.

[17] T. Mikolov, M. Karafiat, L. Burget, C. Jan, and K. Sanjeev, "Recurrent neural network based language model," in *Proceedings of the 11th Annual Conference of the International Speech Communication Association, INTERSPEECH 2010*, pp. 1045–1048, Chiba, Japan, September 2010.

[18] S. Lai, L. Xu, K. Liu, and J. Zhao, "Recurrent convolutional neural networks for text classification," *Association for the Advancement of Artificial Intelligence*, vol. 333, pp. 2267–2273, 2015.

[19] D. Tang, B. Qin, and T. Liu, "Document modeling with gated recurrent neural network for sentiment classification," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 1422–1432, Lisbon, Portugal, September 2015.

[20] L. Yu, W. Zhang, J. Wang, and Y. Yu, "SeqGAN: sequence generative adversarial nets with policy gradient," in *Proceedings of the National Conference on Artificial Intelligence*, pp. 2852–2858, San Francisco, CA, USA, August 2017.

[21] B. Bennett and A. Cohn, "MultiDimensional multi-modal logics as a framework for spatio-temporal reasoning," *Applied Intelligence*, vol. 1, no. 1, p. 1, 2002.

[22] A. Morales, I. Navarrete, and G. Sciavicco, "Using temporal logic for spatial reasoning: temporalized propositional neighborhood logic," *Eurocast*, vol. 1, no. 1, pp. 313–320, 2007.

[23] F. Mohsen, J. Wang, and K. Al-Sabahi, "A hierarchical self-attentive neural extractive summarizer via reinforcement learning (HSASRL),," *Applied Intelligence*, vol. 2, no. 1, p. 1, 2020.

[24] D. Zhao, J. Wang, H. Lin et al., "Sentence representation with manifold learning for biomedical texts," *Knowledge-Based Systems*, vol. 218, no. 4, pp. 106869–106881, 2021.

[25] W. Xu, S. Li, and Y. Lu, "Usr-mtl: an unsupervised sentence representation learning framework with multi-task learning," *Applied Intelligence*, vol. 1, no. 1, pp. 1–16, 2020.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 8519708, 23 pages
https://doi.org/10.1155/2022/8519708

Research Article

A Novel UAV Path Planning Algorithm Based on Double-Dynamic Biogeography-Based Learning Particle Swarm Optimization

Yisheng Ji,[1] Xinchao Zhao,[1] and Junling Hao[2]

[1]*School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*School of Statistics, University of International Business and Economics, Beijing 100029, China*

Correspondence should be addressed to Xinchao Zhao; zhaoxc@bupt.edu.cn

Received 12 November 2021; Accepted 14 December 2021; Published 29 January 2022

Academic Editor: Jianhui Lv

Particle swarm optimization (PSO), one of the classical path planning algorithms, has been considered for unmanned aerial vehicle (UAV) path planning more frequently in recent years. A large amount of studies on UAV path planning based on modified PSO have been reported. However, most UAV path planning algorithms still optimize only one kind terrain problem which is mountain terrain. At the same time, many modified PSO algorithms also have some problems, such as insufficient convergence and unsatisfactory efficiency. In this paper, six kinds of terrain functions of UAV path planning are proposed to simulate real-world application. The terrain functions contain city, village without houses, village with houses, mountainous area without houses, mountainous area with houses, and mountainous area with a huge building. Inspired by CLPSO and BLPSO, we proposed a new double-dynamic biogeography-based learning particle swarm optimization (DDBLPSO) algorithm to solve these problems. The double-dynamic biogeography-based learning strategy replacing the traditional learning mechanism from the personal and global best particles is used to select the learning particles. In this strategy, each particle will learn from the better one of two selected particles which are not worse than itself. However, one random component of particle will replaced by corresponding component of other particle if all components of the particle only learn from itself. In this way, particles sufficiently learn from better objects and maintain the ability of jumping out of local optimality. The superiority of our algorithm is verified with four relevant algorithms, a PSO variant, and a BBO variant on the benchmark suite of CEC2015. Real-world application demonstrates that the algorithm we proposed outperforms four relevant algorithms, a PSO variant, and a BBO variant both in small-scale problems and large-scale problems. This paper shows a good application of our novel algorithm.

1. Introduction

UAV path planning is designed in a space which represents environment experienced by UAV during its flight. The path is a link formed by the combination of all the points and the lines connecting them, and each path is a solution. Let population $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_\mathbf{N}]$, where N is population size and X_i is the ith solution. Every path has m points, and it does not include the starting and the end points. $\mathbf{X_i} = [x_{i1}, y_{i1}, z_{i1}; \ldots; x_{iy}, y_{iy}, z_{iy}; \ldots; x_{im}, y_{im}, z_{im}]$, where (x_{iy}, y_{iy}, z_{iy}) is the coordinate of the jth point of the ith path. Solution is an $m \times 3$ matrix, and population is an $m \times 3 \times N$ tensor. The space is divided into $m + 1$ equal parts along X axis which ensures that the flight path goes from the starting point to the end point without forming a circle during halfway.

UAV path planning is a multiobjective constraint problem, which is described as follows:

$$\begin{aligned} &\min f_i(\mathbf{X}), \quad i = 1, 2, \ldots, m, \\ &c_j(\mathbf{X}) \le 0, \quad j = 1, 2, \ldots, p, \end{aligned} \tag{1}$$

where $f_i(\mathbf{X}), i = 1, 2, \ldots, m$, are objective functions and $c_j(\mathbf{X}) \le 0, j = 1, 2, \ldots, p$, are constraints. We transformed the multiobjective constrained optimization into single-objective unconstrained optimization in order to simplify the problem. All objectives and constraints are multiplied by a weight coefficient and then added. Finally, the objective

cost function is defined as equation (2) aiming to minimize the total cost.

$$F = \sum_{i=1}^{m} W_i \times f_i + \sum_{j=1}^{p} W_{j+m} \times c_j, \tag{2}$$

where W_i $(i = 1, 2, \ldots, m + p)$ are weight coefficients. In this paper, UAV aims to fly as short distance as possible, fly as low as possible, try not to cross danger zones as possible, and try not to cross the ground and buildings as possible. So, three objectives and one constraint are selected which are described as follows:

$$\begin{aligned} F = {} & W_1 \times f_{\text{length}} + W_2 \times f_{\text{altitude}} \\ & + W_3 \times f_{\text{danger}} + W_4 \times c_{\text{collision}}, \end{aligned} \tag{3}$$

where $W_1, W_2, W_3, W_4 \in (0, 1]$ are weight coefficients, f_{length} is a path function penalizing the longer paths, f_{altitude} is an altitude function penalizing higher average altitudes, f_{danger} is a danger function penalizing the paths going through danger zones, and $c_{\text{collision}}$ is a collision function penalizing the paths colliding with ground or buildings. f_{length}, f_{altitude}, f_{danger} are objective functions, and $c_{\text{collision}}$ is a constraint function.

Path function is used to penalize longer paths which are modelled as equation (4) in this paper.

$$f_{\text{length}} = 1 - \frac{L_{P1P2}}{L_{\text{traj}}}, \tag{4}$$

where L_{P1P2} is the length of the line segment which connects the starting point $P1$ and the end point $P2$ and L_{traj} is the actual length of the trajectory. This means that L_{traj} is the length of a broken line connecting all points from the starting point to the end point. It is easy to get that $f_{\text{length}} \in (0, 1]$.

Altitude function is used to penalize higher average altitudes which is modelled as equation (5) in this paper.

$$f_{\text{altitude}} = \frac{\left| A_{\text{traj}} - Z_{\min} \right|}{Z_{\max} - Z_{\min}}, \tag{5}$$

where A_{traj} is the average altitude and Z_{min} and $Z_{\max}$ are the upper and lower limits of the elevation in the search space. However, we usually choose a very low height rather than zero as the optimal choice which means that $Z_{\min}$ is a small positive number. It is easy to get $f_{\text{altitude}} \in (0, 1]$.

Danger function is used to penalize the paths going through danger zones which is modelled as equation (6) in this paper.

$$f_{\text{danger}} = \frac{L_{\text{danger}}}{\sum_{i=1}^{n} d_i}, \tag{6}$$

where L_{danger} is the sum of the length of the subsections of the trajectory which go through danger zones, d_i is the diameter of the ith danger zone, and n is the number of danger zones. In this paper, danger zone usually means that radar can detect and is described as a hemispheric region. If $f_{\text{danger}} > 1$, f_{danger} is set to be 1 to ensure $f_{\text{danger}} \in (0, 1]$.

Collision function is used to penalize the paths colliding with ground or buildings which is modelled as follows:

$$c_{\text{collision}} = \begin{cases} 0, & L_{\text{infeasible}} = 0, \\ P + \dfrac{L_{\text{infeasible}}}{L_{\text{traj}}}, & L_{\text{infeasible}} \neq 0, \end{cases} \tag{7}$$

where $L_{\text{infeasible}}$ is the total length of the subsections of the trajectory which travels below the ground or buildings, P is the penalty constant, and L_{traj} is the length of the trajectory. It is easy to say that $c_{\text{collision}} \in \{0\} \bigcup [P, P + 1]$. If cost function $F(\mathbf{X_i})$ is greater than P, it means that $F(\mathbf{X_i})$ is an infeasible solution.

There are many modified PSO algorithms, and some of them have been applied to UAV path planning [1–10]. However, many PSO variants still have some problems like insufficient convergence and accuracy and unsatisfactory efficiency. In this paper, we proposed a novel double-dynamic biogeography-based learning particle swarm optimization (DDBLPSO) algorithm to solve these problems. The algorithm not only considers the problem of insufficient convergence and accuracy and unsatisfactory efficiency for global optimization but also can be applied to UAV path planning. The main contributions of our paper are summarized as follows.

(i) A novel double-dynamic biogeography-based learning strategy replacing the traditional learning mechanism is proposed. This strategy makes the utmost of the advantages of particles which only learn from particles that are not worse than themselves.

(ii) Six kinds of terrain functions are designed in UAV path planning.

(iii) The experiment results show that DDBLPSO demonstrates the best performance on UAV path planning with four other representative algorithms, a PSO variant, and a BBO variant in CEC2015 benchmark functions and UAV path planning.

The remainder of the paper is organized as follows. Section 2 gives the literature review. Section 3 introduces several existing algorithms such as PSO, BBO, CLPSO, and BLPSO. Section 4 shows the details of the proposed algorithm DDBLPSO. Section 5 presents simulation results for global optimization. Application in UAV path planning is elaborated in Section 6. Section 7 draws the conclusions.

2. Literature Review

Unmanned aerial vehicle (UAV) has been applied to many areas with the rapid progress of science and technology. Evolutionary algorithm-based methods solving the UAV path planning problem are always a hot topic. In [1], the authors proposed an evolutionary algorithm based on offline/online path planning for UAV navigation. Phase angle-encoded and quantum-behaved particle swarm optimization was proposed and applied to three-dimensional route

planning for UAV [2]. Pehlivanoglu [3] introduced a new vibrational genetic algorithm which was enhanced by Voronoi diagram for path planning of autonomous UAV. An improved constrained differential evolution algorithm was introduced for UAV global route planning [4]. The algorithm demonstrates a good performance in terms of the solution quality, robustness, and the constraint-handling ability. A method was proposed to compare the planner performance by jointly employing several general and problem-specific quality indexes, which takes into account the complexity and particularities of the problem [5]. In [6], a novel predator-prey pigeon-inspired optimization (PPPIO) was proposed to solve the three-dimension path planning problem for unmanned combat air vehicle (UCAV) in dynamic environment. This algorithm mainly focuses on optimizing the flight route considering different types of constraints under complex combating environment. The modified wolf pack search (WPS) was applied to compute the quasi-optimal trajectories for the rotor wing UAVs in the complex three-dimensional (3D) spaces including the real and fake 3D spaces [7]. In [8], the authors proposed a constrained adaptive multiobjective differential evolution algorithm for bistatic SAR path planning. It generates multiple feasible paths for the UAV receiver with different trade-offs between navigation for UAV and bistatic SAR imaging performance. The authors compared genetic algorithm and particle swarm optimization for real-time UAV path planning in [9].

Particle swarm optimization (PSO) [10] is one of the most commonly used algorithms to solve the problems of UAV path planning. In recent years, many improved PSO variants have emerged one after another. In [11], the authors put forward an adaptive particle swarm optimization (APSO) algorithm, in which two main steps are conducted to adaptively adjust the parameters when the swarm lies in a different evolutionary state (exploration, exploitation, convergence, and jumping out) in each generation. Then, an elitist learning strategy was used when the evolutionary state was classified as convergence state. Nickabadi et al.[12] studied on adaptive inertia weight and introduced a novel particle swarm optimization algorithm. Numerical experiments show that this algorithm is quite effective in adapting the value of w in the dynamic and static environments. Li and Yao [13] presented a new cooperative coevolving particle swarm optimization (CCPSO) algorithm to scale up particle swarm optimization algorithms in solving large-scale optimization problems (up to 2000 real-valued variables). Comprehensive learning particle swarm optimization (CLPSO) uses a learning strategy whereby all other particles' historical best information is used to update a particle's velocity [14]. In [15], the authors proposed a novel biogeography-based optimization algorithm with momentum migration and taxonomic mutation to deal with problems whose function values change dramatically or barely. In [16], the authors proposed an enhanced particle swarm optimization algorithm (pkPSO) by combining k-nearest neighbours (k-NN) with pattern search (PS). At the same time, the cooperative effect of k-NN and PS strategies was verified. A novel particle swarm optimization algorithm was proposed for parameter determination and feature selection of support vector machines [17]. A novel PSO-GA-based hybrid training algorithm with Adam optimization which performs well in training artificial neural networks was introduced by Yadav and Anubhav [18]. In [19], the authors presented a cellular learning automata-based bare bones PSO algorithm with maximum likelihood rotated mutations. The PSO technique was used to identify the uncertain physical parameters of a real vehicle ETC system [20]. In [21], the authors proposed a new particle swarm optimization algorithm with simple PID-based strategy, which has good global optimization ability. Biogeography-based optimization (BBO) is an algorithm that simulates biological migration to search optimal solution [22]. Biogeography-based learning strategy is a nice method for optimization. In [23], biogeography-based learning particle swarm optimization (BLPSO) uses a learning strategy based on migration of biogeography-based optimization and outperforms other representative algorithms. In [24], the authors introduced a hybrid differential evolution with biogeography-based optimization for global numerical optimization.

3. Relevant Algorithms

3.1. Particle Swarm Optimization (PSO). Particle swarm optimization is a classical evolutionary computing technique which was firstly proposed by Eberhart and Kennedy [10]. Inspired from the study of the predation behaviour of birds, the main idea of PSO algorithm is sharing cooperation and information among individuals to find the optimal solution. PSO simulates birds in a flock by designing a massless particle, which only has velocity and position. Velocity represents how fast and which direction the particle moves along. Position represents where the particle is. For each particle, only the personal best experience and the global best experience of the entire swarm can be learned. Let $x_i = (x_{i1}, x_{i2}, \ldots, x_{i\,D})^T$ and $v_i = (v_{i1}, v_{i2}, \ldots, v_{i\,D})^T$ denote the position and velocity of particle $i\,(i = \{1, 2, \ldots, N\})$, respectively, where D is dimension of the initial space and N is population size. Let $\text{pbest}_i = (\text{pbest}_{i1}, \text{pbest}_{i2}, \ldots, \text{pbest}_{i\,D})^T$ and $\text{gbest} = (\text{gbest}_1, \text{gbest}_2, \ldots, \text{gbest}_D)^T$ be the personal best position of particle i and the global best position of the whole swarm. The update of velocity and position of the particle is indicated as follows:

$$\begin{aligned} v_{\text{id}} &= w * v_{\text{id}} + c_1 * \text{rand}_1(0,1) * (\text{pbest}_{\text{id}} - x_{id}) \\ &\quad + c_2 * \text{rand}_2(0,1) * (\text{gbest}_d - x_{id}), \\ x_{id} &= x_{id} + v_{id}, \end{aligned} \tag{8}$$

where $i = 1, 2, \ldots, N$; $d = 1, 2, \ldots, D$; w is the inertia weight; c_1 and c_2 are the acceleration coefficients; and $\text{rand}_1(0,1)$ and $\text{rand}_2(0,1) \in [0,1]$ are uniform random numbers. The algorithm sets the maximum velocity to prevent the velocity from getting too large. The particle will be set as the border of initial region if it flies out of initial region.

3.2. Biogeography-Based Optimization (BBO). In [22], the author proposed a new algorithm named biogeography-based optimization which solves optimization problems by

simulating biological migration. In BBO, solution is also called as habitat, fitness of solution is called as habitat suitability index (HSI), and component of solution is called as suitable index variable (SIV). The main body of BBO is migration operation and mutation operation.

Each habitat $x_i\,(i = 1, 2, \ldots, N)$ has two parameters to describe the migration, which are an immigration rate λ_i and an emigration rate μ_i. Both parameters are closely related to HSI. For high HSI habitat, there will be a high trend of outward migration. At this time, the emigration rate is high and the immigration rate is low due to the pressure of species competition. However, for low HSI habitat, the opposite is true. Assuming habitat x_i currently accommodates S_i species, $S_{\max}$ is the maximum number of species. Immigration rate λ_i and emigration rate μ_i are usually described as follows:

$$\begin{cases} \lambda_i = I * \left(1 - \dfrac{S_i}{S_{\max}}\right), \\ \mu_i = E * \dfrac{S_i}{S_{\max}}, \end{cases} \tag{9}$$

where I is the maximum immigration rate and E is maximum emigration rate. The component $x_{\mathrm{id}}\,(d = 1, 2, \ldots, D)$ of habitat x_i will immigrate depend on immigration rate λ_i. The habitat x_j is selected depending on emigration rate μ_j.

The properties of the habitat such as HSI and the number of species may changes due to unexpected events. At the same time, mutation rate depends on species probability. According to biogeography, when the number of species in habitat is too large or too small, species probability is low. When the number of species in habitat is moderate, species probability is high. Mutation rate m_i is described as follows:

$$m_i = m_{\max} * \left(1 - \frac{P_i}{P_{\max}}\right), \tag{10}$$

where P_i is the species probability of x_i decided by the number of species S_i, $P_{\max}$ is the maximum species probability, and $m_{\max}$ is the maximum mutation rate.

3.3. Comprehensive Learning Particle Swarm Optimization (CLPSO). Comprehensive learning particle swarm optimization [14] adopts the strategy of comprehensive learning to select the objects to learn instead of learning from itself and the global optimal individual. The velocity updating equation in CLPSO is defined as follows:

$$v_{\mathrm{id}} = w * v_{\mathrm{id}} + c * \mathrm{rand}(0, 1) * \left(\mathrm{pbest}_{f_i(d),d} - x_{id}\right), \tag{11}$$

where f_i defines which particles' pbests that the particle i should follow and $\mathrm{rand}(0, 1) \in [0, 1]$ is a uniform random number. The inertia weight w in CLPSO is a linear attenuation coefficient. CLPSO assigns a learning probability Pc_i for each particle i using the following equation:

$$Pc_i = 0.05 + 0.45 * \frac{\exp(10(i-1)/N - 1) - 1}{\exp(10) - 1}. \tag{12}$$

For each solution x_i, it will learn from many particles instead of only two particles. Each component of particle i learn from itself or other particle depending on learning probability Pc_i. A random component of particle i will learn from other particle if all its components learn from itself. The higher fitness value a solution has, the greater possibility that particle will be learned.

3.4. Biogeography-Based Learning Particle Swarm Optimization (BLPSO). Biogeography-based learning particle swarm optimization [23] adopts the strategy of migration operation of BBO to select the objects to learn. Each component of particle i learns from itself or a selected particle j depending on immigration rate λ_i. The selected particle j depends on emigration rate μ_j. A random component of particle i will learn from other particle if all components of particle i learn from itself. The higher the fitness value of a solution is, the smaller the immigration rate and the greater the migration rate of the particle are. It means that the better particle has larger probability to be learned and the worse particle has the larger probability to learn from others. The velocity updating equation in BLPSO is the same as equation (11). The learning probability Pc_i for particle i and the selected probability Ps_j for particle j can be expressed as follows:

$$\mathrm{Pc}_i = \lambda_i, \tag{13}$$

$$\mathrm{Ps}_j = \frac{\mu_j}{\sum_{k=1}^{N} \mu_k}, \tag{14}$$

It should be noted that learning probability Pc is different from selected probability Ps. Particle i has a probability of learning probability Pc_i to learn from other particles. At the same time, particle j has a selected probability Ps_{ij} to be selected as an object which particle i learn from. The relationship between learning probability and selected probability is shown in Figure 1.

It is worth mentioning that it uses the quadratic migration model instead of the linear migration model. The quadratic migration model is as follows:

$$\begin{cases} \lambda_i = I * \left(1 - \dfrac{S_i}{S_{\max}}\right)^2, \\ \mu_i = E * \left(\dfrac{S_i}{S_{\max}}\right)^2. \end{cases} \tag{15}$$

The differences of CLPSO and BLPSO are learning probability and the selection strategy of the learning particle.

4. Double-Dynamic Biogeography-Based Learning Particle Swarm Optimization (DDBLPSO)

Although many promising PSO variants have emerged, they still have some problems like insufficient convergence and accuracy and unsatisfactory efficiency. These cause them to

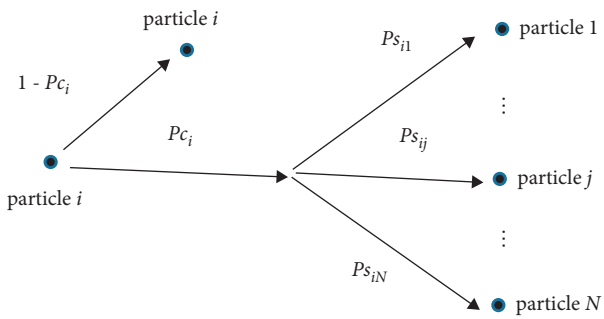


Figure 1: Relationship between learning probability and selected probability.

perform poorly on some complicated functions such as multimodal functions, hybrid functions, composition functions, and real-world problems. Regardless of CLPSO or BLPSO, there is a probability that particle learns from those particles with smaller fitness than itself. These strategies lead some superior particles are not taken full advantage of their superiority and result a low convergence. A double-dynamic selection strategy based on biogeography-based learning is proposed in this paper, in which particles learn from the better one of two dynamic roulette winners. This strategy guides the particle to learn from others that are not worse than itself. In this way, the advantages of the superior particles will be greatly exploited.

4.1. Dynamic Biogeography-Based Learning. Original biogeography-based learning strategy guides particles to learn from the promising particles. However, most particles still have a certain probabilities of learning from worse particles. For each particle, dynamic biogeography-based learning strategy only guides it to learn from other particles not worse than itself. All the particles are sorted in ascending order based on fitness. Particle i will learn from a selected particle if a uniform random number generated in [0,1] is smaller than predefined λ_i. The selected probability Ps_{ij} of particle j being selected by particle i is defined as follows:

$$\mathrm{Ps}_{ij} = \frac{\mu_j}{\sum_{k=i}^{N} \mu_k}, \tag{16}$$

where $j \geq i$. The greater the index number is, the higher fitness the particle has. The learning object particle j is selected with a dynamic roulette which constitutes even more competitive individuals $\mu_i, \mu_{i+1}, \ldots, \mu_N$. In this way, particle will not learn from the worse particles which can reduce the possibility for the particles getting worse and worse. Generally speaking, on the one hand, dynamic biogeography-based learning strategy keeps the worse particles larger exploration abilities to learn from others. On the other hand, the strategy gives the better particles fewer objectives to learn from. It should be pointed out that the best solution only can learn from itself.

4.2. Double-Dynamic Biogeography-Based Learning. Original biogeography-based learning strategy only has one selected particle to learn while original comprehensive learning strategy has two selected particles and then chooses the better one. The strategy has two selected particles, which has a greater possibility to have a high fitness offspring than other strategy only has one selected particle. Inspired by comprehensive learning strategy, double-dynamic biogeography-based learning also has two selected particles and leaves the better one to ensure the high rate of convergence. This strategy allows particles to adequately learn from better particles. For each component x_{id}, the learning objective $x_{Li,d}$ is selected by using the following equation:

$$x_{\mathrm{Li},d} = \begin{cases} x_{\mathrm{ad}}, & \text{if fitness of } x_a \text{ is greater than fitness of } x_b, \\ x_{\mathrm{bd}}, & \text{if fitness of } x_a \text{ is smaller than fitness of } x_b, \end{cases} \tag{17}$$

where x_a and x_b are two particles which are selected by dynamic roulette and x_{ad} and x_{bd} represent the dth component of x_a and x_b, respectively.

4.3. Mutation. According to above strategies, although the particles will fly to a better area, particles are easy to fall into local optimum. One random component of particle will be replaced by other particle's component if the particle only learns from itself. Only high fitness particles have large probability to learn from themselves. There are two reasons for this phenomenon. One is that a high fitness particle has a larger immigration rate. The other is that a high fitness particle only has few particles better than itself. This strategy helps particles to jump out of local optimum. If all components of x_i only learn from themselves, the learning objective x_{Li} is set as follows:

$$x_{\mathrm{Li},j} = x_{l,j}, \quad \text{if } x_{\mathrm{Li}} == x_i, \tag{18}$$

where j and l are random numbers and $l \in \{1, 2, \ldots, i-1, i+1, \ldots, N\}$, $j \in \{1, 2, \ldots, D\}$.

In the early stage of DDBLPSO, the algorithmic search area is large because initial particles evenly distribute in the search area. In the last stage of DDBLPSO, the algorithmic search area becomes smaller and smaller because most particles tend to be very close together. However, the better particle will have the greater probability to learn from other particle's past optimal position because of the mutation mechanism. It prevents the algorithm from falling into local optimum.

4.4. Algorithm Process. Double-dynamic biogeography-based learning particle swarm optimization is proposed based on above strategies. The process of DDBLPSO is described in Algorithm 1.

DDBLPSO, CLPSO, and BLPSO have some commonalities and differences. It is necessary to illustrate the differences among them. Two main differences are summarized as follows.

(i) In CLPSO, learning objective is selected uniformly and randomly. In BLPSO, learning objective is selected by roulette based on equation (14). In

```
Initial maximum velocity v_max, upper bound x_max and lower bound x_min of the initial region, population x, velocity v, inertia
  weight w, acceleration coefficient c, the maximal number of function evaluations maxFES;
Record personal best position pbest and global best position gbest;
while FES < maxFES do
   Sort solutions in ascending order based on fitness;
   Calculate immigration rate λ_i and emigration rate μ_i (i = 1, 2, ..., N), update inertia weight w, and let learning population
  xL = zeros(size(x));
   for i = 1 to N
      for d = 1 to D
         If rand < λ_i
            Select pbest_a and pbest_b (a, b ∈ {1, 2, ..., N}) with a dynamic roulette;
            if fitness of x_a is greater than fitness of x_b
               xL_id = pbest_ad;
            else
               xL_id = pbest_bd;
            end if
         else
            xL_id = pbest_id
         end if
      end for
      if all(xL_i == pbest_i)
         Randomly select j ≠ i (j ∈ {1, 2, ..., N}) and l (l ∈ {1, 2, ..., D});
         xL_id = pbest_id;
      end if
   end for
   v = w * v + c * rand(size(x)). * (xL − x).;
   v(v > v_max) = v_max; v(v < − v_max) = −v_max;
   x = x + v;
   x(x > x_max) = x_max; x(x < x_min) = x_min;
   Update personal best position pbest and global best position gbest;
end while
Output the final result.
```

Algorithm 1: Double-dynamic biogeography-based learning particle swarm optimization (DDBLPSO).

DDBLPSO, learning objective is selected by the dynamic roulette based on equation (16).

(ii) In BLPSO, each component of a particle only has one learning objective. However, in CLPSO and DDBLPSO, each component of a particle has two learning objectives and then selects the better one.

5. Numerical Simulations for Global Optimization

5.1. Test Functions and Parameter Settings. CEC2015 [25] benchmark functions are used to verify the performance of DDBLPSO. CEC2015 benchmark suite is simply introduced as follows. $f1$ and $f2$ are unimodal functions, $f3-5$ are multimodal functions, $f6-8$ are hybrid functions and $f9-15$ are composition functions.

CLPSO [14] is a classical algorithm proposed based on PSO [10], and BLPSO [23] is an algorithm proposed based on BBO [22] and PSO. Inspired by CLPSO and BLPSO, we proposed DDBLPSO. DDBLPSO has an obvious relationship with BBO, PSO, CLPSO, and BLPSO, so we selected these four algorithms for comparison experiments. At the same time, we also selected two other representative algorithms which are PBSPSO [21] and DEBBO [24] for comparison experiments. We conducted four sets of numerical experiments which are labelled as experiment one, experiment two, experiment three, and experiment four. DDBLPSO is compared with four relevant algorithms (PSO, BBO, CLPSO, and BLSPSO), a PSO variant (PBSPSO), and a BBO variant (DEBBO). Design of the four experiments is shown in Table 1. The maximal number of function evaluations (maxFES) is set as $10000 * D$. Search range is $[-100, 100]^D$, and 30 independent runs are conducted in MATLAB 2017b. Other parameters are shown in Table 2.

5.2. Experimental Results and Discussion. Results of experiment one, experiment two, experiment three, and experiment four are statistically shown in Tables 3–6, respectively. Tables 7–10 show the ranks of five algorithms according to the Friedman test of experiment one, experiment two, experiment three, and experiment four, respectively. Min, Mean, Median, and Std indicate the minimum function error value, the mean function error value, the median function error value, and the standard deviation of error values, respectively. In order to exhibit the evolution trend of five algorithms more vividly, converging curves of the average best fitness of functions in experiment one and experiment three are shown in Figures 2 and 3. The converging

Table 1: Design of experiment one, experiment two, experiment three, and experiment four.

Experiment	Population size N	Points number mp	Comparison algorithm
Experiment one	30	30	PSO, BBO, CLPSO, BLSPSO
Experiment two	100	100	PSO, BBO, CLPSO, BLSPSO
Experiment three	30	30	PBSPSO, DEBBO
Experiment four	100	100	PBSPSO, DEBBO

Table 2: Parameters of seven comparison algorithms.

Algorithm	Parameters
PSO	Inertia weight $w = 0.55$, acceleration coefficient $c_1 = c_2 = 2.0$, maximum velocity $v_{\max} = 5.0$
BBO	Maximum immigration rate $I = 1$, maximum emigration rate $E = 1$, maximum mutation rate $m_{\max} = 0.005$
CLPSO	Inertia weight w is linearly decrease from 0.9 to 0.2, acceleration coefficient $c = 2.0$, maximum velocity $v_{\max} = 5.0$
BLPSO	Inertia weight w is linearly decrease from 0.9 to 0.2, acceleration coefficient $c = 2.0$, maximum velocity $v_{\max} = 5.0$, maximum immigration rate $I = 1$, maximum emigration rate $E = 1$
PBSPSO	Inertia weight $w = 0.55$, acceleration coefficient $c = 2.0$, maximum velocity $v_{\max} = 5.0$, decay term $\alpha = 0.9$, the weight of the derivative term $K_d = 0.03$
DEPSO	Maximum immigration rate $I = 1$, maximum emigration rate $E = 1$, maximum mutation rate $m_{\max} = 0.005$, crossover rate CR = 0.5, difference coefficient $F_X = 0.7$
DDBLPSO	Inertia weight w is linearly decrease from 0.9 to 0.2, acceleration coefficient $c = 2.0$, maximum velocity $v_{\max} = 5.0$, maximum immigration rate $I = 1$, maximum emigration rate $E = 1$

Table 3: Performance comparison of five algorithms in experiment one.

	BBO	PSO	CLPSO	BLPSO	DDBLPSO
$f1(x)$					
Min	1.6583 E+06	1.7321 E+06	7.7577 E+05	5.6816 E+05	2.0171 E+059
Mean	4.1619 E+06	1.1235 E+07	1.5641 E+06	1.3061 E+06	.9974 E+058
Median	4.2666 E+06	9.4901 E+06	1.5160 E+06	1.0850 E+06	0404E+05
Std	1.2125 E+06	8.1069 E+06	3.9988 E+05	6.4276 E+05	7.7346 E+05
$f2(x)$					
Min	5.9570 E+07	2.1247 E+06	8.6937 E+04	1.0573 E+03	**2.0057 E+02**
Mean	1.0746 E+08	6.4067 E+06	4.1390 E+05	7.6477 E+03	**3.0520 E+03**
Median	1.0441 E+08	5.5692 E+06	3.5350 E+05	6.7947 E+03	**1.3179 E+03**
Std	2.2645 E+07	3.1246 E+06	2.9045 E+05	5.3207 E+03	**4.2472 E+03**
$f3(x)$					
Min	3.2033E+02	**3.2014E+02**	3.2118 E+02	3.2109 E+02	3.2104 E+02
Mean	3.2093E+02	**3.2026E+02**	3.2134 E+02	3.2131 E+02	3.2128 E+02
Median	3.2092E+02	**3.2026E+02**	3.2135 E+02	3.2131 E+02	3.2129 E+02
Std	3.1412E−01	**6.9821E-02**	7.9537E-02	9.3207E-02	1.1575E-01
$f4(x)$					
Min	5.5276 E+02	4.3839 E+02	4.1498 E+02	4.1592 E+02	4.0995 E+02
Mean	5.8999 E+02	4.5826 E+02	4.2232 E+02	4.2498 E+02	4.1878 E+02
Median	5.8819 E+02	4.5863 E+02	4.2201 E+02	4.2388 E+02	4.1840 E+02
Std	2.3338 E+01	1.1451 E+01	**4.1135 E+00**	6.1503 E+00	5.2623 E+00
$f5(x)$					
Min	2.7940 E+03	1.8947 E+03	1.2380 E+03	1.2387 E+03	1.3481 E+03
Mean	4.2481 E+03	2.6178 E+03	1.9363 E+03	1.6544 E+03	1.7612 E+03
Median	4.1462 E+03	2.5958 E+03	1.9451 E+03	1.5976 E+03	1.7362 E+03
Std	7.4098 E+02	4.0091 E+02	3.1450 E+02	2.9163 E+02	2.5296 E+02
$f6(x)$					
Min	8.5892 E+04	4.7158 E+05	1.3661 E+05	3.4194 E+04	2.7682 E+04
Mean	3.3191 E+05	5.7133 E+06	3.0127 E+05	1.5594 E+05	1.1685 E+05
Median	2.8357 E+05	5.4783 E+06	2.7053 E+05	1.5365 E+05	9.0779 E+04
Std	1.8363 E+05	3.7012 E+06	1.1982E+05	8.9082 E+04	8.6045 E+04
$f7(x)$					
Min	7.0813 E+02	7.0836 E+02	7.0626 E+02	7.0570 E+02	7.0517 E+02
Mean	7.1215 E+02	7.1799 E+02	7.1138 E+02	7.0934 E+02	7.0754 E+02

Table 3: Continued.

	BBO	PSO	CLPSO	BLPSO	DDBLPSO
Median	7.1023 E+02	7.1372 E+02	7.1242 E+02	7.0795 E+02	7.0586 E+02
Std	4.0321 E+00	1.7801 E+01	2.8718 E+00	3.1576 E+00	2.8994 E+00
$f8(x)$					
Min	2.9403 E+04	3.7955 E+04	2.1194 E+04	1.1270 E+04	1.1820 E+04
Mean	1.1430 E+05	1.5521 E+06	6.0038 E+04	4.7864 E+04	4.5841 E+04
Median	9.0699 E+04	1.1775 E+06	5.8822 E+04	4.5294 E+04	3.1882 E+04
Std	9.6038 E+04	1.4589 E+06	2.6934 E+04	2.8916 E+04	3.5138 E+04
$f9(x)$					
Min	1.0045 E+03	1.0046 E+03	1.0033 E+03	1.0032 E+03	1.0024 E+03
Mean	1.0740 E+03	1.0059 E+03	1.0040 E+03	1.0076 E+03	1.0037 E+03
Median	1.0064 E+03	1.0058 E+03	1.0040 E+03	1.0038 E+03	1.0037 E+03
Std	1.3863 E+02	7.2713E−01	2.7210E−01	2.0779 E+01	5.1535E−01
$f10(x)$					
Min	2.0119 E+04	4.0609 E+05	7.5343 E+04	2.9240 E+04	2.1702 E+04
Mean	5.0981 E+05	2.5287 E+06	2.0769 E+05	2.0526 E+05	1.0437 E+05
Median	7.9737 E+04	1.7787 E+06	2.1072 E+05	1.1130 E+05	7.1822 E+04
Std	9.4063 E+05	2.2614 E+06	7.6381 E+04	4.0970 E+05	1.0132 E+05
$f11(x)$					
Min	1.4042 E+03	1.8057 E+03	1.4035 E+03	1.4015 E+03	1.4013 E+03
Mean	1.6772 E+03	1.9459 E+03	1.4617 E+03	1.5100 E+03	1.4753 E+03
Median	1.4285 E+03	1.9471 E+03	1.4056 E+03	1.4031 E+03	1.4021 E+03
Std	3.1792 E+02	6.2006E+01	1.1830 E+02	1.5473 E+02	1.1947E+02
$f12(x)$					
Min	1.3100 E+03	1.3084 E+03	1.3040 E+03	1.3049 E+03	1.3040 E+03
Mean	1.3214 E+03	1.3111 E+03	1.3058 E+03	1.3066 E+03	1.3071 E+03
Median	1.3139 E+03	1.3110 E+03	1.3055 E+03	1.3064 E+03	1.3075 E+03
Std	2.1131 E+01	1.5834 E+00	9.8335E-01	1.1492 E+00	1.4291 E+00
$f13(x)$					
Min	1.4175 E+03	1.3933 E+03	1.4352 E+03	1.4144 E+03	1.3995 E+03
Mean	1.4252 E+03	1.4078 E+03	1.4420 E+03	1.4259 E+03	1.4173 E+03
Median	1.4244 E+03	1.4086 E+03	1.4416 E+03	1.4254 E+03	1.4164 E+03
Std	4.8717 E+00	6.6629 E+00	3.4109 E+00	6.0813 E+00	7.2068E+00
$f14(x)$					
Min	3.7298 E+03	3.2799 E+04	3.2503 E+04	3.2503 E+04	1.5000 E+03
Mean	3.7381 E+04	3.4675 E+04	3.3058 E+04	3.3235 E+04	3.2591 E+04
Median	3.8788 E+04	3.4793 E+04	3.2651 E+04	3.2651 E+04	3.3648 E+04
Std	1.0738 E+04	9.1682 E+02	7.4020 E+02	1.0085 E+03	5.9283 E+03
$f15(x)$					
Min	1.6042 E+03	1.6007 E+03	1.6000 E+03	1.6000 E+03	1.6000 E+03
Mean	1.6105 E+03	1.6012 E+03	1.6001 E+03	1.6000 E+03	1.6004 E+03
Median	1.6127 E+03	1.6011 E+03	1.6001 E+03	1.6000 E+03	1.6000 E+03
Std	5.5047 E+00	1.9659E-01	3.2672E-02	4.7803E-03	2.3806 E+00

curves of functions in experiment two and experiment four are not shown due to their similarities with experiment one and experiment three, respectively. The boldface in Tables 3–6 indicates the best experimental results in experiment one to four, respectively.

Unimodal functions are used to explore the convergence rate of the optimization problem. Multimodal functions are used to explore the ability to jump out of local optimum. Hybrid functions consider that in the real-world optimization problems, different subcomponents of the variables may have different properties. Composition functions consider that in the real-world optimization problems, different simple problems composite a more complex problem. This is a challenge for algorithmic convergence rate. Most of the multimodal functions contain a number of local optima, which may lead to premature convergence of traditional algorithms. It is difficult for traditional algorithms to locate the global optimum for these functions.

It is easy to observe that DDBLPSO shows its dominance over other competitors on unimodal functions and hybrid functions from Table 3. However, DDBLPSO performs less on multimodal functions and composition functions. This means that DDBLPSO has much greater convergence rate than its competitors and performs slightly better than its competitors in jumping out of local optimum. The double-dynamic biogeography-based learning strategy is the main reason for this phenomenon. This

Table 4: Performance comparison of five algorithms in experiment two.

	BBO	PSO	CLPSO	BLPSO	DDBLPSO
$f1(x)$					
Min	1.6872 E+07	1.2995 E+07	1.2592 E+07	9.6047 E+06	**8.8196 E+06**
Mean	3.3932 E+07	2.4028 E+07	2.0317 E+07	1.2281 E+07	**1.1151 E+07**
Median	3.3742 E+07	2.3245 E+07	2.0301 E+07	1.2062E+07	**1.1181 E+07**
Std	5.2575 E+06	6.4139 E+06	2.7277 E+06	1.4734 E+06	**1.0671 E+06**
$f2(x)$					
Min	5.5282 E+08	**1.3908 E + 06**	2.1847 E+08	9.6042 E+07	2.3631 E+07
Mean	6.4517 E+08	**2.2934 E + 06**	2.5524 E+08	1.2074E+08	3.2844 E+07
Median	6.3717 E+08	**2.1154 E + 06**	2.5419 E+08	1.2030E+08	3.2473 E+07
Std	5.2834 E+07	**5.7400 E + 05**	1.7095 E+07	1.3223 E+07	4.2897 E+06
$f3(x)$					
Min	3.2133 E+02	**3.2079E + 02**	3.2134 E+02	3.2139 E+02	3.2136 E+02
Mean	3.2145 E+02	**3.2085E + 02**	3.2148 E+02	3.2146 E+02	3.2145 E+02
Median	3.2147 E+02	**3.2086E + 02**	3.2148 E+02	3.2145 E+02	3.2145 E+02
Std	5.8776E-02	**3.7493 E - 02**	3.9590E-02	4.0692E-02	3.9328E-02
$f4(x)$					
Min	9.1048 E+02	5.0448 E+02	5.9808 E+02	5.2056E+02	**4.8020 E+02**
Mean	1.0131 E+03	7.5271 E+02	6.2837 E+02	5.4878 E+02	**5.0111 E+02**
Median	1.0112 E+03	8.3702 E+02	6.2944 E+02	5.4789 E+02	**4.9888 E+02**
Std	4.6330 E+01	1.5583 E+02	1.3892 E+01	1.3776 E+01	**1.0271 E+01**
$f5(x)$					
Min	1.2581 E+04	1.7639 E+04	8.4258 E+03	5.5564 E+03	**4.7234 E+03**
Mean	1.4282 E+04	1.9160 E+04	9.5008 E+03	7.1715 E+03	**5.5794 E+03**
Median	1.4090 E+04	1.9176 E+04	9.3825 E+03	7.2019E+03	**5.3980 E+03**
Std	9.6948 E+02	8.2684 E+02	6.2791 E+02	6.5159 E+02	**6.1390 E+02**
$f6(x)$					
Min	1.8440 E+06	7.3864 E+06	2.3964 E+06	1.6318 E+06	**1.3833 E+06**
Mean	3.7058 E+06	1.3158 E+07	3.3193 E+06	2.1834 E+06	**1.9272 E +06**
Median	3.6446 E+06	1.2457 E+07	3.2200 E+06	2.2014E+06	**1.9522 E+06**
Std	9.8093 E+05	4.1777 E+06	5.9871 E+05	3.0671 E+05	**2.6129 E+05**
$f7(x)$					
Min	7.3969 E+02	8.1342 E+02	7.9195 E+02	7.3534 E+02	**7.3350 E+02**
Mean	**7.8438 E+02**	8.4624 E+02	8.2456 E+02	7.9700 E+02	8.0765 E+02
Median	**7.7023 E+02**	8.5004 E+02	8.2193 E+02	8.0104 E+02	8.0020 E+02
Std	4.1385 E+01	2.0194 E+01	**1.8298 E+01**	3.2732 E+01	2.8476 E+01
$f8(x)$					
Min	1.2878 E+06	4.1243 E+06	1.1840 E+06	**5.0520 E+05**	5.1323 E+05
Mean	1.9175 E+06	1.0590 E+07	1.9332 E+06	8.9139 E+05	**6.9234 E+05**
Median	1.8616 E+06	1.0212 E+07	1.9618 E+06	8.9816 E+05	**6.8432 E+05**
Std	3.7847 E+05	3.8045 E+06	3.2658 E+05	2.2895 E+05	**1.3364 E+05**
$f9(x)$					
Min	1.0125 E+03	1.0110 E+03	1.0072 E+03	1.0069 E+03	**1.0068 E+03**
Mean	1.1742 E+03	1.0138 E+03	1.0079 E+03	1.0079 E+03	**1.0077 E+03**
Median	1.0414 E+03	1.0138 E+03	1.0078 E+03	1.0079 E+03	**1.0076 E+03**
Std	2.8267 E+02	1.2870 E+00	**5.2094E-01**	5.5162E-01	5.2679E-01
$f10(x)$					
Min	2.0038 E+06	2.1053 E+06	2.7821 E+05	4.7205 E+04	**1.4355 E+04**
Mean	3.8010 E+06	3.7765 E+06	4.9280 E+05	5.5835 E+04	**1.6831 E+04**
Median	3.5081 E+06	3.3635 E+06	4.9422 E+05	5.6250 E+04	**1.6362 E+04**
Std	1.5857 E+06	1.2352 E+06	9.5961 E+04	5.1752 E+03	**2.1156 E+03**
$f11(x)$					
Min	1.4231 E+03	2.5629 E+03	1.4173 E+03	**1.4100 E+03**	1.4102 E+03
Mean	2.1106 E+03	2.8176 E+03	1.7756 E+03	**1.4136 E+03**	1.6127 E+03
Median	1.4395 E+03	2.8267 E+03	1.4220 E+03	1.4139 E+03	**1.4126 E+03**
Std	7.9489 E+02	1.1589 E+02	6.5399 E+02	**1.5746 E+00**	4.5786 E+02
$f12(x)$					
Min	1.3561 E+03	1.3177 E+03	1.3137 E+03	1.3126 E+03	**1.3111 E+03**
Mean	1.3821 E+03	1.3202 E+03	1.3154 E+03	1.3153 E+03	**1.3147 E+03**

Table 4: Continued.

	BBO	PSO	CLPSO	BLPSO	DDBLPSO
Median	1.3811 E+03	1.3202 E+03	1.3155 E+03	1.3154 E+03	**1.3149 E+03**
Std	1.3010 E+01	1.1930E+00	**8.0780E-01**	1.1071 E+00	1.3710 E+00
$f13(x)$					
Min	1.7453 E+03	**1.6946 E+03**	1.7725 E+03	1.7411 E+03	1.7344 E+03
Mean	1.7600 E+03	**1.7135 E+03**	1.7825 E+03	1.7534 E+03	1.7481 E+03
Median	1.7602 E+03	**1.7137 E+03**	1.7833 E+03	1.7540 E+03	1.7487 E+03
Std	7.6554 E+00	7.6113 E+00	**5.0447 E+00**	5.9510 E+00	6.3159 E+00
$f14(x)$					
Min	1.5092 E+04	1.1046 E+05	1.1030 E+05	5.0142 E+03	**3.1265 E+03**
Mean	1.5690 E+05	1.1051 E+05	1.1094 E+05	**3.2405 E+04**	1.1036 E+05
Median	1.5755 E+05	1.1051 E+05	1.1032 E+05	**5.5990 E+03**	1.1770 E+05
Std	4.4260 E+04	**4.0365 E+01**	1.8805 E+03	4.7720 E+04	2.9479 E+04
$f15(x)$					
Min	1.6086 E+03	**1.6010 E+03**	1.6063 E+03	1.6048 E+03	1.6031 E+03
Mean	1.6159 E+03	1.6041 E+03	1.6066 E+03	1.6052 E+03	**1.6037 E+03**
Median	1.6157 E+03	1.6042 E+03	1.6066 E+03	1.6052 E+03	**1.6038 E+03**
Std	3.5097 E+00	1.6093 E+00	**1.4693E−01**	1.9835E−01	3.1219E−01

Table 5: Performance comparison of three algorithms in experiment three.

	PBSPSO	DEBBO	DDBLPSO
$f1(x)$			
Min	1.8592 E+06	4.6232 E+05	**2.0171 E+05**
Mean	4.0977 E+06	2.0454 E+06	**9.9974 E+05**
Median	3.8062 E+06	1.6998 E+06	**8.0404 E+05**
Std	1.4323 E+06	1.3320 E+06	**7.7346 E+05**
$f2(x)$			
Min	8.0625 E+07	2.0358 E+02	**2.0057 E+02**
Mean	1.2518 E+08	4.0561 E+03	**3.0520 E+03**
Median	1.3096 E+08	**1.1378 E+03**	1.3179 E+03
Std	2.6223 E+07	4.5224 E+03	**4.2472 E+03**
$f3(x)$			
Min	3.2058E+02	**3.2026E+02**	3.2104 E+02
Mean	3.2103 E+02	**3.2034E+02**	3.2128 E+02
Median	3.2106 E+02	**3.2035E+02**	3.2129 E+02
Std	2.5954E−01	**4.7224E-02**	1.1575E−01
$f4(x)$			
Min	5.3829 E+02	4.2006E+02	**4.0995 E+02**
Mean	5.9001 E+02	4.3346 E+02	**4.1878 E+02**
Median	5.9107 E+02	4.3369 E+02	**4.1840 E+02**
Std	2.6829 E+01	7.5401 E+00	**5.2623 E+00**
$f5(x)$			
Min	3.2276 E+03	2.4576 E+03	**1.3481 E+03**
Mean	4.1617 E+03	3.1351 E+03	**1.7612 E+03**
Median	4.1973E+03	3.1853 E+03	**1.7362 E+03**
Std	5.1598 E+02	3.1687 E+02	**2.5296 E+02**
$f6(x)$			
Min	9.2791 E+04	1.1225 E+05	**2.7682 E+04**
Mean	3.3549 E+05	3.7463 E+05	**1.1685 E+05**
Median	3.1758 E+05	2.9917 E+05	**9.0779 E+04**
Std	1.8022 E+05	2.4642 E+05	**8.6045 E+04**
$f7(x)$			
Min	7.0838 E+02	7.0717 E+02	**7.0517 E+02**
Mean	7.1403 E+02	7.1034 E+02	**7.0754 E+02**
Median	7.1384 E+02	7.1015 E+02	**7.0586 E+02**
Std	4.4317 E+00	**1.6748 E+00**	2.8994 E+00

Table 5: Continued.

	PBSPSO	DEBBO	DDBLPSO
$f8(x)$			
Min	3.1112 E+04	**5.3259 E+03**	1.1820 E+04
Mean	1.4377 E+05	7.7449 E+04	**4.5841 E+04**
Median	1.1602 E+05	4.8418 E+04	**3.1882 E+04**
Std	1.0130 E+05	1.0253 E+05	**3.5138 E+04**
$f9(x)$			
Min	1.0043 E+03	**1.0022 E+03**	1.0024 E+03
Mean	1.0677 E+03	**1.0031 E+03**	1.0037 E+03
Median	1.0061 E+03	**1.0025 E+03**	1.0037 E+03
Std	1.2610 E+02	2.0847 E+00	**5.1535E-01**
$f10(x)$			
Min	2.0553 E+04	**1.1722 E+04**	2.1702 E+04
Mean	1.6465 E+05	1.2081E+05	**1.0437 E+05**
Median	1.1912E+05	8.6437 E+04	**7.1822 E+04**
Std	1.3543 E+05	1.0541 E+05	**1.0132 E+05**
$f11(x)$			
Min	1.4049 E+03	1.5319 E+03	**1.4013 E+03**
Mean	1.6647 E+03	1.6533 E+03	**1.4753 E+03**
Median	1.4082 E+03	1.6380 E+03	**1.4021 E+03**
Std	3.0497 E+02	**6.1809 E+01**	1.1947E+02
$f12(x)$			
Min	1.3108 E+03	1.3052 E+03	**1.3040 E+03**
Mean	1.3152 E+03	**1.3061 E+03**	1.3071 E+03
Median	1.3139 E+03	**1.3058 E+03**	1.3075 E+03
Std	4.6593 E+00	**1.1863 E+00**	1.4291 E+00
$f13(x)$			
Min	1.4153 E+03	**1.3928 E+03**	1.3995 E+03
Mean	1.4281 E+03	**1.3999 E +03**	1.4173 E+03
Median	1.4289 E+03	**1.3996 E+03**	1.4164 E+03
Std	5.1713 E+00	**4.3428 E+00**	7.2068E+00
$f14(x)$			
Min	3.2641 E+04	3.2795 E+04	**1.5000 E+03**
Mean	3.6568 E+04	3.5561 E+04	**3.2591 E+04**
Median	3.7333 E+04	3.5346 E+04	**3.3648 E+04**
Std	2.0662 E+03	**9.2865 E+02**	5.9283 E+03
$f15(x)$			
Min	1.6034 E+03	**1.6000 E+03**	**1.6000 E+03**
Mean	1.6076 E+03	**1.6000 E+03**	1.6004 E+03
Median	1.6048 E+03	**1.6000 E+03**	**1.6000 E+03**
Std	4.8441 E+00	**1.0394E-02**	2.3806 E+00

strategy guides particle to learn from other particles not worse than itself, which leads a greater convergence rate of DDBLPSO. At the same time, DDBLPSO has a slightly better ability of jumping out of local optimum than other competitors because of simple mutation and complexity of functions. One the other hand, learning from the better particles means that the exploration ability of algorithm will be reduced slightly. For this reason, DDBLPSO does not have such eye-catching performance on multimodal functions and composition functions. Table 7 shows that DDBLPSO attains the best rank, BLPSO attains the second, and CLPSO attains the third, followed by BBO and PSO. It is obvious that DDBLPSO is the most competitive algorithm in this group of experimental comparison.

Table 4 reveals that DDBLPSO performs much better than other four algorithms on $f1$, $f4$, $f5$, $f6$, $f8$, $f9$, $f10$, and $f12$. DDBLPSO cannot achieve obvious dominance on each kind of functions due to the problems of being different types and large scale. On the whole, DDBLPSO still performs best. Table 8 presents the same ranking result as Table 7. Figure 2 shows that DDBLPSO has the strongest evolving trend on $f1$, $f2$, $f4$, $f6$, $f7$, $f8$, $f10$, and $f14$. DDBLPSO is also the most competitive algorithm in this group of experimental comparison with relatively large-scale problems.

Generally speaking, it is observed from Tables 3, 4, 7 and 9 and Figure 2 that DDBLPSO achieves the best performance compared with its competitors.

Table 5 shows that DDBLPSO beats PBSPSO and DEBBO on most functions, and Table 9 shows that DDBLPSO attains the best rank, DEBBO attains the second, and PBSPSO attains the third. Table 6 reveals that DDBLPSO

Table 6: Performance comparison of three algorithms in experiment four.

	PBSPSO	DEBBO	DDBLPSO
$f1(x)$			
Min	2.8403 E+07	1.2931 E+08	**8.8196 E+06**
Mean	3.5153 E+07	1.7284 E+08	**1.1151 E+07**
Median	3.4872 E+07	1.7385 E+08	**1.1181 E+07**
Std	4.0420 E+06	2.4193 E+07	**1.0671 E+06**
$f2(x)$			
Min	5.3463 E+08	**2.0005 E+02**	2.3631 E+07
Mean	1.2518 E+08	**1.6110 E+03**	3.2844 E+07
Median	1.3096 E+08	**1.0518 E+03**	3.2473 E+07
Std	2.6223 E+07	**1.6068 E+03**	4.2897 E+06
$f3(x)$			
Min	3.2131 E+02	**3.2073E+02**	3.2136 E+0
Mean	3.2144 E+02	**3.2080E+02**	3.2145 E+02
Median	3.2144 E+02	**3.2083E+02**	3.2145 E+02
Std	6.3151E−02	4.2976E−02	**3.9328 E−02**
$f4(x)$			
Min	9.1895 E+02	7.2243 E+02	**4.8020 E+02**
Mean	1.0124 E+03	7.6390 E+02	**5.0111 E+02**
Median	1.0051 E+03	7.5774 E+02	**4.9888 E+02**
Std	4.9600 E+01	2.5379 E+01	**1.0271 E+01**
$f5(x)$			
Min	1.2411 E+04	1.6916 E+04	**4.7234 E+03**
Mean	1.4988 E+04	1.7798 E+04	**5.5794 E+03**
Median	1.4942 E+04	1.7938 E+04	**5.3980 E+03**
Std	1.2888 E+03	7.2976 E+02	**6.1390 E+02**
$f6(x)$			
Min	2.1414 E+06	2.3652 E+07	**1.3833 E+06**
Mean	3.6713 E+06	3.3346 E+07	**1.9272 E+06**
Median	3.5486 E+06	3.4022 E+07	**1.9522 E+06**
Std	9.3569 E+05	5.1752 E+06	**2.6129 E+05**
$f7(x)$			
Min	7.4025 E+02	8.4813 E+02	**7.3350 E+02**
Mean	**7.8440 E+02**	8.4978 E+02	8.0765 E+02
Median	**7.7255 E+02**	8.4967 E+02	8.0020 E+02
Std	3.6466 E+01	**8.3256E−01**	2.8476 E+01
$f8(x)$			
Min	1.2536 E+06	1.0239 E+07	**5.1323 E+05**
Mean	1.9625 E+06	1.7527 E+07	**6.9234 E+05**
Median	1.9220 E+06	1.7616 E+07	**6.8432 E+05**
Std	3.8544 E+05	4.3476 E+06	**1.3364 E+05**
$f9(x)$			
Min	1.0124 E+03	**1.0073 E+03**	1.0068 E+03
Mean	1.2043E+03	**1.0176 E+03**	1.0077 E+03
Median	1.0155 E+03	**1.0175 E+03**	1.0076 E+03
Std	3.1718 E+02	**2.2394E−01**	5.2679E−01
$f10(x)$			
Min	1.6908 E+06	**5.3905 E+03**	1.4355 E+04
Mean	3.7009 E+06	**6.7693 E+03**	1.6831 E+04
Median	3.3489 E+06	**6.4983 E+03**	1.6362 E+04
Std	1.7822 E+06	**1.0052 E+03**	2.1156 E+03
$f11(x)$			
Min	1.4281 E+03	1.7210 E+03	**1.4102 E+03**
Mean	2.1024 E+03	2.8620 E+03	**1.6127 E+03**
Median	1.4369 E+03	3.1818 E+03	**1.4126 E+03**
Std	8.3776 E+02	7.4285 E+02	**4.5786 E+02**
$f12(x)$			
Min	1.3512 E+03	1.3172 E+03	**1.3111 E+03**
Mean	1.3720 E+03	1.3176 E+03	**1.3147 E+03**

Table 6: Continued.

	PBSPSO	DEBBO	DDBLPSO
Median	1.3688 E+03	1.3176 E+03	**1.3149 E+03**
Std	1.6991 E+01	**2.6835E−01**	1.3710 E+00
$f13(x)$			
Min	1.7474 E+03	**1.6929 E+03**	1.7344 E+03
Mean	1.7625 E+03	**1.7138 E+03**	1.7481 E+03
Median	1.7631 E+03	**1.7158 E+03**	1.7487 E+03
Std	7.3313 E+00	7.9928 E+00	**6.3159 E+00**
$f14(x)$			
Min	1.5392 E+04	1.1024 E+05	**3.1265 E+03**
Mean	1.4095 E+05	**1.1027 E+05**	1.1036 E+05
Median	1.4401 E+05	**1.1027 E+05**	1.1770 E+05
Std	2.7916 E+04	**1.9162 E+01**	2.9479 E+04
$f15(x)$			
Min	1.6084 E+03	**1.6000 E+03**	1.6031 E+03
Mean	1.6140 E+03	**1.6000 E+03**	1.6037 E+03
Median	1.6138 E+03	**1.6000 E+03**	1.6038 E+03
Std	3.8763 E+00	**2.3967E−02**	3.1219E−01

Table 7: Ranks of five algorithms according to Friedman test in experiment one.

	PSO	BBO	CLPSO	BLPSO	DDBLPSO
Friedman rank	4.15	3.93	2.63	2.44	1.85
Final rank	5	4	3	2	1

Table 8: Ranks of five algorithms according to Friedman test in experiment two.

	PSO	BBO	CLPSO	BLPSO	DDBLPSO
Friedman rank	4.29	3.48	3.21	2.37	1.65
Final rank	5	4	3	2	1

Table 9: Ranks of three algorithms according to Friedman test in experiment three.

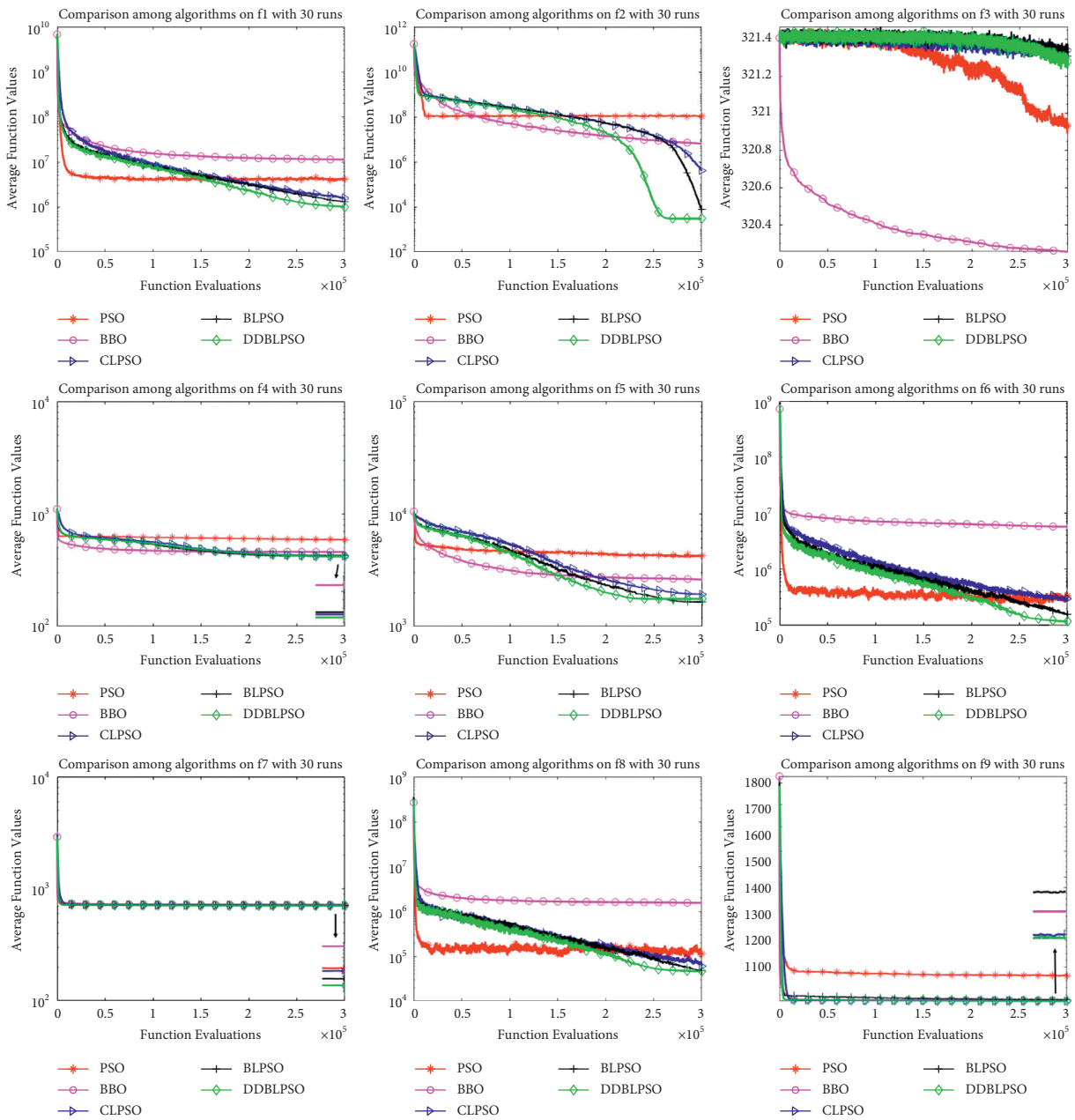
	PBSPSO	DEBBO	DDBLPSO
Friedman rank	2.78	1.73	1.48
Final rank	3	2	1

Table 10: Ranks of three algorithms according to Friedman test in experiment four.

	PBSPSO	DEBBO	DDBLPSO
Friedman rank	2.52	1.95	1.53
Final rank	3	2	1

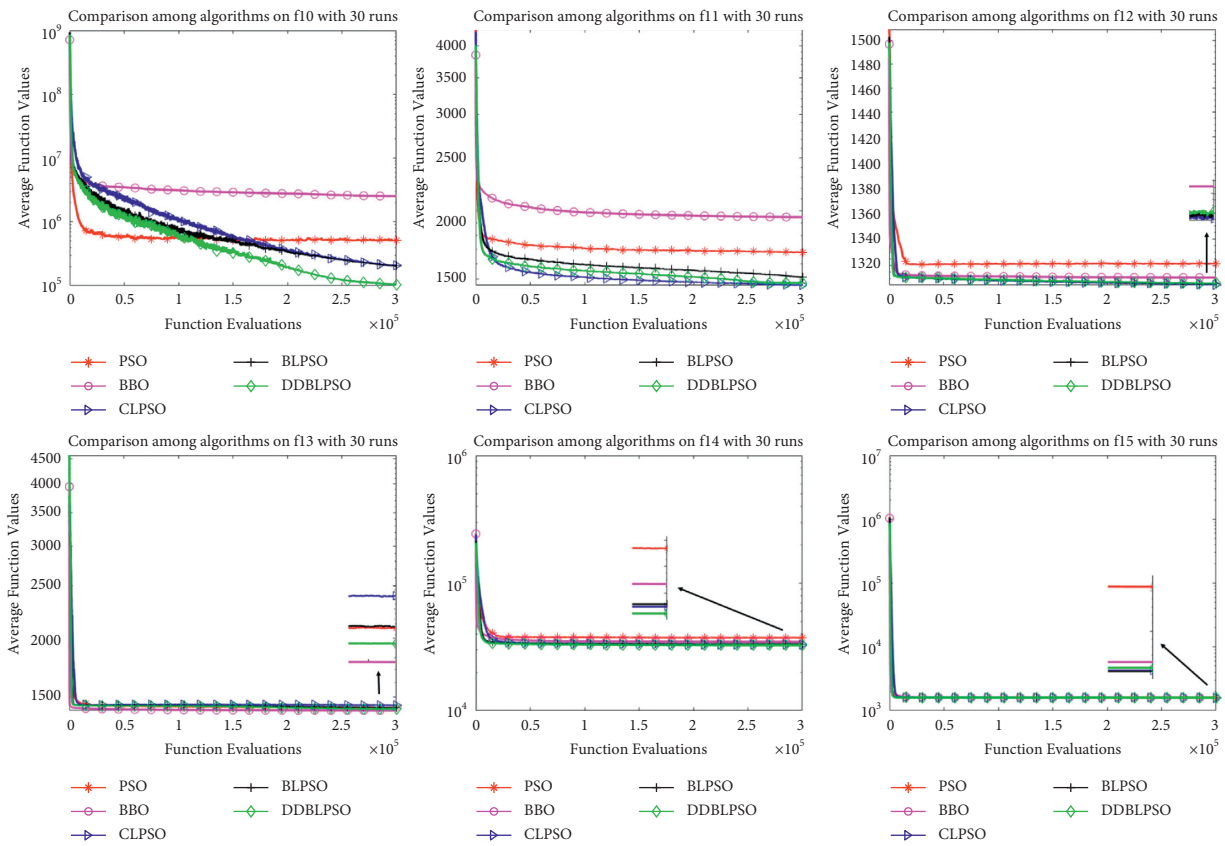
Table 11: Design of experiment five, experiment six, experiment seven, and experiment eight.

Experiment	Population size N	Points number mp	Comparison algorithm
Experiment five	30	14	PSO, BBO, CLPSO, BLSPSO
Experiment six	100	49	PSO, BBO, CLPSO, BLSPSO
Experiment seven	30	14	PBSPSO, DEBBO
Experiment eight	100	49	PBSPSO, DEBBO



(a)

Figure 2: Continued.



(b)

Figure 2: Converging curves of the average best fitness in experiment one.

performs much better than other two algorithms on $f1, f4, f5, f6, f8, f11$, and $f12$. Table 10 presents the same ranking result as Table 9. Figure 3 shows that DDBLPSO has the strongest evolving trend on $f1, f2, f4, f5, f6, f7, f8, f10, f11$, and $f14$. Generally speaking, it is observed from Tables 5, 6, 9 and 10 and Figure 3 that DDBLPSO beats a PSO variant and a BBO variant.

6. Applications in UAV Path Planning

6.1. Six Kinds of Terrain Functions. City, village without houses, village with houses, mountainous area without houses, mountainous area with houses, and mountainous area with a huge building are the most common terrains in UAV path planning. Six kinds of terrain functions are designed depending on these conditions. City has flat ground and tall buildings. Village without houses means uneven ground. Village with houses means uneven ground and low buildings. Mountainous area without houses means uneven ground and mountains. Mountainous area with houses has uneven ground, mountains, and low buildings. Mountainous area with a huge building has uneven ground, mountains, and a huge building. (x', y') is the coordinate of any point on the plane. The details of six terrain functions are as follows.

The terrain function of city f_{terrain1} is defined as follows:

$$f_{\text{terrain1}} = \begin{cases} T_i, & \text{if } (x', y') \in \Omega_i, i = 1, 2, \ldots, T, \\ 0, & \text{if } (x', y') \in \Omega_i, i = 1, 2, \ldots, T, \end{cases} \tag{19}$$

where T_i is the height of the tall building i, Ω_i is the area occupied by the tall building i, and T is the number of tall buildings.

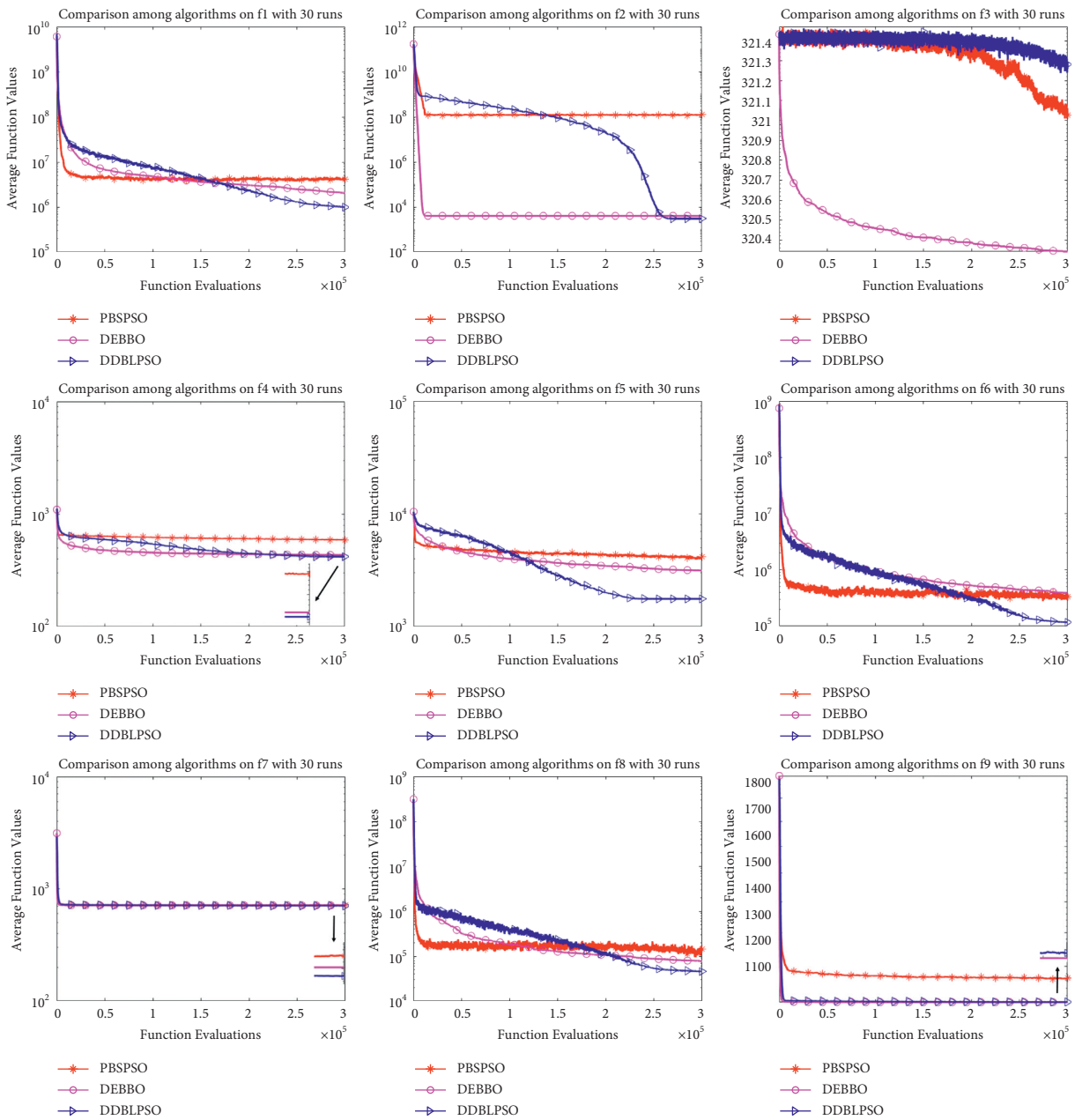
The terrain function of village without houses f_{terrain2} is defined as follows:

$$\begin{aligned} f_{\text{terrain2}} = {} & \sin(x' + a_1) + a_2 * \sin(y') + \cos(y' + a_3) \\ & + a_4 * \cos(x') \\ & + \sin\left(\sqrt{x'^2 + y'^2} + a_5\right) + a_6 * \cos\left(a_6 * \sqrt{x'^2 + y'^2}\right), \end{aligned} \tag{20}$$

where a_1, a_2, a_3, a_4, a_5, and a_6 are terrain parameters, which decide the uneven degree of the terrain.

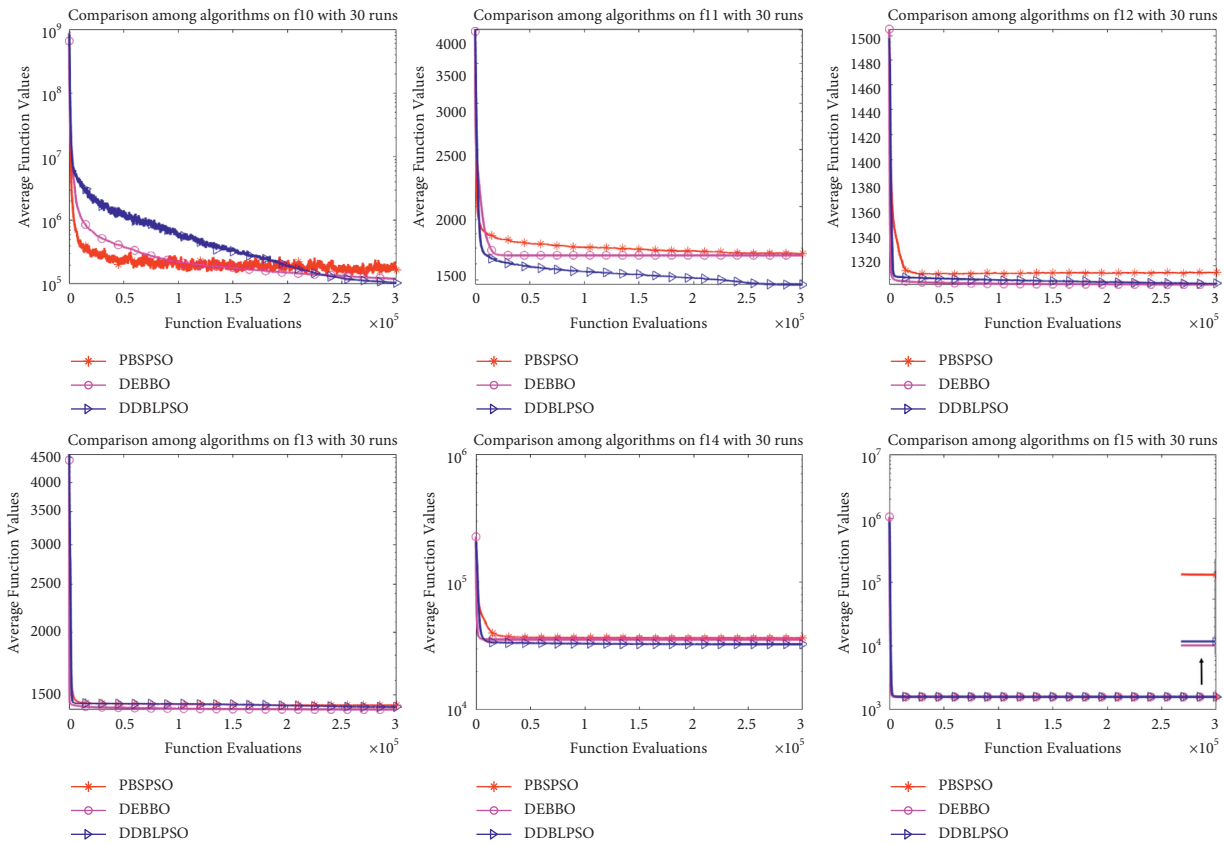
The terrain function of village with houses f_{terrain3} is defined as follows:

$$f_{\text{terrain3}} = \begin{cases} \max\{S_i, f_{\text{terrain2}}\}, & \text{if } (x', y') \in \Gamma_i, i = 1, 2, \ldots, S, \\ f_{\text{terrain2}}, & \text{if } (x', y') \notin \Gamma_i, i = 1, 2, \ldots, S, \end{cases} \tag{21}$$



(a)

Figure 3: Continued.



(b)

Figure 3: Converging curves of the average best fitness in experiment three.

Table 12: Performance comparison of five algorithms in experiment five.

	BBO	PSO	CLPSO	BLPSO	DDBLPSO
costfun1					
Min	0.333786	0.361611	0.333033	0.333038	**0.332735**
Mean	0.354893	0.408034	0.348755	0.347850	**0.338909**
Median	0.348119	0.401855	0.346445	0.340743	**0.335515**
Std	0.020209	0.035911	0.019765	0.020109	**0.005750**
costfun2					
Min	0.333514	0.358298	0.332679	0.332662	**0.332363**
Mean	0.338082	0.400207	0.333074	0.333387	**0.332846**
Median	0.336249	0.392837	0.333044	0.333043	**0.332717**
Std	0.005453	0.030524	0.000334	**0.000680**	0.000376
costfun3					
Min	0.333434	0.363159	0.332662	0.332662	**0.332470**
Mean	0.338407	0.396183	0.333221	0.333186	**0.332957**
Median	0.336943	0.385252	0.333044	0.333043	**0.332687**
Std	0.004669	0.028417	0.000525	**0.000418**	0.001214
costfun4					
Min	0.333579	0.354768	0.332679	0.332679	**0.332454**
Mean	0.336364	0.400745	0.333237	0.333581	**0.333045**
Median	0.335141	0.396060	0.333043	0.333043	**0.332773**
Std	0.003654	0.032782	0.001182	0.001667	**0.001155**
costfun5					
Min	0.333159	0.358298	0.332679	0.332662	**0.332422**
Mean	0.337872	0.399360	0.333108	0.333139	**0.332759**

Table 12: Continued.

	BBO	PSO	CLPSO	BLPSO	DDBLPSO
Median	0.336249	0.391714	0.333043	0.333043	**0.332678**
Std	0.005734	0.030770	0.000442	0.000558	**0.000374**
costfun6					
Min	0.334377	0.362532	0.332445	0.332692	**0.332414**
Mean	0.340639	0.395483	0.333808	0.333623	**0.333442**
Median	0.337852	0.386125	0.334080	0.333738	**0.333598**
Std	0.008596	0.026440	0.000625	**0.000484**	0.000673

Table 13: Performance comparison of five algorithms in experiment six.

	BBO	PSO	CLPSO	BLPSO	DDBLPSO
costfun1					
Min	0.409522	0.452563	0.407823	0.390225	**0.384962**
Mean	0.485857	0.539211	0.423850	0.417825	**0.417211**
Median	0.478369	0.545726	0.414287	0.407063	**0.394661**
Std	0.047892	0.037861	**0.020093**	0.026866	0.037308
costfun2					
Min	0.380857	0.464534	0.375254	0.366452	**0.363554**
Mean	0.388581	0.529762	0.380145	0.371376	**0.368538**
Median	0.388856	0.531348	0.380259	0.371162	**0.368096**
Std	0.003426	0.035597	**0.002097**	0.002428	0.002839
costfun3					
Min	0.380579	0.464561	0.376183	0.366838	**0.363259**
Mean	0.390465	0.530220	0.380478	0.370835	**0.369487**
Median	0.389378	0.526718	0.381116	0.370512	**0.368922**
Std	0.004844	0.042232	0.003074	**0.001831**	0.003726
costfun4					
Min	0.382432	0.472434	0.374150	0.367105	**0.362359**
Mean	0.390443	0.523642	0.380130	0.370614	**0.368372**
Median	0.390181	0.522455	0.380235	0.370568	**0.368463**
Std	0.003990	0.030173	0.002600	**0.001943**	0.002333
costfun5					
Min	0.381921	0.436542	0.375358	0.367693	**0.364425**
Mean	0.389006	0.524626	0.380941	0.371797	**0.368500**
Median	0.389857	0.531651	0.380743	0.371605	**0.367672**
Std	0.003858	0.031882	**0.002198**	0.003142	0.003034
costfun6					
Min	0.382586	0.459559	0.377788	0.368291	**0.363963**
Mean	0.392547	0.517871	0.383203	0.372213	**0.369134**
Median	0.391984	0.523437	0.383597	0.371327	**0.369174**
Std	0.006611	0.033132	0.003701	0.002866	**0.002664**

where S_i is the height of the low building i, Γ_i is the area occupied by the low building i, and S is the number of low buildings.

The terrain function of mountainous area without houses f_{terrain4} is defined as follows:

$$M_i = R_i * \exp\left[-\left(\frac{x' - P_i}{K_i}\right)^2 - \left(\frac{y' - Q_i}{K_i}\right)^2\right], \quad i = 1, 2, \ldots, M,$$
$$f_{\text{terrain4}} = \max\{M_i, f_{\text{terrain2}}\}, \tag{22}$$

where M_i is the mountain function, R_i, R_i, Q_i, and K_i describe the position and the shape of mountain i, and M is the number of mountains.

The terrain function of mountainous area with houses f_{terrain5} is defined as follows:

$$f_{\text{terrain5}} = \begin{cases} \max\{S_i, f_{\text{terrain4}}\}, & \text{if } (x', y') \in \Gamma_i, i = 1, 2, \ldots, S, \\ f_{\text{terrain4}}, & \text{if } (x', y') \notin \Gamma_i, i = 1, 2, \ldots, S, \end{cases} \tag{23}$$

where S_i is the height of the low building i, Γ_i is the area occupied by the low building i, and S is the number of low buildings.

The terrain function of mountainous area with a huge building f_{terrain6} is defined as follows:

$$f_{\text{terrain6}} = \begin{cases} \max\{\text{Huge}, f_{\text{terrain4}}\}, & \text{if } (x', y') \in \Omega', \\ f_{\text{terrain4}}, & \text{if } (x', y') \notin \Omega', \end{cases} \tag{24}$$

Table 14: Performance comparison of three algorithms in experiment seven.

	PBSPSO	DEBBO	DDBLPSO
costfun1			
Min	0.333917	0.332859	**0.332735**
Mean	0.351685	0.341414	**0.338909**
Median	0.340660	0.335595	**0.335515**
Std	0.025157	0.010353	**0.005750**
costfun2			
Min	0.333801	**0.332310**	0.332363
Mean	0.338538	0.339693	**0.332846**
Median	0.337244	0.339361	**0.332717**
Std	0.004749	0.007640	**0.000376**
costfun3			
Min	0.333631	**0.332355**	0.332470
Mean	0.335685	0.340312	**0.332957**
Median	0.334800	0.338883	**0.332687**
Std	0.002172	0.006869	**0.001214**
costfun4			
Min	0.333443	0.332611	**0.332454**
Mean	0.336772	0.339582	**0.333045**
Median	0.334911	0.336338	**0.332773**
Std	0.005649	0.010488	**0.001155**
Costfun5			
Min	0.333852	**0.332310**	0.332422
Mean	0.338636	0.339687	**0.332759**
Median	0.337244	0.339196	**0.332678**
Std	0.004892	0.007473	**0.000374**
costfun6			
Min	0.334083	0.333078	**0.332414**
Mean	0.337138	0.342282	**0.333442**
Median	0.336396	0.339118	**0.333598**
Std	0.002739	0.012053	**0.000673**

where Huge is the height of the huge building and Ω' is the area occupied by the huge building.

6.2. Parameter Settings. Cost functions with city, village without houses, village with houses, mountainous area without houses, mountainous area with houses, and mountainous area with a huge building are labelled as costfun1 – costfun6. Four sets of UAV path planning experiments are conducted which are labelled as experiment five, experiment six, experiment seven, and experiment eight. Design of the four experiments are shown in Table 11. The boldface in Tables 12–15 indicates the best experimental results in experiment five to eight, respectively. Points number means the number of points in each path without starting point and end point. Dimension is equal to twice the point number. The path is divided into mp + 1 small lines by mp points from the starting point to the end point. To simplify the calculation, if a point is in the danger zone, it is assumed that the halves of both small lines before and after and adjacent to the point are also in the danger zone. It is the same for a point under the ground or under a building. UAV search space is $[-50, 50]*[-50, 50]*[0, 25]$, terrain parameters $a_1 = 10$, $a_2 = 1$, $a_3 = 20$, $a_4 = 0.8$, $a_5 = 30$, $a_6 = 0.9$, weight coefficients $W_1 = W_2 = W_3 = W_4 = 1$, penalty constant $P = 3$, and a small positive number $Z_{\min} = 0.1$. Starting point and end point coordinates are (−50,−50, 0.1) and (50, 50, 0.1), respectively. The maximum velocity along Y axis and Z axis is 2.5 and 0.625, respectively. The details of tall, low, and huge buildings are given in Table 16. The detailed parameter settings of mountains are given in Table 17. Danger areas are designed as equation (25). Other parameter settings of algorithms are the same as in Section 5.

$$\begin{aligned}(x' + 25)^2 + (y' + 25)^2 + z'^2 &= 10^2,\\ (x' - 25)^2 + (y' - 25)^2 + z'^2 &= 10^2.\end{aligned} \tag{25}$$

6.3. Results and Discussion of UAV Path Planning. Tables 12–15 show the UAV path planning results of six kinds of terrains in experiment five, experiment six, experiment seven, and experiment eight, respectively. The ranks of five algorithms according to Friedman test of experiment three and experiment four are shown in Tables 18–21, respectively. Min, Mean, Median, and Std indicate the minimum function error value, the mean function error value, the median function error value, and the standard deviation of error values, respectively. In order to exhibit the evolution trend of five algorithms more vividly, the converging curves of the average best fitness in experiment five and experiment seven

Table 15: Performance comparison of three algorithms in experiment eight.

	PBSPSO	DEBBO	DDBLPSO
costfun1			
Min	0.408638	0.740178	**0.384962**
Mean	0.507099	0.855135	**0.417211**
Median	0.500030	0.857994	**0.394661**
Std	0.057598	**0.031987**	0.037308
costfun2			
Min	0.383974	0.373691	**0.363554**
Mean	0.392406	0.400978	**0.368538**
Median	0.392086	0.398095	**0.368096**
Std	0.005132	0.016330	**0.002839**
costfun3			
Min	0.383351	**0.362185**	0.363259
Mean	0.394280	0.402638	**0.369487**
Median	0.393606	0.402676	**0.368922**
Std	0.004780	0.020725	**0.003726**
costfun4			
Min	0.384865	0.368213	**0.362359**
Mean	0.393153	0.407062	**0.368372**
Median	0.392543	0.406902	**0.368463**
Std	0.005216	0.020424	**0.002333**
costfun5			
Min	0.385425	**0.363051**	0.364425
Mean	0.392956	0.409671	**0.368500**
Median	0.393330	0.414376	**0.367672**
Std	0.004038	0.019970	**0.003034**
costfun6			
Min	0.386878	0.379041	**0.363963**
Mean	0.393760	0.407583	**0.369134**
Median	0.394691	0.408277	**0.369174**
Std	0.003657	0.016136	**0.002664**

Table 16: Parameter settings of tall, short, and huge buildings.

Tall buildings	[−20,−10] ∗ [5,15] ∗ [0,16] [6,15] ∗ [−8,3] ∗ [0,14]	[10,20] ∗ [5,15] ∗ [0,15] [−14,-3] ∗ [−20,-10] ∗ [0,14]	[−15,−6] ∗ [−8,3] ∗ [0,15] [3,14] ∗ [−2,−10] ∗ [0,16]
Short buildings	[−20,−10] ∗ [5,15] ∗ [0,6] [6,15] ∗ [−8,3] ∗ [0,4]	[10,20] ∗ [5,15] ∗ [0,5] [−14,−3] ∗ [−20,−10] ∗ [0,4]	[−15,−6] ∗ [−8,3] ∗ [0,5] [3,14] ∗ [−20,−10] ∗ [0,6]
Huge building	[−45,45] ∗ [−45,45] ∗ [0,20]		

Table 17: Parameter settings of tall, short, and huge buildings.

Mountains	$R_1 = 20$, $P_1 = 0, Q_1 = 30$, $K_1 = 5$ $R_3 = 21$, $P_3 = 30$, $Q_3 = 0$, $K_3 = 3$	$R_2 = 18$, $P_2 = 0$, $Q_2 = -30$, $K_2 = 2$ $R_4 = 19$, $P_4 = -30$, $Q_4 = 0$, $K_4 = 4$

are shown in Figures 4 and 5, respectively. The converging curves of six kinds of terrain functions in experiment six and experiment eight are not shown due to the similar converging behaviours with experiment five and experiment seven, respectively.

Table 12 shows that DDBLPSO is the most competitive algorithm for the relatively small-scale problems with 14 points. It is easy to see that DDBLPSO attains the best rank and BLPSO is the second, followed by CLPSO, PSO, and BBO in Table 18. For the relatively large-scale problems with 49 points, Tables 13 and 19 reveal that DDBLPSO achieves the best results in experiment six. Table 19 shows that DDBLPSO attains the best rank, BLPSO is the second, and CLPSO is the third, followed by PSO and BBO. Furthermore, the experimental results indicate that it is much better than other competitors. However, Tables 12 and 13 show that DDBLPSO has relatively larger Std, which indicates a possible enhancement on the convergence capability of DDBLPSO. Figure 4 shows that DDBLPSO has the strongest evolving trend on all the terrain functions for

Table 18: Ranks of five algorithms according to Friedman test in experiment five.

	PSO	BBO	CLPSO	BLPSO	DDBLPSO
Friedman rank	4.00	5.00	2.42	2.38	1.21
Final rank	4	5	3	2	1

Table 19: Ranks of five algorithms according to Friedman test in experiment six.

	PSO	BBO	CLPSO	BLPSO	DDBLPSO
Friedman rank	4.04	4.96	2.71	1.96	1.33
Final rank	4	5	3	2	1

Table 20: Ranks of three algorithms according to Friedman test in experiment seven.

	PBSPSO	DEBBO	DDBLPSO
Friedman rank	2.38	2.50	1.13
Final rank	2	3	1

Table 21: Ranks of three algorithms according to Friedman test in experiment eight.

	PBSPSO	DEBBO	DDBLPSO
Friedman rank	2.25	2.63	1.13
Final rank	2	3	1

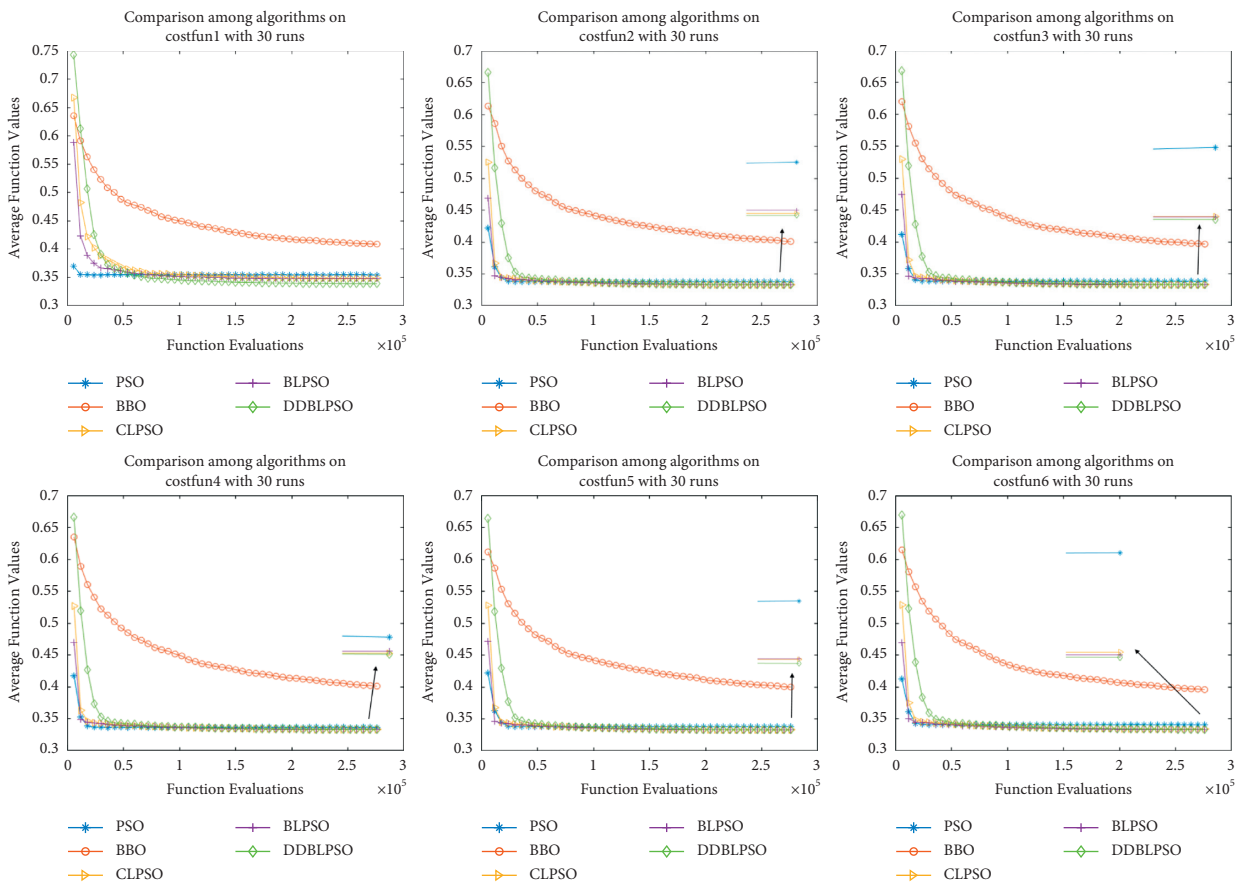


Figure 4: Converging curves of different terrain functions in experiment five.

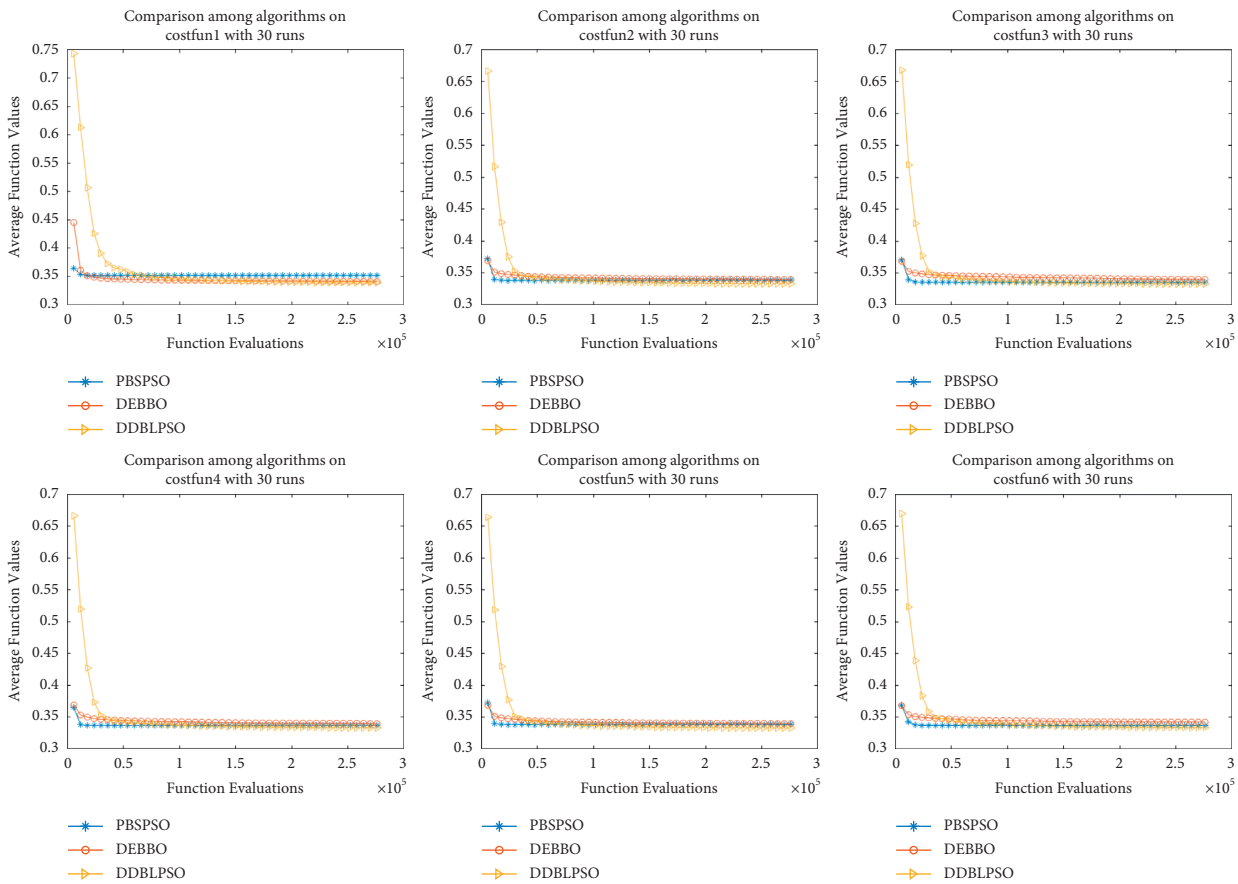


Figure 5: Converging curves of different terrain functions in experiment seven.

the relatively large-scale problem in experiment five. Generally speaking, it can be observed from Tables 12, 13, 18, 19 and Figure 4 that DDBLPSO performs very encouraging in UAV path planning, but there is still room for improvement in terms of Std.

It is easy to see that DDBLPSO beats DEBBO and PBSPSO for the relatively small-scale problems with 14 points in Table 14. Table 15 shows that DDBLPSO attains the best rank, DEBBO is the second, and PBSPSO is the third in small-scale problems of UAV path planning. Figure 5 shows that DDBLPSO has the stronger evolving trend on all the terrain functions than DEBBO and PBSPSO in experiment seven. Tables 17 and 21 reveal the same results in large-scale problems.

7. Conclusions

This paper presents a novel UAV path planning algorithm based on double-dynamic biogeography-based learning particle swarm optimization. DDBLPSO adopts double-dynamic biogeography-based learning strategy. Under this strategy, particle only learns from itself or other even better individuals. Six terrain functions are presented to simulate the most six common terrains in UAV path planning. Computational experiments and simulations demonstrate the performance advantages of the algorithm both in global optimization and UAV path planning. Especially for the relatively large-scale problems, it performs even more competitive.

Our algorithm has the potential to be applied to other problems such as oil exploration, scheduling, and deep learning. The UAV path planning problem also can be extended to multiobjective or constrained optimization problems to be suitable for more complex situations.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Authors' Contributions

Yisheng Ji was responsible for novelty construction and verification, simulation design, and manuscript writing. Xinchao Zhao was responsible for research motivation, simulation design, and analysis. Junling Hao read through the manuscript and gave some suggestions on structure and representation.

Acknowledgments

The authors are grateful to Prof. Xingquan Zuo from School of Computer Science, Beijing University of Posts and Telecommunications, for his suggestions on research motivation and simulation. This research was supported by the Beijing Natural Science Foundation (1202020) and National Natural Science Foundation of China (61973042).

References

[1] I. K. Nikolos, K. P. Valavanis, N. C. Tsourveloudis, and A. N. Kostaras, "Evolutionary algorithm based offline/online path planner for uav navigation," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 33, no. 6, pp. 898–912, 2003.

[2] Y. Fu, M. Ding, and C. Zhou, "Phase Angle-Encoded and Quantum-Behaved Particle Swarm Optimization Applied to Three-Dimensional Route Planning for UAV," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 42, no. 2, pp. 511–526, 2012.

[3] Y. V. Pehlivanoglu, "A new vibrational genetic algorithm enhanced with a Voronoi diagram for path planning of autonomous UAV," *Aerospace Science and Technology*, vol. 16, no. 1, pp. 47–55, 2012.

[4] X. Zhang and H. Duan, "An improved constrained differential evolution algorithm for unmanned aerial vehicle global route planning," *Applied Soft Computing*, vol. 26, pp. 270–284, 2015.

[5] E. Besada-Portas, L. de la Torre, A. Moreno, and J. L. Risco-Martín, "On the performance comparison of multi-objective evolutionary UAV path planners," *Information Sciences*, vol. 238, pp. 111–125, 2013.

[6] B. Zhang and H. Duan, "Three-Dimensional Path Planning for Uninhabited Combat Aerial Vehicle Based on Predator-Prey Pigeon-Inspired Optimization in Dynamic Environment," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 14, no. 1, pp. 97–107, 2017.

[7] YB Chen, YS Mei, and JQ Yu, "Three-dimensional unmanned aerial vehicle path planning using modified wolf pack search algorithm[J]," *Neurocomputing*, vol. 266, pp. 445–457, 2017.

[8] Z. Sun, J. Wu, J. Yang, Y. Huang, C. Li, and D. Li, "Path Planning for GEO-UAV Bistatic SAR Using Constrained Adaptive Multiobjective Differential Evolution," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 11, pp. 6444–6457, 2016.

[9] V. Roberge, M. Tarbouchi, and G. Labonte, "Comparison of Parallel Genetic Algorithm and Particle Swarm Optimization for Real-Time UAV Path Planning," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 132–141, 2013.

[10] RC Eberchart and J Kennedy, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks IEEE*, pp. 1942–1948, WA, Australia, December 1995.

[11] ZH Zhi-Hui Zhan, J Jun Zhang, Y Yun Li, and H. S.-H. Chung, "Adaptive Particle Swarm Optimization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 6, pp. 1362–1381, 2009.

[12] A. Nickabadi, M. M. Ebadzadeh, and R. Safabakhsh, "A novel particle swarm optimization algorithm with adaptive inertia weight," *Applied Soft Computing*, vol. 11, no. 4, pp. 3658–3670, 2011.

[13] XD Xiaodong Li and X Xin Yao, "Cooperatively Coevolving Particle Swarms for Large Scale Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 2, pp. 210–224, 2012.

[14] J. J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 281–295, 2006.

[15] X. Zhao, Y. Ji, and J. Hao, "A Novel Biogeography-based Optimization Algorithm with Momentum Migration and Taxonomic Mutation," in *The Eleventh International Conference on Swarm Intelligence (ICSI 2020)*, vol. 12145, pp. 83–93, Belgrade, Serbia, 2020.

[16] X. Zhao, W. Lin, J. Hao, X. Zuo, and J. Yuan, "Clustering and pattern search for enhancing particle swarm optimization with Euclidean spatial neighborhood search," *Neurocomputing*, vol. 171, pp. 966–981, 2016.

[17] S.-W. Lin, K.-C. Ying, S.-C. Chen, and Z.-J. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines," *Expert Systems with Applications*, vol. 35, no. 4, pp. 1817–1824, 2008.

[18] R. K. Yadav and Anubhav, "PSO-GA based hybrid with Adam Optimization for ANN training with application in Medical Diagnosis," *Cognitive Systems Research*, vol. 64, pp. 191–199, 2020.

[19] R. Vafashoar and M. R. Meybodi, "Cellular learning automata based bare bones PSO with maximum likelihood rotated mutations," *Swarm and Evolutionary Computation*, vol. 44, pp. 680–694, 2019.

[20] G. Li and X. Jiao, "Synthesis and validation of finite time servo control with PSO identification for automotive electronic throttle," *Nonlinear Dynamics*, vol. 90, no. 2, pp. 1165–1177, 2017.

[21] Z Xiang, D Ji, H Zhang, W Hongrun, and L Yuanxiang, "A simple PID-based strategy for particle swarm optimization algorithm," *Information Sciences*, vol. 502, 2019.

[22] D. Simon, "Biogeography-Based Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 6, pp. 702–713, 2008.

[23] X. Chen, H. Tianfield, C. Mei, W. Du, and G. Liu, "Biogeography-based learning particle swarm optimization," *Soft Computing*, vol. 21, no. 24, pp. 7519–7541, 2017.

[24] W. Gong, Z. Cai, and C. X. Ling, "DE/BBO: a hybrid differential evolution with biogeography-based optimization for global numerical optimization," *Soft Computing*, vol. 15, no. 4, pp. 645–665, 2010.

[25] JJ Liang, BY Qu, PN Suganthan, and Q Chen, *Problem Definitions and Evaluation Criteria for the CEC 2015 Competition on Learning-based Real-Parameter Single Objective Optimization*, Zhengzhou University (China) and Nanyang Technological University, Singapore), 2014.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9159130, 18 pages
https://doi.org/10.1155/2022/9159130

Research Article

A Modified Whale Optimization Algorithm and Its Application in Seismic Inversion Problem

Xiaodan Liang,[1] **Siwen Xu**,[1] **Yang Liu**,[2] **and Liling Sun**[1]

[1]*Tiangong University, Tianjin 300387, China*
[2]*Liaoning Inspection Examination and Certification Centre, Shenyang 110000, China*

Correspondence should be addressed to Liling Sun; sunliling2019@outlook.com

Received 13 October 2021; Accepted 30 December 2021; Published 24 January 2022

Academic Editor: Jianhui Lv

The whale optimization algorithm (WOA) is a popular swarm intelligence algorithm which simulates the hunting behavior of humpback whales. WOA has the deficiency of easily falling into the local optimal solutions. In order to overcome the weakness of the WOA, a modified variant of WOA called OCDWOA is proposed. There are four main operators introduced into the OCDWOA to enhance the search performance of WOA. The operators include opposition-based learning method, nonlinear parameter design, density peak clustering strategy, and differential evolution. The proposed algorithm is tested on 19 optimization benchmark functions and a seismic inversion problem. OCDWOA is compared with the classical WOA and three typical variants of WOA. The results demonstrate that OCDWOA outperforms the compared algorithms in terms of obtaining the global optimal solution.

1. Introduction

In recent years, the nature-inspired algorithms have attracted lots of attention from researchers. Swarm intelligence (SI) algorithm is a main kind of nature-inspired algorithms. Swarm intelligence algorithms are designed by simulating the behaviors of animals or plants. The widely used SI algorithms include particle swarm optimization (PSO) [1], Ant colony optimization (ACO) [2], artificial bee colony algorithm (ABC) [3], grey wolf optimization (GWO) [4], wolf pack algorithm (WPA) [5], brain storm optimization [6], and whale optimization algorithm (WOA) [7].

The whale optimization algorithm (WOA) is proposed in 2016. WOA searches the global optimal solution by simulating the foraging behaviors of the whales, including encircling prey, researching prey, and attacking prey. Many studies about WOA have been presented. Generally, these studies can be classified into two categories. One is to apply applying the WOA to solve some practical issues. The other one is to improve the performance of WOA with some strategies. A modified WOA named chaotic WOA (CWOA) [8] was proposed by Kaur and Arora in 2017. The CWOA introduces chaos theory into WOA. The main parameter of WOA is tuned by chaotic map to control exploration and exploitation. In [9], Lévy flight trajectory-based WOA (LWOA) is presented. The Lévy flight trajectory is used to increase the diversity of the population and enhance the performance of jumping out of local optimal solutions. The balanced variant of WOA (BWOA) proposed in [10] introduced Lévy flight and chaotic local search into WOA to enhance the inclusive exploratory and neighborhood-informed capacities of the conventional technique. Chen [11] proposed a modified WOA based on adaptive convergence and Lévy features (IWOA). The adaptive convergence operator is used to balance the local and global optimization ability and the Lévy flight mechanism is introduced to improve global searching ability. For solving the large-scale global optimization problems, a modified self-adaptive WOA (MWOA) [12] is proposed. A nonlinear dynamic strategy is applied to update the control parameter and

balance the exploration and exploitation abilities. The Lévy-flight strategy is utilized to enhance the search performance and a quadratic interpolation method is adopted to improve the exploitation ability of the algorithm. For solving high-dimensional function optimization problems, A WOA based on Lamarckian learning (WOALam) [13] brought in the evolutionary theory of Lamarck to enhance the local search. Trivedi et al. [14] proposed an adaptive whale optimization algorithm (AWOA) with a randomization technique. The results show that the performance of AWOA is better than the performance of the standard WOA. Sun et al. proposed a whale optimization algorithm based on quadratic interpolation (QIWOA) [15]. A modified exploration process and quadratic interpolation are used to improve the exploitation ability and the search accuracy. A reinforced variant of WOA named RDWOA [16] was proposed in 2019. A strategy of random spare or random replacement is used to enhance the convergence speed. In addition, a strategy of double adaptive weight is introduced to improve the exploration ability in the early stage and exploitative ability in the later stage. Trivedi et al. [17] proposed a hybrid PSO-WOA. The PSO operator is used in exploitation phase and the basic WOA operator is adopted to explore in uncertain environments. The experimental results reveal the effectiveness of PSO-WOA compared to standard PSO and WOA. Wei et al. [18] proposed a modified WOA based on different searching paths and perceptual disturbance (PDWOA). Several spiral curves have been tested and the results show that the equal-pitch of Archimedean spiral curve is more capable than other types of spiral curve. Mafarja and Mirjalili [19] proposed a hybrid SA-WOA algorithm. In order to enhance the exploitation of WOA, the simulated annealing (SA) algorithm is introduced into the algorithm by searching the most promising regions located by basic WOA. An improved whale optimization algorithm (IWOA) [20] is proposed in 2016. Inertia weight, as a new parameter, is introduced to adjust the best solution in each iteration. Abdel-Basset et al. [21] proposed a memetic algorithm combining the WOA with a Tabu Search for solving quadratic assignment problems.

In addition, WOA has been applied to many engineering problems. Oliva et al. [22] used an improved chaotic whale optimization algorithm to estimate the parameters of solar cells. The design problem involves the solution of the complex nonlinear and multimodal objective functions. The proposed algorithm solves the design problem with the chaotic map and automatically adapt parameters. Zhang et al. [23] proposed an enhanced whale optimization algorithm (EWOA) to optimize the flow shop construction scheme. The enhanced WOA with Lévy flight is used to improve the robustness and efficiency of the existing construction stage and zone optimization approaches. Bentouati et al. [24] proposed a new variant WOA with pattern search algorithm (PS). The results prove the effectiveness and the performance of the proposed algorithm for solving the optimal power flow problem. Abd Elaziz and Oliva [25] proposed an enhanced WOA based on opposition (OBWOA). The opposition-based learning is used to enhance the exploration of the search space. OBWOA has the exploration abilities for different benchmark optimization functions and has been applied to estimate three different diode models of solar cells. Aljarah et al. [26] applied WOA to neural networks training. The proposed algorithm has been tested on 20 datasets. The results showed that the proposed method can obtain the best solutions and outperform the other compared algorithms on convergence speed. Medani et al. [27] proposed a modified WOA to solve the optimal reactive power dispatch problem. The results show that the proposed algorithm is more desirable than two comparative algorithms.

Prestack seismic inversion technology is an important seismic inversion technology to solve the problems of reservoir lithology and fluid prediction in oil and gas exploration and development. Physical properties obtained by general prestack inversion methods provide many detailed subsurface characteristics, but the method is not mature and limited by the high resolution, low speed, and poor stability of prestack inversion. Currently, this technology is in the research stage. Some inevitable technology problems should be solved before its large-scale applications [28, 29].

Many researches on solving the problems of prestack inversion by optimization algorithms have been studied. Xia et al. [30] and Sen [31] improved the speed and stability of prestack seismic waveform inversion by using the step inversion method. Simulated annealing (SA) and genetic algorithm (GA) have been successfully used in prestack seismic waveform inversion. Ingber and Rosen [32] conducted a comparative study on SA and GA. The experiment results proved that both algorithms could achieve global optimization and the speed of fast simulated annealing (VFSA) was the fastest. Liu et al. [33] proposed a new adaptive genetic algorithm for seismic inversion to overcome slow convergence rate and falling into local optimal solutions. Wu et al. [34] proposed a parameter inversion method based on improved differential evolution algorithm. The proposed method was effective in solving parameter inversion problems of prestack seismic data.

WOA has a good convergence rate, but the performance of WOA in finding the global optimal solution of the multimodal problems with many local optimal solutions is not ideal. In order to enhance the global search ability and solve the prestack seismic waveform inversion, a modified whale optimization algorithm named OCDWOA is proposed in this paper. There are four operators used in the proposed algorithm. Nonlinear parameter design is introduced to balance the ratio of exploitation phase and exploration phase. Density peak clustering strategy is adopted to decompose population into numbers of groups for enhancing the diversity of population. Opposition-based learning method and differential evolution are used to improve the search performance. Experiments on 19 optimization benchmark functions and a seismic inversion problem indicate the effectiveness of the proposed algorithm. The proposed algorithm OCDWOA represents outstandingly the experiments. The results demonstrate that the proposed algorithm OCDWOA can solve the seismic inversion problem successfully.

The main contributions are summarized as follows:

(1) The opposition-based learning method adopted by OCDWOA could avoid the solutions trapping into local optimal, which improves the ability of global search.
(2) The nonlinear parameter design of a in OCDWOA could achieve the global exploration in early stage and local exploitation in late stage.
(3) Density peak clustering strategy is adopted to decompose population into numbers of groups, which enhances the diversity of population.
(4) The proposed OCDWOA has an outstanding performance of solving seismic inversion problem, which verifies its universality for solving practical engineering problems.

The rest of this paper is organized as follows. Section 2 describes the WOA algorithm. Section 3 explains the proposed algorithm and the operators. The experimental results and comparisons are shown in Section 4. Finally, in Section 5, the conclusions are presented.

2. Whale Optimization Algorithm

WOA is a kind of population-based meta-heuristic algorithm. WOA simulates the hunting mode of humpback whales, which includes encircling prey, bubble-net attacking method, and search for prey.

2.1. Encircling Prey. WOA supposes that the optimal solution with the best fitness in current population is the prey. All the other solutions will encircle the prey and update themselves toward the optimal solution. The mathematical model of encircling behavior is as follows:

$$\vec{D} = \left|\vec{C} \cdot \vec{X}^{*}(t) - \vec{X}(t)\right|, \tag{1}$$

$$\vec{X}(t+1) = \vec{X}^{*}(t) - \vec{A} \cdot \vec{D}, \tag{2}$$

where t indicates the number of the current iteration, X is the current solution, X^* is the best solution in the population, and $\vec{A}$ and $\vec{C}$ are the coefficient vectors calculated as follows:

$$\begin{aligned} \vec{A} &= 2\vec{a} \cdot \vec{r} - \vec{a}, \\ \vec{C} &= 2\vec{r}, \end{aligned} \tag{3}$$

where $\vec{a}$ is linearly decreased from 2 to 0 over the iterations and $\vec{r}$ is a random vector in [0, 1].

2.2. Bubble-Net Attacking Method. In order to simulate the behavior of bubble-net attacking, shrinking encircling mechanism and spiral updating position are designed as follows.

2.2.1. Shrinking Encircling Mechanism. The mathematical model of shrinking encircling mechanism is defined as equations (1) and (2). Because of $\vec{a}$ is linearly decreased from 2 to 0, the value of $\vec{A}$ is in the range of [−1, 1]. The current solution can update itself toward the best solution with a random step.

2.2.2. Spiral Updating Position. The new position of current solution updated with the best solution is obtained as follows:

$$\vec{D}' = \left|\vec{X}^{*}(t) - \vec{X}(t)\right|, \tag{4}$$

$$\vec{X}(t+1) = \vec{D}' \cdot e^{\mathrm{bl}} \cdot \cos(2\pi l) + X^{*}(t), \tag{5}$$

where b is constant for defining logarithmic spiral and l is randomly generated number in the range [−1, 1]. To balance the two models, a selection parameter p is set. The selection parameter p in WOA is 0.5; there is a probability of 50% to choose either the shrinking encircling mechanism or the spiral updating position model. The finial mathematical model of bubble-net attacking is as follows:

$$\vec{X}(t+1) = \begin{cases} \vec{X}^{*}(t) - \vec{A} \cdot \vec{D}, & \text{if } p < 0.5, \\ \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + X^{*}(t), & \text{if } p \geq 0.5, \end{cases} \tag{6}$$

where p is a random number in [0, 1].

2.3. Search for Prey. In nature, whales will go away from each other and randomly search for a new prey. In order to simulate the search behavior, a random search operation is used in the exploration phase when $|A| \geq 1$. The current solution will update with a random solution in the current population instead of the best solution. The mathematical model is as follows:

$$\begin{aligned} \vec{D} &= \left|\vec{C} \cdot \vec{X}_{\mathrm{rand}}(t) - \vec{X}(t)\right|, \\ \vec{X}(t+1) &= \vec{X}_{\mathrm{rand}}(t) - \vec{A} \cdot \vec{D}, \end{aligned} \tag{7}$$

where $\vec{X}_{\mathrm{rand}}$ is a random solution in the current population.

The pseudocode of the WOA is shown as Algorithm 1.

3. The Proposed Approach

This section introduces a modified WOA named OCDWOA. Four strategies are introduced into the WOA. Opposition-based learning method (OBL) is used to initialize population and randomly update solutions in each iteration. Nonlinear parameter design is introduced to balance the ratio of exploitation phase and exploration phase. Density peak clustering strategy is adopted to decompose population into numbers of groups for enhancing the diversity of population. The effect of differential evolution (DE) is to exchange information between groups and improve the ability of jumping out of the local optimal solutions.

3.1. OCDWOA. The pseudocode of the OCDWOA algorithm is shown as Algorithm 2.

```
Input: Population size N
The halting criterion
Output: The best solution X*
Generate the initial population X_i (i = 1, 2, ..., N)
Calculate the fitness for each solution in the population
Find the best solution X* with the best fitness
while the halting criterion is not satisfied do
  Update a
  for i = 1 to N do
    Update A, C, l and p
    for j = 1 to m do
      if p < 0.5 then
        if (|A| < 1) then
          D = |C · X*(j) − X_i(j)|
          X_i(j) = X*(j) − A · D
        else if (|A| ≥ 1) then
          Select a random solution X_rand
          D = |C · X_rand(j) − X_i(j)|
          X_i(j) = X_rand(j) − A · D
        end if
      if p ≥ 0.5 then
        D' = |X*(j) − X_i(j)|
        X_i(j) = D' · e^(bl) · cos(2πl) + X*(j)
      end if
    end for
    Evaluate the fitness for X_i
  end for
  Find the best solution X*
end while
```

Algorithm 1: WOA.

Opposition-based learning is used to initialize the population (see lines 1–6 of Algorithm 2) and to estimate the opposition solutions (see lines 39–46 of Algorithm 2). The fitness of each solution in the population is calculated and the best solution is found (see lines 7-8 in Algorithm 2). The parameters a, CR, A, C, l, and p are updated (see lines 10 and 13 in Algorithm 2). The population is decomposed into k groups by density peak clustering algorithm (see line 11 of Algorithm 2). The mutation operator of DE is introduced into the proposed algorithm (see lines 20–23 of Algorithm 2). A solution X_i is updated by r_1 and r_2 under mutation operator, where r_1 is from the group including X_i and r_2 is from the groups excluding X_i (see lines 14-15 of Algorithm 2). With the increase of iterations, the operation probability of DE declines. The basic three operators of WOA are used in lines 25-26, 29–31, and 34-35, respectively. The encircling prey is used in lines 25-26, the search for prey is introduced in lines 29–31, and the spiral updating position is applied in lines 34-35.

3.2. Opposition-Based Learning Method. It has been tested that opposition-based learning (OBL) method proposed by Rahnamayan et al. [35] is helpful for avoiding trapping into local optimal solutions.

A solution X_i in d-dimension search space is denoted as $X_i^* = \{X_{i,1}, X_{i,2}, X_{i,3}, \ldots, X_{i,d}\}$. The opposition point X_i^* of X_i can be calculated as

$$X'_{i,j} = \left(X_{ub,j} + X_{lb,j}\right) - X_{i,j}, \tag{8}$$

where $X_{ub,j}$ and $X_{lb,j}$ denote the upper and lower bound of j-th dimension.

OBL is used in the initialization step and solution update step at the end of each iteration. In the initialization step, OBL can generate a new solution on opposition side and enhance the diversity in the search space. In solution update step at the end of each iteration, OBL can generate a new solution on opposition side and choose the better solution between the new solution and the original solution. OBL can help WOA improve the search performance and jump out the local optimal solutions.

3.3. Nonlinear Parameter Design. An exponential function $\vec{a}$ is introduced into the modified WOA. The value of coefficient vector $\vec{A}$ controls the proportion of exploitation phase and exploration phase. The nonlinear parameter design $\vec{a}$ is designed as follows [36]:

$$a = a_{\min} + \left(a_{\max} - a_{\min}\right) \times e^{-2t/\text{Max\_iter}}, \tag{9}$$

where $a_{\min}$ and $a_{\max}$ denote the minimum and maximum of $\vec{a}$, t is the number of the current iteration, and Max_iter is the maximum iteration.

A large value of $\vec{a}$ generates a large step size in exploitation phase, which can improve the searching speed. A

```
Input: Population size N
       Group size k
       The halting criterion
Output: The best solution X*
Generate the initial population X_i (i = 1, 2, ..., N)
for i = 1 to N do
  if (rand < 0.6)
    Generate new solution with OBL
  end if
end for
Calculate the fitness for each solution in the population
Find the best solution X* with the best fitness
while the halting criterion is not satisfied do
  Update a and CR
  Decompose population into k groups by density peak clustering algorithm
  for i = 1 to N do
    Update A, C, l and p
    R_1 is the group with X_i and R_2 are the groups without X_i
    Select randomly r_1 from R_1 and r_2 from R_2
    j_rand = randint (1, m)
    for j = 1 to m do
      if p < 0.5 then
        if (|A| < 1) then
          if rand (0, 1) ≤ CR and j == j_rand then
            Update F
            Select randomly r_1 from R_1 and r_2 from R_2
            U_i(j) = X*(j) + F[X_{r_2}(j) − X_{r_1}(j)]
          Else
            D = |C · X*(j) − X_i(j)|
            X_i(j) = X*(j) − A · D
          end if
        elseif (|A| ≥ 1) then
          Select a random solution X_rand
          D = |C · X_rand(j) − X_i(j)|
          X_i(j) = X_rand(j) − A · D
        end if
      if p ≥ 0.5 then
        D' = |X*(j) − X_i(j)|
        X_i(j) = D' · e^{bl} · cos(2πl) + X*(j)
      end if
    end if
  end for
  for i = 1 to N do
    if (rand < 0.6)
      Generate new solution O_i with OBL
      if O_i is better than X_i then
        X_i = O_i
      end if
    end if
  end for
  Find the best solution X* with the best fitness
end while
```

Algorithm 2: OCDWOA.

small step size with a small value of $\vec{a}$ can be helpful for the exploration phase.

3.4. Density Peak Clustering. A density-based clustering strategy named density peak clustering algorithm [37] is used in the proposed WOA. The density peak clustering algorithm defines two important parameters for each solution i: the density ρ_i and distance δ_i. A scatter graph called decision diagram is constructed based on these two parameters. The decision diagram could decide the center of each clustering group and assign the remaining solutions to each clustering center.

The mathematical model of the density ρ_i of the ith solution used in this paper is as follows:

$$\rho_i = \sum_j e^{-\left(d_{ij}/d_c\right)^2}, \tag{10}$$

where j represents the j-th solution in current population, and $i \neq j$, d_{ij} is the Euclidean distance between i-th and j-th solutions, and d_c is a pre-set cut-off distance. A moderate value of d_c could make the average number of neighbors around 1% to 2% of the population size.

The distance parameter δ_i denotes the minimum distance among the i-th solution and the j-th solution, where the local density of j-th solution is higher than that of the i-th solution. The equation of δ_i is as follows:

$$\delta_i = \min_{j:\ \rho_j > \rho_i} \left(d_{ij}\right). \tag{11}$$

Figure 1(a) shows the distribution of 28 solutions in two dimensions and Figure 1(b) is the decision diagram obtained by these 28 solutions. In Figure 1(b), the horizontal axis is ρ and the vertical axis is δ. The solutions with both large values of ρ and δ are defined as the clustering centers. In Figure 2, the 1st and 10th solutions with both large density and distance will be selected as the clustering centers. The 26th, 27th, and 28th solutions have large distance and low density, and the 7th and 8th solutions have large density and low distance. Therefore, these solutions cannot be selected as the clustering centers.

According to the rules of selecting the clustering centers, a modified equation of calculating quality of solutions is designed as follows:

$$Q_i = \rho_i \cdot \delta_i. \tag{12}$$

In this paper, a modified model of cut-off distance d_c is introduced [38]. The transition distance Dis is calculated as follows:

$$\mathrm{Dis} = 1 - e^{-d_{ij}^2/2}. \tag{13}$$

The cut-off distance d_c is calculated as follows:

$$d_c = \frac{\mathrm{Dis}_{\max} + \mathrm{Dis}_{\min}}{2}, \tag{14}$$

where $\mathrm{Dis}_{\max}$ and $\mathrm{Dis}_{\min}$ are the maximum and minimum of transition distances in the current iteration.

3.5. Differential Evolution. For enhancing abilities of searching and jumping out of the local optimal solutions, the DE operator is introduced in this paper. The crossover and mutation operators in differential evolution (DE) [39] are calculated as follows:

$$U_i(j) = X^*(j) + F \times \left[X_{r_2}(j) - X_{r_1}(j)\right], \tag{15}$$

where X_{r_1} and X_{r_2} are the j-th dimension of the r_1-th and r_2-th solutions, respectively, U is the offspring, and F is the zoom factor. CR is the crossover probability updated as an exponential function:

$$\mathrm{CR} = 0.9 \times e^{-3t/\mathrm{Max\_iter}}. \tag{16}$$

4. Experiment Results and Discussion

4.1. Parameter Settings and Benchmark Functions. In this section, in order to analyze the performance of the proposed algorithm, 19 benchmark functions [40] F_1–F_{19} have been used. The expressions of all the functions are given in Table 1. F_1–F_7 are the unimodal functions with only one global optimal solution. F_8–F_{12} are the multimodal functions with many local optimal solutions. These functions can test the performance of jumping out the local optimal solutions. F_{13}–F_{19} are the fixed-dimension multimodal functions, which can evaluate the convergence property of an algorithm.

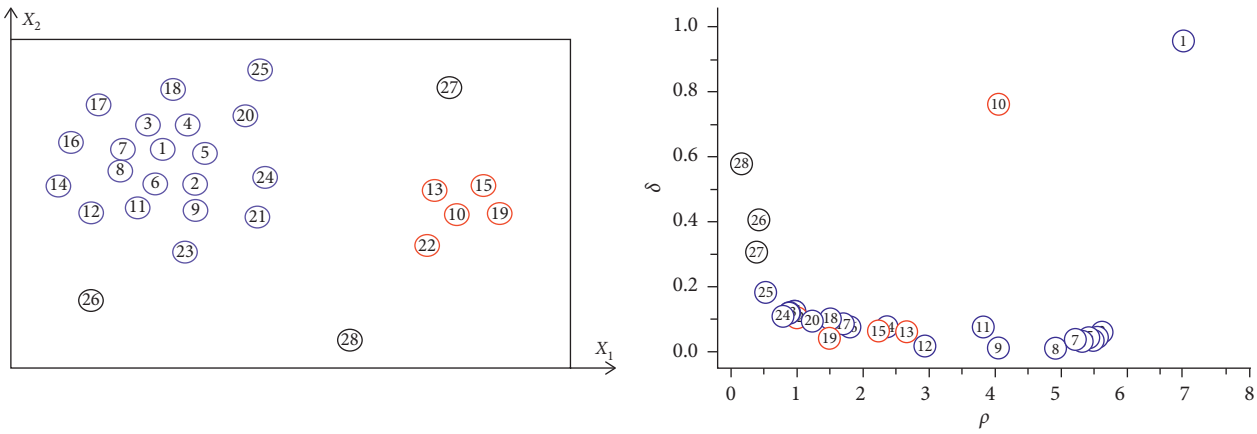
The proposed OCDWOA is compared with the classical WOA and its typical variants, such as OWOA [41], OEWOA [42], and IWOA [43]. The parameter b is set to 1 for all the algorithms. The parameter $\vec{a}$ decreases linearly from 2 to 0 for WOA, OWOA, and IWOA. For OEWOA, the parameter $\vec{a}$ nonlinearly decreases from 2 to 0. For IWOA, the crossover probability CR is set to 0.9, and the zoom factor F is random within (0.2, 0.8).

For all the algorithms, population size is 30 and maximum iteration is 500. The experiment is repeated 30 times for each benchmark function.

4.2. Experiment Results. The statistics of the experimental results is demonstrated in Table 2. The best fitness, mean fitness, and standard deviation of OCDWOA and each compared algorithm are given.

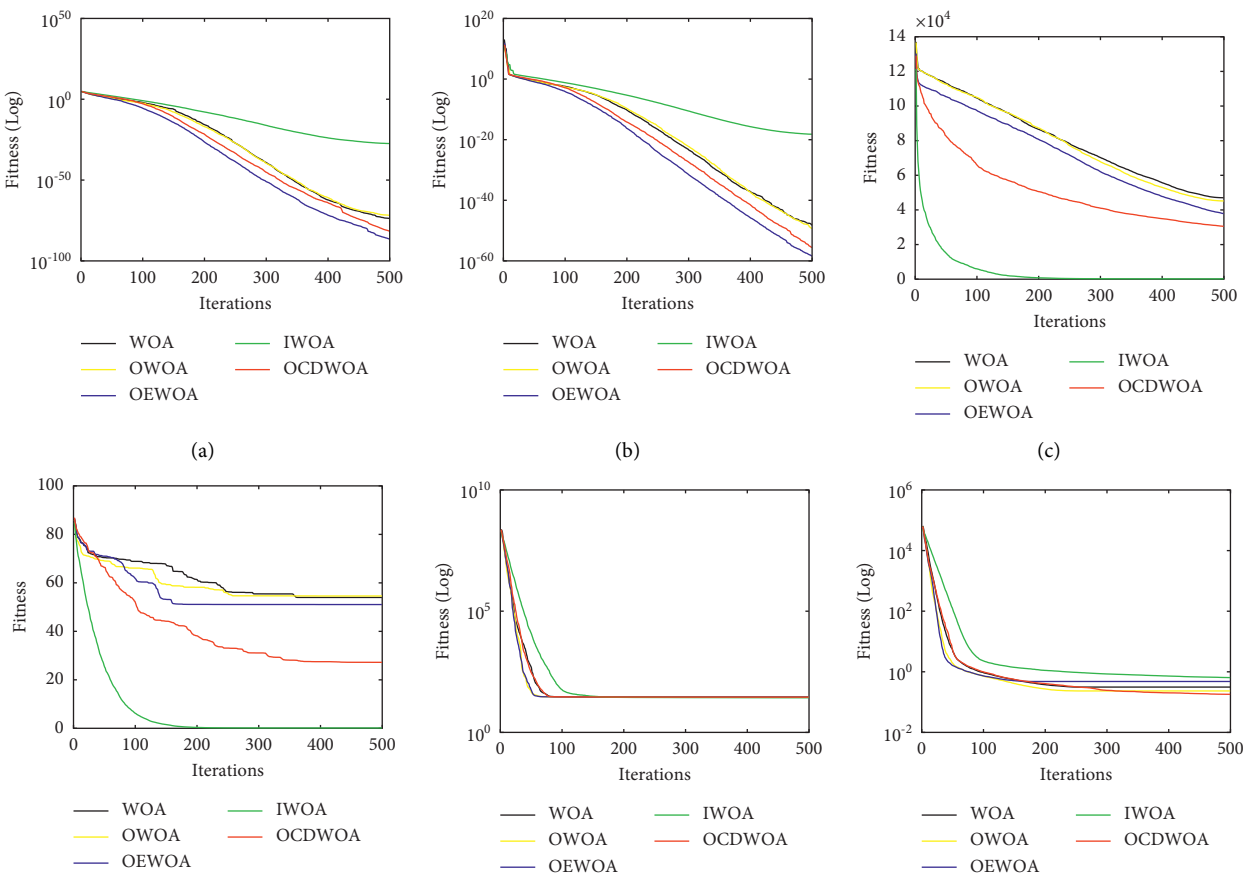
For F_6, F_9, F_{11}, F_{12}, F_{13}, F_{15}, F_{17}, and F_{18}, the performance of OCDWOA is the absolutely best. For F_{14}, the results obtained by all the variants of WOA are comparable. For F_1, F_2, F_3, F_4, F_5, F_{10}, and F_{16}, OCDWOA obtains the second rank. For the simple functions F_1 and F_2 without any local optimal solution, OCDWOA is worse than OEWOA. The DE operator used in OCDWOA can enhance the search performance. The convergence speed would be reduced, though the operation probability of DE declines in exploration phase. For the nonseparable unimodal functions F_3, F_4, and F_5, OCDWOA is only worse than IWOA. For the multimodal function F_{16}, the result reveals the positive influence of DE.

For F_7 and F_8, the performance of OCDWOA is similar to ones of OWOA and OEWOA. For F10, OCDWOA gets the second rank. For F_8 and F_{10}, IWOA can get the optimal solution. However, IWOA fails in solving the function F_7. The mean value and standard deviation value of IWOA are also not ideal. It shows that the DE operator is helpful for jumping out the local optimal solutions, but the stability may be impaired for the multimodal functions with easy or few local optimal solutions. In OCDWOA, the clustering operator and reducing the operation probability of DE can drop off the influence of DE. The DE operator can improve the performance of jump out the local optimal solutions. The DE operator and OBL can enhance the search performance of algorithm. Dynamical adjustment of $\vec{a}$ can avoid premature convergence. The clustering operator enhances information exchange between solutions.



(a) (b)

Figure 1: (a) Solution distribution in two dimensions. (b) Decision diagram for each sample point.



(a) (b) (c) (d) (e) (f)

Figure 2: Continued.

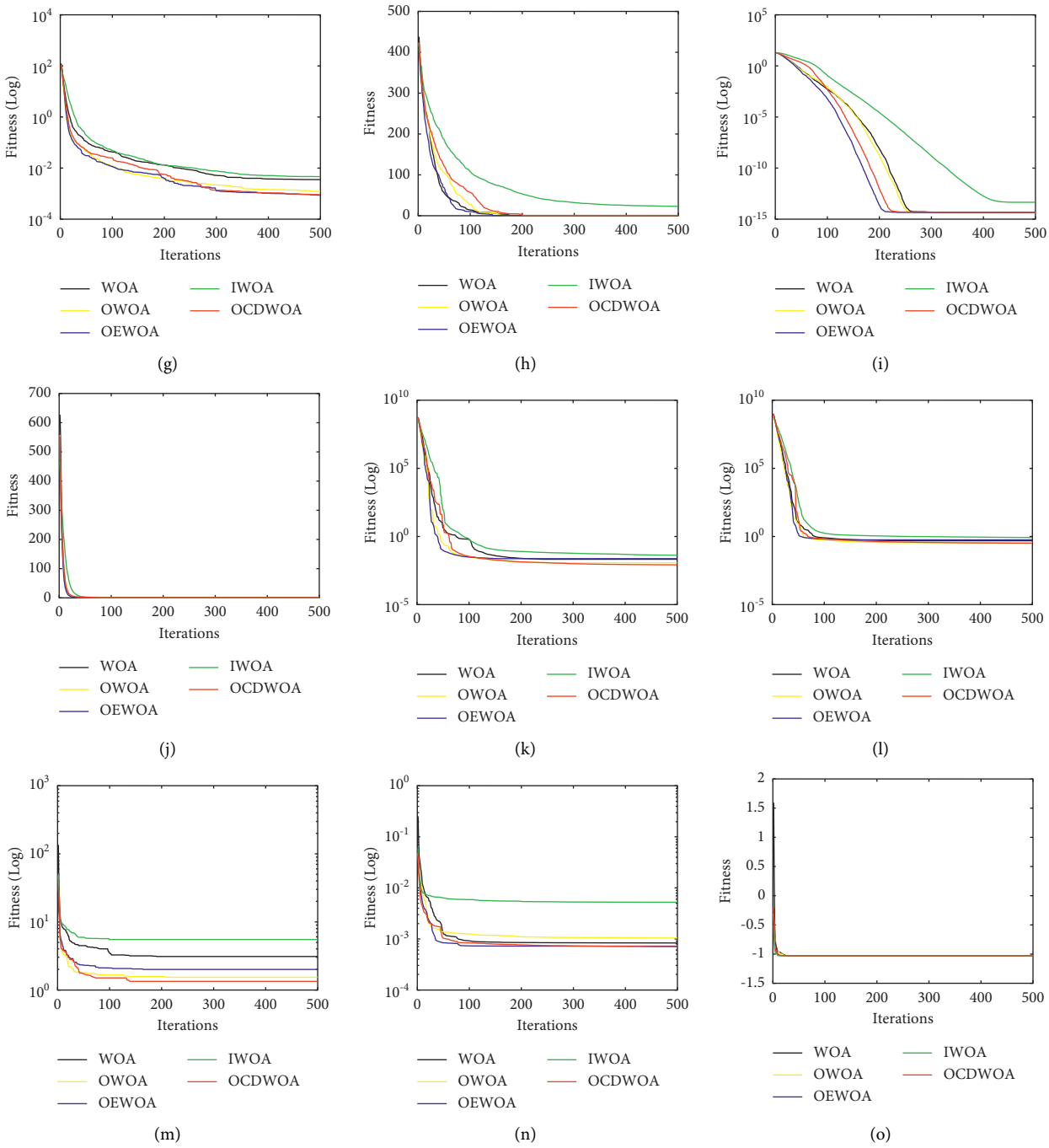


Figure 2: Continued.

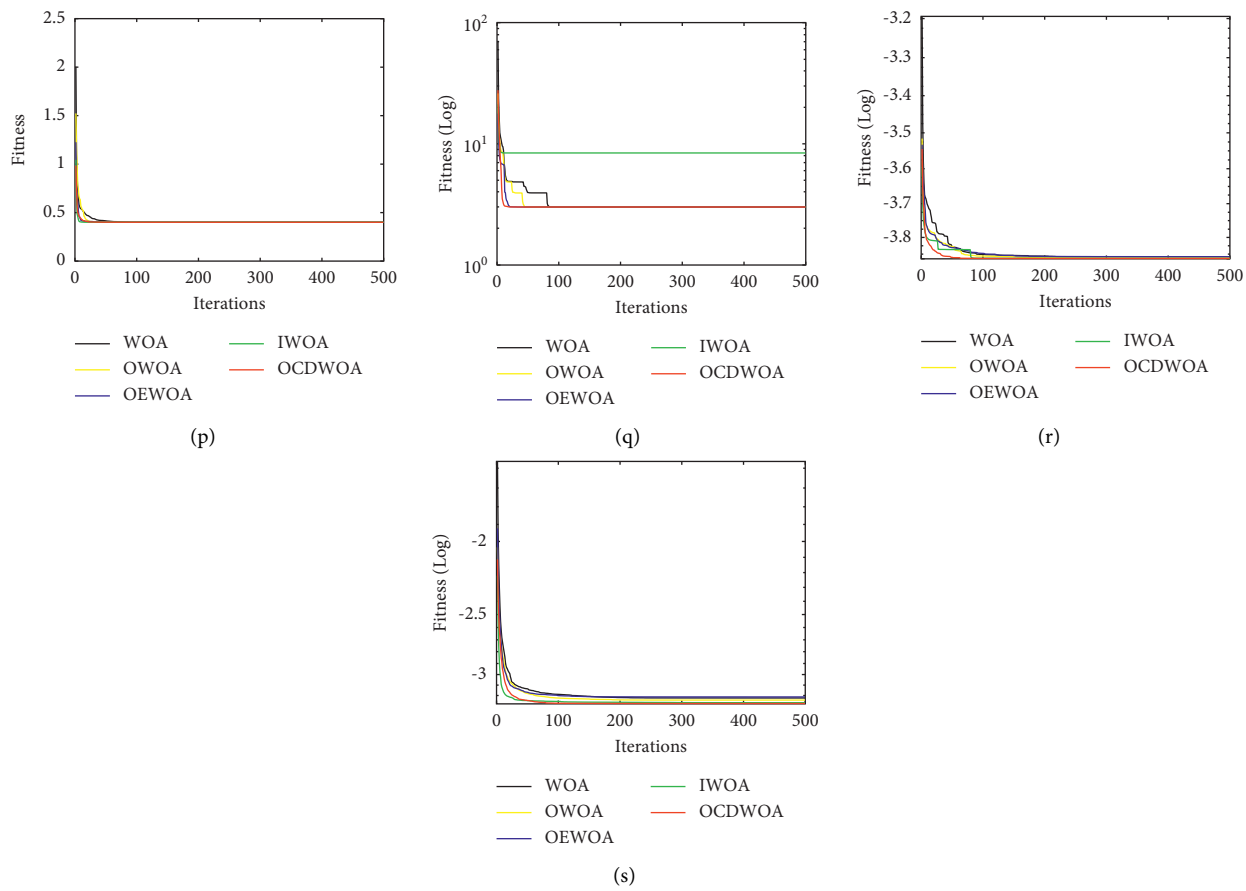


Figure 2: Convergence graphs of OCDWOA with respect to other algorithms. (a) F_1. (b) F_2. (c) F_3. (d) F_4. (e) F_5. (f) F_6. (g) F_7. (h) F_8. (i) F_9. (j) F_{10}. (k) F_{11}. (l) F_{12}. (m) F_{13}. (n) F_{14}. (o) F_{15}. (p) F_{16}. (q) F_{17}. (r) F_{18}. (s) F_{19}.

The results analyzed by the Wilcoxon rank-sum test are shown in Table 3. In the Wilcoxon rank-sum test, a significance level is set to 0.05. In Table 3, "+," "−," and "=" indicate that the result of the proposed algorithm OCDWOA is significantly better, significantly worse, and statistically similar to that of the compared algorithm, respectively. For most test functions, the performance of OCDWOA is significantly better or statistically similar to the other variants of WOA.

The convergence curves and boxplots are shown in Figures 2 and 3, respectively.

In Figure 2, for most functions, the convergence curves of OCDWOA are better or statistically similar to ones of the other variants of WOA. For F_3 and F_4, the convergence curves of OCDWOA are stabler than those of IWOA, but the convergence rates of OCDWOA are slower than those of IWOA. In OCDWOA, the solutions selected in the crossover and mutation operators of DE are based on the clustering. The clustering operator reduces the randomness of the DE operator. The stability of OCDWOA has been improved.

In Figure 3, for most functions, the stability of OCDWOA is better or statistically similar to the stability of the other variants of WOA. Only for F_3 and F_4, the stability of OCDWOA is worse than the stability of IWOA.

4.3. Application of Seismic Inversion Problems

4.3.1. Seismic Inversion Problem. Seismic inversion is the key technology of quantitative reservoir prediction and prestack waveform inversion [44] is a newly developed seismic inversion technique. In prestack waveform inversion technique, the complex wave propagation effect is considered and the complete waveform information is fully used. The prestack waveform inversion method assumes that each CMP/CDP point of underground medium layer structure with local one-dimensional or level uses the reflectivity method forward prestack data [45]. The prestack waveform inversion method searches the best fitting with the actual prestack seismic data of underground through genetic algorithm, simulated annealing, and other global optimization algorithms [46]. At present, there are many problems in prestack waveform inversion technology that need to be further studied and the resolution of prestack waveform inversion technology needs to be improved. The existing optimization algorithms used in prestack waveform inversion are mainly traditional intelligent optimization algorithms. These algorithms are prone to premature convergence and local optimal solution when dealing with large data volume, highly nonlinear, multiparameter, and multiextremum prestack waveform inversion problems.

Table 1: Description of benchmark functions.

Function	Dim	Range	$f_{\min}$
Unimodal functions			
$F_1 = \sum_{i=1}^{n} x_i^2$	30	[−100, 100]	0
$F_2 = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	30	[−10, 10]	0
$F_3 = \sum_{i=1}^{n} (\sum_{j=1}^{i} x_j)^2$	30	[−100, 100]	0
$F_4 = \max_i \{\lvert x_i \rvert, 1 \le i \le n\}$	30	[−100, 100]	0
$F_5 = \sum_{i=1}^{n-1} [100 (x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	[−30, 30]	0
$F_6 = \sum_{i=1}^{n} (\lvert x_i + 0.5 \rvert)^2$	30	[−100, 100]	0
$F_7 = \sum_{i=1}^{n} i x_i^4 + \text{random}[0, 1)$	30	[−1.28, 1.28]	0
Multimodal functions			
$F_8 = \sum_{i=1}^{n} [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	[−5.12.5.12]	0
$F_9 = -20 \exp(-0.2\sqrt{1/n} \sum_{i=1}^{n} x_i^2) - \exp[1/n \sum_{i=1}^{n} \cos(2\pi x_i)] + 20 + e$	30	[−32, 32]	0
$F_{10} = 1/4000 \sum_{i=1}^{n} x_i^2 + \prod_{i=1}^{n} \cos(x_i/\sqrt{i}) + 1$	30	[−600, 600]	0
$F_{11} = \pi/n \left\{10 \sin(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1}) + (y_n - 1)^2]\right\} + \sum_{i=1}^{n} u(x_i, 10, 100, 4)$ $y_i = 1 + x_i + 1/4 u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m & a < x_i \\ 0 & -a < x_i < a \\ k(-x_i - a)^m & x_i < -a \end{cases}$	30	[−50, 50]	0
$F_{12} = 0.1\left\{\sin^2(3\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2 \left\{1 + \sin^2(3\pi x_{i+1}) + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)]\right\}\right\} + \sum_{i=1}^{n} u(x_i, 5, 100, 4)$	30	[−50, 50]	0
Fixed-dimension multimodal functions			
$F_{13} = (1/500 + \sum_{j=1}^{25} 1/j + \sum_{i=1}^{2} (x_i - a_{ij})^6)^{-1}$	2	[−65, 65]	0
$F_{14} = \sum_{i=1}^{11} [a_i + x_i (b_i^2 + b_i x_2)/b_i^2 + b_i x_3 + x_4]^2$	4	[−5, 5]	0.00030
$F_{15} = 4x_1^2 - 2.1x_1^4 + 1/3x_1^6 + x_1 x_2 - 4x_2^2 + 4x_2^4$	2	[−5, 5]	−1.0316
$F_{16} = (x_2 - 5.1/4\pi^2 x_1^2 + 5/\pi x_1 - 6)^2 + 10(1 - 1/8\pi)\cos x_1 + 10$	2	[−5, 5]	0.398
$F_{17} = [1 + (x_1 + x_2 + 1)^2 (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1 x_2 + 3x_2^2)] \times [30 + (2x_1 - 3x_2)^2 \times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1 x_2 + 27x_2^2)]$	2	[−2, 2]	3
$F_{18} = -\sum_{i=1}^{4} c_i \exp[-\sum_{j=1}^{3} a_{ij} (x_j - p_{ij})^2]$	3	[1, 3]	−3.86
$F_{19} = -\sum_{i=1}^{4} c_i \exp[-\sum_{j=1}^{6} a_{ij} (x_j - p_{ij})^2]$	6	[0, 3]	−3.12

Table 2: Results of OCDWOA and compared algorithms.

Function	Algorithm	Best	Mean	Std
F_1	WOA	$3.45E-84$	$1.84E-74$	$6.85E-74$
	OWOA	$3.67E-87$	$1.54E-72$	$6.15E-72$
	OEWOA	**8.10E − 104**	**3.41E − 87**	**1.61E − 86**
	IWOA	$3.88E-31$	$4.18E-28$	$1.52E-27$
	OCDWOA	$1.16E-94$	$2.68E-82$	$8.18E-82$
F_2	WOA	$3.76E-59$	$4.29E-49$	$2.30E-48$
	OWOA	$8.72E-58$	$3.71E-50$	$9.41E-50$
	OEWOA	**8.88E − 68**	**2.26E − 59**	**1.15E − 58**
	IWOA	$6.02E-20$	$6.36E-19$	$6.72E-19$
	OCDWOA	$1.85E-63$	$1.86E-56$	$4.92E-56$
F_3	WOA	29331.39043	46962.31228	11654.1269
	OWOA	23102.63585	45115.1214	9775.63579
	OEWOA	10896.08787	37765.42249	12462.65209
	IWOA	**0.0204919**	**4.761001851**	**12.20526468**
	OCDWOA	9123.146589	30378.88003	11250.945
F_4	WOA	2.845944535	53.98492009	28.64604813
	OWOA	0.221683283	54.59717151	28.7671405
	OEWOA	0.395500666	51.05649771	30.09654391
	IWOA	**2.53E − 05**	**1.43E − 03**	**1.46E − 03**
	OCDWOA	0.091259942	27.15705548	22.72468237
F_5	WOA	27.05096579	27.98206458	0.492345074
	OWOA	27.0462713	27.8628119	**0.3574981**
	OEWOA	27.47961584	28.17255965	0.377986162
	IWOA	**24.72608991**	**26.21998458**	1.236007685
	OCDWOA	26.94330849	27.75236194	0.519791973
F_6	WOA	0.039352423	0.308186276	0.178197178
	OWOA	0.058900257	0.229444501	**0.138575416**
	OEWOA	0.094203593	0.471932521	0.211010397
	IWOA	0.132553897	0.634850188	0.363709611
	OCDWOA	**0.011880743**	**0.177746121**	0.202309313
F_7	WOA	$7.31E-05$	$3.49E-03$	$4.08E-03$
	OWOA	$9.68E-06$	$1.20E-03$	$2.15E-03$
	OEWOA	$1.78E-05$	**8.51E − 04**	**9.84E − 04**
	IWOA	0.001170766	0.004571304	0.003016255
	OCDWOA	**1.19E − 05**	$8.98E-04$	$1.21E-03$
F_8	WOA	0	$5.68E-15$	$2.29E-14$
	OWOA	0	0	0
	OEWOA	0	0	0
	IWOA	0	22.36681071	24.91175078
	OCDWOA	0	0	0
F_9	WOA	$8.88E-16$	$4.32E-15$	$2.55E-15$
	OWOA	$8.88E-16$	**3.85E − 15**	$2.48E-15$
	OEWOA	$8.88E-16$	$4.20E-15$	$2.27E-15$
	IWOA	$2.22E-14$	$4.29E-14$	$1.27E-14$
	OCDWOA	$8.88E-16$	$4.44E-15$	**2.09E − 15**
F_{10}	WOA	0	0.014239143	0.055902557
	OWOA	0	0.006369305	0.034886118
	OEWOA	**0**	**0**	**0**
	IWOA	0	0.005601519	0.009188179
	OCDWOA	0	0.00307175	0.016824667
F_{11}	WOA	0.003303307	0.020703883	0.013561528
	OWOA	0.002886391	0.011523424	0.008134223
	OEWOA	0.008832644	0.022931786	0.008957016
	IWOA	0.01289871	0.04121751	0.023964971
	OCDWOA	**0.000984744**	**0.008086227**	**0.00817336**

Table 2: Continued.

Function	Algorithm	Best	Mean	Std
F_{12}	WOA	0.091759692	0.476972639	0.2731912
	OWOA	0.071293381	0.343978384	0.176667484
	OEWOA	0.264598322	0.558788251	0.219290838
	IWOA	0.360052384	0.826186996	0.250307701
	OCDWOA	**0.008972727**	**0.315016992**	**0.165937507**
F_{13}	WOA	0.998003838	3.087707817	3.643013869
	OWOA	0.998003838	1.527492833	0.770156088
	OEWOA	0.998003838	1.991064676	1.244254244
	IWOA	0.998003838	5.491627348	4.859439425
	OCDWOA	0.998003838	**1.328819151**	**0.7056337**
F_{14}	WOA	0.000309354	0.000754701	0.000591433
	OWOA	0.000308046	0.000692785	0.000506369
	OEWOA	0.000308042	**0.000599412**	0.000433827
	IWOA	**0.000307494**	0.005763864	0.008959844
	OCDWOA	0.000308015	0.00068318	**0.000396378**
F_{15}	WOA	−1.031628453	−1.031628453	$7.57E-10$
	OWOA	−1.031628453	−1.031628453	$1.10E-09$
	OEWOA	−1.031628453	−1.031628453	$5.29E-10$
	IWOA	−1.031628453	−1.031628453	$2.29E-10$
	OCDWOA	−1.031628453	−1.031628453	**3.68E − 11**
F_{16}	WOA	0.397887358	0.397894412	$1.60E-05$
	OWOA	0.397887358	0.397899721	$2.59E-05$
	OEWOA	0.397887358	0.397916992	0.000107977
	IWOA	0.397887358	**0.397887585**	**5.13E − 07**
	OCDWOA	0.397887358	0.397888516	$2.38E-06$
F_{17}	WOA	3.000000001	3.000043872	$7.74E-05$
	OWOA	3.00000002	3.000044635	0.000152146
	OEWOA	3.000000001	3.000123665	0.000226823
	IWOA	3	8.400000042	20.55035867
	OCDWOA	3.00000001	**3.000006355**	**1.60E − 05**
F_{18}	WOA	−3.862688519	−3.857262065	0.007143618
	OWOA	−3.862781863	**−3.861815423**	0.003572799
	OEWOA	−3.862782142	−3.85726697	0.019011496
	IWOA	−3.862782147	−3.862521555	0.001426916
	OCDWOA	−3.862782147	−3.862775916	**9.47E − 06**
F_{19}	WOA	−3.321580972	−3.223597877	0.10220952
	OWOA	−3.321545137	−3.244143349	0.085289027
	OEWOA	**−3.320833167**	**−3.213034803**	0.091087411
	IWOA	−3.32199476	−3.270086944	0.060385879
	OCDWOA	−3.321994702	−3.278324784	**0.058345799**

The bold values mean the best results.

Table 3: The Wilcoxon rank-sum test results of OCDWOA and compared algorithms.

	OCDWOA-WOA	OCDWOA-OWOA	OCDWOA-OEWOA	OCDWOA-IWOA
Comparison	+/=/−	+/=/−	+/=/−	+/=/−
Result	14/5/0	10/9/0	11/6/2	10/4/5

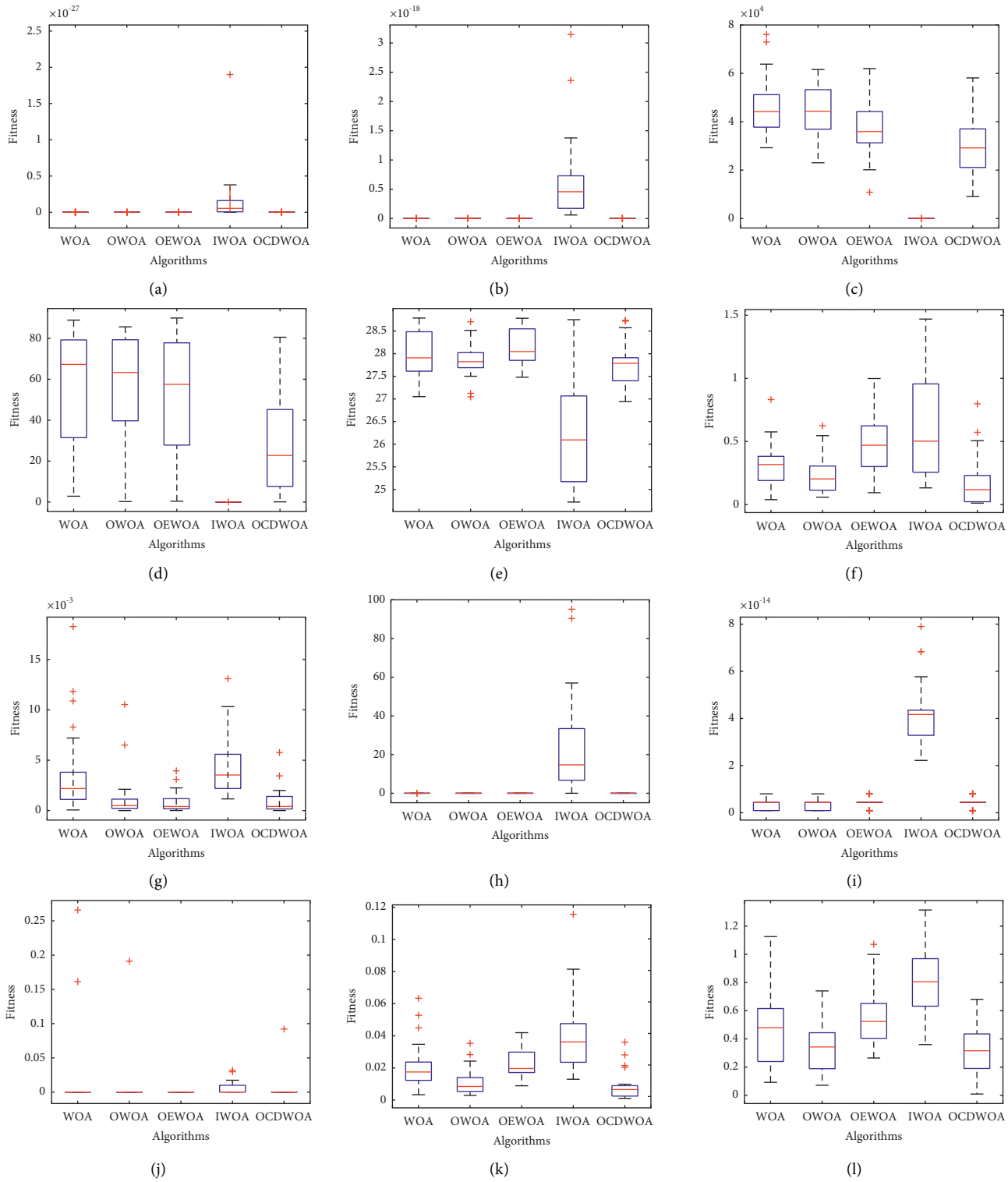


Figure 3: Continued.

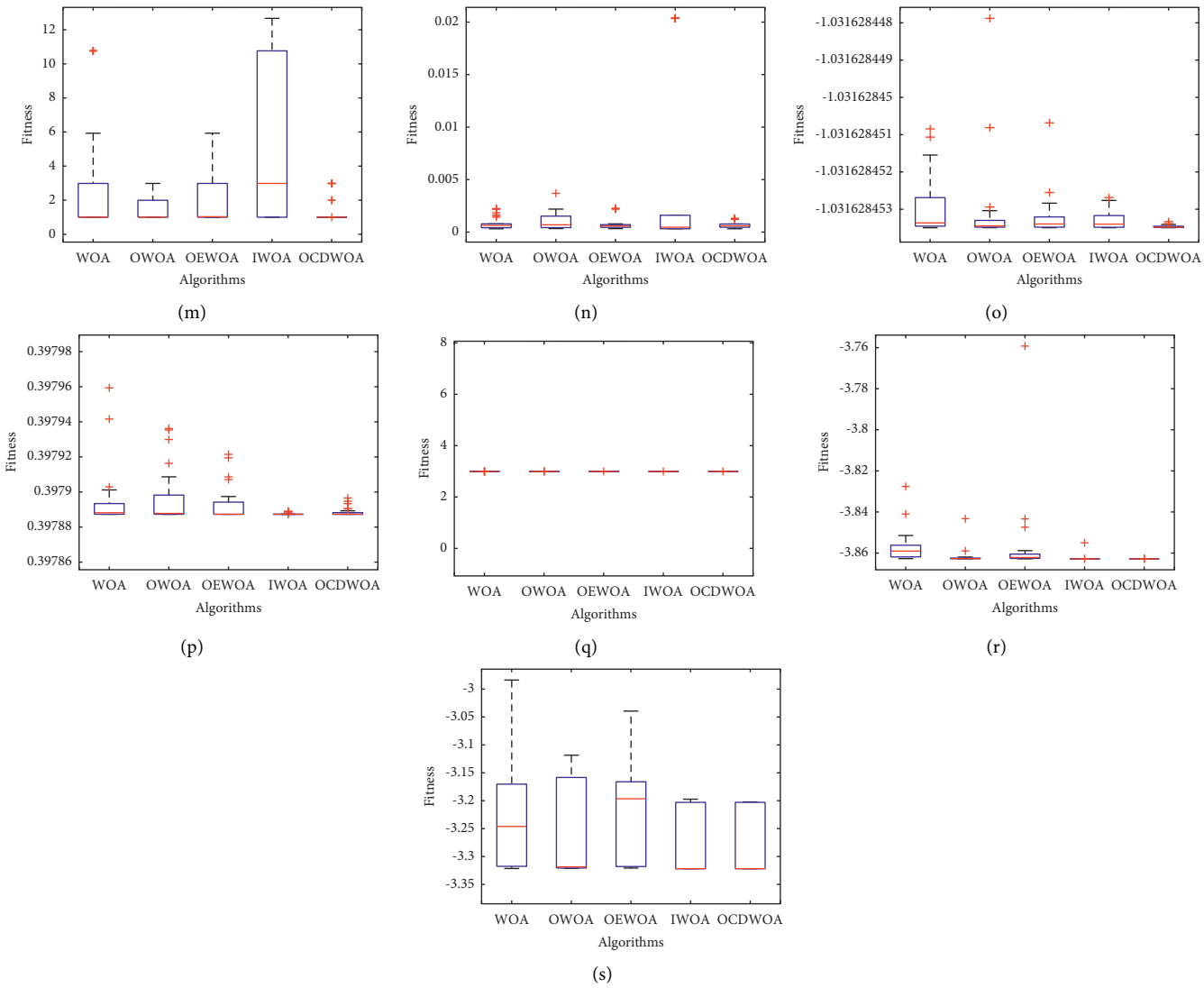


Figure 3: Boxplots of OCDWOA with respect to other algorithms. (a) F_1. (b) F_2. (c) F_3. (d) F_4. (e) F_5. (f) F_6. (g) F_7. (h) F_8. (i) F_9. (j) F_{10}. (k) F_{11}. (l) F_{12}. (m) F_{13}. (n) F_{15}. (o) F_{16}. (p) F_{17}. (q) F_{18}. (r) F_{19}. (s) F_{19}.

Table 4: Values of the theoretical geological model.

	V_P (m·s^{-1})	V_S (m·s^{-1})	ρ (g·cm^{-3})
Layer 1	1980	851	2
Layer 2	2185	1605	2.2
Layer 3	1983	856	2
Layer 4	2200	1622	2.2
Layer 5	2437	1510	2.3
Layer 6	2663	1380	2.4
Layer 7	1983	851	2
Layer 8	2437	1501	2.3
Layer 9	1980	856	2
Layer 10	2656	1380	2.4
Layer 11	1983	851	2

Table 5: Results of different algorithms.

Problem	Algorithm	Best	Mean	Std
Seismic inversion	WOA	0.007178754	0.0124127116	0.0038698430
	OWOA	0.006168262	0.0119509070	0.0053173170
	OEWOA	0.006901579	0.0112028011	0.0035417924
	IWOA	0.006136581	0.0128976065	0.0041582246
	OCDWOA	**0.002662229**	**0.0055975119**	**0.0019518514**

The bold values mean the best results.

In this paper, a theoretical geological model with 11 layers is established [47]. Each layer in the model contains 3 parameters: velocity of longitudinal wave (V_P), velocity of shear wave (V_S), and density (ρ). The values of model are shown in Table 4.

In this paper, the reflection coefficients of the model will be calculated by Zoeppritz equation as follows:

$$\begin{bmatrix} R_P \\ R_S \\ T_P \\ T_S \end{bmatrix} = \begin{bmatrix} -\sin\theta_1 & -\cos\phi_1 & \sin\theta_2 & \cos\phi_2 \\ \cos\theta_1 & -\sin\phi_1 & \cos\theta_2 & -\sin\phi_2 \\ \sin 2\theta_1 & \frac{V_{P1}}{V_{S1}}\cos 2\phi_1 & \frac{\rho_2 V_{S2}^2 V_{P1}}{\rho_1 V_{S1}^2 V_{P2}}\cos 2\phi_1 & \frac{\rho_2 V_{S2} V_{P1}}{\rho_1 V_{S1}^2}\cos 2\phi_2 \\ -\cos 2\phi_1 & \frac{V_{S1}}{V_{P1}}\sin 2\phi_1 & \frac{\rho_2 V_{P2}}{\rho_1 V_{P1}}\cos 2\phi_2 & -\frac{\rho_2 V_{S2}}{\rho_1 V_{P1}}\sin 2\phi_2 \end{bmatrix}^{-1} \begin{bmatrix} \sin\theta_1 \\ \cos\theta_1 \\ \sin 2\theta_1 \\ \cos 2\phi_1 \end{bmatrix}, \tag{17}$$

where V_{P1}, V_{S1}, and ρ_1 are the parameters of the up-layer, V_{P2}, V_{S2}, and ρ_2 are the parameters of the low-layer, θ_1 and θ_2 are the incident angle and the refraction angle of the P-wave, respectively, and Φ_1 and Φ_2 are the incident angle and the refraction angle of the S-wave, respectively.

Φ_1, θ_2, and Φ_2 are calculated as follows, respectively:

$$\begin{aligned} \phi_1 &= \operatorname{arc}\left(\sin(\theta_1) \times \frac{V_{S1}}{V_{P1}}\right), \\ \theta_2 &= \operatorname{arc}\left(\sin(\theta_1) \times \frac{V_{S2}}{V_{P1}}\right), \\ \phi_2 &= \operatorname{arc}\left(\sin(\theta_1) \times \frac{V_{P2}}{V_{P1}}\right). \end{aligned} \tag{18}$$

Incident angles θ_1 of 50 reflection coefficients are from 1° to 50°. The 50 seismic records are obtained by convolving the reflection coefficients with the Ricker wavelet with 35 Hz as follows:

$$S = R * Rw, \tag{19}$$

where S is the seismic record, R is the reflection coefficient, and Rw is the ricker wavelet (35 Hz).

The 50 seismic records will be divided into three groups. The three groups are small incident angle group (Group 1), medium incident angle group (Group 2), and large incident angle group (Group 3). Group 1 consists of the seismic records whose incident angles are from 1° to 15°. Group 2 consists of the seismic records whose incident angles are from 16° to 32°. Group 3 consists of the seismic records whose incident angles are from 33° to 50°. Superimposed seismic records S_1, S_2, and S_3 are obtained by synthesizing three groups of seismic records, respectively. The superimposed seismic records $S_1(X)$, $S_2(X)$, and $S_3(X)$ of the solution X are obtained the same way as the model.

The fitness of the solution X is expressed as follows:

$$F_{SI}(X) = \frac{1}{3}\left(\|S_1(X) - S_1\|_2 + \|S_2(X) - S_2\|_2 + \|S_3(X) - S_3\|_2\right). \tag{20}$$

4.3.2. Experiment Results. The experiment is repeated 10 times for each algorithm. Table 5 shows the best fitness, mean fitness, and standard deviation of test results on seismic inversion problem. The convergence curves and boxplots are shown in Figure 4. From the test results, the performance of OCDWOA for seismic inversion problem is the best.

In Table 5, the best fitness, mean fitness, and standard deviation of OCDWOA are all obviously smaller than those of the comparative algorithms. In Figure 4(a), the convergence rate of OCDWOA is better than those of the comparative algorithms. The convergence curve of OCDWOA is

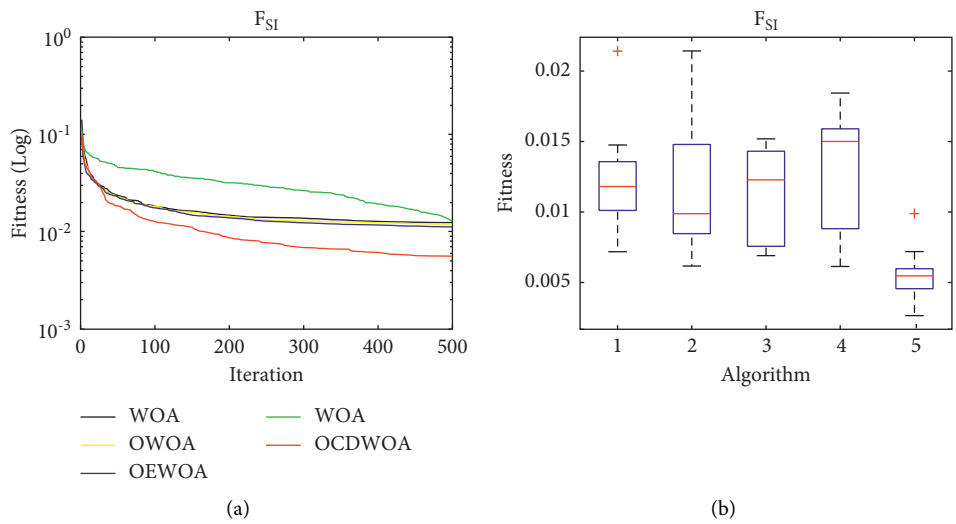


Figure 4: Results of OCDWOA with respect to other algorithms. (a) Convergence graph of the algorithms. (b) Boxplots of the algorithms.

stable. In Figure 4(b), the stability of OCDWOA is the best among the comparative algorithms.

5. Conclusions

In this paper, a modified WOA named OCDWOA is proposed. In order to balance the exploitation phase and exploration phase, a nonlinear parameter design is used. OBL, DE operator, and a density-based clustering strategy are introduced into the proposed algorithm to improve the performance of the search phase and the ability of jumping out of the local optimal solution. The proposed algorithm is tested on 19 optimization benchmark functions and a seismic inversion problem. The proposed algorithm is compared with WOA and the other three variants of WOA. Experimental results show that the proposed algorithm can enhance the performance of WOA and is better than other compared algorithms.

Our future work will pay attention to the following two points: first, we would like to research the relation between the parameters and the performance of the OCDWOA. Second, we would like to apply the OCDWOA to machine learning.

Data Availability

Some or all data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

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

Acknowledgments

This work was supported by National Natural Science Foundation of China under Grants 41772123, 61772365, 61802280, and 61806143.

References

[1] G. Venter and S. S. Jaroslaw, "Particle swarm optimization," *AIAA Journal*, vol. 41, no. 8, pp. 129–132, 2003.

[2] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28–39, 2006.

[3] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (abc) algorithm," *Journal of Global Optimization*, vol. 39, no. 3, pp. 459–471, 2007.

[4] A. Sm, B. Smm, and A. Al, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.

[5] S. Wu and F.-M. Zhang, "Wolf pack algorithm for unconstrained global optimization," *Mathematical Problems in Engineering*, vol. 2014, Article ID 465082, 17 pages, 2014.

[6] L. Ma, S. Cheng, and Y. Shi, "Enhancing learning efficiency of brain storm optimization via orthogonal learning design," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 11, pp. 6723–6742, 2021.

[7] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.

[8] G. Kaur and S. Arora, "Chaotic whale optimization algorithm," *Journal of Computational Design and Engineering*, vol. 5, no. 3, pp. 275–284, 2018.

[9] Y. Zhou, Y. Ling, and Q. Luo, "Lévy flight trajectory-based whale optimization algorithm for global optimization," *IEEE Access*, vol. 5, pp. 6168–6186, 2017.

[10] H. Chen, Y. Xu, M. Wang, and X. Zhao, "A balanced whale optimization algorithm for constrained engineering design problems," *Applied Mathematical Modelling*, vol. 71, no. JUL, pp. 45–59, 2019.

[11] X. Chen, "Research on new adaptive whale algorithm," *IEEE Access*, vol. 8, Article ID 90165, 2020.

[12] Y. Sun, X. Wang, Y. Chen, and Z. Liu, "A modified whale optimization algorithm for large-scale global optimization problems," *Expert Systems with Applications*, vol. 114, no. DEC, pp. 563–577, 2018.

[13] Q. Zhang and L. Liu, "Whale optimization algorithm based on lamarckian learning for global optimization problems," *IEEE Access*, vol. 7, 2019.

[14] I. N. Trivedi, J. Pradeep, J. Narottam, K. Arvind, and L. Dilip, "Novel adaptive whale optimization algorithm for global optimization," *Indian Journal of Science and Technology*, vol. 9, no. 38, pp. 319–326, 2016.

[15] Y. Sun, T. Yang, and Z. Liu, "A whale optimization algorithm based on quadratic interpolation for high-dimensional global optimization problems," *Applied Soft Computing*, vol. 85, Article ID 105744, 2019.

[16] H. Chen, C. Yang, A. A. Heidari, and X. Zhao, "An efficient double adaptive random spare reinforced whale optimization algorithm," *Expert Systems with Applications*, vol. 154, Article ID 113018, 2019.

[17] I. N. Trivedi, P. Jangir, A. Kumar, N. Jangir, and R. Totlani, "A novel hybrid PSO-WOA algorithm for global numerical functions optimization," in *Advances in Computer and Computational Sciences*, pp. 53–60, Springer, Berlin, Germany, 2018.

[18] S. Z. Wei, W. J. Sheng, and W. Xian, "An improved whale optimization algorithm based on different searching paths and perceptual disturbance," *Symmetry Plus*, vol. 10, no. 6, 2018.

[19] M. M Mafarja and S. Mirjalili, "Hybrid whale optimization algorithm with simulated annealing for feature selection," *Neurocomputing*, vol. 260, pp. 302–312, 2017.

[20] Y. X. Ting and H. U. H. Bai, "A whale optimization algorithm with inertia weight," *WSEAS Transactions on Computers*, vol. 15, pp. 319–326, 2016.

[21] M. Abdel-Basset, G. Manogaran, D. El-Shahat, and S. Mirjalili, "Integrating the whale algorithm with Tabu search for quadratic assignment problem: a new approach for locating hospital departments," *Applied Soft Computing*, vol. 73, pp. 530–546, 2018.

[22] D. Oliva, M. E. Abd Elaziz, and A. E. Hassanien, "Parameter estimation of photovoltaic cells using an improved chaotic whale optimization algorithm," *Applied Energy*, vol. 200, pp. 141–154, 2017.

[23] J. Zhang, D. Zhong, M. Zhao, J. Yu, and F. Lv, "An optimization model for construction stage and zone plans of rockfill dams based on the enhanced whale optimization algorithm," *Energies*, vol. 12, no. 3, 2019.

[24] B. Bentouati, L. Chaib, and S. Chettih, "A hybrid whale algorithm and pattern search technique for optimal power flow problem," in *Proceedings of the 2016 8th international conference on modelling, identification and control (ICMIC)*, pp. 1048–1053, IEEE, Algiers, Algeria, November 2016.

[25] M. Abd Elaziz and D. Oliva, "Parameter estimation of solar cells diode models by an improved opposition-based whale optimization algorithm," *Energy Conversion and Management*, vol. 171, no. 1083-186, pp. 1843–1859, 2018.

[26] I. Aljarah, H. Faris, and S. Mirjalili, "Optimizing connection weights in neural networks using the whale optimization algorithm," *Soft Computing*, vol. 22, no. 1, pp. 1–15, 2018.

[27] K. b. o. Medani, S. Sayah, and A. Bekrar, "Whale optimization algorithm based optimal reactive power dispatch: a case study of the Algerian power system," *Electric Power Systems Research*, vol. 163, pp. 696–705, 2018.

[28] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Transactions on Cybernetics*, 2021.

[29] L. Ma, N. Li, Y. Guo et al., "Learning to optimize: reference vector reinforcement learning adaption to constrained many-objective optimization of industrial copper burdening system," *IEEE Transactions on Cybernetics*, 2021.

[30] G. Xia, M. K. Sen, and P. L. Stoffa, "1-D elastic waveform inversion: a divide-and-conquer approach," *Geophysics*, vol. 63, no. 5, pp. 1670–1684, 1998.

[31] M. K. Sen, "Pre-stack wavefrom inversion: current status and future directions," in *Proceedings of the 7th International Congress of the Brazilian Geophysical Society*, European Association of Geoscientists & Engineers, Bahia, Brazil, October 2001, https://www.earthdoc.org/content/papers/10.3997/2214-4609-pdb.217.141.

[32] L. Ingber and B. Rosen, "Genetic algorithms and very fast simulated reannealing: a comparison," *Mathematical & Computer Modeling*, vol. 16, no. 11, 1992.

[33] Liu, W. A. N. G. Sixiu, H. U. Deli, and Bin, "Pre-stack seismic waveform inversion based on adaptive genetic algorithm," *Global Geology*, vol. 22, no. 3, pp. 50–60, 2019.

[34] Q. Wu, Z. Zhu, and X. Yan, "Research on the parameter inversion problem of prestack seismic data based on improved differential evolution algorithm," *Cluster Computing*, vol. 20, no. 4, pp. 2881–2890, 2017.

[35] S. Rahnamayan, H. R. Tizhoosh, and M. M. A. Salama, "Opposition-based differential evolution," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 1, pp. 64–79, 2008.

[36] S. Mirjalili, S. M. Mirjalili, S. Saremi, and S. Mirjalili, "Whale optimization algorithm: theory, literature review, and application in designing photonic crystal filters," *Nature-Inspired Optimizers*, pp. 219–238, 2020.

[37] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[38] J. Cai, H. Wei, H. Yang, and X. Zhao, "A novel clustering algorithm based on DPC and PSO," *IEEE Access*, vol. 8, Article ID 88200, 2020.

[39] K. V. Price, "Differential evolution," in *Handbook of Optimization*, pp. 187–214, Springer, Berlin, Heidelberg, 2013.

[40] P. N. Suganthan, N. Hansen, J. J. Liang et al., "Problem Definitions and Evaluation Criteria for the CEC 2005 Special Session on Real-Parameter Optimization," *Nanyang Technol. University*, Singapore, 2005.

[41] H. S. Alamri, Y. A. Alsariera, and K. Z. Zamli, "Opposition-based whale optimization algorithm," *Advanced Science Letters*, vol. 24, no. 10, pp. 7461–7464, 2018.

[42] R. Salgotra, U. Singh, and S. Saha, "On some improved versions of whale optimization algorithm," *Arabian Journal for Science and Engineering*, vol. 44, no. 11, pp. 9653–9691, 2019.

[43] S. Mostafa Bozorgi and S. Yazdani, "IWOA: an improved whale optimization algorithm for optimization problems," *Journal of Computational Design and Engineering*, vol. 6, no. 3, pp. 243–259, 2019.

[44] J. Pafeng, S. Mallick, and H. S. Sharma, "3-D prestack waveform inversion-a real data example," in *Proceedings of the 2015 SEG Annual Meeting*, OnePetro, New Orleans, LO, USA, October 2015, https://onepetro.org/SEGAM/conferences/browse-by-conference.

[45] J. P. Tchuindjang, S. Mallick, and H. Sharma, "Reservoir characterization using prestack waveform inversion," in *Proceedings of the SEG Technical Program Expanded Abstracts 2016*, pp. 3597–3602, Society of Exploration Geophysicists, Dallas, TX, USA, September 2016.

[46] J. Pafeng, S. Mallick, and H. Sharma, "Prestack waveform inversion of three-dimensional seismic data - an example from the Rock Springs Uplift, Wyoming, USA," *Geophysics*, vol. 82, no. 1, pp. B1–B12, 2017.

[47] H. D. Huang, Y. C. Wang, F. Guo, and C. H. Liu, "High precision pre-stack inversion algorithm based on Zoeppritz equations," *OG*, vol. 48, no. 5, pp. 740–749, 2013.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5309556, 7 pages
https://doi.org/10.1155/2022/5309556

Research Article

Integration Mechanism of Heterogeneous Foreign Language Education Resources Based on Time Series Analysis in IIoT

Hongyue Jin

Jilin Normal University, Siping 136000, China

Correspondence should be addressed to Hongyue Jin; 1808023@jlnu.edu.cn

Received 6 December 2021; Accepted 31 December 2021; Published 21 January 2022

Academic Editor: Jianhui Lv

Industrial Internet of Things (IIoT) has attracted much attention from global researchers and has been applied into many fields, such as medical treatment, transportation, and education. This paper pays attention to an IIoT-oriented education problem and gives the corresponding solution. Heterogeneous educational resources have multisource target data, so it is necessary to integrate the repetitive data and data with the same attributes. However, due to the poor tracking effect of the model constructed by traditional methods, the mining technology loses a part of the data characteristics and affects the multisource foreign language education data integration. So this article studies the integration mechanism of foreign language heterogeneous educational resources based on time series analysis. The mechanism adopts a data cleaning and fusion method based on the time series similarity measurement. This method uses approximate symbol aggregation, European algorithm, and similar sequences with adjusted similarity weights to complete the data cleaning of foreign language heterogeneous educational resources. After that, it uses multiple heterogeneous data fusion algorithms to complete data integration. Experiments with foreign language education resources at all levels in a certain city show that the mechanism can detect abnormal data of foreign language education resources, fill in vacant data, reduce data redundancy, and integrate heterogeneous data. After the data are cleaned by multisource heterogeneous data fusion algorithm, the credibility of the measurement data is reflected, and the mean absolute percentage error is only 6.25%. The data quality is improved as a whole, and it provides reliable basic data for the application of foreign language education resources.

1. Introduction

In recent years, Industrial Internet of Things (IIoT) has been widely applied with the explosive increase of mobile devices and cloud platforms. In fact, during the process of building smart city, IIoT plays an important role, for example, improving the level of medical treatment, the efficiency of transportation, and the quality of education. Especially during the period of COVID-19, online education has become more are more significant. Therefore, the IIoT-oriented education resource allocation issue has attracted much attention from global researchers.

With the advancement of the construction of foreign language education information, most schools, education departments, and related institutions have established their own foreign language education and teaching resource systems, but such resource systems have been established due to different establishment periods and lack of unified technical specifications and education resources. The unified understanding of utilization has led to repeated investment in hardware facilities, repeated development of software platforms, and repeated construction of online courses, resulting in heterogeneity such as uneven distribution of digital foreign language education resources, low standardization, and difficulty in integration and sharing. The problem seriously hinders the effective use and reasonable distribution of digital foreign language education resources. In view of the current heterogeneity of foreign language digital education resources and the need for the rational allocation and effective use of educational resources in the industrialization of foreign language digital education, the research and realization of the integration of foreign

language heterogeneous educational resources have become the current digital educational industrialization project. Only by establishing a reasonable development strategy for digital foreign language education resources can we better promote the construction of digital foreign language education resources, give full play to its maximum benefits, and serve the construction of educational information [1].

Data cleaning is a method used to detect and eliminate errors and inconsistencies in data [2]. In recent years, researchers have proposed a variety of data cleaning technologies to improve data quality, such as missing data attribution, object repeated detection, anomaly detection, logical error detection, and data inconsistency detection [3]. However, current data cleaning methods have high computational complexity and inaccurate detection of missing data. In 2017, the University of California, Riverside, Keogh, group [4] by based on changes in the field of feature representation, symbol, and piecewise linear representation to complete the feature extraction, but as a result of data in the process of feature extracting large dimensions, high computational complexity does not have scale invariance defects for multivariate time series data cleaning effect.

Similarity measurement technology is the basis of sequence data analysis. Sequence analysis can reflect the characteristics and relationships between data and judge data outliers based on the relationships mined, attracting a large number of scholars to conduct in-depth research. At present, in addition to the longest common substate distance and edit distance, similarity measurement methods mainly include Euclidean distance [5], dynamic time warping distance, and singular value decomposition and point distribution-based methods. The commonly used methods for data correction include the interpolation model [6], random replacement model, mean replacement model, and regression model [7]. When it comes to the identification of abnormal data, there may be a problem that the quoted "correct" sequence does not exist.

Data integration is essentially the collaborative processing of data from multiple parties to achieve the purpose of reducing redundancy, comprehensive complementation, and capturing collaborative information. This technology has become a research hotspot in the fields of data processing, target recognition, situation assessment, and intelligent decision making. In [8], Yu et al. studied multisensor data integration technology based on statistics and artificial intelligence (AI) methods; in [9], Lai et al. studied the organization and management of multisource heterogeneous data in mobile geographic information systems and established a multisource heterogeneous data fusion model; in [10], Premkumar and Ganesh combined wireless sensor network and data fusion technology and proposed a Kalman filter batch estimation fusion algorithm; in [11], Lasheng and Yiquang studied a massive multisource heterogeneous data fusion method in the Internet of Things environment and successfully applied it in the process of target positioning and tracking; in [12], Zhang et al. studied the intelligent maintenance decision-making architecture of the high-speed rail signal system based on heterogeneous data fusion, which improved the accuracy and effectiveness of decision making; in [13], Wen et al. studied many aspects of the digital mine construction process. Source heterogeneous data fusion technology ensures the safety, stability, and efficiency of the basic information platform in the construction of digital mines.

In view of this, in order to improve the quality of foreign language education resource data and support the large-scale collection and storage of data, this article focuses on the data cleaning and fusion in the construction of foreign language heterogeneous resource integration system and conducts a preliminary analysis of the characteristics of foreign language education multisource data. And a practical value-based foreign language education resource data cleaning and fusion algorithm based on time-series similarity measurement is proposed, and experiments have shown that it can achieve a better cleaning and fusion effect.

The rest paper is structured as follows. In Section 2, the integration of foreign language education resources is studied. In Section 3, the time series analysis-based integration mechanism of heterogeneous foreign language education resources is studied. The experimental results are presented in Section 4, and Section 5 concludes this paper.

2. The Integration of Foreign Language Education Resources

At present, the data of foreign language education resources present the characteristics of "two" (diversified data sources and data types) and "two" (high heterogeneity dimension and high overall value of data), and its greatest value lies in the realization of cross system and cross platform data exchange and sharing. The integrated application of big data in total foreign language education aims to break the "data island," establish the data governance system of foreign language education resources, and form the total data assets of foreign language education in the smart city ecosystem. The outstanding problems existing in the data integration of foreign language education resources are as follows: no unified data standard, data source is not clear, out of sync data exchange, data storage and scattered in disorder; this series of factors has resulted in unfavorable situations of low data quality, chaotic data flow, insufficient data sharing, and unsmooth data lifecycle management, which greatly restricts the height that foreign language education big data-assisted smart application terminals can achieve [14, 15].

The integration and application of foreign language resource education data focus on three aspects of "management + governance + application," and the key problems are mainly reflected in the following three aspects:

(1) Business data cannot effectively follow a unified data standard. Data standards regulate the consistency and accuracy standards of data used and exchanged within and outside regions at all levels, restrict the normative documents of data standards, and carry out data standardization control and data standard management organization to provide data for

foreign language education resources at all levels. The platform provides a unified data definition standard and logical model. However, due to the different construction ages and different structural levels of various platforms, they were not defined in accordance with a unified data standard at the initial stage of construction, which brought a lot of inconvenience to the exchange and sharing of foreign language education resource data. In response to such problems, we should first start with the top-level design of informatization and intelligent information services, formulate unified data standards, establish a scientific and standardized data application assessment and evaluation mechanism, and carry out transformations in stages and steps; source data carries out all aspects of data cleaning.

(2) The data source of foreign language education resources is not unique, and the data flow is unreasonable. The producer of the data must determine the focal point. The focal point is the uniqueness of the data source. The content of the data cannot be maintained by multiple systems at the same time; otherwise, the uniqueness and accuracy of the data source cannot be guaranteed. The flow of data is aimed at achieving exchange and sharing, and the public data platform (public data pool) completes the cross-business data interaction [16]. This type of problem requires the establishment of relevant organizational structures through administrative management methods, determining the authority for data generation, clarifying the data responsible unit, constructing a data flow relationship table, and providing a complete data flow for the data source connected to the system and the data interface released by the system. The unified combing of foreign language education resource data application requirements is managed and completed.

(3) The quality of business data is not high, and there are certain phenomena of "lack of data" and "wrong data." Data quality describes the applicability of the data, that is, the suitability of the data to meet the needs of users. Data quality measures data through multiple dimensions such as completeness, consistency, accuracy, timeliness, and legitimacy. In the business platform, data quality provides clean and structured data for it. It is a necessary prerequisite for the data platform to develop data products, provide data services, and play the value of big data. It is also a key factor in the management of foreign language education data assets at all levels and regions. Currently, data quality is generally not high in all levels and regions. On the one hand, it is necessary to improve data quality through in-depth data governance (analysis, correlation, cleaning, and exchange of multisource heterogeneous data), and on the other hand, it is necessary to establish data quality and improve the process and assessment system.

3. Integration Mechanism of Heterogeneous Foreign Language Education Resources Based on Time Series Analysis

In view of the lack of data and wrong data in the foreign language education resources described above, this paper proposes a data cleaning method based on time-series similarity measurement to detect abnormal data and fill in missing data in foreign language materials.

The data cleaning and fusion process is mainly divided into four steps: first, the approximate symbol aggregation algorithm is used to discretize and symbolize the foreign language resource data; second, the Euclidean distance algorithm is used to calculate the similarity between the symbol sequences; then, it is fitted according to the similar sequence. The curve of foreign language data completes the identification and correction of abnormal data and the filling of missing data; finally, the cleaned data are fused.

3.1. Approximate Symbol Aggregation Algorithm. In recent years, the symbolic aggregation approximation (SAX) algorithm is a new method of discretizing time series data. The basic idea of this method is to convert numerical time series data into discrete symbol sequences [17]. Through the specified mapping rules, the SAX algorithm can weaken the influence of abnormal and missing data in the time series on the local fluctuations and can also generate a smaller-sized symbol and nondigital sequences, which can improve further aggregation efficiency and strengthen the comparison of similarity in the later stage.

SAX is a sequence of equal-length partition based on the piecewise aggregate approximation (PAA). If the partition length is long, there may be a large internal difference, and the mean value is equal. The key point improvement method can be adopted to achieve the purpose. But the algorithm complexity is improved. In this paper, SAX is used to reduce a time series of arbitrary length n to a string of length N ($N < 26$), usually with English sentences of no less than 26 letters.

SAX first converts the data to PAA representation, reduces the time sequence from n dimension to N dimension, then maps all PAA coefficients to m equal probability intervals, and the last SAX symbolizes the PAA representation into a discrete string. The following is a brief execution process of SAX on the original time series $X = \{x_1, x_2, \ldots, x_n\}$.

(1) Performance of normalization processing: normalization is to convert the average value of each time series to 0 and the standard deviation to 1, which is expressed as $C = \{c_1, c_2, \ldots, c_n\}$. The i element is

$$C_i = \frac{x_i - \mu(X)}{\delta} \ (i = 1, 2, \ldots, n), \tag{1}$$

where μ is the average value of the original time series; δ is the standard deviation.

(2) The dimensionality of normalized sequence C is reduced by PAA to reduce the original time series

vector of n dimension to N dimension. In the process of dimensionality reduction, the N dimensional time series $C = \{C_1, C_2, \ldots, C_N\}$. The i element in C is calculated as

$$\begin{cases} \overline{C}_i = \dfrac{1}{t} \displaystyle\sum_{t(i-1)+1}^{ti} c_j \\ j = t(i-1) \end{cases}, \tag{2}$$

where $\overline{C}_i$ is the mean value of the original time series vector divided into N segments; $t = n/N$ is called the compression rate; $1/t$ is the interval length of each segment.

After converting the time series set $\overline{C}$ into PAA, it is further converted into a discrete symbol form; that is, elements in the PAA representation form of the time series are mapped to equal probability symbols. Since the normalized time series has a highly characteristic Gaussian distribution, the "break point" β is determined by looking up the Gaussian distribution statistics table, thereby generating m equal sizes, that is, regions with the same probability distribution, where β is a series of the ordered list of values, and all β areas are $1/m$.

After querying and comparing the breakpoint β, the time series collection C is transformed into the string collection $\widehat{C}$, namely,

$$\widehat{C}_j = P_j, \text{if}, \quad \beta_{j-1} \le \overline{C}_j \le \beta_j, \tag{3}$$

where P is the alphabet; the j element of the N dimensional time series C is between β_{j-1} and β_j; the j element of the alphabet A can be expressed as the j element in the string $\widehat{C}$.

3.2. Similarity Measurement. Euclidean distance is one of the most widely used algorithms in similarity measurement. In the application process, the sequence to be compared is required to have the corresponding length and point, and the difference between the two sequences corresponds to each other [18]. The Euclidean distance can quickly calculate the similarity of SAX symbolic expressions with low computational complexity. The greater the distance between the SAX expressions of two foreign language education data sequences, the lower the similarity. Therefore, the similarity of the two foreign language education data time series curves is

$$S(Q, C) = \left[\sqrt{\sum_{i=1}^{n} \left(q_i - c_j\right)^2} \right]^{-1}, \tag{4}$$

where Q and C are the two time series, respectively; q_i is the i point of the Q sequence; c_j is the C point of the j sequence.

3.3. Similarity Curve Adjustment. After approximate aggregation of symbols and similarity measurement of the time series, ω similar time series A and similar time series SXA to be cleaned up are obtained, where X is the original series and ω is usually 30 time series in a month. The similar time series A is adjusted by the weighted adjustment method (fitted curve algorithm) to obtain the reference curve $\widehat{X}$ relative to the original time series X. If there is a missing value in the original time series, it is filled with the value of the corresponding point in the reference curve.

Judge whether a word in the data is abnormal by comparing within the reference curve, which is calculated by weighting similar letters in all foreign language education resources. This article uses an improved maximum threshold method to determine whether a foreign language word is abnormal data, and this method uses a more accurate weighted average of similar time series to calculate, and calculates x_k of the threshold δ_k, that is,

$$\delta_k = \max\left(A - \widehat{x}_k\right). \tag{5}$$

If x_k does not meet the following criteria, it is considered abnormal data.

$$\begin{cases} x_k > \widehat{x}_k - \delta_k, \\ x_k < \widehat{x}_k + \delta_k. \end{cases} \tag{6}$$

3.4. Integration of Foreign Language Heterogeneous Educational Resource Data. A foreign language multisource heterogeneous data integration focuses on computing structured and comparable foreign language heterogeneous data, with the goal of improving data quality and obtaining more significant data characteristics. Kalman filter is an efficient recursive filter, which uses a series of data obtained from the data measurement process to estimate the state vector of the dynamic system, which is more effective for heterogeneous structured data [19]. In this paper, the concept of information pair is introduced into the Kalman filter algorithm, that is, the distributed Kalman algorithm. The cleaned data are exchanged and merged with the adjacent data sequence. The information matrices X1 and X2 used are respectively

$$\begin{cases} \mathbf{M}_{k|k}^i = \left(\mathbf{P}_{k|k}^i\right)^{-1}, \\ \mathbf{A}_{k|k}^i = \mathbf{M}_{k|k}^i \widehat{x}_{k|k}^i, \end{cases} \tag{7}$$

where $\mathbf{P}_{k|k}$ is the posterior estimated covariance matrix at time k and x^i is the estimated state value at time k. The recursive form of the distributed Kalman filter is

$$\begin{cases} \mathbf{M}_{k|k}^i = \mathbf{M}_{k-1|k}^i + \left(\mathbf{Q}^i\right)^{\mathrm{T}} \left(\mathbf{R}^i\right)^{-1} \mathbf{Q}_k^i, \\ \mathbf{A}_{k|k}^i = \mathbf{A}_{k-1|k}^i + \left(\mathbf{Q}^i\right)^{\mathrm{T}} \left(\mathbf{R}^i\right)^{-1} \mathbf{Q}_k^i, \end{cases} \tag{8}$$

where Q and R are the covariance matrices of system noise and observation noise, respectively. In order to improve the accuracy of local fusion, the total dataset is N, and each series P_i can send its local posterior covariance k/k to the j adjacent series j and perform data fusion with local posterior covariance matrix P_j of j, and the fusion calculation is, respectively,

Table 1: Comparison of three data fusion methods.

Test dataset	Data residual rate (%)		
	Unfused data	Kalman algorithm	This paper
60	100	38	10
120	100	41	12
180	100	47	15
240	100	53	18
300	100	57	21
360	100	63	23
...	...	...	...
1440	100	95	36

$$\begin{cases} \mathbf{M}_{k|k}^{i,j} = \Pi_{i,j}\mathbf{M}_{k|k}^{i} + \sum_{i,j\in N} \mathbf{M}_{k|k}^{i}, \\ \mathbf{A}_{k|k}^{i,j} = \Pi_{i,j}\mathbf{A}_{k|k}^{i} + \sum_{i,j\in N} \mathbf{A}_{k|k}^{i}, \end{cases} \tag{9}$$

where $\Pi_{i,j}$ is the combined weight and positive value, satisfying any node, $\Pi_{i,j}$ can be obtained from the following formula:

$$\Pi_{i,j} + \sum_{i,j\in N} \Pi_{i,j} = 1. \tag{10}$$

The multisource heterogeneous data fusion algorithm flow is as follows:

(1) Initialize data, and time series $i, j \in N, N$ are data space

(2) Observe the status of the integrated data and update the information pair of the time series X1 through equation (8)

(3) The information pair of the time series i is transmitted to the adjacent time series j. If the time series j data are completed and safe, the time series j receives the information pair sent from the time series i

(4) Through formulas (9) and (10), the local data pair and the information pair of adjacent data are used for data fusion to obtain the fused information pair

(5) Update the local filter value

(6) Return to Step (2)

4. Experimental Results and Analysis

Section 3 describes the SAX method, similarity measure method, and data cleaning and data fusion method used in this experiment, respectively. In this paper, foreign language resources at all levels in a city were selected as the experimental dataset, and data cleaning and integration experiments were carried out on the experimental dataset. The experimental data consisted of 16 series of three data types mentioned above, with a collection interval of 1 min and a sampling frequency of 10 kHz for multisource data.

In order to prove the effectiveness of the multisource heterogeneous data fusion algorithm proposed in this paper, in addition to the traditional Kalman filtering algorithm [20], this paper also compares the case of not adding a filtering algorithm and draws corresponding conclusions. Table 1 shows the ratio of effective data obtained after data fusion using three data fusion methods. The three methods are the distributed Kalman filter algorithm without data fusion, the Kalman filter algorithm, and the edge calculation. The Kalman algorithm of each fused iteration of the dataset and the neighbor nodes are smaller, and the effect is better. With the rapid increase in the size of the dataset, the integration of data efficiency declines rapidly. This is because there are many series, large resource consumption, larger dataset, low data fusion ratio, and high data redundancy. The multisource heterogeneous data fusion algorithm can effectively reduce redundant data information so as to get closer to the actual effective data value. Compared with the unfused data method, it reduces a lot of resource consumption and can align different data more effectively and resolve feature conflicts between data.

Table 2: Results of data cleaning of heterogeneous foreign language education resources.

Algorithm	Wrong comparison	
	MAPE (%)	RMSE
Gaussian filtering	18.23	1998.04
Wavelet threshold	12.34	1634.98
This paper	6.25	1139.42

Table 3: Comparison of data integration results of heterogeneous foreign language education resources.

Algorithm	Wrong comparison		
	RMSE	MARE	t test
Kalman algorithm	2.76	0.028	2.97
This paper	2.12	0.013	2.16

Select the data sequence with no more than 1% gap as the experimental set, and set the data to be empty according to a random ratio, then use the method in this article to clean the data, compare the gap prediction value with the original value, and calculate the average absolute percentage error and mean square root error and standard root mean square error. Mean absolute percentage error (MAPE) is the average absolute value of the relative error ratio, which can reflect the credibility of the measurement data, namely,

$$E_{\text{MAPE}} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%. \tag{11}$$

In addition to the traditional Gaussian filter algorithm [21], this paper also compares the data cleaning of wavelet threshold value, as shown in Table 2. The average absolute percentage error deviation of the algorithm proposed in this paper is 6.25%, which shows that the prediction results in this paper are relatively accurate, and the root mean square error is within the numerical range of the standard point, which meets the requirements.

At the same time, the error analysis of the integrated foreign language education data indicators is carried out,

and the accuracy of the algorithm in this paper is evaluated according to the RMSE, MAPE, and t test [22]. Table 3 reflects the RMSE, MAPE, and t-test results of the two algorithms from a quantitative perspective. From the analysis of the data results, the error of the algorithm in this paper tends to be flat or reduced, which proves that the algorithm in this paper has good performance and can meet the accuracy requirements.

5. Conclusions

In the era of AI education, building a reusable and sharable educational data model is one of the urgent problems to be solved in the development of foreign language education. The model needs to standardize the multisource and heterogeneous foreign language education data in the AI education environment, so as to achieve a high degree of sharing of heterogeneous foreign language data. In this paper, an IIoT-oriented environment is considered, and a method based on time-series similarity measurement is proposed to clean and integrate heterogeneous data of foreign language education resources. SAX by piecewise reduce the dimension of time series data, so as to achieve the aim of reducing noise and is calculated using Euclidean distance similarity measurement method, and in smaller time complexity to find a similar set of time series, using the maximum threshold method to detect outliers, the reference curve is obtained by the weighted adjustment method, according to the reference curve filling vacant values. It is more accurate than the traditional method using the maximum threshold of the average value of similar days. The data after cleaning adopt the multisource heterogeneous data fusion algorithm. Through data fusion between adjacent series, redundant data can be better fused to ensure data quality and provide more practical value for the high sharing of heterogeneous foreign language education resources.

This research still has a lot of work that needs to be further explored, such as the way to increase the integration of different types of foreign language education resources and to integrate different types of foreign language education resources such as online teaching, online vocational training, online examination, and synchronized online education guidance. The same platform is the future research and development direction of the foreign language education resource integration platform. I believe that foreign language education can be improved through different angles and ways to improve its fairness, universality, and sharing. This is a beautiful goal that we all have been working on together.

Data Availability

All data used to support the findings of the study are included within the article.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This paper was supported by the 13th Five-Year Social Science Research Project of the Education Department of Jilin Province (no. JJKH20200438SK) and the Social Science Fund Project of Jilin Province (no. 2020C113).

References

[1] P. Xia, "Application scenario of artificial intelligence technology in higher education," in *Proceedings of the 2019 International Conference on Applications and Techniques in Cyber Intelligence*, Huainan, China, June 2019.

[2] T. Odia, S. Misra, and A. Adewumi, "Evaluation of hadoop/mapreduce framework migration tools," in *Proceedings of the Asia-Pacific World congress on Computer Science and Engineering*, Nadi, Fiji, November 2015.

[3] Q. Ma, S. Tian, J. Wei, J. Wang, and W. W. Y. Ng, "Attention-based spatio-temporal dependence learning network," *Information Sciences*, vol. 503, pp. 92–108, 2019.

[4] F. Karim, S. Majumdar, H. Darabi, and S. Harford, "Multivariate LSTM-FCNs for time series classification," *Neural Networks*, vol. 116, pp. 237–245, 2019.

[5] F. Karim, S. Majumdar, H. Darabi, and S. Chen, "LSTM fully convolutional networks for time series classification," *IEEE Access*, vol. 6, no. 99, pp. 1662–1669, 2017.

[6] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, June 2018.

[7] Y. Gu, W. Lu, L. Qin, M. Li, and Z. Shao, "Short-term prediction of lane-level traffic speeds: a fusion deep learning model," *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 1–16, 2019.

[8] H. Yu, N. Rao, and I. S. Dhillon, "Temporal regularized matrix factorization for high-dimensional time series prediction," in *Proceedings of the Neural Information Processing Systems*, Barcelona, Spain, December 2016.

[9] G. Lai, W. Chang, Y. Yang, and H. Liu, "Modeling long and short term temporal patterns with deep neural networks," in *Proceedings of the Fourty First International ACM sigir conference on research and development in information retrieval*, Ann Arbor MI USA, July 2018.

[10] M. S. Premkumar and S. H. Ganesh, "A median based external initial centroid selection method for k-means clustering," in *Proceedings of the World Congress on Computing and Communication Technologies*, pp. 143–146, Tiruchirappalli, India, February 2017.

[11] C. Lasheng and L. Yuqiang, "Improved Initial clustering center selection algorithm for K-means," in *Proceedings of the 2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications(SPA)*, pp. 275–279, Poznan, Poland, September 2017.

[12] R. Zhang, J. Huang, and T. Kumar, "Preventive leak detection for high pressure gas transmission networks," in *Proceedings of the AAAI Workshop*, pp. 64–70, San Francisco, C A, U S A, February 2017.

[13] Q. Wen, J. Gao, X. Song, L. Sun, H. Xu, and S. Zhu, "RobustSTL: a robust seasonal-trend decomposition algorithm for long time series," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Honolulu, H I, USA, February 2019.

[14] D. Rajan and J. J. Thiagarajan, "A generative modeling approach to limited channel ECG classification," in *Proceedings*

of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2571–2574, Honolulu, H I, USA, July 2018.

[15] Z. Che, Y. Cheng, S. Zhai, Z. Sun, and Y. Liu, "Boosting deep learning risk prediction with generative adversarial networks for electronic health records," in *Proceedings of the IEEE International Conference on Data Mining*, New Orleans, LA, USA, November 2017.

[16] H. Wang, C. Luo, and X. Wang, "Synchronization and identification of nonlinear systems by using a novel self-evolving interval type-2 fuzzy LSTM-neural network," *Engineering Applications of Artificial Intelligence*, vol. 81, no. 7, pp. 79–93, 2019.

[17] S. Lin and G. C. Runger, "GCRNN: group-constrained convolutional recurrent neural network," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 10, pp. 4709–4718, 2018.

[18] Z. Gong, H. Chen, B. Yuan, and X. Yao, "Multiobjective learning in the model space for time series classification," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 918–932, 2019.

[19] J. Wang, T. Sun, B. Liu, Y. Cao, and H. Zhu, "CLVSA: a convolutional LSTM based variational sequence-to-sequence model with attention for predicting trends of financial markets," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Macao China, August 2019.

[20] F. Amato, M. Laib, F. Guignard, and M. F. Kanevski, "Analysis of air pollution time series using complexity-invariant distance and information measures," *Physica A: Statistical Mechanics and Its Applications*, vol. 547, Article ID 124391, 2020.

[21] X. Zhao, M. Ji, N. Zhang, and P. Shang, "permutation transition entropy: measuring the dynamical complexity of financial time series," *Chaos Solitons and Fractals*, vol. 139, Article ID 109962, 2020.

[22] X. Guo, Y. Wang, N. Zhou, and X. Zhu, "Optimal weighted two-sample t-test with partially paired data in a unified framework," *Journal of Applied Statistics*, vol. 48, no. 6, pp. 961–976, 2021.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 9668631, 10 pages
https://doi.org/10.1155/2022/9668631

Research Article

A Novel Online Education Reform Model Based on Risky Decision-Making under the Situation of Internet Plus

Yao Lv

Shenyang University, Shenyang 110044, China

Correspondence should be addressed to Yao Lv; lvyao@syu.edu.cn

Received 16 November 2021; Revised 15 December 2021; Accepted 17 December 2021; Published 11 January 2022

Academic Editor: Jianhui Lv

In the new situation of Internet plus, information technology has been widely applied in education, and hence online education has attracted wide attention from all walks of life. Today's society is a risk society, and risk is everywhere. Online education reform is also risky, which is determined by many reasons. Some risks will cause certain losses to the online education reform, so based on risky decision-making, it is necessary to carry out online education reform under the new situation of Internet plus. At first, the risky decision-making in online education reform is analyzed, which is the risk of online education reform in risk society and the allocation logic of online education reform. Then, taking interval type-2 fuzzy logic (IT2FL) as the information environment, this study proposes the optimal risky decision-making method based on IT2FL utility functions, IT2FL entropy, and risk preference factor of online education reform to solve the multipath risky decision-making problem of online education reform. Finally, the experimental results show that, in the risky decision-making model, the decision-maker's risk preference has an impact on the path weight and the ranking of the scheme, and the idea has a certain reference role for risky decision-making. Compared with the three benchmarks, the proposed method has the fewest ranking time with the same ranking results.

1. Introduction

Society today is a risk society, from Ulrich Beck, a recognizable German author in environmental sociology [1]. At the same time, he argues that a risk society is a disaster society [2]. There is a danger that the anomaly will become the norm. In this way, the risk is universal and objective in today's society, which is a fact we should face squarely. On the other hand, as risks may cause disasters to human society, they should be prevented and avoided. As an artificial social practice, online education reform also has inevitable risks. Facing the risks in online education reform, what kind of attitude should we hold and what kind of actions should we take? A clear understanding of the risks in education reform is helpful to reduce the errors caused by the risks and enhance the effectiveness of online education reform, so the research on the risks in online education reform has a certain theoretical and practical significance [3–5]. Education reform in a risky social environment is risky decision-making. The logic of risk distribution is equalization, but the vulnerable groups in the current society will bear more risks. Therefore, we must consider the online education reform within the perspective of risk, call for decision-making ethics, strengthen institutional prevention, and establish the decision-making system of online education reform with multiple participation.

The new situation of Internet plus drives the reform of online education. Internet plus emphasizes the integration of the Internet and traditional industries [6, 7]. With the continuous promotion and progress of Internet plus, Internet plus education has become one of the new development directions in education [8, 9]. The common feature of Internet plus education is to make up for the deficiency of teaching by relying on online courses and making use of the advantages of convenient access to resources and flexible teaching methods. However, the disadvantage is that schools or colleges generally do not have a clear position on the role of online education. In addition, the variety of online education courses increases the difficulty for teachers to choose. In particular, each course has obvious differences in content and the proportion of practice, leading to the independent setting of education based on Internet plus

according to the characteristics of the course, so as to better meet the needs of teachers and students.

The online education reform has brought an unprecedented impact on education, including the following four aspects: (i) Educational resources are transformed from segmentation to sharing. The idea of sharing educational resources arises from the rise of massive open online courses (MOOC) [10]. Apart from MOOC, as another important phenomenon, Open Educational Resources (OER) is another important trend of education development in recent years [11, 12]. No matter MOOC or OER, the revolutionary change it brings is that education breaks through the traditional time and spatial limitation and solves the unbalanced distribution of educational resources. (ii) The shift in the form of learning from linear to nonlinear: linear learning is planned and purposeful learning based on the logic and sequence of subject knowledge within a certain time range. Linear learning is the most important form of student learning. However, the disadvantages of this model are also obvious. Its prominent disadvantage is that the learning process is step-by-step, which cannot take into account the personalized and diversified learning needs of students. Different from linear learning, nonlinear learning is not in accordance with the unified learning plan. But according to the students' personalized learning differences, students choose the learning content, learning process, and learning way [13–15]. (iii) The course reform changes from structured course to unstructured course. (iv) Educational technology changes from auxiliary means to deep integration with education. Given the above, it is necessary to reform online education based on risky decision-making under the new situation of Internet plus.

Accordingly, the main contributions of this paper are summarized as follows: (i) the risky decision-making in online education reform is studied, which is the risk of online education reform in risk society and the allocation logic of online education reform. (ii) The optimal risky decision-making method based on interval type-2 fuzzy logic (IT2FL) utility functions, IT2FL entropy, and risk preference factor of online education reform is proposed to solve the multipath risky decision-making problem of online education reform.

The rest of this paper is organized as follows. Section 2 reviews the related work. In Section 3, risky decision-making in online education reform is studied. In Section 4, the risky decision-making of IT2FL and entropy in novel online education reform mode is proposed, and the risk preference factor is introduced. The experimental results are shown in Section 5. Section 6 concludes this paper.

2. Related Work

The research on risky decision-making has put forward different views in people's rational judgment and choice. There are some risky decision-making strategies have been proposed. In [16], the authors explored the effects of optimism on self-framing and risky decision-making. In [17], a normal distribution-based interval number risky decision-making method was proposed to rank schemes. In [18], the author determined whether the risky decision-making of medicine users was increased. In [19], the authors examined the effects of physiological and combined mental stress on decision-making under risk and whether risk-taking differed between women and men.

With the continuous development of information technology, vigorously promoting "Internet plus education" and sharing high-quality education and teaching resources can be regarded as an effective means to solve the current problems of basic education. In [20], the mode of ideological and political education under the Internet plus environment was proposed to ensure the effectiveness of education. In [21], the Internet plus education-based innovative personnel training mode was studied. In [22], a novel smart learning-based education paradigm was proposed to enhance the teaching effects. In [23], the authors analyzed the transformation of film education mode in Internet plus. In [24], innovative methods were proposed for improving the talent cultivation of software engineering under the perspective of Internet plus.

IT2FL has become a hot issue in current academic research. In [25], the IT2FL controller design was proposed to control chaos and associated instability in a nonlinear dynamical power system. In [26], a novel calculation-effective IT2FL controller was designed. In [27], an IT2FL mutual subset fuzzy neural inference system was proposed. In [28], the IT2FL method was used for the route planning problem. In [29], an integrated ranking algorithm GRAP is proposed to solve decision-making problems by combining grey relational analysis, rank-sum, and preference ranking organization method enrichment evaluation methods. In [30], a rough set-based ranking algorithm is proposed to deal with the decision-making problem. In [31], a possibility degree-based decision-making method is proposed.

3. Risky Decision-Making in Online Education Reform

3.1. Risk of Online Education Reform in Risk Society. The risks of modern society are the results of the increasing extremes of modernization and economization. The online education reform may be risky in any social context. The reform is an attempt to replace the old paradigm with a new one, but the new paradigm is not superior to the old model in all respects, and this "replacing the old with the new" often comes at the expense of what is good in the old model. Due to the existence of social differences and different interest groups, a reform may harm the interests of another group while benefiting one group, leading to contradictions and conflicts. The reform is also risky due to the deficiency and error of people's subjective understanding. In addition, online education reform is an exploratory activity, and the implementation of reform is a dynamic process, which means that uncertainty and risk are inherent dimensions of education reform. Particularly in the risk society, the risks of online education reform are even greater.

3.1.1. Excessive Trust and Reliance on Education Experts. Beck, Giddens, and other scholars believe that risk in modern society is different from that in traditional society. The source of risk is no longer ignorance but knowledge,

which is mainly brought by decision-making. Because modern society is a society ruled by science and technology, experts of all walks of life have gained people's trust, and the decisions made by these experts with professional knowledge make people feel safe. However, social risks come out when science and technology develop uncontrollably and are used irrationally.

The current online education reform in China is also facing a similar situation. Important education reforms are planned and implemented by some experts whose views influence education policy. On the other hand, due to the special knowledge on an educational part by education experts, officials in education departments, teachers, and people all believe in the views of education experts, and the views of experts have been become an important standard to measure the fact.

Although education experts are people who have more research on education, their understanding is not always correct and comprehensive. At first, the consequences of reform cannot be fully predicted. The uncertainty, concealment, lag, and sudden characteristics of risks make it impossible for education experts to predict when formulating reform plans. Then, due to the limitation of knowledge, experts may have wrong understanding and judgment. Many experts who participate in online education reform are scholars who study advanced knowledge in colleges or scientific research institutions. Their alienation from the front-line educational practice makes them underestimate the complexity of practice. Finally, experts' understanding of "what is ideal education" may be different from teachers' understanding, so online education reform may be planned against the will of the general public. In addition, experts may carry out sensationalist reforms for their own purposes (e.g., seeking prestige, political achievements, and economic benefits) regardless of facts. In this case, people's trust and reliance on experts and their expertise to some extent breed the risk of online education reform.

3.1.2. Risk Transmission and Intensified Diffusion. The development of communication tools, especially the popularization of the Internet, and the accelerated flow of information around the world have led to the alienation of the social structure into a network virtual society in a certain sense, which makes the spread and diffusion of risk present an overwhelming trend. Giddens pointed out that one of the characteristics of modernity is that the far-reaching events and actions continue to affect our lives, and this impact is still increasing. In Giddens's view, the high extension of modern society in the space-time structure makes people can only rely on the symbolic system and expert system to obtain information and reach a consensus. However, if there are problems in the symbolic system and expert system for providing and interpreting information, the society may fall into high tension and risk brought by emergencies.

At present, China has also entered a globalized and information-based society. We can see that a local online education reform step will spread to the whole country and even the world in an instant. In the process of information transmission, people's different understanding, processing, and even misinformation, coupled with the rendering of the media, may expand tension and risks. The spread of a large amount of information is actually creating and expanding risks to some extent because people who lack professional knowledge do not know the truth, and uncontrolled and excessive information will only bring them more confusion and panic.

3.2. The Allocation Logic of Online Education Reform. Beck emphasizes that the transformation of risk society means the transformation of new contradiction and distribution mode. In a risk society, the allocation logic of risk is different from that of wealth. The original hierarchical allocation logic of wealth will be disrupted, and finally, the situation of risk equalization will emerge; that is, "poverty is hierarchical, but smog is democratic." On the other hand, Beck states that the risk is still allocated by the hierarchy in fact. We are in an era of overlapping industrial society and risk society, and the two allocation logics play a powerful role in it. Wealth accumulates mainly at the top, while risk accumulates at the bottom. Poverty absorbs a lot of risks, and wealth obtains security and avoids risk. That is to say, in current society, the risk does not eliminate the hierarchy but depends on it, and social stratification plays a filtering or enlarging effect of risk.

In China, the ability and potential of people of different professions and education classes to deal with the risks of online education reform are obviously quite different. In view of the unemployment risk caused by the expansion of higher education and the financial crisis, families with strong economic capital, cultural capital, and social capital will obviously have a greater ability to avoid the unemployment risk for their children. The online education reform is planned by the powerful groups in society, but the risks are more shouldered by the vulnerable groups in society, which is obviously unreasonable and contrary to our goal of building a harmonious society. Moreover, the harm of risk will eventually spread to the whole society. Therefore, actions must be taken to prevent and reduce the risks brought by online education reform.

4. The Risky Decision-Making of Interval Type-2 Fuzzy Logic and Entropy in Novel Online Education Reform Mode

To sum up, online education reform presents fuzziness and uncertainty, so risky decision-making in a fuzzy environment is attracting more and more attention from experts and scholars. This section takes interval type-2 fuzzy logic (IT2FL) as the information environment and proposes two IT2FL utility functions based on the cut set. These two utility functions effectively extract all information of IT2FL, which is conducive to reducing the decision-making error of online education reform. The risky decision-making model of IT2FL and entropy in novel online education reform mode based on utility function, entropy, and risk preference factor is constructed to observe the influence of decision-makers'

risk preference on attribute weight and decision scheme ordering by solving the model.

4.1. Two IT2FL Utility Functions. Based on the idea of the cut set, this section proposes two utility functions of IT2FL to measure the advantages and disadvantages of IT2FL. The larger the utility function $U(P)$ is, the better P is.

Definition 1. (T2FS). P is a type-2 fuzzy set (T2FS) in domain X; if for any $x \in X$ and $u \in J_X$, there is $\mu_P(x,u) \in X(\forall x \in X, u \in J_X \in [0,1])$.

Assume that $P = (P^U, P^L)$ is an IT2FL, where P^U and P^L are the upper bound and lower bound of IT2FL, respectively. Then, P can be defined as follows:

$$P = \begin{pmatrix} \left(p_1^U\right), \left(p_2^U\right), \left(p_3^U\right), \left(p_4^U\right); H\left(P^U\right) \\ \left(p_1^L\right), \left(p_2^L\right), \left(p_3^L\right), \left(p_4^L\right); H\left(P^L\right) \end{pmatrix}, \tag{1}$$

where $0 \le p_1^* \le p_2^* \le p_3^* \le p_4^* \le 1$ and $0 \le H(P^L) \le H(P^U) \le 1$. $*$ represents L or U, and P^U and P^L are the Interval Type-1 Fuzzy Set (IT1FS), respectively. The membership of p_2^* and p_3^* is equal to $H(P^*)$. Particularly, when $p_2^* = p_3^*$, the IT2FL degenerates into an interval type-2 triangular fuzzy set and meets the relation $p_1^U \le p_1^L \le p_4^L \le p_4^U$.

Let $\overline{\mu_P}(x)$ and $\underline{\mu_P}(x)$ denote the upper membership function and α cut set-based IT2FL P. Two IT2FL utility functions $U_1(P)$ and $U_2(P)$ based on cut set can be defined as follows:

$$U_1(P) = \begin{pmatrix} \dfrac{1}{p_4^U - p_1^U} \displaystyle\int_0^{h+} \int_{p_1^U}^{p_4^U} \left(\dfrac{P_{(\alpha)l}^U + P_{(\alpha)r}^U}{2} + \log \dfrac{x}{2}\left(P_{(\alpha)l}^U + P_{(\alpha)r}^U\right) \right) \overline{\mu_P}(x)\mathrm{d}x\mathrm{d}\alpha + \\ \dfrac{1}{p_4^L - p_1^L} \displaystyle\int_0^{h+} \int_{p_1^L}^{p_4^L} \left(\dfrac{P_{(\beta)l}^L + P_{(\beta)r}^L}{2} + \log \dfrac{y}{2}\left(P_{(\beta)l}^L + P_{(\beta)r}^L\right) \right) \underline{\mu_P}(y)\mathrm{d}y\mathrm{d}\beta \end{pmatrix}, \tag{2}$$

$$U_2(P) = \begin{pmatrix} \dfrac{1}{p_4^U - p_1^U} \displaystyle\int_0^{h+} \int_{p_1^U}^{p_4^U} \left(\dfrac{P_{(\alpha)l}^U + P_{(\alpha)r}^U}{2} + \dfrac{x}{2}\left(1 - \log\left(P_{(\alpha)l}^U + P_{(\alpha)r}^U\right)\right) \right) \overline{\mu_P}(x)\mathrm{d}x\mathrm{d}\alpha + \\ \dfrac{1}{p_4^L - p_1^L} \displaystyle\int_0^{h+} \int_{p_1^L}^{p_4^L} \left(\dfrac{P_{(\beta)l}^L + P_{(\beta)r}^L}{2} + \dfrac{y}{2}\left(1 - \log\left(P_{(\beta)l}^L + P_{(\beta)r}^L\right)\right) \right) \underline{\mu_P}(y)\mathrm{d}y\mathrm{d}\beta \end{pmatrix}, \tag{3}$$

where $h^+ = H(P^U)$ and $h^- = H(P^L)$ represent the membership values of the middle two parameters in the upper membership and lower membership, respectively. l and r represent the left cut point and right cut point under the cut set, respectively.

Both utility functions meet the following two properties:

(1) For any fuzzy set P IT2FL, $U_i(P) \in [0,4]$ $(i = 1,2)$

(2) $U_i(P^+) = 4$ and $U_i(P^-) = 0$, where $P^+ = ((1,1,1,1;1),(1,1,1,1;1))$ and $P^- = ((0,0,0,0;0),(0,0,0,0;0))$

For any fuzzy set P and Q of IT2FL, the utility functions of P and Q have three partial order relations, which are listed as follows:

(1) If $U_i(P) < U_i(Q)$, then P is inferior to Q, which can be expressed as $P < Q$

(2) If $U_i(P) > U_i(Q)$, then P is better than Q, which can be expressed as $P > Q$

(3) If $U_i(P) = U_i(Q)$, then P is equal to Q, which can be expressed as $P \cong Q$

4.2. IT2FL Entropy. Entropy is a measure of the uncertainty of things. In this section, the IT2FL entropy is proposed to measure the uncertainty of IT2FL. Furthermore, the three new uncertainty measures of IT2FL are proposed to describe the uncertainty of IT2FL, which are fuzziness measure (ζ), hesitation measure (θ), and interval measure (δ).

$$\zeta_P = \sqrt{\frac{\left(\zeta_P^U\right)^2 + \left(\zeta_P^L\right)^2}{2}}, \tag{4}$$

$$\theta_P = \sqrt{\frac{\left(\theta_P^U\right)^2 + \left(\theta_P^L\right)^2}{2}}. \tag{5}$$

The interval measure (δ) of IT2FL can be defined as follows:

$$\delta_P = \delta_P^U - \delta_P^L = \int_{p_1^U}^{p_4^U} \overline{\mu_P}(x)\mathrm{d}x - \int_{p_1^U}^{p_4^U} \underline{\mu_P}(x)\mathrm{d}x, \tag{6}$$

where $\delta_P^U = \int_{p_1^U}^{p_4^U} \overline{\mu_P}(x)\mathrm{d}x$ and $\delta_P^L = \int_{p_1^U}^{p_4^U} \underline{\mu_P}(x)\mathrm{d}x$ represent the areas of upper membership and lower membership, respectively. The larger the area difference between the two is, the larger the interval measure will be. It is easy to prove that $\zeta_P = \zeta_P^C \in [0,(1/2)]$, $\theta_P = \theta_P^C \in [0,(1/2)]$, and $\delta_P = \delta_P^C \in [0,1]$.

Supposing that $P = (P^U, P^L)$ is a fuzzy set of IT2FL, the entropy $E(P)$ of the fuzzy set P of IT2FL based on the above three uncertain measures is defined as follows:

$$E(P) = -\frac{2}{3\ \ln\ 2}\begin{pmatrix} \zeta_P \ln\ \zeta_P + (1-\zeta_P)\ln(1-\zeta_P)+ \\ \theta_P \ln\ \theta_P + (1-\theta_P)\ln(1-\theta_P)+ \\ \frac{\delta_P}{2}\ln\frac{\delta_P}{2} + \left(1-\frac{\delta_P}{2}\right)\ln\left(1-\frac{\delta_P}{2}\right) \end{pmatrix}. \quad (7)$$

For any fuzzy set P of IT2FL, $E(P)$ meets the following four properties:

(1) If and only if P is a definable set, then $E(P) = 0$; that is, $\zeta_P = \theta_P = \delta_P = 0$

(2) $E(P) = f(\zeta_P, \theta_P, \delta_P)$ is a continuous real-valued function, which increases at larger ζ_P, θ_P, and δ_P

(3) For any fuzzy set P of IT2FL, $0 \le E(P) \le 1$

(4) $E(P) = E(P^C)$

The uncertainty measure of IT2FL is analyzed from three aspects, and the entropy of IT2FL is proposed based on the uncertainty measure, which makes the entropy equation more scientific and reasonable.

4.3. Risk Preference Factor of Online Education Reform. This section introduces the risk preference factor to reflect the different risk attitudes of decision-makers during the risky decision-making of online education reform. Risk preference is the degree of the decision-maker's preference for risk, and its uncertainty is difficult to measure. Therefore, for this uncertainty, the decision-maker's risk attitude and tendency are the concrete embodiment of risk preference. According to the different risk preference of online education reform, it can be divided into online education reform risk aversion type, online education reform relative risk aversion type, online education reform risk-neutral type, online education reform relative risk preference type, and online education reform risk preference type. Therefore, according to different risk attitudes of decision-makers, the risk preference factors are set as follows:

$$\Phi(x) = \begin{cases} 1, & \text{online education reform risk preference type,} \\ \Phi_1, & \text{online education reform relative risk preference type } \Phi_1 \in (0,1), \\ 0, & \text{online education reform risk neutral type,} \\ \Phi_2, & \text{online education reform relative risk aversion type } \Phi_2 \in (-1,0), \\ -1, & \text{online education reform risk aversion type.} \end{cases} \quad (8)$$

4.4. Online Education Reform Optimal Decision-Making Method. Taking IT2FL as the information environment, this section proposes the optimal risky decision-making method based on IT2FL utility functions, IT2FL entropy, and risk preference factor of online education reform to solve the multipath risky decision-making problem of online education reform. Suppose that there are m paths Path_j $(1 \le j \le m)$ and n online education scheme sets x_i $(1 \le i \le n)$. The weight of the path $\omega = (\omega_1, \omega_2, \ldots, \omega_m)$, $\omega_j \ge 0$, and $\sum_{j=1}^{m} \omega_j = 1$. Let $P = (p_{ij})_{n\times m}$ be the risky decision-making matrix of online education reform. Each of p_{ij} is the fuzzy set of IT2FL, which represents the decision value of decision-maker for online education reform scheme x_i under path Path_j. Path types are generally divided into benefit-type and cost-type. When the path types are different, $P = (p_{ij})_{n\times m}$ needs to be normalized to obtain the normalized risky decision-making matrix $R = (r_{ij})_{n\times m}$, where r_{ij} is the normalized form of p_{ij}. To be specific, the online education reform optimal decision-making process can be summarized as follows:

(i) *Step 1.* Normalize the original decision-making matrix.

$$r_{ij} = \begin{cases} p_{ij}, & j \text{ is the benefit} - \text{type path,} \\ p_{ij}^C, & j \text{ is the cost} - \text{type path.} \end{cases} \quad (9)$$

(ii) *Step 2.* Calculate the utility functions $U_1 = (r_{ij})_{n\times m}$, $U_2 = (r_{ij})_{n\times m}$ and IT2FL entropy $E = (r_{ij})_{n\times m}$ of the normalized risky decision-making matrix of online education reform.

(iii) *Step 3.* An optimal linear programming model based on IT2FL utility functions, IT2FL entropy, and risk preference factor of online education reform is constructed to solve the optimal path weight. There are two cases of path weight, which are fully unknown and partially known.

(iv) *Step 4.* The ordered weighted averaging (OWA) operator [32] is used to aggregate the path weight and utility function of each scheme to obtain the comprehensive utility value $U(x_i)$, and the maximum value is the optimal scheme of online education reform.

5. Experiment and Result Analysis

5.1. Setup. The experiment is running on a computer with Intel i9-10850K, CPU 3.6 GHz, and 32 GB RAM 3333 MHz. The decision-maker selects one from the following five ways to reform online education, including live steaming interaction (x_1), online-merge-offline (x_2), modular reconstruction of course resources (x_3), MOOC (x_4), and home-school linkage (x_5). The teaching quality (Path_1), online platform supervision (Path_2), strategic positioning (Path_3), and development path (Path_4) are used to evaluate the risky decision-making of online education reform. The evaluation results are scaled by very low

(VL), low (L), relatively low (RL), medium (M), relatively high (RH), high (H), and very high (VH). As seen from Table 1, each scale corresponds to a fuzzy set of IT2FL, and the evaluation results are listed in Table 2.

5.2. Multipath Risky Decision-Making of Online Education Reform. The specific steps of multipath risky decision-making of online education reform are summarized as follows:

(i) *Step 1.* Let $P = (p_{ij})_{5\times4} = R = (r_{ij})_{5\times4}$.

(ii) *Step 2.* Equations (2) to (6) are used to calculate utility function $U_t(r_{ij})_{5\times4}$ and IT2FL entropy $E_t(r_{ij})_{5\times4}$, which are expressed as follows:

$$U_1(r_{ij})_{5\times4} = \begin{bmatrix} 1.5995 & 1.1432 & 1.4995 & 0.3561 \\ 1.4124 & 0.7152 & 1.4995 & 1.4995 \\ 1.8829 & 0.3561 & 1.4995 & 0.7152 \\ 1.4995 & 0.0717 & 1.8872 & 0.7152 \\ 1.4995 & 1.1314 & 0.7152 & 1.1314 \end{bmatrix},$$

$$U_2(r_{ij})_{5\times4} = \begin{bmatrix} 1.6056 & 1.2735 & 1.6056 & 0.5737 \\ 1.2735 & 0.8599 & 1.6056 & 1.6056 \\ 1.6480 & 0.5737 & 1.6056 & 0.8599 \\ 1.6056 & 0.2046 & 1.8460 & 0.8599 \\ 1.6056 & 1.2735 & 0.8599 & 0.8599 \end{bmatrix}, \quad (10)$$

$$E(r_{ij})_{5\times4} = \begin{bmatrix} 0.6722 & 0.6242 & 0.6722 & 0.6242 \\ 0.6242 & 0.6276 & 0.6722 & 0.6722 \\ 0.5479 & 0.6242 & 0.6722 & 0.6276 \\ 0.6722 & 0.6722 & 0.5479 & 0.6276 \\ 0.6722 & 0.6242 & 0.6276 & 0.6242 \end{bmatrix}.$$

(iii) *Step 3.* When the path weight is fully unknown and the utility function is $U_1(r_{ij})_{5\times4}$, the path weight values and changing images under different risks are shown in Table 3 and Figure 1, respectively.

(iv) When the path weight is fully unknown and the utility function is $U_2(r_{ij})_{5\times4}$, the path weight values and changing images under different risks are shown in Table 4 and Figure 2, respectively.

(v) As can be seen from Figures 1 and 2, when the decision-maker's risk preference Φ changes from -1 to 1, ω_1 and ω_3 gradually decrease, while ω_2 and ω_4 gradually increase.

(vi) *Step 4.* The OWA operator is used to aggregate the path weights and utility functions under different paths in each scheme to obtain the comprehensive utility values $U_1(x_i)$ and $U_2(x_i)$, as shown in Tables 5 and 6.

As can be seen from Tables 3–6, as Φ of decision-makers' risk preference changes from -1 to 1, when the utility function is $U_1(r_{ij})_{5\times4}$ and $U_2(r_{ij})_{5\times4}$, respectively, the weights ω_1 and ω_3 gradually decrease, while the ω_2 and ω_4 gradually increase. When the utility function is $U_1(r_{ij})_{5\times4}$ and the decision-maker's risk preference Φ changes from 1 to 0, the ranking is $x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$. When the risk preference Φ changes from 0.5 to 1, the ranking changes to $x_2 \succ x_1 \succ x_5 \succ x_3 \succ x_4$. When the utility function is $U_1(r_{ij})_{5\times4}$ and the risk preference changes from -1 to 1, the ranking changes to $x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$.

Table 1: Scale and the corresponding fuzzy set of IT2FL.

Scale	The corresponding fuzzy set of IT2FL
VL	$\{(0, 0.05, 0.05, 0.1; 1), (0, 0, 0.05, 0.05; 0.9)\}$
L	$\{(0, 0.05, 0.1, 0.2; 1), (0, 0.1, 0.1, 0.2; 0.9)\}$
RL	$\{(0.2, 0.3, 0.3, 0.5; 1), (0.2, 0.3, 0.3, 0.4; 0.9)\}$
M	$\{(0.4, 0.5, 0.5, 0.7; 1), (0.4, 0.5, 0.5, 0.6; 0.9)\}$
RH	$\{(0.6, 0.7, 0.7, 0.9; 1), (0.6, 0.7, 0.7, 0.8; 0.9)\}$
H	$\{(0.8, 0.9, 0.9, 1; 1), (0.9, 0.9, 0.9, 1; 0.9)\}$
VH	$\{(0.9, 1, 1, 1; 1), (0.95, 1, 1, 1; 0.9)\}$

When the path weight is partially known and the utility function is $U_1(r_{ij})_{5\times4}$ and $U_2(r_{ij})_{5\times4}$, the path weight obtained under different risk preferences of decision-makers is the same; that is, $\omega = (0.3, 0.1, 0.4, 0.2)$. Therefore, when the utility function is $U_1(r_{ij})_{5\times4}$ and $U_2(r_{ij})_{5\times4}$, respectively, the utility value and ranking result of the scheme can be obtained by using the OWA operator, as shown in Table 7.

5.3. Comparison Analysis

5.3.1. Ranking Results. In order to explore the validity of the proposed method in this paper, the ranking results of the proposed method are compared with those of the existing methods. Suppose that the path weight is set as $\omega = (0.3, 0.1, 0.4, 0.2)$.

(1) In [29], an integrated ranking algorithm GRAP is proposed to solve decision-making problems by combining grey relational analysis, rank-sum, and preference ranking organization method enrichment evaluation methods. $R(x_1) = 1.6425$, $R(x_2) = 1.8726$, $R(x_3) = 1.6893$, $R(x_4) = 1.4192$, and $R(x_5) = 1.7368$. So the ranking of the scheme is $x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$.

(2) In [30], a rough set-based ranking algorithm is proposed to deal with the decision-making problem. Based on this, the ranking is summarized as follows: $R(x_1) = 4.1216$, $R(x_2) = 4.6738$, $R(x_3) = 3.5724$, $R(x_4) = 2.9980$, and $R(x_5) = 4.2907$. So the ranking of the scheme is $x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$.

(3) In [31], a method for solving multicriteria decision-making problem is proposed to deal with evaluating and ranking alternatives from the best to the worst with respect to decision-maker's preferences, and the possibility degree matrix PE is defined as follows:

$$\mathrm{PE} = \begin{bmatrix} 0.5001 & 0.4778 & 0.5087 & 0.5424 & 0.4926 \\ 0.5311 & 0.5001 & 0.5192 & 0.5453 & 0.5057 \\ 0.4933 & 0.4908 & 0.5001 & 0.5146 & 0.4890 \\ 0.4975 & 0.4626 & 0.4863 & 0.5001 & 0.4270 \\ 0.5083 & 0.4952 & 0.5192 & 0.5208 & 0.5001 \end{bmatrix}. \quad (11)$$

Table 2: Evaluation results.

	$Path_1$	$Path_2$	$Path_3$	$Path_4$
Live steaming interaction (x_1)	H	RH	H	RL
MOOC (x_2)	RH	M	H	H
Online-merge-offline (x_3)	VH	RL	H	M
Modular reconstruction of course resources (x_4)	H	L	VH	M
Home-school linkage (x_5)	H	RH	M	RH

Table 3: Path weights of different risk preferences under $U_1(r_{ij})_{5\times4}$.

Φ	−1	−0.5	0	0.5	1
ω_1	0.4551	0.3707	0.3357	0.3171	0.3040
ω_2	0.0202	0.1080	0.1442	0.1701	0.1875
ω_3	0.4037	0.3446	0.3111	0.3023	0.2966
ω_4	0.1312	0.1727	0.1999	0.2046	0.2117

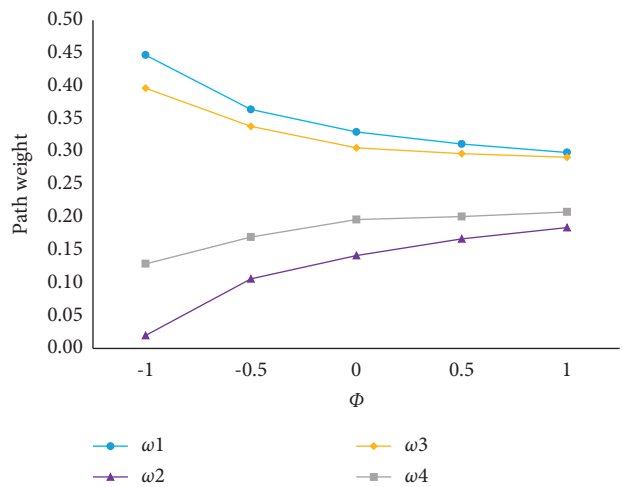


Figure 1: The changing of fully unknown path weight under different risk preferences with $U_1(r_{ij})_{5\times4}$.

Table 4: Path weights of different risk preferences under $U_2(r_{ij})_{5\times4}$.

Φ	−1	−0.5	0	0.5	1
ω_1	0.3999	0.3545	0.3149	0.3093	0.2937
ω_2	0.0785	0.1307	0.1696	0.1874	0.1960
ω_3	0.3676	0.3253	0.3050	0.2929	0.2851
ω_4	0.1568	0.1932	0.2087	0.2158	0.2250

Based on the possibility degree matrix PE, the ranking vector $r = (0.1002, 0.1038, 0.0985, 0.0937, 0.1095)$ can be obtained. So the ranking of the scheme is $x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$. Therefore, Table 8 shows the ranking results of all methods.

It can be seen from Table 8 that the ranking results of the four methods are basically the same, which indicates that the method proposed in this paper is scientific and effective.

5.3.2. Ranking Time. As can be seen from Figure 3, the proposed method has a lower ranking time with the increasing number of known path weights. This is because the risky decision-making ranking method proposed in this paper has some advantages. At first, based on the cut set, this paper proposes two IT2FL utility functions, which effectively extract all the information of IT2FL and comprehensively consider each parameter and membership, making the ranking of IT2FL more scientific, real, and effective. Then,

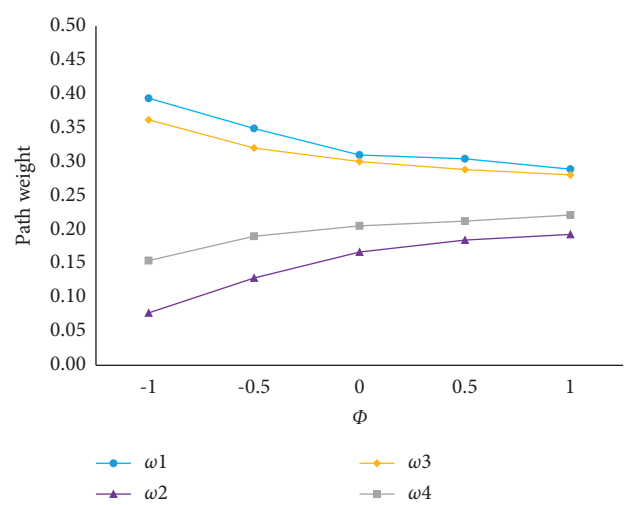


Figure 2: The changing of fully unknown path weight under different risk preferences with $U_2(r_{ij})_{5\times4}$.

Table 5: The scheme ordering of different risk preferences under $U_1(r_{ij})_{5\times4}$.

	−1	−0.5	0	0.5	1
$U_1(x_1)$	1.2693	1.2268	1.2162	1.2053	1.1976
$U_1(x_2)$	1.3081	1.2858	1.2851	1.2674	1.2713
$U_1(x_3)$	1.2146	1.1815	1.1792	1.1682	1.1514
$U_1(x_4)$	1.1882	1.1354	1.1153	1.1002	1.0983
$U_1(x_5)$	1.2996	1.2319	1.2169	1.1986	1.1883
Ranking	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$

Table 6: The scheme ordering of different risk preferences under $U_2(r_{ij})_{5\times4}$.

	−1	−0.5	0	0.5	1
$U_1(x_1)$	1.3067	1.2826	1.2775	1.2659	1.2645
$U_1(x_2)$	1.3644	1.3515	1.3505	1.3487	1.3406
$U_1(x_3)$	1.2970	1.2568	1.2478	1.2491	1.2398
$U_1(x_4)$	1.2286	1.1945	1.1802	1.1717	1.1694
$U_1(x_5)$	1.3367	1.3076	1.2961	1.2836	1.2776
Ranking	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$

Table 7: The ranking of partially known path weight.

Online education reform scheme	$U_1(x_1)$	$U_1(x_2)$
x_1	1.1701	1.2551
x_2	1.2359	1.3223
x_3	1.0817	1.1856
x_4	1.0246	1.1198
x_5	1.1939	1.2897
Ranking	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$

Table 8: The comparison of ranking results of four methods.

Method	Ranking
GRAP	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$
Rough set-based ranking	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$
Possibility degree	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$
This paper	$x_2 \succ x_5 \succ x_1 \succ x_3 \succ x_4$

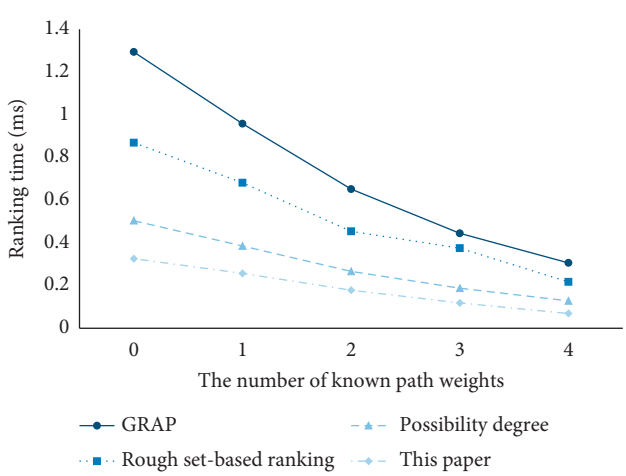


Figure 3: The comparison of ranking time.

the IT2FL entropy solves the problem of the different entropy values of complementary fuzzy sets. Finally, the risk preference factor of online education reform is introduced.

6. Conclusions

Online education reform has become an important direction of contemporary education development. In risk society, the reform process is also a continuous process, gradually transferring the teacher-centered teaching mode to the student-centered teaching mode. This study proposes two IT2FL utility functions based on the cut set. These two utility functions effectively extract all information of IT2FL, which is conducive to reducing the decision-making error of online education reform. In addition, the IT2FL entropy is proposed to measure the uncertainty of IT2FL. Furthermore, the risk preference factor is introduced to reflect the different risk attitudes of decision-makers during the risky decision-making of online education reform. The experimental results reveal that the proposed risky decision-making method has good validity in online education reform.

This paper selects five ways to reform online education, which has some limitations on risky decision-making, while there are many factors affecting the results of online education reform. In future, the joint multiattribute and multipath method will be used to reform online education reform on risky decision-making to obtain better results.

Data Availability

All data used to support the findings of the study are included within the article.

Conflicts of Interest

The author declares no conflicts of interest in this paper.

Acknowledgments

This work was supported by the General Reform Research Project of Graduate Education and Teaching at Shenyang University (Research on Graduate Education Promoting Economic Development) (Grant no. 2021-7-1-59).

References

[1] M. N. Hasan, "Techno-environmental risks and ecological modernisation in "double-risk" societies: reconceptualising Ulrich Beck's risk society thesis," *Local Environment*, vol. 23, no. 3, pp. 258–275, 2018.

[2] A. P. Di Floristella, "Dealing with natural disasters Risk society and ASEAN: a new approach to disaster management," *The Pacific Review*, vol. 29, no. 2, pp. 283–305, 2016.

[3] Y. Li, "Discussion on the teaching reform of massive open online courses in Universities," in *Proceedings of the 3rd International Conference on Economics, Management, Law and Education (EMLE 2017)*, vol. 32, pp. 621–625, Zhengzhou, China, November 2017.

[4] S. Le and Y. Wen, "Study on the organic integration of physical education and online education," in *Proceedings of the 2016 International Conference on Education, Management and Computing Technology (ICEMCT-16)*, vol. 59, pp. 297–300, Hangzhou, China, January 2016.

[5] X. Yuqing, "Study on the reform of "online and offline" mixed music education model under the 5G of the times," *International Journal of Electrical Engineering Education*, 2021.

[6] W. Jun, "Study on the reform of business English talents cultivation mode under the background of Internet plus," in *Proceedings of the 2018 International Workshop on Advances in Social Sciences (IWASS 2018)*, pp. 42–46, Hongkong, China, December 2018.

[7] M. Zhao, "Research on the innovation model of the service of convenient store under the Internet plus," in *Proceedings of the 2017 4th ICMIBI International Conference on Training, Education, and Management (ICMIBI-TEM 2017)*, vol. 88, pp. 233–237, Dubai, UAE, September 2017.

[8] D. Zhang, "Analysis on possibilities and challenges of realization of education fairness under the background of Internet plus," in *Proceedings of the 2016 International Conference on Arts, Design and Contemporary Education*, vol. 64, pp. 1392–1395, Moscow, Russia, May 2016.

[9] H. Li, Q. Xu, and B. Song, "The development of higher education in the era of "Internet plus"" in *Proceedings of the 4th International Conference on Education, Management, Arts, Economics and Social SciencE (ICEMAESS 2017)*, vol. 172, pp. 293–296, Sanya, China, November 2017.

[10] W. Wu and Q. Bai, "Why do the MOOC learners drop out of the school? Based on the investigation of MOOC learners on some Chinese MOOC platforms," in *Proceedings of the 2018 First International Cognitive Cities Conference (IC3 2018)*, pp. 299–304, Okinawa, Japan, August 2018.

[11] O. V. Semenikhina, M. G. Drushlyak, Y. A. Bondarenko, S. M. Kondratiuk, and I. M. Ionova, "Open educational resources as a trend of modern education," in *Proceedings of the 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 779–782, Opatija, Croatia, May 2019.

[12] M. Caeiro-Rodriguez, M. Llamas-Nistal, M. Fernandez-Iglesias, F. Mikic-Fonte, and M. Lama-Penín, "Supporting real open educational resources in edu-AREA different views about open educational resources," in *Proceedings of the Frontiers in Education Conference (FIE)*, pp. 1925–1932, El Paso, TX, USA, October 2015.

[13] E. Shuklina, "NON-LINEAR model OF higher education: towards the methodological problem OF the research strategy," in *Proceedings of the 9th International Conference of Education, Research and Innovation*, pp. 1016–1021, Seville, Spain, November 2016.

[14] P. Ambarova, "Educational simulacra as a consequence OF the NON-linear dynamics OF the Russian higher education," in *Proceedings of the 9th International Conference of Education, Research and Innovation*, pp. 5953–5959, Seville, Spain, November 2016.

[15] H. Jeon, D. Salinas, and D. P. Baker, "Non-linear education gradient across the nutrition transition: mothers' overweight and the population education transition," *Public Health Nutrition*, vol. 18, no. 17, pp. 3172–3182, 2015.

[16] R. Zhang, L. Zhao, L. Wu et al., "The effects of optimism on self-framing and risky decision making," *Social Behavior and Personality: An International Journal*, vol. 48, no. 10, pp. 1–9, 2020.

[17] S. Fu, X.-L. Qu, Y.-Z. Xiao, H.-J. Zhou, and G.-B. Fan, "Risky multi-attribute decision-making method based on the interval number of normal distribution," *Symmetry*, vol. 12, no. 2, Article ID 264, 2020.

[18] S. Chen, P. Yang, T. Chen, H. Su, H. Jiang, and M. Zhao, "Risky decision-making in individuals with substance use disorder: a meta-analysis and meta-regression review," *Psychopharmacology*, vol. 237, no. 7, pp. 1893–1908, 2020.

[19] J. Nowacki, H. R. Heekeren, C. E. Deuter et al., "Decision making in response to physiological and combined physiological and psychosocial stress," *Behavioral Neuroscience*, vol. 133, no. 1, pp. 59–67, 2019.

[20] Z. Dong, "The effective mode of ideological and political education in "Internet plus" environment," in *Proceedings of the 2018 3rd International Social Sciences and Education Conference (ISSEC 2018)*, pp. 343–346, Xiamen, China, June 2018.

[21] Y. Gao and S. Zhang, "Research on innovative talent cultivation mode of "Internet plus education"" in *Proceedings of the Asia-Pacific Social Science and Modern Education Conference (SSME 2018)*, vol. 193, pp. 196–200, Shanghai, China, June 2018.

[22] F. Shen, L. Ye, X. Ma, and W. Zhong, "Flipped classroom based on Internet plus education and smart learning," in *Proceedings of the 2018 2nd International Conference on Education, Economics and Management Research (ICEEMR 2018)*, vol. 182, pp. 74–77, Singapore, January 2018.

[23] J. Xu, "Research on the mode and concept of film education in the era of "Internet plus"" in *Proceedings of the 2018 3rd International Conference on Education & Education Research (EDUER 2018)*, pp. 145–148, Wuhan, China, November 2018.

[24] H. Rong, S. Li, S. Xiao, N. Bian, and H. Li, "Innovative talent cultivation pattern of software engineering under Internet plus," in *Proceedings of the 14th International Conference on Computer Science and Education (ICCSE 2019)*, pp. 180–183, Toronto, Canada, August 2019.

[25] S. P. Nangrani, A. R. Singh, and A. Chandan, "Chaos driven instability control using interval type-2 fuzzy logic controller for better performance," *Journal of Intelligent and Fuzzy Systems*, vol. 34, no. 3, pp. 1491–1501, 2018.

[26] A. Hailemichael, S. M. Salaken, A. Karimoddini, A. Homaifar, K. Abbas, and S. Nahavandi, "Developing a computationally effective interval type-2 TSK fuzzy logic Controller1," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 2, pp. 1915–1928, 2020.

[27] V. Sumati and C. Patvardhan, "Interval type-2 mutual subsethood fuzzy neural inference system (IT2MSFuNIS)," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 1, pp. 203–215, 2018.

[28] T. Zhao, Y. Xiang, S. Dian, R. Guo, and S. Li, "Hierarchical interval type-2 fuzzy path planning based on genetic

optimization," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 1, pp. 937–948, 2020.

[29] N. Deepa, B. Prabadevi, and G. Srivastava, "Integrated ranking algorithm for efficient decision making," *International Journal of Information Technology and Decision Making*, vol. 20, no. 02, pp. 597–618, 2021.

[30] X. Ma and J. Zhan, "Applications of rough soft sets to BL-algebras and corresponding decision making methods," *Journal of Intelligent and Fuzzy Systems*, vol. 34, no. 1, pp. 645–658, 2018.

[31] S. Samanta and D. K. Jana, "A multi-item transportation problem with mode of transportation preference by MCDM method in interval type-2 fuzzy environment," *Neural Computing & Applications*, vol. 31, no. 2, pp. 605–617, 2019.

[32] W. Zeng, D. Li, and Y. Gu, "Monotonic argument-dependent OWA operators," *International Journal of Intelligent Systems*, vol. 33, no. 8, pp. 1639–1659, 2018.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 6839179, 12 pages
https://doi.org/10.1155/2021/6839179

Research Article

Extracting Function-Driven Tracing Characteristics for Optimized SVM Classification

Ming Wan,[1] **Xinlu Xu**,[1] **Yan Song**,[2] **Quanliang Li**,[1] **and Jiawei Li**[1]

[1]*School of Information, Liaoning University, Shenyang 110036, China*
[2]*School of Physics, Liaoning University, Shenyang 110036, China*

Correspondence should be addressed to Yan Song; song.yan@lnu.edu.cn

Received 10 November 2021; Revised 8 December 2021; Accepted 11 December 2021; Published 26 December 2021

Academic Editor: Jianhui Lv

Due to its openness and simplicity, Modbus TCP has wide applications to facilitate the actual management and control in industrial wireless fields. However, its potential security vulnerabilities can also create lots of complicated information security challenges, which are increasingly threatening the availability of industrial real-time traffic delivery. Although anomaly detection has been recognized as a workable security measure to identify attacks, the critical step to successfully extract data characteristics is an extremely difficult task. In this paper, we focus on the continuous control mode in industrial processes and propose a control tracing feature algorithm to extract the function-driven tracing characteristics from Modbus TCP data traffic. Furthermore, this algorithm can flexibly integrate the time factor with critical functional operations and adequately describe the dynamic control change of technological processes. To closely cooperate with this algorithm, one optimized SVM (support vector machine) classifier is introduced as the practicable decision engine. By designing one applicable attack mode, we develop an in-depth and meticulous analysis on the decision accuracy, and all experimental results clearly explain that the extracted features can strongly reflect the changing pattern of continuous functional operations, and the proposed algorithm can effectively cooperate with the optimized SVM classifier to distinguish abnormal Modbus TCP data traffic.

1. Introduction

Modbus TCP, which is regarded as one representative industrial communication protocol, has been widely applied in various critical infrastructures, including power generation, steel rolling, oil refinery, gas purification, and so on. In particular, with the rapid development of IIoT (industrial internet of things), many industrial wireless fields have collaborated Modbus TCP with other wireless communication protocols to perform various control, management, and monitoring activities of mobile devices or endpoints [1–3]. As a typical case, one wireless gateway can successfully apply Modbus TCP to realize industrial real-time traffic delivery for mobile sensors or actuators. Generally, Modbus TCP is used to accomplish information interactions between different field devices and facilitates the actual management and control in industrial processes [4, 5]. Furthermore, Modbus TCP defines the master/slave communication style, which exploits the request/response way to achieve the data exchange between master and slave stations [6]. For example, when one master station sends one read or write request message, the slave station builds a response message to perform different functional operations. Additionally, one slave station cannot actively send any request to one master station, and all slave stations are forbidden to communicate with each other. According to its protocol specification, Modbus TCP belongs to one real-time communication protocol implemented on the application layer, and its essence is to embed the Modbus frame into the data field of TCP protocol. To be more specific, function code is an important protocol field in the basic Modbus frame structure, which not only can represent the main function of the request message but also can clearly point out the specific type of functional operation to control field devices [7, 8]. In other words, all key functional operations in the whole industrial process can be determined by a range of different function codes, and this design can improve the efficiency of industrial production by simplifying process control and

management. However, it can also be exploited by malicious adversaries to launch targeted attacks due to the potential security vulnerabilities of Modbus TCP. In traditional applications, ICSs (industrial control systems) are physically isolated, and all nonessential external accesses can be disabled. As a consequence, the original design of Modbus TCP misses the privacy and security concerns and only depends on the basic TCP security mechanism. With the tight integration between IT (information technology) and OT (operational technology), the potential security vulnerabilities of Modbus TCP are gradually exposed due to multiple external threats. Currently, many scientists have carried out tremendous researches on Modbus TCP security issues [4, 9–11], mainly including unauthenticated access, function code abuse, unencrypted data delivery, data tampering and replaying, and so on. For example, when one adversary gains access to industrial control systems based on Modbus TCP, it not only can intercept all critical device states and production data but also can insert some abnormal function codes to destroy the regular technological processes. That is to say, any adversary can falsify one normal Modbus TCP request with one malicious function code to change the basic control logic and perform disruptive operations on industrial field devices [12, 13]. To some extent, the potential security vulnerabilities of Modbus TCP have become a breakthrough to attack key control components, whose failures may cause serious production accidents.

In most security measures, anomaly detection has been widely acknowledged and researched by both academia and industry because it can effectively identify known and unknown attacks without impacting ICS's availability [14]. Especially, anomaly detection based on machine learning algorithms can enhance the popular application of artificial intelligence technologies in the domain of information security and achieve relatively good effect in security practices [15–18]. However, before applying anomaly detection, an essential prerequisite is to design an excellent feature generation and extraction algorithm that must inherit the main characteristics from the original data [15, 19]. Actually, the working quality of these algorithms can directly affect the performance of anomaly detection, and one excellent algorithm can enhance the detection efficiency to some extent. Different from the high-dimensional and complex data in traditional IT systems, industrial data traffic presents relatively simple characteristics and trends due to the periodic operating process and stable system structure [20]. In most industrial production activities, process control is an important component, which can indirectly reflect some technological process by using a series of consecutive functional operations. Correspondingly, they are the most direct connection with the changing laws of function codes in Modbus TCP data traffic [21]. From this point of view, this paper proposes a control tracing feature algorithm, which extracts function-driven tracing characteristics by analyzing the continuous control mode from Modbus TCP data traffic. Moreover, this algorithm not only takes into account the time factor caused by the time intervals between every two consecutive functional operations but also associates with the critical characteristics of sequential control predefined by the technological process. That is, this algorithm can flexibly integrate the time factor with critical functional operations and adequately describe the dynamic control change of technological process. In order to closely cooperate with this algorithm, one optimized SVM (support vector machine) classifier is introduced as the practicable decision engine, which is applied in combination with three different intelligent optimization algorithms, including PSO (particle swarm optimization), GA (genetic algorithm), and GRID. At last, we design one applicable attack mode to evaluate SVM's decision accuracy, and our main purpose includes the following two aspects: on the one hand, based on the function-driven tracing characteristics expressed from consecutive function codes, we prove that the extracted features can strongly reflect the changing pattern of continuous functional operations; on the other hand, compared with different intelligent optimization algorithms, we prove that the optimized SVM classifier can effectively cooperate with the proposed feature algorithm to distinguish abnormal Modbus TCP data traffic.

The main contributions of this paper are summarized as follows:

Firstly, according to the continuous control mode in the industrial process, we propose a novel control tracing feature algorithm, and this algorithm can effectively extract function-driven tracing characteristics, mainly including the probability characteristic and time parameter of operational event sequences.

Secondly, in order to closely cooperate with the proposed feature algorithm, we select the SVM classifier as the practicable decision engine and discuss the most appropriate optimization algorithm to strengthen SVM's decision ability.

Thirdly, in the experimental evaluation, we design one attack mode to simulate one possible attack, which fully exploits the security vulnerabilities of unauthenticated access and function code abuse. Also, we give a whole scale analysis on the effects of critical parameter in the proposed feature algorithm.

2. Function-Driven Tracing Characteristics Description and Extraction

In our control tracing feature algorithm, we extract the function-driven tracing characteristics from Modbus TCP data traffic by performing the following steps: firstly, we preprocess Modbus TCP data traffic to extract operational event sequences that consist of a series of consecutive function codes; secondly, based on PST (probabilistic suffix tree) [22, 23], we further perform the control tracing analysis in industrial processes and build a functional association tree; finally, according to the probability of control tracing and time parameter, we obtain the simplified feature samples by the vector calculation.

2.1. Operational Event Sequence Description. From the Modbus TCP data traffic over a period of time T, we extract the corresponding function code of each packet in chronological order and define an operational event sequence as

$O_i = f_1^i, f_2^i, f_3^i, \ldots, f_m^i$ that is composed of m consecutive function codes. Here, we regard a function code as an operational event, and f_j^i represents the j_{th} $(j \le m)$ operational event in the operational event sequence O_i. Additionally, all operational event sequences construct the operational event sequence set $O = \{O_1, O_2, O_3, \ldots, O_n\}$. Because each O_i $(i \in [1, n])$ is composed of m consecutive function codes, its corresponding time interval T_i $(i \in [1, n])$ also includes the time intervals of m consecutive function codes, and $T = \sum_{i=1}^{n} T_i$. Due to the disparate time intervals of any two consecutive function codes, we also have $T_i \neq T_{i+1}$ when $\forall i \in [1, n-1]$.

2.2. Functional Association Tree Construction. In general, PST can describe the probabilistic characteristics of sequence sets and present a k-ary tree whose nodes arrange regularly. Furthermore, each leaf node can denote a tracing path from the root node to this leaf node, and the L-depth tree involves L nodes, which store L symbols. As it describes, the jumping of two adjacent nodes can be expressed by the conditional probability of some prior symbols. Actually, the consecutive functional operations in industrial processes can be considered as a tracing path from the root node to some leaf node, and this tracing path is defined as the functional control tracing in this paper. Here, each leaf node involves a functional control tracing sequence, and the value of each leaf node corresponds to one probability value of its operational event sequence. Similarly, the jumping of two adjacent nodes can depict the changing of two consecutive function codes in functional operations. Figure 1 shows the construction process of the functional association tree. In this tree, the tree depth is L, which indirectly expresses L functional operations. Moreover, we also assume all control tracing sequences include v different function codes $f_a, f_b, f_c, f_d, \ldots, f_v$. The detailed process is described below:

(1) Create the tree root.

(2) According to v different function codes, create v different branches. In each branch, there is a leaf node that includes two variables, such as $\{f_b', p(f_b)\}$ of the first leaf node f_b in the figure. Here, f_b' represents the initial control tracing sequence, and $p(f_b)$ represents the corresponding probability of this sequence, namely, the jumping probability from the root node to this leaf node (indicated by the dotted line in the figure).

(3) Under every first leaf node, create v different branches. In each branch, there is a leaf node that includes two variables, such as $\{f_b' f_a', p(f_b f_a)\}$ of the second leaf node f_a under the first leaf node f_b. Here, $f_b' f_a'$ represents the control tracing sequence of this leaf node, and $p(f_b f_a)$ represents the corresponding probability of this sequence. Additionally, the jumping probability from the first leaf node f_b to this leaf node (indicated by the dotted line in the figure) can be expressed by the conditional probability $p(f_a | f_b)$.

(4) Under every second leaf node, create v different branches. In each branch, there is a leaf node that includes two variables, such as $\{f_b' f_a f_c', p(f_b f_a f_c)\}$ of the third leaf node f_c under the second leaf node f_a. Similarly, $f_b' f_a f_c'$ represents the functional control tracing sequence of this leaf node, and $p(f_b f_a f_c)$ represents the corresponding probability of this sequence. Additionally, the jumping probability from the second leaf node f_a to this leaf node (indicated by the dotted line in the figure) can be expressed by the conditional probability $p(f_c | f_b f_a)$.

(5) In a similar fashion, the L_{th} leaf node is further created. As shown in the figure, suppose the $(L-1)_{\text{th}}$ leaf node is f_d and its next leaf node is f_v. So the corresponding variables are $\{f_b' f_a f_c, \ldots, f_d f_v', p(f_b f_a f_c, \ldots, f_d f_v)\}$. Here, $f_b' f_a f_c, \ldots, f_d f_v'$ represents the control tracing sequence of this leaf node, and $p(f_b f_a f_c, \ldots, f_d f_v)$ represents the corresponding probability of this sequence. Additionally, the jumping probability from the $(L-1)_{\text{th}}$ leaf node f_d to this leaf node (indicated by the dotted line in the figure) can be expressed by the conditional probability $p(f_v | f_b f_a f_c, \ldots, f_d)$.

(6) After all L leaf nodes are created, the whole functional association tree is built successfully.

According to the above construction process, the probability of each control tracing sequence can be calculated by the conditional probability from the root node to the last leaf node.

Suppose the control tracing sequence is $s = f_1, f_2, \ldots, f_L$, then

$$\begin{aligned} p(s) &= p(f_1) \times p(f_2 | f_1) \times \cdots \times p(f_L | f_1 f_2, \ldots, f_{L-1}) \\ &= \prod_{j=1}^{L} p\left(f_j | f_1 f_2, \ldots, f_{j-1}\right), \end{aligned} \tag{1}$$

where $p(f_j | f_1 f_2, \ldots, f_{j-1})$ $(j \in [1, L])$ represents the probability of the next function code f_j when the prior $j-1$ function codes are $f_1 f_2, \ldots, f_{j-1}$ in the operational event sequence.

For example, as indicated by the dotted line in Figure 1, if the control tracing sequence of L_{th} leaf node is $s_L = f_b, f_a, f_c, \ldots, f_d, f_v$; the corresponding probability of this sequence can be calculated by

$$\begin{aligned} p(s_L) = p(f_b) \times p(f_a | f_b) \times p(f_c | f_b f_a) \times \cdots \\ \times p(f_v | f_b f_a f_c, \ldots, f_d). \end{aligned} \tag{2}$$

3. Feature Factor Selection and Feature Value Calculation

In the feature factor selection, we cannot only consider a unilateral role of the control tracing sequence's probability. Through our careful observation, the time intervals between

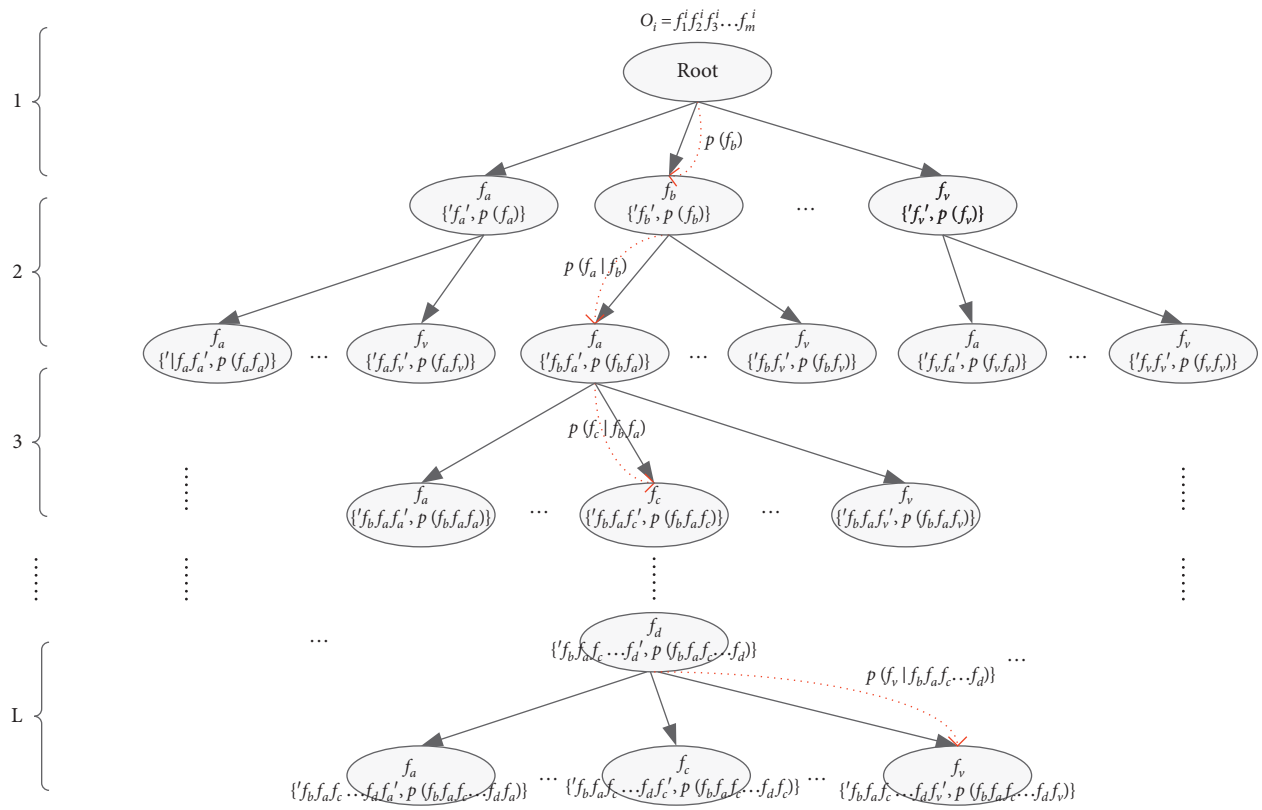


Figure 1: Construction process of functional association tree.

any two consecutive functional operations in each control tracing sequence are different. In other words, the functional operations in industrial processes are closely associated with the time parameter. In order to establish a relationship between the control tracing sequence's probability and the time parameter, we make the following setting and calculation:

According to the tracing path in the functional association tree (supposing the control tracing sequence is $s = f_1, f_2, \ldots, f_L$), we construct the probability vector of functional operation $\overrightarrow{S} = (p(f_1), p(f_1 f_2), \ldots, p(f_1 f_2, \ldots, f_L))$, which contains L different variables. Moreover, in the tracing path of s, it also takes some time intervals to implement the jumping of two adjacent leaf nodes. In this case, we can construct the corresponding the interval vector $\overrightarrow{\tau} = (\tau_1, \tau_2, \ldots, \tau_L)$. Here, $\tau_{j+1} = \sum_1^j t_j$ $(j \in [1, L-1])$, and t_j represents the time interval between two consecutive functional operations in s. From the above analysis, we regard the probability vector $\overrightarrow{S}$ and the interval vector $\overrightarrow{\tau}$ as two feature factors and calculate the feature value by the dot product operation of these two feature factors as follows:

$$x = \overrightarrow{S} \cdot \overrightarrow{\tau}. \tag{3}$$

In summary, if we suppose the operational event sequence O_i involves h different functional operations (namely, h different function codes) and the tree depth is L, we can calculate h^L feature values. That is, there are h^L feature values in each feature sample.

3.1. Dimension Reduction of Feature Sample. In the above-mentioned design, the oversize dimension of the feature sample may become a drawback, and the dimension of the feature sample keeps increasing exponentially with the growth of tree depth. For example, if we suppose one industrial process involves four different functional operations and the tree depth is 6, each feature sample can consist of $4^6 = 4096$ feature values. In practice, the oversize dimension of the feature sample may have a harmful influence on SVM's decision ability because it can cost lots of computing and storage resources. However, due to the limited behaviors and states in each industrial process, the combination number of functional operations may take on an evident downtrend. As a result, the number of different control tracing sequences in one operational event sequence set is also limited and is far less than h^L. More specifically, we can reduce the number of leaf nodes in the functional association tree and the dimension of the feature sample in the following ways: firstly, we rearrange all operational event sequences in chronological order and determine the number r of all control tracing sequences, here $r \ll h^L$; secondly, for each operational event sequence O_i, we build the corresponding functional association tree by using these r control tracing sequences; and finally, we

calculate the feature sample $X_i = (x_1^i, x_2^i, \ldots, x_r^i)$ for each operational event sequence O_i, and the dimension of each feature sample changes to r.

4. Optimized SVM Decision Engine

In the design of the decision engine, we select the classical SVM classifier to match with the above control tracing feature algorithm. Furthermore, by using the above feature samples, this decision engine can train a mathematical model to recognize the statistical deviation for the observed Modbus TCP data traffic. Simultaneously, in order to strengthen its decision ability, we use three different intelligent optimization algorithms to optimize its key parameters respectively, and our ultimate goal is to obtain one optimal decision engine that can effectively cooperate with the proposed feature algorithm.

4.1. SVM Classifier. SVM classifier [18, 24, 25], which is a representative machine learning method based on the statistical learning theory, explores structural risk minimization to improve its adaptive generation ability. Essentially, the SVM classifier belongs to the binary classification, and its core idea is to search for one optimal hyperplane that can make a big effort to separate two categories of samples. For the undivided linear sample space, the SVM classifier can use the nonlinear mapping algorithm to change the original data space to the high-dimensional feature space and further carry out the linear analysis on the nonlinear characteristics of samples in this space.

In particular, by introducing the relaxation factor and the penalty factor C, SVM is always designed to resolve the quadratic programming problem, which can be further simplified by Lagrange function into

$$\begin{cases} \max\limits_{\alpha} \quad \sum\limits_{i=1}^{n} \alpha_i - \dfrac{1}{2} \sum\limits_{j=1}^{n} \sum\limits_{i=1}^{n} \alpha_i \alpha_j y_i y_j k\left(x_i, x_j\right) \\ \\ \text{s.t.} \quad \sum\limits_{i=1}^{n} \alpha_i y_i = 0, \quad 0 \le \alpha_i \le C, i = 1, 2, \ldots, n, \end{cases} \tag{4}$$

where $k(x_i, x_j)$ is the kernel function, typically including RBF (radial basis function) kernel, polynomial kernel, linear kernel, and so on. In our decision engine, we select RBF kernel as the final kernel function because it has lower complexity to realize the nonlinear mapping by using fewer kernel parameters. RBF kernel can be calculated by

$$k(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right), \tag{5}$$

where σ is the kernel parameter, which represents the radial action limit of this function. Finally, the decision function can be calculated by

$$\begin{cases} f(x) = \operatorname{sgn}\left(\sum\limits_{i=1}^{n} \alpha_i y_i k(x_i, x) + b\right), \\ \text{s.t.}\, b = y_j - \sum\limits_{i=1}^{n} y_i \alpha_i k\left(x_i, x_j\right), \quad j \in \left\{j | 0 < \alpha_j < C\right\}. \end{cases} \tag{6}$$

In practical applications, the performance of SVM classifier can be largely influenced by the penalty factor C and the kernel function's parameters, such as RBF's kernel parameter σ. Furthermore, the penalty factor can affect the margin maximization and sample misclassification, and the kernel parameter can affect the data distribution and the separated hyperplane in the high-dimensional feature space. Different from the experimental selection and settings, some intelligent optimization algorithms have been successfully applied to optimize these parameters and achieve relatively satisfying effects.

4.2. Parameter Optimization Selection. It is universally acknowledged that intelligent optimization algorithms have been widely used in all kinds of information computing fields, and they can further enhance learning efficiencies and working performances by solving single- or many-objective optimization problems [26, 27]. To effectively cooperate with the proposed feature algorithm, the SVM classifier must be optimized for the maximum performance benefit. However, the applicability of different intelligent optimization algorithms may show great variation due to their distinct generalization abilities and fitting degrees. In our decision engine, we select three representative intelligent optimization algorithms to compare different optimization effects, and these three algorithms have been shown to positively affect SVM's decision ability by many researchers. Moreover, these three algorithms are PSO [28, 29], GA [30], and GRID [31], and the training process of the SVM decision engine based on parameter optimization is presented in Figure 2. In each optimization algorithm, the fitness value can be obtained by using the current SVM's decision accuracy, and the optimal parameters can be chosen from the global optimums in all iterations. Actually, for the same training samples, each optimization algorithm can generate a group of optimal parameters, which can be further compared to obtain the optimal SVM decision engine.

5. Experimental Analysis and Discussion

In our experiments, we use the captured Modbus TCP data traffic in [32] to carry out the proposed feature algorithm and introduce the decision accuracy as one evaluation indicator to compare SVM's decision abilities under three optimization algorithms. Practically speaking, our main purpose is to explore the matching degree between the control tracing feature algorithm and SVM classifier and try to find a workable intelligent optimization algorithm to obtain the optimal SVM decision engine. Additionally, we define one applicable attack mode to facilitate the evaluation process, and this attack can continuously send a stream of malicious function codes to disturb normal functional

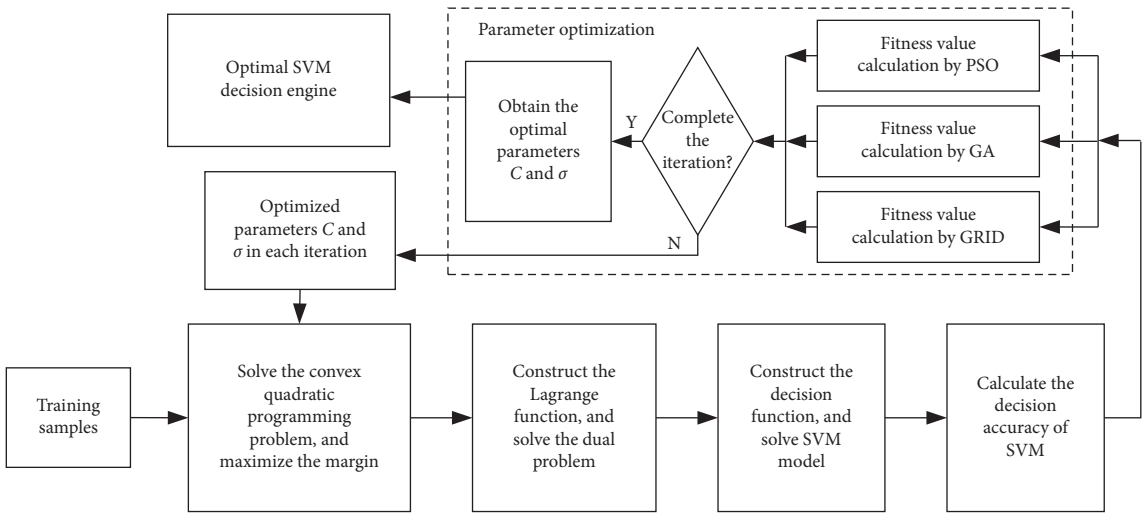


Figure 2: Training process of SVM decision engine based on parameter optimization.

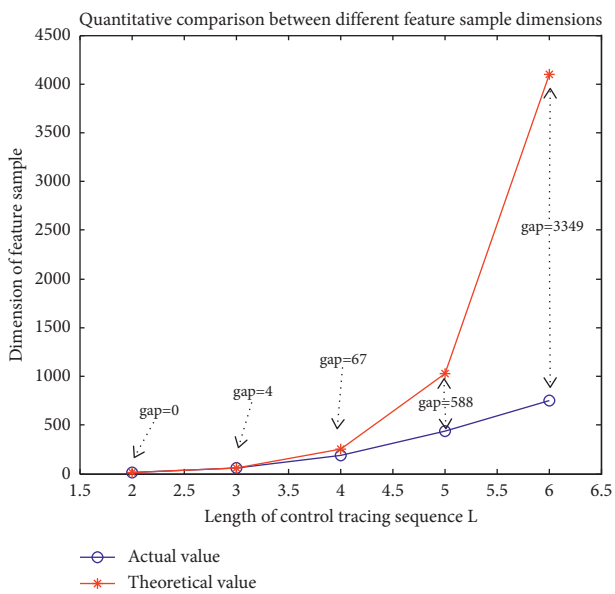


Figure 3: Compared feature sample dimensions under different lengths of control tracing sequence.

operations. In practice, this attack mode possesses a certain level of implementation possibility because it fully exploits the security vulnerabilities of unauthenticated access and function code abuse. Without loss of generality, we also define four different attack powers by changing the number of malicious function codes, whose percentages in each operational event sequence are 2.58%, 3.87%, 5.16%, and 6.45% for light-power attacks, medium-power attacks, heavy-power attacks, and fatal-power attacks, respectively. Furthermore, we discuss the change of SVM's decision abilities under different attack powers.

5.1. Certifiable Dimension Reduction. Based on the basic technological process, we define the length of one operational event sequence as 155, namely $m = 155$ consecutive function codes in each operational event sequence. After the preprocessing, we can obtain a total of 287 original operational event sequences and corresponding intervals. Similarly, we can further get the same number of normal feature samples calculated by the control tracing feature algorithm. However, an extremely important problem is that the large dimension of the feature sample may waste lots of computing resources and affect the decision ability when the SVM decision engine is applied. In theory, each feature sample may consist of h^L feature values at the extreme, and the corresponding dimension may become a very large value. Fortunately, our algorithm can reduce the dimension of the feature sample by determining the number of all the control tracing sequences in actual applications. Figure 3 gives the compared feature sample dimensions under different lengths of control tracing sequence. From this figure, we can see that with the length of the control tracing sequence increasing, the gap between the actual dimension

Table 1: Optimized results under three intelligent optimization algorithms.

	PSO	GA	GRID
C	100,000	19,271.29	24,833.50
σ	0.0478	0.0199	0.0136
Training time (s)	1,283.67	757.99	1,385.14
Training accuracy (%)	99.10	94.93	94.81

and theoretical dimension of the feature sample grows larger and larger. For instance, when the lengths of control tracing sequences are 3, 4, and 5, the corresponding gaps reach 4, 57, and 588, respectively. The above quantitative comparison fully illustrates that our algorithm can produce the desired results on dimension reduction.

5.2. Decision Results and Analysis. As discussed above, we further set the length of the control tracing sequence as 3 (namely, the tree depth $L = 3$), and the corresponding dimension of the feature sample is 60. Based on the obtained feature samples, we use the offline way to train three optimal SVM decision engines, whose main parameters are optimized by PSO, GA, and GRID. Moreover, Table 1 presents the optimized results under these three intelligent optimization algorithms. From this table, we can see that three optimization algorithms can achieve different groups of penalty factor C and kernel parameter σ due to their own optimization mechanisms, and the training results of the SVM decision engine are not identical under different key parameters. More specifically, the SVM decision engine optimized by PSO exhibits the highest training accuracy that can reach 99.10%, and this value is about 4 percentage points higher than the ones optimized by GA and GRID. Additionally, although the training time has failed to meet the expectation of real-time training and optimization, we can ignore its impact by using the way of offline training and online detection.

Based on the defined attack mode, we generate 600 test feature samples in every experiment, mainly including 200 normal feature samples and 400 malicious feature samples. Furthermore, we record the number of feature samples that are correctly classified and calculate its percentage as the decision accuracy. For each level of attack power, we conduct 10 different experiments whose malicious samples are generated by forging and inserting abnormal function codes and present the change of SVM's decision accuracies. Figure 4 compares the decision accuracies of three optimized SVM decision engines under four different attack powers, and Table 2 gives the corresponding average decision accuracies. From Table 2, we can see that three optimization algorithms can fulfill their mandates to achieve three promising SVM decision engines, whose average decision accuracies do not differ significantly under each attack power. As a whole, the SVM decision engine optimized by PSO yields the best decision ability because it has the highest decision accuracy under each attack power. More specifically, its average decision accuracies to detect light- and fatal-power attacks can reach 87.82% and 97.40%, respectively, which are 1 and 0.5 percentage points higher than the lowest one of the SVM decision engine optimized by GA. Additionally, as the attack power increases, the average decision accuracies for all optimization algorithms become larger. In other words, when the number of malicious function codes in each operational event sequence increases, the possibility to successfully detect abnormal functional operations by the optimized SVM decision engine can also be enhanced gradually. For the decision time, the average values of the SVM decision engine optimized by PSO can reach 0.01968 s, which is short enough to meet the real-time demand in industrial production. From the above experimental results and analysis, we can conclude that on the one hand, the proposed feature algorithm can adequately reflect the changing pattern of continuous functional operations, and one slight change of functional operations may bring a great difference of feature values that can be used to improve decision accuracies; on the other hand, SVM classifier can be selected as one applicable decision engine to cooperate with the proposed feature algorithm, and the PSO optimization algorithm can offer the most brilliant contribution to SVM's decision ability by comparing with other representative optimization algorithms.

In practice, the fluctuation of decision accuracy can indirectly reflect the stability of SVM decision ability because the great fluctuation of decision accuracy may cause relatively low decision accuracy in certain circumstances. For each attack power, the larger the divergence between the decision accuracy in each experiment and the average decision accuracy is, the greater the fluctuation of decision accuracy is, the poorer the stability of SVM decision ability becomes. In the related theories of probability statistics, the standard deviation, which describes the dispersion degree between individuals, is most commonly used as the measurement of statistical dispersion. Essentially, the larger standard deviation represents the greater dispersion. In our evaluation, we introduce the standard deviation to further analyze different stabilities for three optimized SVM decision engines, and Figure 5 depicts the corresponding standard deviations of decision accuracies under four different attack powers. As shown in Figures 4 and 5, with the increase of attack power, the standard deviations of decision accuracies for all optimized SVM decision engines gradually decrease. That is to say, the divergence between the decision accuracy in each experiment and the average decision accuracy becomes smaller and smaller, and the stability of each optimized SVM decision ability gets better and better. However, from the experimental results in Table 2, although the SVM decision engine optimized by PSO has the most outstanding decision ability, its standard deviation under light-power attacks is the largest, and the corresponding stability is far from satisfactory. Differently, the other SVM

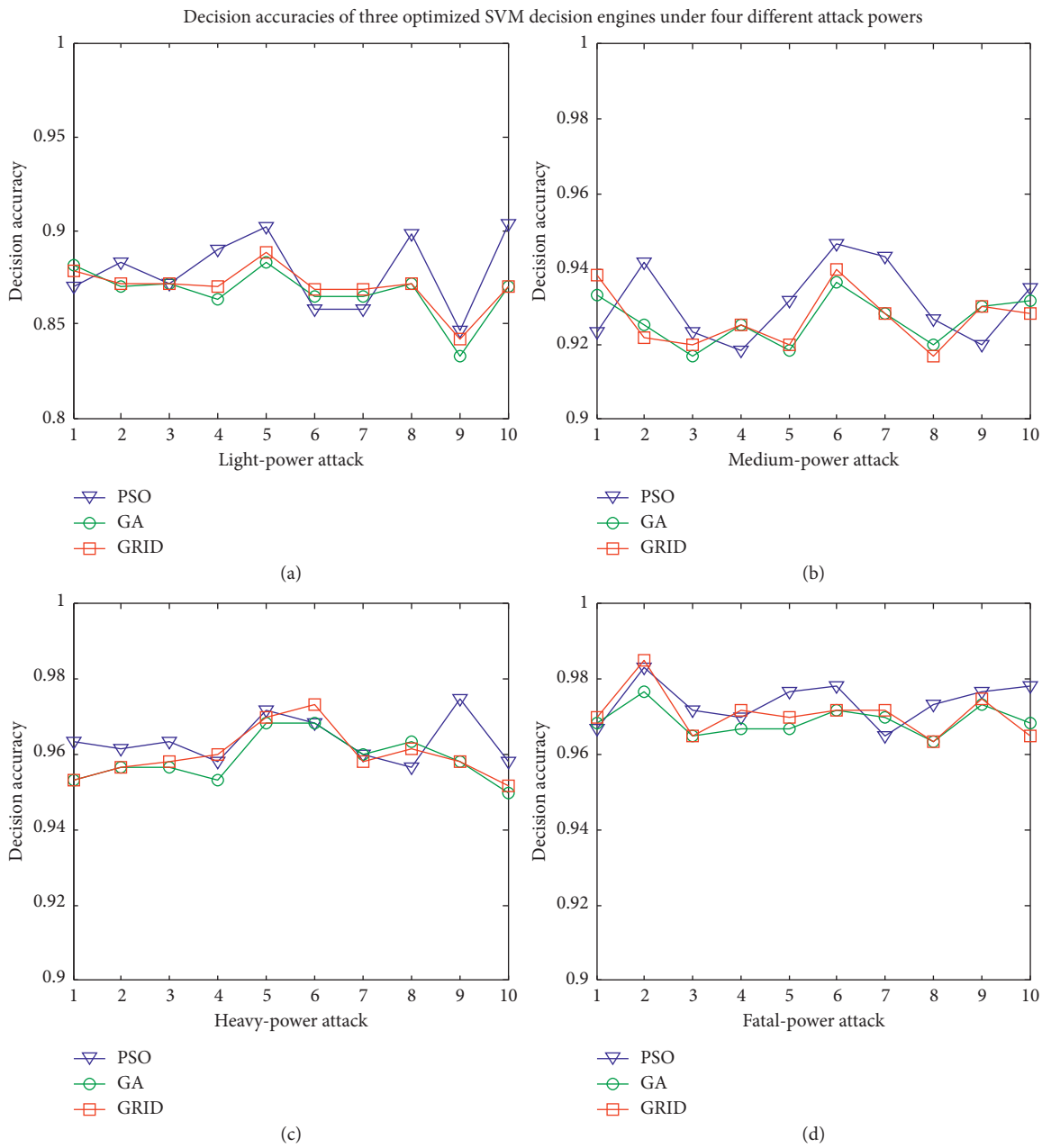


Figure 4: Decision accuracies of three optimized SVM decision engines under four different attack powers: (a) light-power attack, (b) medium-power attack, (c) heavy-power attack, and (d) fatal-power attack.

decision engines have relatively high decision stability, but their decision abilities are slightly inferior. In practical applications, we should select an appropriate intelligent optimization algorithm to optimize the SVM classifier according to the specific scenarios and requirements.

5.3. Influence of Control Tracing Sequence Length. In the proposed control tracing feature algorithm, the length L of the control tracing sequence is a preconfigured variable, which may impact the decision ability of the SVM decision engine. On the one hand, it is actually easy to understand that the decision time will undoubtedly increase with the growth of L because the dimension of the feature sample becomes higher. On the other hand, the decision accuracy can also change in different degrees when L increases or decreases. In our experiments, we select $L = 2$ to illustrate this point because the dimension of the feature sample can get greater than the length of the operational event sequence if $L > 3$. Figure 6 depicts the decision accuracy comparison of three optimized SVM decision engines under different lengths of control tracing sequence, and every average decision accuracy in this figure is also calculated by 10 different experiments. According to the direct presentation of this figure, we can draw the following conclusions: for one thing, when $L = 2$, the SVM decision

Table 2: Average decision accuracies of three optimized SVM decision engines under four different attack powers.

Attack power	Average detection accuracy (%)		
	PSO-SVM	GA-SVM	GRID-SVM
Light-power attack	87.82	86.75	87.00
Medium-power attack	93.10	92.65	92.68
Heavy-power attack	96.37	95.88	96.02
Fatal-power attack	97.40	96.90	97.08

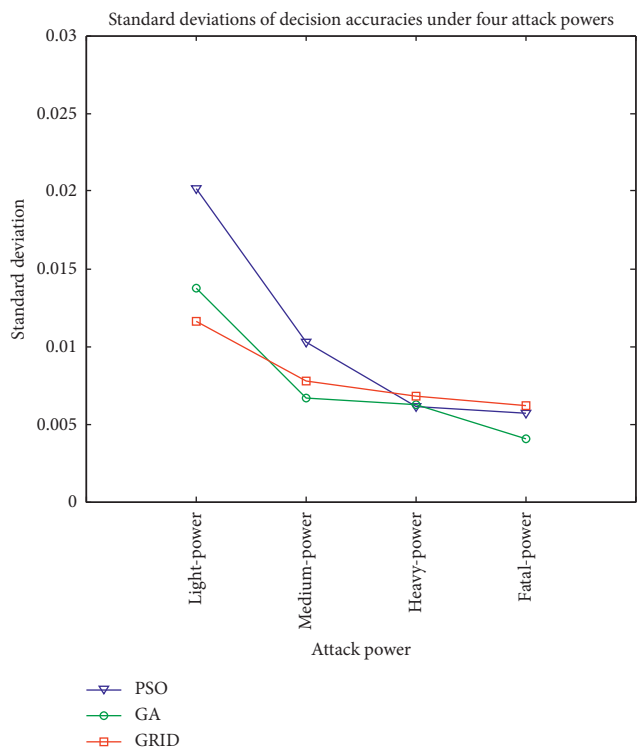


Figure 5: Standard deviations of decision accuracies for three optimized SVM decision engines under four different attack powers.

engine optimized by PSO still has the best decision ability by comparing the average decision accuracies of three optimized SVM decision engines; for another thing, all average decision accuracies of three optimized SVM decision engines are improved to one higher level when the length of control tracing sequence increases under each attack power. To sum up, if we select an SVM classifier as the applicable decision engine, we should design one appropriate L under comprehensive considerations of decision accuracy and computation efficiency.

5.4. Effect of Changed Time Parameter. In our control tracing feature algorithm, the time parameter (interval vector $\overrightarrow{\tau}$) is selected as an important feature factor to calculate the final feature value, and the change of time parameter can initiate dynamic modifications to the feature sample. Without loss of generality, when one baleful adversary wants to launch the given attack, the time interval between two consecutive functional operations will change as well. In order to illustrate the effect of the changed time parameter, we also perform 10 different experiments to compare the average decision accuracies, and the corresponding attack power belongs to the fatal-power level. In these experiments, we develop two hypothetical scenarios: one is that the given attack can generate changed time parameters, and the other is that the given attack cannot cause the change of time parameter. Additionally, we choose the SVM decision engine optimized by PSO as the experimental object because other optimized SVM decision engines will have a similar presence. Figure 7 shows its decision accuracy comparison under the changed time parameter and the unchanged time parameter. Intuitively, the average decision accuracy under the unchanged time parameter is only 95.22%, which is less than the one under the changed time parameter. According to the compared results, we recognize that the changed time parameters can improve the average decision accuracy of the optimized SVM decision engine to some degree. In other words, the introduced time parameter in our control tracing feature algorithm can provide an effective role in improving the decision ability of the SVM decision engine.

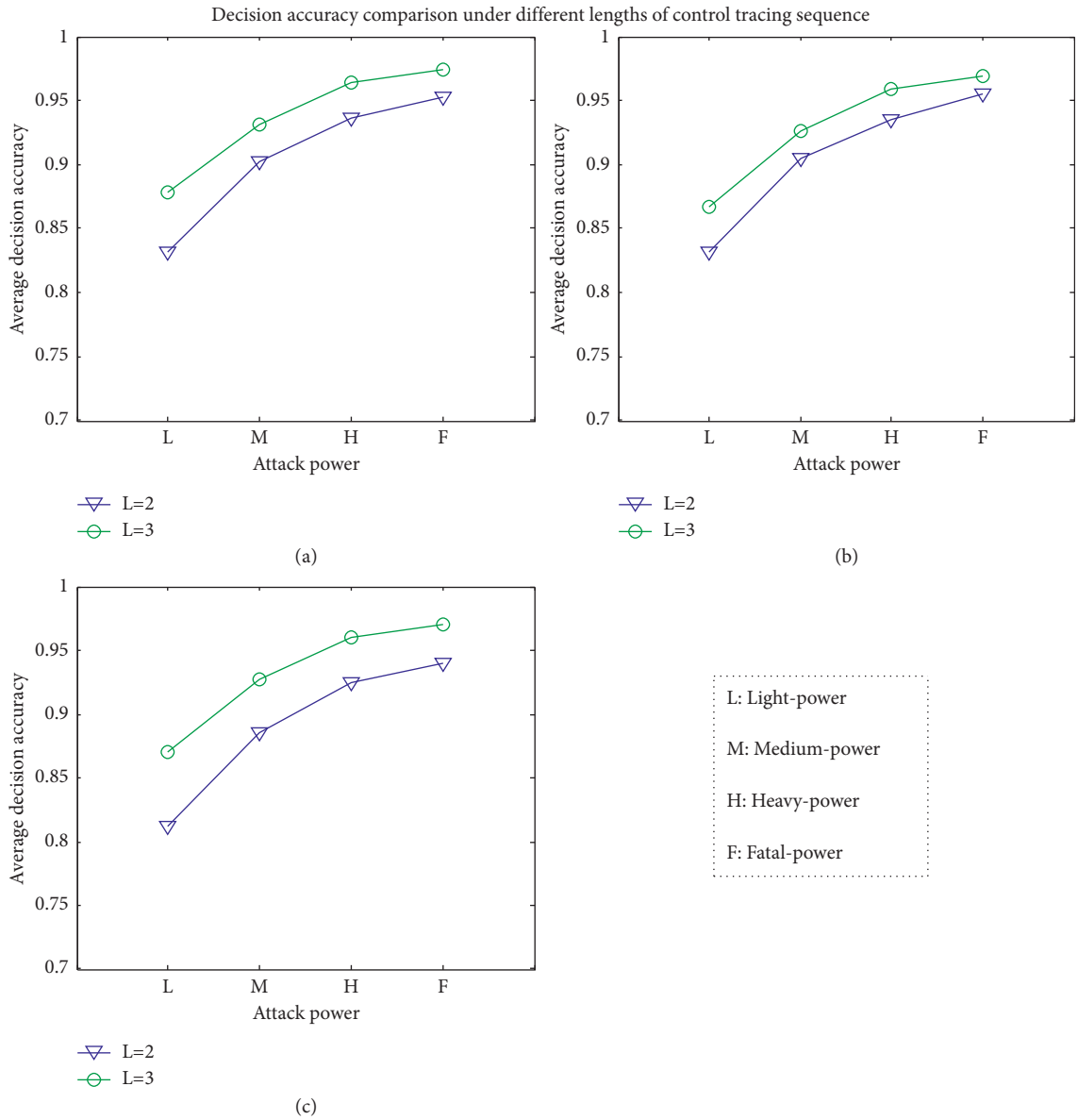


Figure 6: Decision accuracy comparison of three optimized SVM decision engines under different lengths of control tracing sequence: (a) PSO-SVM, (b) GA-SVM, and (c) GRID-SVM.

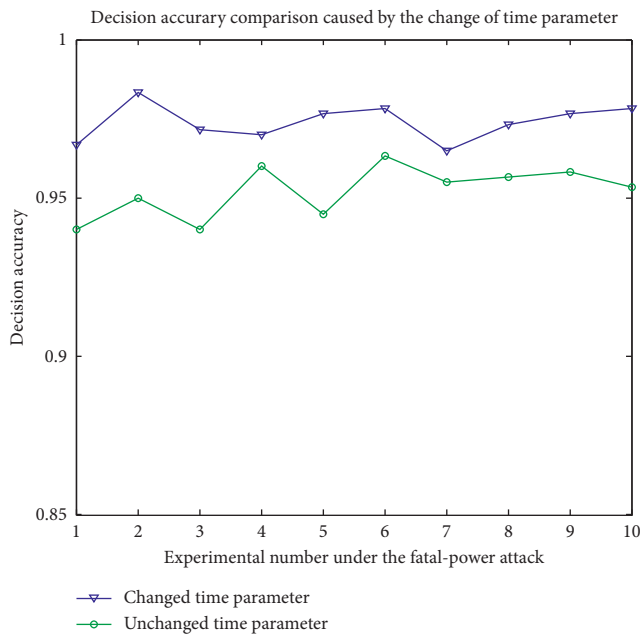


Figure 7: Decision accuracy comparison of SVM decision engine optimized by PSO under the change of time parameter.

6. Conclusions

Focusing on the continuous control mode in industrial processes, this paper proposes a control tracing feature algorithm to extract function-driven tracing characteristics from Modbus TCP data traffic. Moreover, this algorithm not only takes into account the time factor caused by the time intervals between every two consecutive functional operations but also associates with the critical characteristics of sequential control predefined by the technological process. In order to effectively cooperate with this algorithm, this paper introduces the classic SVM classifier as one practicable decision engine, and three intelligent optimization algorithms are compared to optimize its main parameters. Finally, the experimental evaluation shows that on the one hand, this paper designs one applicable attack mode to analyze SVM's decision accuracy, and the experimental results prove that the proposed feature algorithm can adequately reflect the changing pattern of continuous functional operations, and the PSO optimization algorithm can offer the most brilliant contribution to SVM's decision ability; on the other hand, this paper gives the quantitative discussion on the impacts of decision accuracy caused by the length of control tracing sequence and the time parameter, which can play a strong part in improving the decision ability of SVM decision engine.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

This work was supported by the National Key R & D Plan, Research and Development of Intelligent Factory Industrial Internet System Transmission Performance Optimization Technology and Verification Platform, Ministry of Science and Technology of the People's Republic of China (Grant no. 2018YFB1700103), and the Scientific Research Project of Educational Department of Liaoning Province (Grant no. LJKZ0082).

References

[1] P. R. C. Araujo, R. H. Filho, J. J. C. Rodrigues, J. P. C. M. Oliveira, and S. A. Braga, "Middleware for integration of legacy electrical equipment into smart grid infrastructure using wireless sensor networks," *International Journal of Communication Systems*, vol. 31, no. 1, pp. 1–15, 2018.

[2] S. K. Datta, C. Bonnet, and N. Nikaein, "An IoT gateway centric architecture to provide novel M2M services," in *Proceedings of the 2014 IEEE World Forum on Internet of Things (WF-IoT)*, pp. 514–519, Seoul, Republic of Korea, March 2014.

[3] S. Dodla, L. Mahendra, K. Jaganmohan, R. K. Senthil Kumar, and B. S. Bindhumadhava, "Wireless real-time meter data acquisition system," in *Proceedings of the 2019 IEEE Region 10 Conference (TENCON)*, pp. 997–1002, Kochi, India, December 2019.

[4] A. Volkova, M. Niedermeier, R. Basmadjian, and H. de Meer, "Security challenges in control network protocols: a survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 619–639, 2019.

[5] Y. Si, N. Korada, R. Ayyanar, and Q. Lei, "A high performance communication architecture for a smart micro-grid testbed using customized edge intelligent devices (EIDs) with SPI and Modbus TCP/IP communication protocols," *IEEE Open Journal of Power Electronics*, vol. 2, pp. 2–17, 2021.

[6] M. Cheminod, L. Durante, L. Seno, and A. Valenzano, "Performance evaluation and modeling of an industrial application-layer firewall," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 2159–2170, 2018.

[7] O. Liu, B. Zheng, W. Sun et al., "A data-driven approach for reverse engineering electric power protocols," *Journal of Signal Processing Systems*, vol. 93, no. 7, pp. 769–777, 2021.

[8] M. Faisal, A. A. Cardenas, and A. Wool, "Modeling modbus TCP for intrusion detection," in *Proceedings of the 2016 IEEE Conference on Communications and Network Security (CNS)*, pp. 1–5, Philadelphia, PA, USA, October 2016.

[9] Q. Li, Y. Liu, S. Meng, H. Zhang, H. Shen, and H. Long, "A dynamic taint tracking optimized fuzz testing method based on multi-modal sensor data fusion," *EURASIP Journal on Wireless Communications and Networking*, vol. 202021 pages, 2020.

[10] C. Alcaraz, G. Bernieri, F. Pascucci, J. Lopez, and R. Setola, "Covert channels-based stealth attacks in industry 4.0," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3980–3988, 2019.

[11] J. Luswata, P. Zavarsky, B. Swar, and D. Zvabva, "Analysis of SCADA security using penetration testing: a case study on modbus TCP protocol," in *Proceedings of the 2018 29th Biennial Symposium on Communications (BSC)*, pp. 1–5, Toronto, Canada, June 2018.

[12] L. Rosa, M. Freitas, S. Mazo, E. Monteiro, T. Cruz, and P. Simoes, "A comprehensive security analysis of a SCADA

protocol: from OSINT to mitigation," *IEEE Access*, vol. 7, pp. 42156–42168, 2019.

[13] I. Siniosoglou, P. Radoglou-Grammatikis, G. Efstathopoulos, P. Fouliras, and P. Sarigiannidis, "A unified deep learning anomaly detection and classification approach for smart grid environments," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, pp. 1137–1151, 2021.

[14] M. Wan, J. Li, J. Yao, R. Wang, and H. Luo, "State-based control feature extraction for effective anomaly detection in process industries," *Computers, Materials & Continua*, vol. 63, no. 3, pp. 1415–1431, 2020.

[15] X. Zhou, Y. Hu, W. Liang, J. Ma, and Q. Jin, "Variational LSTM enhanced anomaly detection for industrial big data," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3469–3477, 2021.

[16] I. Ahmed, A. Dagnino, and Y. Ding, "Unsupervised anomaly detection based on minimum spanning tree approximated distance measures and its application to hydropower turbines," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 2, pp. 654–667, 2019.

[17] A. Kavousi-Fard, W. Su, and T. Jin, "A machine-learning-based cyber attack detection model for wireless sensor networks in microgrids," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 650–658, 2021.

[18] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1644–1652, 2017.

[19] W. Yan, L. K. Mestha, and M. Abbaszadeh, "Attack detection for securing cyber physical systems," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8471–8481, 2019.

[20] B. Galloway and G. P. Hancke, "Introduction to industrial control networks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 860–880, 2013.

[21] N. Goldenberg and A. Wool, "Accurate modeling of modbus/TCP for intrusion detection in SCADA systems," *International Journal of Critical Infrastructure Protection*, vol. 6, no. 2, pp. 63–75, 2013.

[22] W. Li, Z. Su, R. Li, K. Zhang, and Q. Xu, "Abnormal crowd traffic detection for crowdsourced indoor positioning in heterogeneous communications networks," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2494–2505, 2020.

[23] B. Chikhaoui, S. Wang, T. Xiong, and H. Pigot, "Pattern-based causal relationships discovery from event sequences for modeling behavioral user profile in ubiquitous environments," *Information Sciences*, vol. 285, pp. 204–222, 2014.

[24] Q. Ma, C. Sun, B. Cui, and X. Jin, "A novel model for anomaly detection in network traffic based on kernel support vector machine," *Computers & Security*, vol. 104, no. 2, Article ID 102215, 2021.

[25] J. Alvarez Cid-Fuentes, C. Szabo, and K. Falkner, "Adaptive performance anomaly detection in distributed systems using online SVMs," *IEEE Transactions on Dependable and Secure Computing*, vol. 17, no. 5, pp. 928–941, 2020.

[26] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Transactions on Cybernetics*, pp. 1–13, 2021.

[27] L. Ma, N. Li, Y. Guo et al., "Learning to optimize: reference vector reinforcement learning adaption to constrained many-objective optimization of industrial copper burdening system," *IEEE Transactions on Cybernetics*, pp. 1–14, 2021.

[28] J. Liu, D. Yang, M. Lian, and M. Li, "Research on intrusion detection based on particle swarm optimization in IoT," *IEEE Access*, vol. 9, pp. 38254–38268, 2021.

[29] L. Ma, S. Cheng, and Y. Shi, "Enhancing learning efficiency of brain storm optimization via orthogonal learning design," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 11, pp. 6723–6742, 2021.

[30] A. Assiri, "Anomaly classification using genetic algorithm-based random forest model for network attack detection," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 767–778, 2021.

[31] P. Zhang, S. Shu, and M. Zhou, "An online fault detection model and strategies based on SVM-grid in clouds," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 2, pp. 445–456, 2018.

[32] M. Wan, W. Shang, and P. Zeng, "Double behavior characteristics for one-class classification anomaly detection in networked control systems," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 12, pp. 3011–3023, 2017.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 1048186, 10 pages
https://doi.org/10.1155/2021/1048186

Research Article

A Novel Model for Large-Scale Online College Learning in Postpandemic Era: AI-Driven Approach

Cong Wang

Jinzhou Medical University, Jinzhou 121000, China

Correspondence should be addressed to Cong Wang; wangcong410@jzmu.edu.cn

Received 25 October 2021; Revised 8 November 2021; Accepted 1 December 2021; Published 14 December 2021

Academic Editor: Jianhui Lv

COVID-19 is a pandemic with a wide reach and explosive magnitude, and the world has been bracing itself for impact. Many have lost their jobs and savings, and many are homeless. For better or worse, COVID-19 has permanently changed our lives. For college students, the pandemic means giving up most of the on-campus experience in the postpandemic era and performing online learning instead. Virtual lessons may become a permanent part of college education. Large-scale online learning typically utilizes interactive live video streaming. In this study, we analyzed a codec and video streaming transmission protocol using artificial intelligence. First, we studied an intraframe prediction optimization algorithm for the H.266 codec based on long short-term memory networks. In terms of video streaming transmission protocols, real-time communication optimization based on Quick UDP Internet connections and Luby Transform codes is proposed to improve the quality of interactive live video streaming. Experimental results demonstrate that the proposed strategy outperforms three benchmarks in terms of video streaming quality, video streaming latency, and average throughput.

1. Introduction

In early 2020, COVID-19 spread across the globe. Major events around the world have been postponed or canceled for fear of COVID-19. Given the characteristics of human-to-human transmission of COVID-19, some countries have proposed measures for combating COVID-19 using unmanned technologies [1, 2]. In addition to robotic buses, unmanned delivery vehicles and sweepers have also been installed. Some tasks involving frequent contact with people have been automated. The pandemic has also boosted the development of automation in different industries such as live streaming commerce and online shopping.

COVID-19 has also upended the traditional model of higher education. As campuses across the country sit empty in the midst of the pandemic, administrators are scrambling to prepare for what comes next. In many ways, the attack that COVID-19 represents on higher education is a straightforward assault on traditional views. This sudden pandemic has thrown the original teaching schedules of colleges into turmoil. With all education now being virtual, online learning has become an opportunity to supply higher quality education. In the early pandemic, nearly all learning moved from the classroom to the Internet, which has accelerated the trend toward online learning and raised fundamental questions regarding online learning. What will be the lasting legacy of COVID-19 in higher education? Will the classroom that we once knew gradually return or could COVID-19 permanently transform how we learn?

In recent times, online learning models have been iterated and upgraded with the help of information technology [3–5]. There are three stages of online learning: video-on-demand- (VOD-) based online learning, live-streaming-based online learning, and real-time interactive online learning. In the late 1990s, online schools were distance education platforms with networks as a medium and videos were viewed through the Internet. This type of platform mainly relies on different forms of recording to record videos of lectures and upload them to the Internet. Students then follow what is happening in the lectures. The wave of online learning began in 2011 with live streaming. Teachers can interact with students to a certain degree

through live streaming and they can also answer student questions online. Live streaming online education has some advantages for online learning compared to the traditional PC era, which reduces the cost of learning to a significant extent and makes learning more fruitful. Live-streaming-based online learning is less engaging than real classes and it is easier to become distracted. However, learning is a real-time and two-way communication process and neither VOD nor live streaming learning can fully meet these two needs. Representing real-time communication scenarios between teachers and students is the biggest hurdle in online learning [6].

Interactive video live streaming is popular for online learning based on its interactivity and involvement, especially for large-scale online college learning in the post-pandemic era. However, given the limitations of data transmission and audio and video (AV) technology, it is difficult to avoid latency in interactive video live streaming, which negatively affects user experiences [7]. Artificial intelligence (AI) has significant promise for Internet applications and can serve as a positive force for the advancement of human society. By combining video live streaming with AI, it can be optimized in terms of coding and decoding (codec) technology, video transmission protocols, and other aspects to improve transmission efficiency and reduce latency to improve the performance of interactive video live streaming for large-scale online college learning.

Accordingly, the main contributions of this study can be summarized as follows. (i) An intraframe prediction optimization algorithm for the H.266 codec based on long short-term memory networks (LSTMs) is proposed. (ii) A real-time communication optimization based on Quick UDP Internet connections (QUIC) and the Luby Transform (LT) code video transmission protocol is proposed to improve the quality of interactive live video streaming.

The remainder of this paper is organized as follows. Section 2 reviews related work. In Section 3, the intraframe prediction optimization algorithm for the H.266 codec based on LSTMs is presented. In Section 4, real-time communication optimization based on the QUIC and LT code video transmission protocol is proposed. Experimental results are presented in Section 5. Section 6 concludes this paper.

2. Related Work

In recent years, many strategies for online learning have been proposed. In [8], the authors presented an enhanced recommendation method called adaptive recommendation based on an online learning style, which implements learning resource adaptation by mining learner behavioral data. In [9], the authors conducted in-depth exploration from the perspective of student self-efficacy by extending the four dimensions of online learning: a sense of effort, sense of control, sense of participation, and sense of environment. In [10], the authors attempted to provide an effective online teaching method and investigated the effects of online competency-based learning and design-based learning on enhancing student learning performance, self-directed learning readiness, and experience of online learning in an online computing course. In [11], the authors examined the impact of personality traits, learning styles, gender, and online course factors (course difficulty, group affiliation, provided materials, etc.) on the academic success of students taking online courses and their overall success rate compared to traditional classes.

For large amounts of video data and high visual quality requirements, video codec technology has been rapidly developed and has become much more mature. In [12], a high-efficiency video codec-based spatial resolution scaling type for mixed resolution coding for frame interleaved multiview videos was proposed. In [13], the authors proposed a novel video transmission strategy that effectively transfers the computational complexity of video coding from the terminal to the cloud environment. In [14], a method for creating representative video sets covering all segments of user videos was proposed. In [15], novel compression tools for inter-/intraprediction, in-loop filtering, and entropy coding were proposed. In [16], the authors presented a model to distinguish regions of interest based on the VP9 video codec. Their model contained high motion-, contrast-, and color-sensitive areas that fit the human visual system. In [17], three new approaches for generating spatial intraprediction signals were supported: a line-wise application for conventional intraprediction modes coupled with a mode-dependent processing order, region-based template matching prediction method, and intraprediction modes based on neural networks. In [18], diamond adaptive rod pattern-search-algorithm-based block-matching motion estimation algorithms were proposed for multistranded codec hardware design to provide a high compression rate with less computational complexity. In [19], the authors proposed a deep-learning-based systematic approach that included an effective convolutional neural network structure, hierarchical training strategy, and video-codec-oriented switchable mechanism.

Several improved QUIC protocols for video transmission have also been investigated. For example, in [20], the authors proposed a novel design for the socket secure (SOCKS) protocol over QUIC (QSOCKS), which improved the browsing experience while enhancing reliability. In [21], the authors presented a security analysis of the QUIC handshake protocol based on symbolic model checking. In [22], the authors proposed a lightweight latency reduction scheme for the QUIC protocol. The proposed scheme calculated the average congestion window, which was utilized as the initial congestion window when a new connection was established. In [23], the authors proposed a modification of the handshaking mechanism for the QUIC protocol to minimize the overhead incurred by control signals and time required to update the congestion window size. In [24], the authors proposed QUIC-EST, which is a transmission scheme that combines the congestion control and multistream features of the recently proposed QUIC transport protocol with an optimized scheduling algorithm to maximize the value of information at the receiver. In [25], the authors proposed QUIC-based secure transport protocol to improve the transport performance of network traffic.

Additionally, many strategies for LT codes have been proposed. In [26], the authors proposed a second coding scheme based on LT codes under inactivation decoding. In [27], the authors proposed a modified version of LT encoding for the delivery phase to take advantage of channel coding. In [28], an improved LT code with a reverse coding framework was designed to reduce the error floor caused by low-degree information nodes.

3. Intraframe Prediction Optimization Algorithm for the H.266 Codec Based on LSTMs

In large-scale online college learning, the main traffic of interactive video live streaming is the same as that of traditional video live streaming. The signal source successively completes the information collection, encoding, and push streaming of streaming media while the terminals complete pull streaming, decoding, and playing.

Versatile video coding, which is also known as H.266, is a relatively new video encoding standard. Codecs are software packages that compress and decompress video files so they can use less space, which allows them to use fewer resources. H.266 can shrink files by up to 50% more than the current H.265 standard, which means that higher-definition streaming media information can be transmitted at a lower bandwidth. Therefore, the upgrading of codec technology can significantly reduce the cost of video storage and transmission, and decreased data transmission can reduce latency and improve overall stability. Therefore, students can receive better video quality under the same network conditions.

Intraframe prediction coding is the core of a video encoder. From H.265 to H.266, the main method for improving intraframe prediction performance is to add an elaborate angle direction scheme. However, there are some limitations to this method of relying solely on increasing the prediction direction. Typically, the brightness and chromaticity of two adjacent pixels are often relatively close in an image, meaning color tends to change gradually. The goal of video coding is to use correlations to compress an image. The larger the prediction error, the more the bits required for coding and the less efficient the video compression.

LSTMs are designed for applications in which the input is an ordered sequence. LSTMs are a type of recurrent neural network (RNN), which are networks that reuse the output from a previous step as the input for the next step. Similar to all neural networks, nodes perform calculations on inputs and return output values. In a recurrent network, this output is then used in combination with the next element as the input for the next step. In an LSTM, nodes are recurrent, but they also have an internal state. Each node uses its internal state as a working memory space, meaning that information can be stored and retrieved. The input value from the previous output and the internal state are both used in node calculations. The results of calculations are used not only to provide output values, but also to update the state. Like any neural network, LSTM nodes have parameters that determine how inputs are used in calculations and they are also known as gates that control the flow of information within nodes.

LTSMs solve the problems of vanishing gradients and gradient explosions caused by RNN networks during training, so they can make better use of the spatial correlations between adjacent pixels to express brightness changes in the prediction direction. Based on our analysis of the variation rules of pixel brightness in the prediction direction, this paper proposes an angle prediction and compensation algorithm based on LSTMs that makes a secondary prediction for prediction errors in the horizontal and vertical angle models of H.266 and aids the standard linear prediction model to improve prediction accuracy.

3.1. Horizontal and Vertical Mode of H.266. The typical equation for H.266 angle prediction is defined as follows:

$$\mathrm{pre}[x, y] = \frac{\{(32 - \omega LI) * \mathrm{ref}[x + iPM + 1] + \omega LI * \mathrm{ref}[x + iPM + 2] + 16\}}{32}, \tag{1}$$

where ωLI is the linear interpolation weight, iPM is the index of the prediction model, $\mathrm{pre}[x, y]$ is the predicted value of a pixel, and $\mathrm{ref}[x]$ is the one-dimensional reference pixel set projected according to the prediction direction.

The horizontal and vertical modes of the angle prediction have typical characteristics. During the prediction process, the predicted values of each row or column are equal to the values of the same reference pixels. When the prediction mode is vertical, the predicted value of the pixel is equal to the value of the reference pixel above the current pixel, which can be defined as follows:

$$\mathrm{pre}[x, y] = \mathrm{ref}[x + 1]. \tag{2}$$

When the prediction mode is horizontal, the predicted value of the pixel is equal to the value of the reference pixel to the left of the current pixel and x on the right side of (2) can be directly replaced with y.

We assume that the prediction mode of the current pixel block is horizontal and the width and height of the pixel block are w and h, respectively. The x-th reference pixel above is denoted by $abo[x]$, the y-th reference pixel on the left is denoted by $\mathrm{lef}[y]$, the pixel value of the x-th column in the y-th row of the current pixel block is denoted by $p[x, y]$, and the predicted value $\mathrm{pre}[x, y] = \mathrm{lef}[y]$. The prediction errors $\mathrm{pe}[x, y]$ can be calculated as follows:

$$pe[x, y] = p[x, y] - \text{pre}[x, y] = p[x, y] - \text{lef}[y]. \quad (3)$$

Given the definitions above, in the horizontal and vertical modes of the H.266 intraframe prediction scheme, the prediction results are directly obtained from the left or above reference pixels. In this method, for coding blocks with complex internal details, the prediction errors are too large and the coding efficiency is reduced. To solve this problem, this study leveraged the ability of LSTMs of expressing sequence data and making a secondary prediction of prediction errors to compensate for the deficiencies of the standard linear prediction model. The core concept of the proposed algorithm is that the prediction error data of the horizontal and vertical modes in the output results of the standard encoder are counted; then the error data are classified according to the size of the coding block, and then the LSTM model is trained. The LSTM model is used to calculate the prediction compensation values in real time and compensate for the original prediction data to improve prediction accuracy.

3.2. Network Architecture. The number of neurons in the input layer of the network is twice the predicted block width (i.e., $2 \times w$) and the number of neurons in the output layer is equal to the predicted block width (i.e., w). The network depth is K and the time step is the height of the coding block (i.e., h).

In the network architecture of this algorithm, the construction method for the input data is one of the keys to the algorithm. Here, the construction method for the input data is introduced by considering the horizontal prediction mode as an example. In the horizontal prediction mode, the prediction error of each row is the difference between the pixel values of each pixel of the row and the left reference pixel. The reference row above the prediction block is also calculated to obtain $h + 1$ rows of prediction errors.

According to the spatial correlation of images and how the horizontal prediction model works, the prediction errors of adjacent rows are also correlated. To improve the prediction effect, for each row to be predicted, the proposed method selects the prediction errors (pe_{n-1} and pe_{n-2}) of two rows above the current row as the inputs for the network. The vertical prediction mode is treated similarly to the horizontal prediction mode and is not described here.

According to the characteristics of error data, the tanh function was selected as the activation function for the input layer in this study and the prediction results were obtained through the hidden layer and output layer. The tanh function is also known as the hyperbolic tangent activation function. The mean squared error (MSE) function was used as the loss function during network training as follows:

$$\text{MSE} = \frac{\sum_{i=0}^{w-1} (pre[i] - y[i])^2}{w}. \quad (4)$$

4. Real-Time Communication Optimization Based on QUIC and the LT Codes Video Transmission Protocol

The current mainstream Internet video transmission protocols are the real-time messaging protocol and hypertext transfer protocol, both of which are based on the transmission control protocol (TCP). The TCP is designed to be complex to provide a reliable transmission service, which increases the latency of data packets and cannot meet the real-time requirements of interactive live video streaming. In contrast, the user datagram protocol (UDP) has a high transmission efficiency and low latency. However, as a result of its unreliability, even if it is directly applied to interactive live videos on the Internet, it cannot meet the requirements for the smooth transmission of streaming media information. Currently, there are many optimization schemes for UDP, including the reliable UDP model and QUIC developed by Google. QUIC is a UDP-based low latency Internet transmission protocol. According to research results, the lag rate of QUIC pull streaming is 57% lower than that of TCP pull streaming. Although the QUIC protocol has the characteristics of zero round-trip time (RTT), low latency, connection migration, and security, it can still be optimized and improved according to the requirements of real-time communication scenarios. We attempted to improve the QUIC protocol for real-time communication. Our improvements include improving the corresponding format of the transmission frame and partial reliability marker, which makes the QUIC more appropriate for a communication scenario that allows moderate packet loss. The improved QUIC protocol is called rt-QUIC, where "rt" stands for "real-time."

4.1. Designing an Improved Frame Format. The packet format of rt-QUIC must transmit a parameter with a maximum frame size of δ and this parameter is used before data stream exchanges at multiple communication ends. In addition to informing each other δ, the communication ends also need to carry a series of parameters, including the frame length, load, and data type that the communication end can receive. When rt-QUIC is used to send "unreliable" packets between multiple communication ends, it generates a new frame, which is sent immediately and may be merged with other frames during transmission. rt-QUIC will do its best to deliver these unreliable merged frames. When the communication end determines that the transmitted data at one end may be lost, it directly informs the end that a frame has been lost. The rt-QUIC packet will incur acknowledgment (ACK) frame latency when a packet is lost and it will not retransmit the packet after the frame is lost but will rearrange the serial number of the packets. The response receiving time is controlled by the response time of the ACK frames and the lost data frames can be received and acknowledged at a later time. To prevent the loss of the tail of the data stream, a reliability mark is added to the end of the stream. When a data stream that the communication end attempts to use has been lost, it will receive a stream of size

zero, which is a supplementary stream to fill in the lost data. Finally, the last byte in each stream is reliably transmitted and the data stream can be delivered as completely as possible.

4.2. Partial Reliability Marker. The design of a partial reliability marker is suitable for AV transmission scenarios. Each frame is marked and only one bit is used to indicate whether it is reliable. Adding an identifier bit may slightly increase the cost of sending and receiving, but this is largely inconsequential. This is mainly because AV streaming media files are often accompanied by codecs in the transmission process. As shown in Figure 1, in the codec process of AV transmission, the AV stream of the sender is first coded and then the coded data are transmitted. In the coding process, the partial reliability bit will not increase the cost of the sender. For the receiver, the file encoded by the sender is transmitted through a peer-to-peer or server connection and the power of the receiving end must be used for decoding. The decoding process rarely causes an increase in processing time. Because the receiving end must judge the reliability marker in a packet and process data streams in different scenarios, overall, the communication service experience is improved.

4.3. LT Codes. In addition to allowing moderate packet loss in the network layer, fountain code technology can be used to realize data integrity under the condition of a small amount of packet loss. Fountain codes are approximations of an ideal digital fountain. A digital fountain allows a client to obtain droplets or encoded message packets from a server and use them to rebuild an encoded file. The code makes it so that the actual packets are received and the order in which they are received does not matter. An ideal digital fountain can generate an infinite supply of droplets from the original data and a receiver can reconstruct an original data file made of N packets once N droplets have been received. Digital fountains are similar to water fountains in that we do not care which drops we get from the fountain as long as we get sufficient drops to fill our water bottle. In 1998, Luby invented the LT code. LT codes rely on an exclusive OR operation to encode and decode a message. LT codes are called rate lists because their encoding algorithm can produce an infinite supply of message droplets in principle. Fountain codes are useful in scenarios where packet loss is likely and in scenarios where a receiver cannot communicate with a sender.

The interactive video live streaming of online college learning considered in this study is based on LT codes combined with rt-QUIC (rt-QUIC-LT), with a sender and receiver. The sender can be a client or server. Likewise, the receiver can also be a client or server. In this study, we assumed that the client is the sender and the server is the receiver. First, the sender loads a file and divides it into data blocks of the same size. The sender then divides each data block into k source characters for LT coding. Finally, the encoded characters and encoded information are packaged and sent to the receiver. After receiving sufficient encoded packets, the receiver unpacks the encoded characters and encoded information, and then LT decoding is performed. After decoding, the transmitted video stream can be recovered.

4.4. Implementation of the Receiver. Online college learning realizes interactive video live streaming based on the rt-QUIC socket, which is based on the connectionless UDP. Therefore, interactive video live streaming between a sender and receiver does not establish a direct connection. The rt-QUIC video transmission scheme has high real-time performance and a fast transmission rate without data confirmation and retransmission mechanisms. Interactive video live streaming for online college learning uses a connectionless socket on the receiver side to communicate with the sender through the following six steps:

Step 1: The receiver first creates and instantiates a socket based on the port number and IP address. The port number was set to 8735 in this study.

Step 2: Each frame of video streaming is added with a partial reliability marker.

Step 3: The Bind() function is called to bind the local address and port.

Step 4: The receiver initializes a receiving thread and waits for data to be received.

Step 5: Interactive video live streaming uses the recvfrom() function to receive data from the sender.

Step 6: The socket is closed and all resources are released.

According to these steps, a flow chart for creating a socket on the receiver side is presented in Figure 2.

4.5. Implementation of the Sender. First, a receiver creates a socket and binds the address to the port of the sender and then creates a receiving thread and waits for the sender to send data. In the meantime, the sender creates a socket and a sending thread, which begins sending data to the specified address until all data have been sent and no connection is established with the receiver. When the sender begins to send data, the receiver begins to receive data and obtains information regarding the transferred video stream from the received data until all data are received.

In interactive video live streaming for online college learning through the sockets of rt-QUIC, LT codes are used to ensure the reliability of video streaming. First, LT encodes a video stream to generate sufficient encoded characters at the sender. The encoded characters and video stream are then packaged and sent to the receiver through sockets. After receiving the data, the receiver unpacks the encoded characters and recovers the video stream using the LT decoder. LT codes can ensure the reliability of video streaming transfer, which is determined by the LT code characteristics. The nonbit-rate characteristic of LT codes enables a sender to generate an infinite number of encoded characters and a receiver to recover a video stream after receiving sufficient

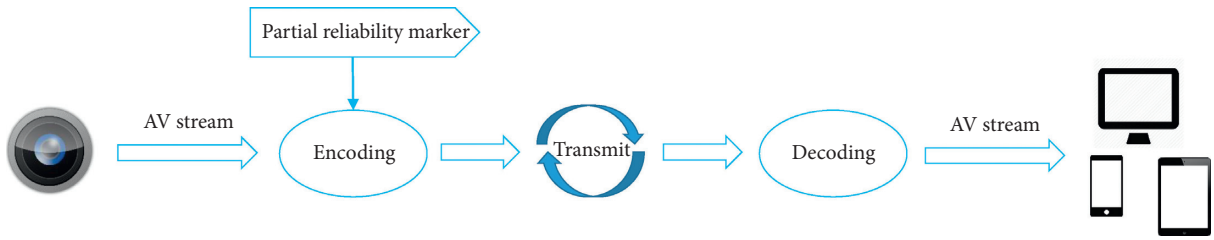


Figure 1: AV streaming codec process.

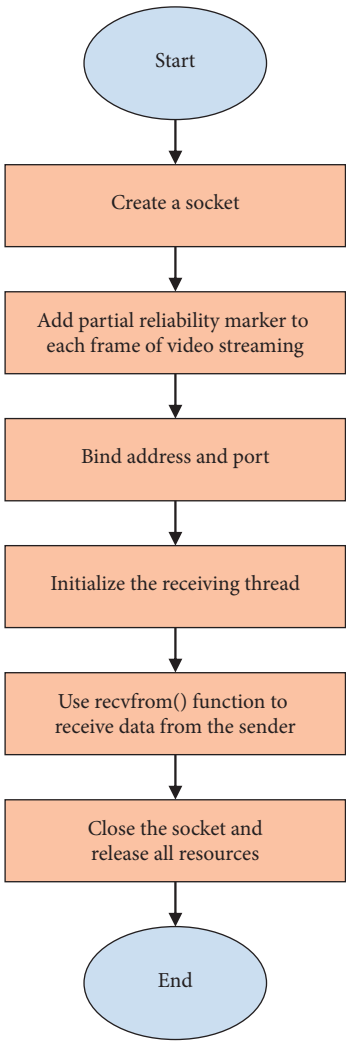


Figure 2: Socket of receiver.

encoded characters. Data lost during video streaming transfer and the sequence of received data do not affect video stream recovery.

5. Experiments and Results Analysis

5.1. Setup. In this study, the encoding, transmission protocol, and transmission efficiency of traffic in interactive video live streaming for online college learning were optimized. Simulations were executed on a computer with an Intel i9-11900KF 3.7 GHz CPU, 32 GB of RAM (3333 MHz), and NVIDIA RTX 3080 Ti, 12GB GDDR6X GPU. To verify the efficiency of the intraprediction optimization algorithm proposed in this paper, the test platform VTM-7.3 for H.266 was used as a test environment. The encoded video stream quality was measured using the peak signal-to-noise ratio (PSNR) and Bjontegaard delta bit rate (BDBR) [29]. Additionally, we compare the proposed rt-QUIC and LT code video transmission to three benchmarks: QUIC, QUIC-EST, and QSOCKS.

5.2. Parameter Settings

5.2.1. Depth of the Network. The size and depth of the network have a significant impact on network performance. Network performance can be effectively improved by increasing network depth. However, not all types of networks can achieve higher performance by using deeper layers. Based on the intraframe prediction data of live video streaming coding, this study compared network performance with different sizes and depths, and LSTM loss function curves for the same size and different depths were obtained. As shown in Figure 3, the model with six hidden layers can achieve the lowest training cost with an increasing number of training epochs. Therefore, this study adopted LSTMs with six hidden layers for training.

5.2.2. Size of Hidden Layers. The size of the input layer of the network was fixed at $2 \times w$ and the size of the output layer was fixed at w. The size of the hidden layer could be adjusted. In [30], a clustering strategy was used to divide a population into several clusters. In this study, network performance with different hidden layer sizes was calculated, as shown in Figure 4. One can see that the network performance improves with a larger size of hidden layer, but this change becomes less obvious when the size is above 48. Therefore, the size of the hidden layer selected in this study was 48.

5.3. Comparative Analysis. Simulations were performed to examine video streaming quality, video streaming latency, average throughput, and online learning quality of experience (QoE). The number of simulations was 2000.

5.3.1. Video Streaming Quality. PSNR is directly proportional to video streaming quality, but it is often impossible to consider both a low bitrate and a high recovery quality simultaneously when measuring the quality of video encoded by the proposed algorithm. Therefore, the change in bitrate

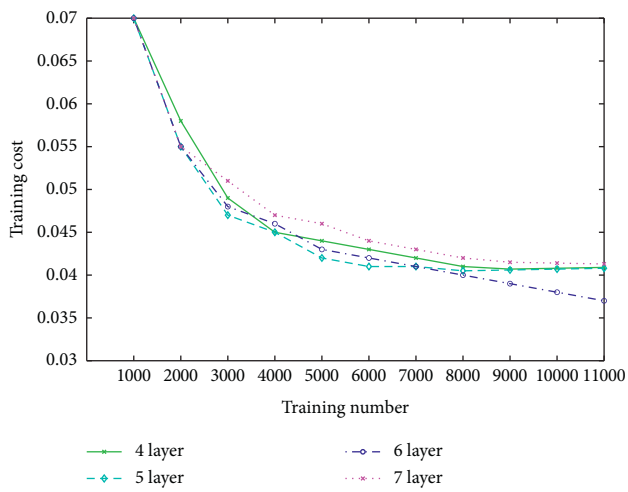


Figure 3: Cost curves of networks with the same size and different depths.

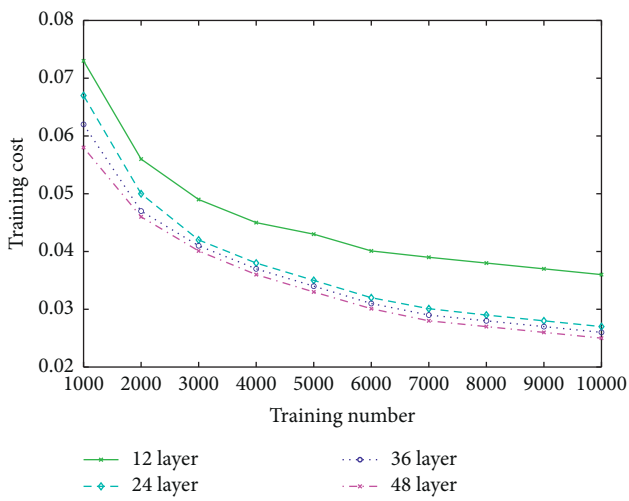


Figure 4: Cost curves of networks with different hidden layer sizes.

after coding was evaluated on the premise of restoring the same video image quality. The PSNR and BDBR of the proposed algorithm and benchmarks with different sequences are listed in Table 1.

In Table 1, one can see that the PSNR of the proposed algorithm is smaller than those of the two benchmarks for different sequences, indicating that the proposed algorithm performs well in terms of video streaming quality. In terms of BDBR, the average bitrate change is less than 1%, meaning that it takes less time to restore the same video streaming quality. For the horizontal and vertical prediction angle models, the LSTM is used to perform secondary prediction, which compensates for the original intraframe prediction results and effectively reduces prediction errors.

5.3.2. Video Streaming Latency. The latency test results under the conditions of network transmission are presented in Table 2. The simulations were driven by video streaming data from the H.266 codec based on LSTMs with resolutions of 240p, 360p, 480p, 720p, 1080p, 1440p, and 2160p.

As shown in Table 2, the transmission latency of the proposed algorithm is lower than those of the two benchmarks before and after decoding. LSTMs solve problems that RNNs cannot handle in terms of long-term dependencies, which is important to this study's large-scale online learning scenario. The hidden layer in a typical RNN has only one state, which is sensitive to short-term inputs. LSTMs add an additional state to hold long-term information. The tanh activation function is used to help regulate video streams flowing through the network.

5.3.3. Average Throughput. The proposed rt-QUIC-LT was compared to three benchmarks. The experimental variables were the RTT and bandwidth. After several rounds of testing, the throughput of the results in each round was averaged as a throughput reference value. In Figure 5, one can see that there are no significant differences between the performances of the four protocols when the RTT is small and the throughput can basically remain near the peak value. However, the throughputs of the three benchmarks decrease with an increase in the RTT. Therefore, it can be inferred that the processing capacity of the three benchmarks is not as high as that of rt-QUIC-LT when network congestion occurs. Because rt-QUIC-LT improves the frame format for real-time communication scenarios, it can be combined with the selective send strategy to achieve smooth and reliable performance in poor network conditions.

When the variable is bandwidth, the experimental performance results are presented in Figure 6, where the average throughput of the four protocols increases with an increase in bandwidth. However, when comparing the four protocols, one can see that rt-QUIC-LT yields its best performance under a given bandwidth more quickly, whereas the three benchmarks reach the optimal state more slowly. It can be inferred that the improvement of the frame format and packet loss strategy plays an important role in rt-QUIC-LT. The rt-QUIC-LT protocol actively loses some retransmitted and meaningless packets, and the transmission cost is relatively low. Therefore, the average throughput of the rt-QUIC-LT protocol is better than those of the benchmarks under the same bandwidth.

Figure 7 presents performance comparisons of the four protocols in real-time communication scenarios. The preset packet loss rate of the test network environment is 1% and the preset latency is 50 ms. According to the simulation results, QUIC-EST and QSOCKS are the first to reach the maximum throughput in the initial stage of connection establishment, but they do not maintain the maximum throughput steadily. In contrast, the average throughput of rt-QUIC-LT shows a steady increase with less fluctuation. Overtime, QUIC-EST and QSOCKS exhibit several periods of lag, leading to low performance in terms of effective throughput. Although the rt-QUIC-LT protocol gives up partial reliability, its stability is enhanced and transmission performance can be maintained at a higher level. LT codes

Table 1: The PSNR and BDBR of the proposed algorithm and benchmarks with different sequences.

Test sequence	fCU-BIDBF		fISP-CU		This paper	
	BDBR (%)	PSNR (dB)	BDBR (%)	PSNR (dB)	BDBR (%)	PSNR (dB)
Cactus	1.48	−0.06	2.56	−0.11	0.66	−0.02
Foodmarket	1.21	−0.03	2.71	−0.08	1.11	−0.04
Catrobot	2.08	−0.13	2.65	−0.15	0.66	−0.08
Arena of Valor	2.46	−0.12	2.38	−0.13	0.32	−0.04
Slide Editing	1.04	−0.08	2.59	−0.09	0.25	−0.05
BQ Mall	1.54	−0.11	1.37	−0.16	0.44	−0.04
Fourpeople	1.99	−0.13	3.69	−0.21	1.17	−0.03
Johnny	1.37	−0.09	3.72	−0.16	1.33	−0.02
Average	1.65	−0.09	2.71	−0.14	0.74	−0.04

Table 2: The comparison of network transmission latency under different resolutions.

Resolution	fCU-BIDBF			fISP-CU			This paper		
	Before encoding (ms)	After decoding (ms)	Time difference (ms)	Before encoding (ms)	After decoding (ms)	Time difference (ms)	Before encoding (ms)	After decoding (ms)	Time difference (ms)
240p	11.05	11.36	0.31	10.54	10.79	0.25	9.81	9.89	0.08
360p	15.17	18.99	3.82	13.38	15.89	2.51	12.25	13.01	0.76
480p	20.96	31.49	10.53	18.94	23.57	4.63	15.67	16.78	1.11
720p	30.45	45.68	15.23	24.56	30.02	5.46	20.08	22.48	2.4
1080p	56.37	75.76	19.39	50.19	58.46	8.27	44.52	47.08	2.56
1440p	180.31	204.34	24.03	120.48	131.64	11.16	100.69	103.96	3.27
2160p	311.72	343.66	31.94	260.49	281.33	20.84	210.77	214.12	3.35
Average	89.43	104.47	15.04	71.23	78.81	7.59	59.11	61.05	1.93

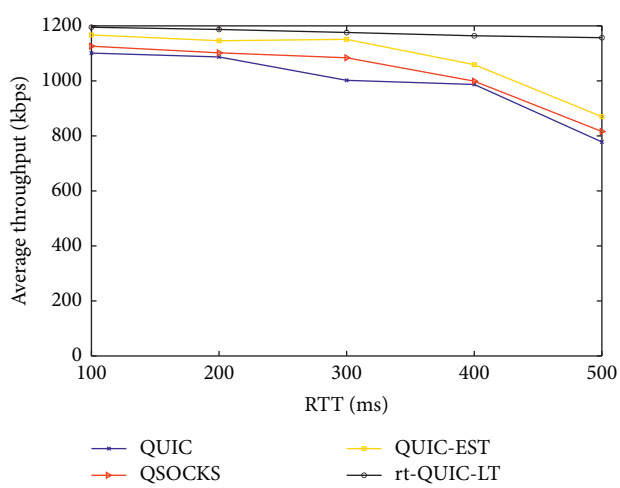


Figure 5: Average throughput of different RTT.

inherit the advantages of fountain code while significantly reducing the complexity of coding and decoding.

The performances of the four protocols in a poor network scenario are now compared. Figure 8 presents a comparison of the results when the network packet loss rate is 10% and the network latency is 250 ms in the test environment. Based on this poor simulated network environment, the results reveal that QUIC-EST and QSOCKS suffer from multiple transmission lag events. Because they are affected by packet loss, their effective throughput fluctuates significantly, resulting in a poor communication experience. In contrast, rt-QUIC-LT has an advantage in this scenario. rt-QUIC-LT takes advantage of its incomplete reliability and drops packets that do not need to be retransmitted to reduce fluctuations in network quality. The improved rt-QUIC-LT can adapt to fluctuations in the network state. Therefore, in the case of poor network conditions, although lag still occurs, the effective throughput does not decrease significantly and the entire communication experience is acceptable. Finally, it can be concluded that rt-QUIC-LT is suitable for real-time AV communication and video streaming transmission scenarios.

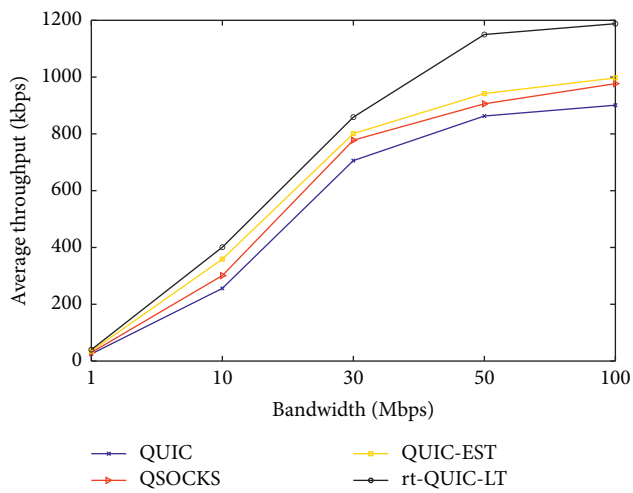


Figure 6: Average throughput of different bandwidths.

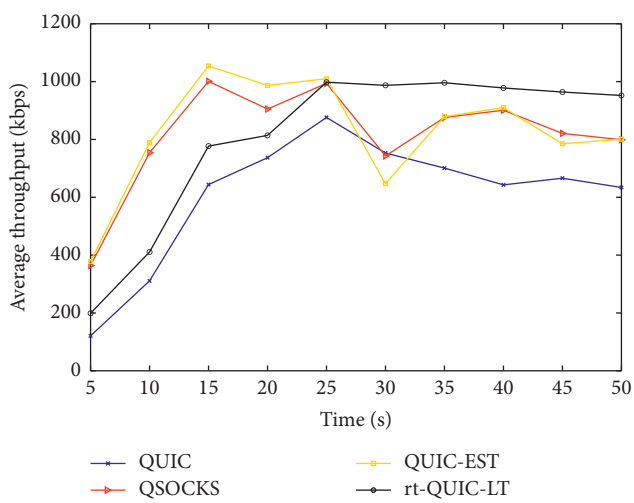


Figure 7: Average throughput under 1% packet loss and 50 ms latency.

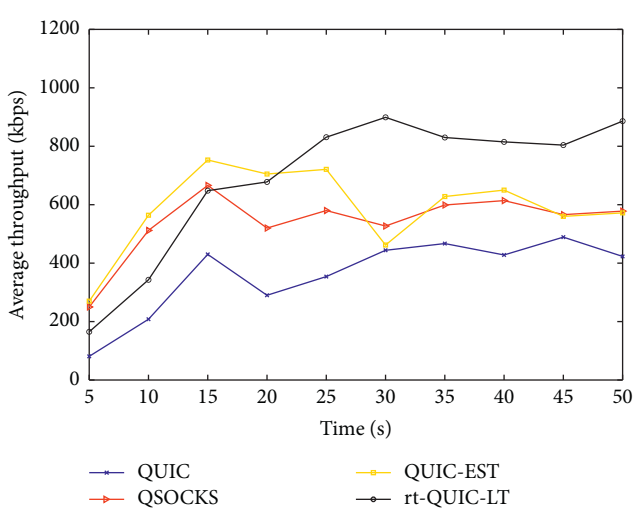


Figure 8: Average throughput under 10% packet loss and 250 ms latency.

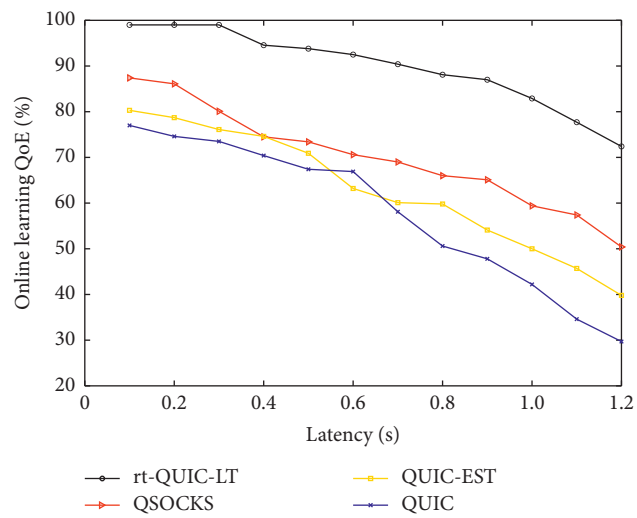


Figure 9: Online learning QoE.

5.3.4. Online Learning QoE. Figure 9 presents the QoE of large-scale online college learning under different network transmission latencies. As shown in Figure 9, rt-QUIC-LT can guarantee better online learning QoE under low latency. However, the online learning QoE decreases with an increase in latency. Even though the online learning QoE of rt-QUIC-LT decreases, it is still higher than those of the three benchmarks. The online learning QoE of the three benchmarks decreases more rapidly at higher latency, which verifies that the real-time communication optimization based on the QUIC and LT code video transmission protocol proposed in this paper can effectively improve the online learning QoE.

6. Conclusions

Interactive video live streaming plays a significant role in large-scale online college learning in the postpandemic era. Regarding video codecs, we proposed an intraframe prediction optimization algorithm for the H.266 codec based on LSTMs. Additionally, real-time communication optimization based on the QUIC and LT code video transmission protocol was proposed to improve the quality of interactive video live streaming. Experimental results demonstrated that the proposed algorithms perform well in terms of video streaming quality, video streaming latency, and average throughput.

The quality of real-time AV transmission is very strict and the realization of online real-time interactive scenes faces four main challenges: high concurrency, high reliability, low latency, and antiweak network characteristics. In addition to low latency, these other three aspects will be studied in the future. We will also enhance the experience of online learning by studying virtual reality and augmented reality.

Data Availability

All the data used to support the findings of the study are included within the article.

Conflicts of Interest

The author declares no conflicts of interest in this paper.

References

[1] O. Ozkan and O. Atli, "Transporting COVID-19 testing specimens by routing unmanned aerial vehicles with range and payload constraints: the case of Istanbul," *Transportation Letters-The International Journal of Transportation Research*, vol. 13, no. 5-6, pp. 4482–4491, 2021.

[2] S. Manigandan, P. K. Thaloor Ramesh, N. T. L. Chi, and K. Brindhadevi, "Early detection of SARS-CoV-2 without human intervention to combat COVID-19 using drone technology," *Aircraft Engineering & Aerospace Technology*, vol. 93, no. 1, pp. 85–88, 2021.

[3] D. Shen, M.-H. Cho, C.-L. Tsai, and R. Marra, "Unpacking online learning experiences: online learning self-efficacy and learning satisfaction," *The Internet and Higher Education*, vol. 19, pp. 10–17, 2013.

[4] W. Holmes, Q. Nguyen, J. Zhang, M. Mavrikis, and B. Rienties, "Learning analytics for learning design in online distance learning," *Distance Education*, vol. 40, no. 3, pp. 309–329, 2019.

[5] S. R. Palmer and D. M. Holt, "Examining student satisfaction with wholly online learning," *Journal of Computer Assisted Learning*, vol. 25, no. 2, pp. 101–113, 2009.

[6] A. W. Cole, L. Lennon, and N. L. Weber, "Student perceptions of online active learning practices and online learning climate predict online course engagement," *Interactive Learning Environments*, vol. 29, no. 5, pp. 866–880, 2021.

[7] T. Zhao, Q. Liu, and C. W. Chen, "QoE in video transmission: a user experience-driven strategy," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 285–302, 2017.

[8] H. Chen, C. Yin, R. Li, W. Rong, Z. Xiong, and B. David, "Enhanced learning resource recommendation based on online learning style model," *Tsinghua Science and Technology*, vol. 25, no. 3, pp. 348–356, 2020.

[9] H. Liu, Y. Ye, and H. Jiang, "Self-efficacy in home-based online learning environments," *Journal of Internet Technology*, vol. 22, no. 3, pp. 557–567, 2021.

[10] C.-W. Tsai, "Applying online competency-based learning and design-based learning to enhance the development of students' skills in using PowerPoint and Word, self-directed learning readiness, and experience of online learning," *Universal Access in the Information Society*, vol. 19, no. 2, pp. 283–294, 2020.

[11] E. Idrizi, S. Filiposka, and V. Trajkovijk, "Analysis of success indicators in online learning," *International Review of Research in Open and Distance Learning*, vol. 22, no. 2, pp. 205–223, 2021.

[12] B. Mallik, A. Sheikh-Akbari, and A.-L. Kor, "Mixed-resolution HEVC based multiview video codec for low bitrate transmission," *Multimedia Tools and Applications*, vol. 78, no. 6, pp. 6701–6720, 2019.

[13] T. C.-W. Lei and F.-S. Tseng, "Light-weight video codec at terminal for video coding in the cloud," *Journal of Signal Processing Systems*, vol. 91, no. 6, pp. 627–639, 2019.

[14] A. V. Zvezdakova, D. L. Kulikov, S. V. Zvezdakov, and D. S. Vatolin, "BSQ-rate: a new approach for video-codec performance comparison and drawbacks of current solutions," *Programming and Computer Software*, vol. 46, no. 3, pp. 183–194, 2020.

[15] K. Choi, J. Chen, M. W. Park et al., "Video codec using flexible block partitioning and advanced prediction, transform and loop filtering technologies," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 5, pp. 1326–1345, 2020.

[16] Z. Zhang, T. Jing, and C. Tian, "Hybrid context modeling and wavelet transform for region of interest detection and rate control based on VP9 video codec," *Concurrency and Computation: Practice and Experience*, vol. 31, no. 10, 2019.

[17] J. Pfaff, H. Schwarz, D. Marpe et al., "Video compression using generalized binary partitioning, trellis coded quantization, perceptually optimized encoding, and advanced prediction and transform coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 5, pp. 1281–1295, 2020.

[18] S. M. Balamurugan and R. Seshasayanan, "An efficient field-programmable gate array-based hardware oriented block motion estimation algorithm based on diamond adaptive rood pattern search algorithm for multi-standard video codec," *Transactions of the Institute of Measurement and Control*, 2021.

[19] D. Ding, L. Kong, and G. Chen, "A switchable deep learning approach for in-loop filtering in video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 7, pp. 1871–1887, 2020.

[20] M. R. Kanagarathinam, S. Singh, S. R. Jayaseelan, M. K. Maheshwari, G. K. Choudhary, and G. Sinha, "QSOCKS: 0-RTT proxification design of SOCKS protocol for QUIC," *IEEE Access*, vol. 8, pp. 145862–145870, 2020.

[21] J. Zhang, L. Yang, X. Gao, G. Tang, J. Zhang, and Q. Wang, "Formal analysis of QUIC handshake protocol using symbolic model checking," *IEEE Access*, vol. 9, pp. 14836–14848, 2021.

[22] J. Jung and D. An, "Access latency reduction in the QUIC protocol based on communication history," *Electronics*, vol. 8, no. 10, 2019.

[23] P. Kharat and M. Kulkarni, "Modified QUIC protocol with congestion control for improved network performance," *IET Communications*, vol. 15, no. 9, pp. 1210–1222, 2021.

[24] F. Chiariotti, A. A. Deshpande, M. Giordani, K. Antonakoglou, T. Mahmoodi, and A. Zanella, "QUIC-EST: a QUIC-enabled scheduling and transmission scheme to maximize VoI with correlated data flows," *IEEE Communications Magazine*, vol. 59, no. 4, pp. 30–36, 2021.

[25] J. Zhang, X. Gao, L. Yang, T. Feng, D. Li, and Q. Wang, "A systematic approach to formal analysis of QUIC handshake protocol using symbolic model checking," *Security and Communication Networks*, vol. 2021, Article ID 1630223, 12 pages, 2021.

[26] A. Severinson, A. Graell Amat, and E. Rosnes, "Block-diagonal and LT codes for distributed computing with straggling servers," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 1739–1753, 2019.

[27] A. H. F. Raouf, J. Abouei, and M. Jaseemuddin, "Delivery phase in cache-based wireless networks with modified LT codes," *Physical Communication*, vol. 42, 2020.

[28] X. Song, N. Cheng, Y. Liao, S. Ni, and T. Lei, "Design and analysis of LT codes with a reverse coding framework," *IEEE Access*, vol. 9, pp. 116552–116563, 2021.

[29] B. Abdallah, F. Belghith, M. A. Ben Ayed, and N. Masmoudi, "Low-complexity QTMT partition based on deep neural network for Versatile Video Coding," *Signal, Image and Video Processing*, vol. 15, no. 6, pp. 1153–1160, 2021.

[30] L. Ma, S. Cheng, and Y. Shi, "Enhancing learning efficiency of brain storm optimization via orthogonal learning design," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 11, pp. 6723–6742, 2021.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 3577541, 7 pages
https://doi.org/10.1155/2021/3577541

Research Article

Neural Network for Intelligent and Efficient Volleyball Passing Training

Bo Liu, Ning Yang, Xiangwei Han, and Chen Liu

Shandong Youth University of Political Science, Jinan 250103, China

Correspondence should be addressed to Ning Yang; 190040@sdyu.edu.cn

Received 17 October 2021; Accepted 6 November 2021; Published 22 November 2021

Academic Editor: Jianhui Lv

Passing is a relatively basic technique in volleyball. In volleyball passing teaching, training the correct passing technique plays a very important role. The correct pass can not only accurately grasp the direction of the ball point and the drop point but also effectively connect the defense and the offense. In order to improve the efficiency and quality of volleyball passing training, improve the precise extraction of sport targets, reduce redundant feature information, and improve the generalization performance and nonlinear fitting capabilities of the algorithm, this paper studies volleyball based on the nested convolutional neural network model and passing training wrong movement detection method. The structure of the convolutional neural network is improved by nesting mlpconv layers, and the Gaussian mixture model is used to effectively and accurately extract the foreground objects in the video. The nested multilayer mlpconv layer automatically learns the deep-level features of the foreground target, and the generated feature map is vectorized and input to the Softmax classifier connected to the fully connected layer for passing wrong behavior detection in volleyball training. Based on the detection of nearly 1,000 athletes' action datasets, the simulation experiment results show that the algorithm reduces the acquisition of redundant information and shortens the calculation time and learning time of the algorithm, and the improved convolutional neural network has generalization performance and nonlinearity. The fitting ability has been improved, and the detection of abnormal volleyball passing behaviors has achieved a higher accuracy rate.

1. Introduction

1.1. Volleyball Passing Technique. For volleyball, passing technique is a very basic technique. To be able to pass the ball accurately, it is necessary to have the corresponding passing technique. The volleyball passing technique seems simple, but it is actually a very delicate and complex technique that requires high wrist strength. In the process of volleyball passing training, after the students learn the movements for the first time, the connection between the nerves and muscles of the main body of the movement is not precise, so wrong movements are often accompanied during the movement, so how to correct them in the process of volleyball passing teaching and preventing wrong moves are the basic requirements for physical education teachers to improve the quality of volleyball teaching. According to relevant investigations and studies, in the past correction of volleyball passing training actions, wrong actions could not be accurately detected, so wrong actions could not be corrected in time [1–3]. Driven by the general trend of international sports, volleyball sports activities continue to enter people's field of vision, and people begin to pay attention to volleyball training [4, 5]. In the actual training process, the volleyball passing technique seems relatively easy, but it is relatively difficult to learn. In the process of passing technique learning, many students often have uncoordinated body movements, and their fingers and wrists are obviously insufficient. It is easier to rub hands when passing volleyball. Many students have different situations such as being afraid of passing. This leads to more difficulties in teaching volleyball passing techniques. In this case, it is required to pass training in volleyball. Instructors in the middle school can effectively point out and correct the wrong actions of the athletes [6].

1.2. The Role of Neural Networks. In recent years, pattern recognition using machine learning to build detectors has been successfully applied in the field of visual detection, such as face recognition and car recognition [7]. For the detection of volleyball players, there are two main challenges: first, the target may be distorted due to the change of posture and scale [8]. Second, the resolution of the image is low, and the target object may be represented less than 200 pixels. At present, a new method for athlete detection in sports videos is the AdaBoost algorithm [9], which firstly extracts Har features from rectangular images of athletes and then uses AdaBoost algorithm to cascade weak classifiers to build strong classifiers [10], achieving good detection performance. However, it needs too many features, the detection speed is slow, and it cannot meet the real-time requirements well. With the further development of research in related fields, convolutional neural networks have been widely used in the process of human action recognition [11] and target detection [12]. It can realize the further processing of the action samples, which leads to the poor applicability of applying this method to the detection of wrong actions in physical education training.

Therefore, in order to make up for the abovementioned deficiencies in the process of volleyball passing training wrong action detection, this paper proposes an improved convolutional neural network crowd abnormal behavior recognition method, which improves the convolutional neural network structure by nesting mlpconv layers and uses a mixture of Gaussian models to effectively and accurately extract foreground targets in volleyball passing videos. The nested multilayer mlpconv layer automatically learns the depth-level features of foreground targets, and the generated feature maps are vectorized and input to the Softmax classification connected to the fully connected layer detectors for abnormal behaviors in volleyball passing training. The simulation experiment results show that the algorithm reduces the acquisition of redundant information and shortens the calculation time and learning time of the algorithm, and the improved convolutional neural network has improved generalization performance and nonlinear fitting ability, and it is abnormal for volleyball passing. Behavior detection achieves a high accuracy rate.

The remaining structure of this paper is organized as follows: Section 2 introduces convolutional neural network architecture. Section 3 gives the detailed design on volleyball passing training scheme. Section 4 reports the rich results. Finally, Section 5 concludes this paper.

2. Convolutional Neural Network Architecture

A convolutional neural network [13] is a hierarchical neural network based on local connections between neurons. Compared with traditional machine learning methods, it has a more complex network structure and more powerful feature learning and expression ability. The visual mode is decomposed into multiple submodes, and the submodes are processed by the feature planes connected step by step, so that the target can be well recognized even with small distortion [14]. The convolutional neural network consists of six different types of convolutional layers. The input layer receives 21×43 grayscale images, and C1 input images are convolved. The layer consists of four feature maps, each sharing an acceptance field and a bias. The S1 layer performs a secondary sampling and local average operation on the image to form four feature maps. The secondary sampling operation reduces the two dimensions of the input and enhances the invariance of image translation, scaling, and deformation. In addition, the output of the mixed feature map combines different features, which helps to extract more complex information. The C2 layer is not completely connected with the S1 layer, and the output image of the S1 layer is convolved in the 3×3 acceptance domain to generate 14 feature maps. The S2 layer has the same function as the S1 layer and consists of 14 feature maps. Each neuron in the N1 layer is connected to the feature map of the S2 layer. The N2 layer is the output layer and is fully connected to the N1 layer. The N2 layer uses a typical sigmoid neuron. After completing feature extraction and input dimensionality reduction, the role of the N1 and N2 layers is to perform output classification. The output of the neurons in the N2 layer is the identifier of the input image for the athlete or nonathlete, -1 for nonathletes and $+1$ for athletes.

3. Proposed Volleyball Passing Training Detection Scheme

3.1. Detection and Extraction of Volleyball Passing Movement Targets. In actual volleyball passing training, especially in outdoor sports, the environment and background of the sports target are constantly changing, and there are often some nontarget small sports in the background of the video image, such as throwing volleyball and shaking. In order to avoid the influence of background transformation and interfering targets on the detection effect of volleyball passes, the researchers proposed a mixture of Gaussian background modeling. If the difference between the gray value of the target area and the background information in the video image is large, the gray histogram of the video image is a double-peak-valley type, where one peak represents the moving target and the other peak represents the background of the image. For more complex images, the resulting grayscale histogram is multimodal, which can be seen as using multiple single Gaussian models to describe the change of a certain pixel over a period of time, which is a mixed Gaussian background modeling. In short, the mixed Gaussian background modeling is to accurately describe the characteristics of the pixels. By setting multiple Gaussian models for each pixel to improve its ability to describe the background, the purpose of accurately describing the image background is achieved, thereby obtaining a complete moving target. The number of Gaussian models usually selected in different documents is between 3 and 5 [15]. For more complex scenes, a larger number of models are chosen to improve the model's ability to describe the background. For simpler scenes, choosing a small number of models can avoid overdepicting the background by the model and causing the loss of moving targets [16].

At present, there are three main methods of passing target detection in volleyball training videos: optical flow method, interframe difference method, and background difference method [17]. In this paper, the Gaussian Mixture Model (GMM) [18] in the background difference method is selected. Compared with other methods of foreground target extraction, this model can not only successfully detect volleyball passing targets but also reduce the influence of small repetitive objects in the background scene on foreground target detection. For the detection of passing targets in volleyball training videos, firstly, Gaussian distribution is used to establish a background model for each pixel, and then background model parameters are automatically updated. Finally, the successful detection and extraction of passing targets in volleyball training videos are realized.

3.1.1. Mixed Gaussian Background Modeling. For any pixel, its historical pixel sequence can be traced as follows:

$$\{x_1, x_2, \ldots, x_t\} = \{F_i(x, y), \quad 1 \le i \le t\}, \tag{1}$$

where $F_i(x, y)$ is the gray value at the i moment.

At time t, the calculation formula of the probability function of the pixel (x, y) is as follows [19]:

$$p(F_t(x, y)) = \sum_{i=1}^{k} H_{i,t} \times \eta\left(F_t(x, y), \mu_{i,t}, \sum_{i,t}\right), \tag{2}$$

where $H_{i,t}$ is the i model weight value at time t; $\mu_{i,t}$ is the mean value of the i Gaussian distribution at time t; $\sum_{i,t}$ is the covariance at time t; and $\eta(F_t(x, y), \mu_{i,t}, \sum_{i,t})$ is the probability density function at time t. The calculation formula is as follows:

$$\eta\left(F_t(x, y), \mu_{i,t}, \sum_{i,t}\right) = \frac{1}{(2\pi)^{\pi/2}\left|\sum_{i,t}\right|^{1/2}} \times \exp\left(-\frac{1}{2}\left(F_t(x, y) - \mu_{i,t}\right)^T \sum_{i,t}^{-1}\left(F_t(x, y) - \mu_{i,t}\right)\right). \tag{3}$$

3.1.2. Update of Gaussian Mixture Model Parameters. The Gaussian distribution of the frame pixel value $F_i(x, y)$ is sorted according to the priority, and formula (4) is satisfied, indicating that the frame pixel value $F_i(x, y)$ matches the Gaussian distribution i successfully, and the parameters of the Gaussian distribution of the frame pixel value $F_i(x, y)$ follow formula (4) and equation (8) is updated; equation (4) is not satisfied, the frame pixel value $F_i(x, y)$ matches the Gaussian distribution unsuccessfully, the parameters of the Gaussian distribution remain unchanged, and the weight value is updated according to equation (9).

$$\left|F_t(x, y) - \mu_{i,t-1}\right| < D \times \sigma_{i,t-1}, \tag{4}$$

$$H_{i,t} = (1 - \alpha)H_{i,t-1} + \alpha, \tag{5}$$

$$\mu = (1 - \beta)\mu_{i,t-1} + \beta F_i(x, y), \tag{6}$$

$$\delta_{i,t}^2 = (1 - \beta)\delta_{i,t-1}^2 + \beta\left(F_t(x, y) - \mu_{i,t-1}\right)^T\left(F_t(x, y) - \mu_{i,t-1}\right), \tag{7}$$

$$\beta = \alpha\eta\left(F_t(x, y) \mid \mu, \delta_{i,t}\right), \tag{8}$$

$$F_{i,t} = (1 - \alpha)F_{i,t-1} + \alpha, \tag{9}$$

where α and β, respectively, represent the learning rate and update rate of the Gaussian mixture model.

3.1.3. Extraction of Volleyball Pass Target by the Gaussian Mixture Model. After the Gaussian mixture model of each pixel is generated, the Gaussian distribution is arranged in descending order according to the value of ψ/μ, and the first B Gaussian distribution is obtained as the background model [20]. The formula is as follows:

$$B = \arg\min \sum_{i-1}^{b} \psi_{i,t} > T, \tag{10}$$

where T is the set threshold.

The first B Gaussian distribution is obtained as the background model, and the current pixel value $F_i(x, y)$ is matched with the generated background. If the current pixel value $F_i(x, y)$ is not successfully matched with the generated background, then the current pixel $F_i(x, y)$ is a good target for volleyball passing. Otherwise, the current pixel $F_i(x, y)$ point is the background point. After the above process, the Gaussian mixture model realizes the detection and extraction of the passing target in the volleyball training video.

3.2. Obtaining Volleyball Pass Characteristic Information

3.2.1. Mlpconv Layer. The mlpconv layer consists of a linear convolutional layer and a multilayer perceptron (MLP), and the input mapping in the local perception field of view corresponds to the feature vector. The mlpconv layer uses

multiple fully connected layers with nonlinear activation functions to extract the feature information of the volleyball pass target, converts the extracted feature information into a feature map, and then uses the feature map as the input of the next layer [21].

The calculation process of the mlpconv layer is as follows:

$$\begin{aligned} g^1_{i,j,k_1} &= \max\left(H^{1T}_{k_1} x_{i,j} + b_{k_1}, 0\right), \\ g^2_{i,j,k_2} &= \max\left(H^{2T}_{k_2} g^1_{i,j} + b_{k_2}, 0\right), \end{aligned} \tag{11}$$

$$g^n_{i,j,k_n} = \max\left(H^{nT}_{k_n} g^{n-1}_{i,j} + b_{k_n}, 0\right), \tag{12}$$

where (i, j) is the position of the pixel in the feature map, $x_{i,j}$ is the input block centered at the pixel point (F, j), k_1, k_2, and k_n are the channel numbers in the feature map, and n is the number of MLP layers.

3.2.2. Batch Normalization Technology. In the neural network learning process, with the changes of the parameters of each layer, especially the algorithm's learning rate and weight initialization, it will take a long time to find a suitable value, which reduces the training speed of the neural network. When using a saturated nonlinear activation function to train a neural network model, the input data will mistakenly enter the saturation region of the activation function, which reduces the convergence speed of the neural network.

Ioffe et al. [22] used BN (Batch Normalization) technology to standardize the input of each layer to solve the above problems. BN technology makes the input data have zero mean and unit variance:

$$\widehat{x}_{i,j,n} = \frac{\widehat{x}_{i,j,n} - E[x_n]}{\sqrt{\mathrm{Var}[x_n]}}. \tag{13}$$

After normalization, the parameters need to be scaled and translated accordingly:

$$g_{i,j,n} = \gamma_n \widehat{x}_{i,j,n} + \beta_n, \tag{14}$$

where $\widehat{x}_{i,j,n}$ is the value of the input data at position (i, j), n is the channel number in the feature map, and γ_n and β_n are the newly introduced zoom and translation parameters in network training.

3.3. Building a Nested Model of the Convolutional Neural Network. The core idea of the convolutional neural network nested model is as follows: the nested network model can automatically learn deep-level features excellently. The deep-level features acquired by this model are mainly local features. When the nested network model obtains the feature information of the moving target, especially in the separation of the target in the background, the local features will play an important role. In addition, the nested network model also has a certain degree of robustness when dealing with drastic changes in the background target.

When the network nested model is trained, first, the weights of the convolutional neural network model is initialized containing a single mlpconv layer, and then the convolutional neural network is trained. The entire training process is over, and the weights of the single mlpconv layer are updated; then, it is connected to the second mlpconv layer. The input of the second mlpconv layer is the output of the first mlpconv layer. The weights of the second mlpconv layer are initialized, and then the convolutional neural network is trained. The whole training process is over, and an update of the weight of the second mlpconv layer is received. When a new mlpconv layer is added, weight initialization, convolutional neural network training, and weight update are performed according to the above process.

In addition, the use of BN technology after the convolution calculation also enables the nonlinear unit to produce a relatively stable distribution and achieve the effect of desaturation. The BN operation is added to the nested mlpconv layer, and the calculation method of the feature map in the model is as follows:

$$\begin{aligned} g^1_{i,j,k_1} &= \max\left(\mathrm{BN}\left(H^{1T}_{k_1} x_{i,j} + b_{k_1}\right), 0\right), \\ g^2_{i,j,k_1} &= \max\left(\mathrm{BN}\left(H^{2T}_{k_2} g^1_{i,j} + b_{k_2}\right), 0\right), \\ g^n_{i,j,k_n} &= \max\left(\mathrm{BN}\left(H^{n^T}_{k_n} g^{n-1}_{i,j} + b_{k_n}\right), 0\right), \end{aligned} \tag{15}$$

where $\mathrm{BN}(g)$ represents the BN layer, (i, j) is the position of the pixel in the feature map, $x_{i,j}$ is the input block centered on the pixel point (i, j), k_1, k_2, and k_n are the channel numbers in the feature map, and n is the MLP layer number.

4. Experiment and Result Analysis

On 3.0 GHz CPU, 64 bit Windows 7 operating system, MATLAB 2016a, and Open CV are used as development tools for simulation experiments. In order to prove the effectiveness of the intelligent and efficient volleyball pass training detection modeling method based on convolutional neural network, an experiment is needed. The experimental objects were collected from the data collection of volleyball passing training of 1,000 students in a physical education college. Two indoor and outdoor scenes and different shooting angles were selected to record different volleyball passing behaviors. The input data is to crop each frame into 80×60 grayscale images. The convolution kernels used in the three convolutional layers of the convolutional neural network model are 9×7, 7×7, and 6×4 scales. The convolution kernels used in the two downsampling layers are all 3×3 scales. The input $80 \times 60 \times 9$ volleyball pass video block is finally transformed into a 128-dimensional feature vector. On this basis, all experimental data are divided into two groups, one group is used for deep convolutional neural network training and the other group is used for experimental testing.

4.1. Experimental Results. During the experiment, the dataset used in this article is a nonpublic dataset. Machine vision technology is used to capture volleyball passing training actions, and the captured results are denoised and enhanced to improve the accuracy of the experimental results. In the simulation experiment, quantitative evaluation

adopts the AUC evaluation index, equal error rate (EER) and running time (Time), and other indexes. This paper selects algorithms that have achieved better recognition rates in the above databases for comparisons, such as TCP model [23], AMDN (double fusion) model [24], Motion Energy model [23], SpatioTemporal Convolutional Neural Network (ST-CNN) model [25], and Commotion model [26]. It can be seen from Table 1 that, using frame-level measurement tests, the algorithm in this paper has an advantage in EER and AUC evaluation indicators, and the algorithm has been improved in terms of time-consuming.

Six methods are used to detect volleyball passing training wrong actions on experimental samples, the error rate of different methods of volleyball passing wrong action detection is compared, and the comparison results are used to measure the comprehensive effectiveness of six different methods for detecting volleyball passing training wrong actions. The comparison results are shown in Table 2. Analyzing Table 2 shows that with the continuous increase of the number of experiments, the detection error rate of the method in this paper for volleyball passing training errors has been maintained at a low level. The average error rate of error detection for volleyball passing training in this paper is about 0.027%, which is lower than other methods. When using the method in this paper to detect the wrong action of volleyball passing training, the error can be controlled within a reasonable area.

In order to verify the effectiveness and robustness of this method in the detection of wrong actions in volleyball passing training, four test indicators, ACC (accuracy rate), TPR (sensitivity), FPR (specificity), and PPV (positive prediction rate), are adopted. For quantitative comparison, the specific calculation formula is described as follows:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}, \tag{16}$$

$$\begin{aligned} \mathrm{TPR} &= \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \\ \mathrm{FPR} &= \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \\ \mathrm{PPV} &= \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \end{aligned} \tag{17}$$

Among them, TP means the number of positive samples, which is actually the number of positive samples; FP means the number of positive samples but actually the number of negative samples; TN means the number of negative samples, which is actually the number of negative samples; FN indicates the number of samples that are judged as negative, but in fact, it is the number of positive samples. The higher the ACC and TPR, the lower the FPR and PPV and the better the detection performance.

Table 3 is the test experiment result of wrong action detection in physical education training. From Table 3, among the six methods of volleyball passing training error detection, the four parameters of ACC, FPR, PPV, and TPR in this method are better than the other five methods, and the detection accuracy of this method has reached more than 95%, Therefore, the experimental results verify the superior performance of the deep convolutional neural network.

4.2. Time Complexity Analysis and Comparison. The time complexity of an algorithm is a function that quantitatively describes the running time of the algorithm.

4.2.1. Time Complexity of a Single Convolutional Layer. Time complexity refers to the computational workload required to execute the algorithm. The time complexity of a single convolutional layer is calculated as follows:

$$\mathrm{Time} \sim O\left(P^2 \cdot Q^2 \cdot C_{\mathrm{in}} \cdot C_{\mathrm{out}}\right), \tag{18}$$

where P represents the side length of each convolution kernel output feature map; Q represents the side length of each convolution kernel; C_{in} represents the number of channels of each convolution kernel, that is, the number of input channels (number of output channels of the previous layer); and C_{out} represents the number of convolution kernels in the convolution layer, that is, the number of output channels.

It can be seen from equation (16) that the time complexity of the convolutional layer is determined by the output feature map area P^2, the convolution kernel area Q^2, the input C_{in}, and the number of output channels C_{out}; the size of the output feature map is in turn determined by the input matrix size X, and the size of the convolution kernel is determined by Q. The and the expression of the side length P of the output feature map is as follows:

$$P = \frac{(X - Q + 2 \times \mathrm{Padding})}{\mathrm{Stride}}. \tag{19}$$

4.2.2. The Overall Time Complexity of the Convolutional Neural Network. The complexity of a single-layer convolutional neural network is calculated by equation (16). The overall time complexity of a convolutional neural network (including multilayer structure) is the sum of the time complexity of each layer. The calculation formula is as follows:

$$\mathrm{Time} \sim O\left(\sum_{l=1}^{D} P_l^2 \cdot Q_l^2 \cdot C_{l-1} \cdot C_l\right), \tag{20}$$

where D represents the number of convolutional layers of the neural network, that is, the network depth; l represents the lth convolutional layer of the neural network; C_1 represents the number of output channels of the lth convolutional layer of the neural network C_{out}, that is, the number of convolution kernels in this layer; the number of input channels of the lth convolutional layer; and X_5 is the number of output channels of the $(l-1)$th convolutional layer.

In terms of time complexity, this article selects the TCP model, AMDN (double fusion) model, motion energy model, spatio-temporal convolutional neural network model, and commotion model to compare with the algorithm in this paper. From equation (12), the time complexity of each algorithm can be calculated. Because the specific parameter data of each algorithm are not clear, this article only calculates which order the time complexity of the algorithm belongs to. The common time complexity

Table 1: AUC and EER used for the frame and pixel-level comparison on the datasets.

Algorithm	EER (%)	AUC	Time (s)
TCP	19.5	0.802	0.34
AMDN	18.6	0.909	0.26
Motion energy	22.1	0.923	0.13
ST-CNN	25.5	0.895	0.52
Commotion	22.8	0.872	0.28
This paper	16.2	0.936	0.12

Table 2: Detection results of wrong actions in physical education training.

Number of experiments/time	Error rate (%)					
	TCP	AMDN	Motion energy	ST-CNN	Commotion	This paper
200	0.092	0.098	0.086	0.132	0.076	0.024
400	0.095	0.108	0.123	0.191	0.085	0.017
600	0.088	0.082	0.092	0.255	0.126	0.036
800	0.097	0.089	0.102	0.234	0.089	0.012
1000	0.086	0.072	0.089	0.145	0.097	0.047

Table 3: Detection results of wrong actions in physical education training.

Test index	ACC	TPR	PPV	FPR
TCP	0.892	0.802	0.142	0.086
AMDN	0.835	0.909	0.264	0.032
Motion energy	0.868	0.923	0.138	0.097
ST-CNN	0.814	0.895	0.129	0.112
Commotion	0.886	0.872	0.282	0.062
This paper	0.976	0.966	0.087	0.013

Table 4: Time complexity of the algorithms.

Algorithm	Time complexity
TCP	$O(n^2)$
AMDN	$O(n^3)$
Motion energy	$O(n^2)$
ST-CNN	$O(n^2)$
Commotion	$O(n^2)$
This paper	$O(n \log n)$

relationship is $O(1) < O(\log n) < O(n) < O(n \log n) < O(n^2) < O(n^3)$.

As shown in Table 4, the time complexity of this algorithm and other algorithms is mostly on the $O(n \log n)$ order. Analyzing the neural network structure model of other algorithms, the output feature map area P^2, the convolution kernel area Q^2, the input C_{in}, and the output channel number C_{out} are all more complicated than the algorithm in this paper. It can be concluded that the algorithm in this paper is better than other algorithms in terms of time complexity.

5. Concluding Remarks

Since the traditional methods cannot accurately obtain the characteristics of the wrong movements in volleyball passing training, resulting in the decrease of detection accuracy, this paper proposes a method for detecting the wrong movements in volleyball passing training based on a convolutional neural network. The convolutional neural network structure is improved by the nested mlpconv layer. The mixed Gaussian model is used to extract the passing target from the volleyball training video sequence effectively and accurately. The mixed Gaussian model shows robustness in the complex scene background, which can not only successfully detect the volleyball passing training target but also reduce the influence of the small repetitive objects in the background scene on the passing target detection. The nested multilayer mlpconv layer automatically learns the deep-level pass features of the volleyball target that has been extracted, and the improved convolutional neural network reduces the acquisition of redundant information. Experiments show that the method can effectively detect the wrong actions of the athletes during the volleyball passing training process, the detection accuracy is high, the detection error can be effectively controlled, and the wrong actions can be accurately judged in time. And the improved convolutional neural network has excellent generalization performance and nonlinear fitting ability.

Data Availability

All data used to support the findings of the study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Social Science Planning and Research Project of Shandong Province (Grant No. 21CTYJ18).

References

[1] X. Y. Zhang, X. Y. Zhou, M. X. Lin, and J. Sun, "ShuffleNet: an extremely efficient convolutional neural network for mobile devices," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6848–6856, Salt Lake City, UT, USA, June 2018.

[2] A. G. Howard, M. L. Zhu, B. Chen et al., "MobileNets: efficient convolutional neural networks for mobile vision applications," 2017, https://arxiv.org/abs/1704.04861v1.

[3] G. Larsson, M. Maire, and G. Shakhnarovich, "FractalNet: Ultra-Deep neural networks without residuals," 2017, https://arxiv.org/abs/1605.07648v4.

[4] D. Y. Han, J. H. Kim, and J. M. Kim, "Deep pyramidal residual networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, pp. 6307–6315, Honolulu, HI, USA, July 2017.

[5] J. Hu, L. Shen, G. Sun, and E. Wu, "Squeeze-and-Excitation networks," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, Salt Lake City, UT, USA, June 2018.

[6] D. Chen, P. Wang, L. Yue, Y. Zhang, and T. Jia, "Anomaly detection in surveillance video based on bidirectional prediction," *Image and Vision Computing*, vol. 98, Article ID 103915, 2020.

[7] J. Carreira and A. Zisserman, "Quo Vadis, action recognition? A new model and the kinetics dataset," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4724–4733, Honolulu, HI, USA, July 2018.

[8] F.-N. Yuan, L. Zhang, S. Jin-Ting, X. Xue, and G. Li, "Theories and applications of auto-encoder neural network- s: a literature survey," *Chinese Journal of Computers*, vol. 42, no. 1, pp. 203–230, 2019.

[9] W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," in *Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6479–6488, IEEE, Salt Lake City, UT, USA, June 2018.

[10] L. Li, "Analysis and data mining of intellectual property using GRNN and SVM," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 139–150, 2020.

[11] M. Ravanbakhsh, M. Nabi, H. Mousavi, E. Sangineto, and N. Sebe, "Plug-and-Play CNN for crowd motion analysis: an application in abnormal event detection," in *Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Tahoe, NV, USA, March 2018.

[12] D. X. Xu, W. Y. Yan, E. Ricci, and N. Sebe, *Detecting Anomalous E- Vents in Videos by Learning Deep Representations of Appearance and motion*, Elsevier Science Inc, Amsterdam, Netherlands, 2017.

[13] T. Chen, C. Hou, Z. Wang, and H. Chen, "Anomaly detection in crowded scenes using motion energy model," *Multimedia Tools and Applications*, vol. 77, no. 3, pp. 1–16, 2017.

[14] L. Ma, S. Cheng, and Y. Shi, "Enhancing learning efficiency of brain storm optimization via orthogonal learning design," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 11, pp. 6723–6742, 2021.

[15] H. Mousavi, M. Nabi, H. Kiani, P. Alessandro, and M. Vittorio, "Crowd motion monito- ring using tracklet-based commotion measure," in *Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP)*, Quebec City, QC, Canada, September 2015.

[16] J. Y. Ma, F. R. Jie, and Y. J. Hu, "Moving target detection method based on improved Gaussian mixture model," in *Proceedings of the Ninth International Conference on Digital Image Processing (ICDIP 2017)*, Hong Kong, China, July 2017.

[17] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on International Conference on Machine Learning—Volume 37*, pp. 448–456, Lille, france, July 2015.

[18] R. T. Ionescu, F. S. Khan, M. Georgescu, and L. Shao, "Object-centric auto-encoders and dummy anomalies for abnormal event detection in video," in *Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7834–7843, Long Beach, CA, USA, December 2019.

[19] J. Xiao, M. Shen, J. Lei, J. Zhou, R. Klette, and H. Sui, "Single image dehazing based on learning of haze layers," *Neurocomputing*, vol. 389, pp. 108–122, 2020.

[20] Y. Zhou, X. Sun, Z. Zha, and W. Zeng, "MiCT: mixed 3D/2D convolutional tube for human action recognition," in *Proceedings of the 2018 IEEE Conference on Computer Vision and Pat- tern Recognition*, pp. 449–458, Salt Lake City, UT, USA, June 2018.

[21] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid net- works for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2117–2125, Honolulu, HI, USA, July 2017.

[22] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: convolutional block Attention module," in *Proceedings of the Computer Vision—ECCV 2018*, pp. 3–19, Munich, Germany, September 2018.

[23] B. Jiang, R. Luo, J. Mao, T. Xiao, and Y. Jiang, "Acquisition of localization confidence for accurate object detection," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 784–799, Munich, Germany, September 2018.

[24] Z. Cai and N. Vasconcelos, "Cascade r-cnn: Delving into high quality object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6154–6162, Salt Lake, UT, USA, June 2018.

[25] Z. Tian, C. Shen, H. Chen, and T. He, "Fcos: Fully convolutional one- stage object detection," in *Proceedings of the IEEE international conference on computer vision*, pp. 9627–9636, Seoul, South Korea, October 2019.

[26] Y. Liu, Y. Wang, S. Wang et al., "CBNet: a novel composite backbone network architecture for object detection," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 7, pp. 11653–11660, 2020.