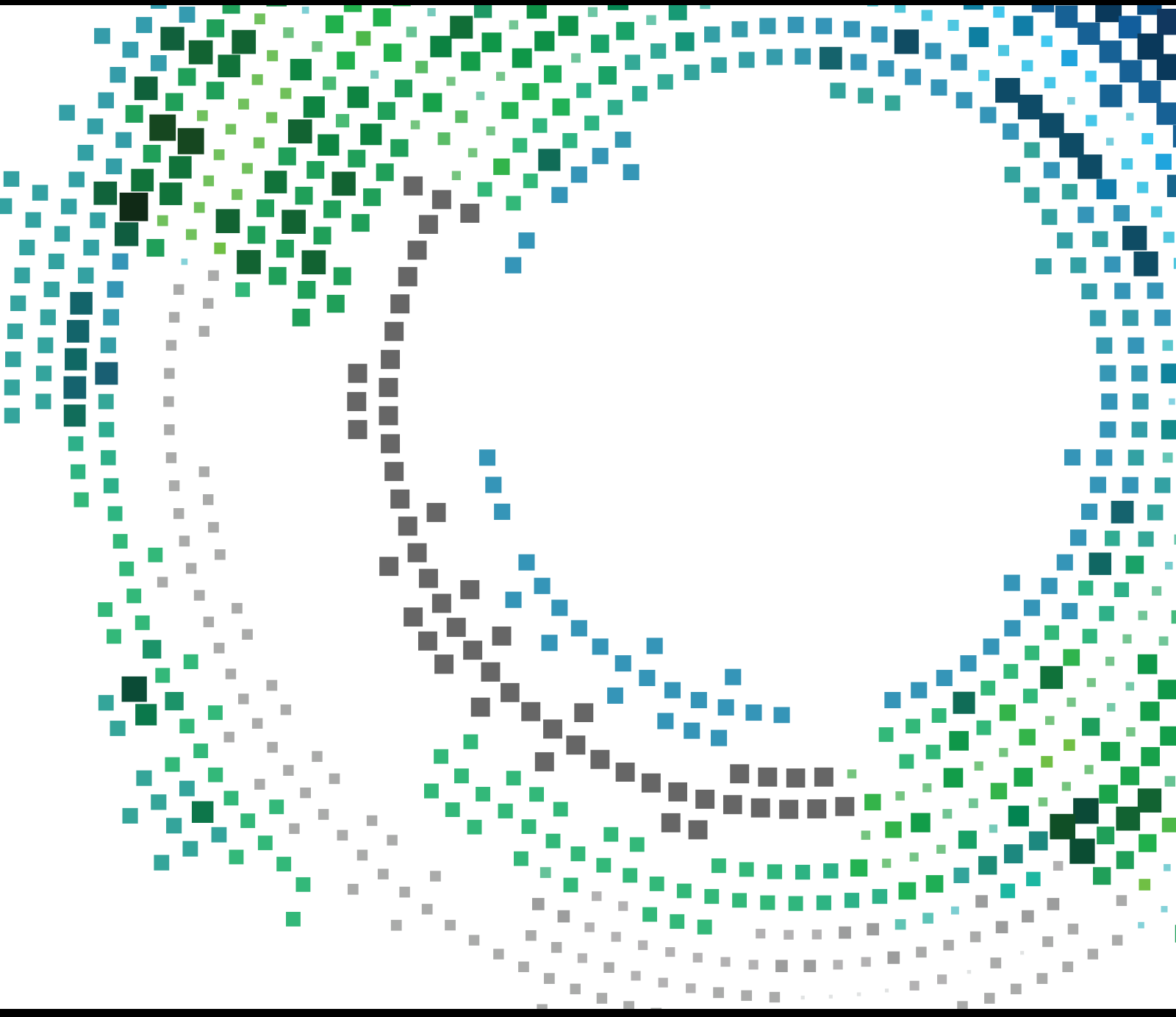
Mobile Information Systems

Artificial Intelligence for Next-Generation Wireless Networks

Lead Guest Editor: Han Wang
Guest Editors: Qinglin Zhao, Lingwei Xu, and Thomas Aaron Gulliver



Artificial Intelligence for Next-Generation Wireless Networks

Mobile Information Systems

Artificial Intelligence for Next-Generation Wireless Networks

Lead Guest Editor: Han Wang
Guest Editors: Qinglin Zhao, Lingwei Xu, and Thomas Aaron Gulliver

Contents

Multiobjective Optimal Dispatching of Smart Grid Based on PSO and SVM
Man Bao, Hongqi Zhang, Hao Wu, Chao Zhang, Zixu Wang, and Xiaohui Zhang
Research Article (9 pages), Article ID 2051773, Volume 2022 (2022)

A Shadow Capture Deep Neural Network for Underwater Forward-Looking Sonar Image Detection
Taowen Xiao, Zijian Cai, Cong Lin, and Qiong Chen
Research Article (10 pages), Article ID 3168464, Volume 2021 (2021)

Energy-Efficient UAV Trajectory Design with Information Freshness Constraint via Deep Reinforcement Learning
Xinmin Li, Jiahui Li, and Dandan Liu
Research Article (9 pages), Article ID 1430512, Volume 2021 (2021)

Adaptive Polymorphic Fusion-Based Fast-Tracking Algorithm in Substations
Wenbin Shi, Jingsheng Lei, Xingli Gan, and Zhongguang Yang
Research Article (18 pages), Article ID 2292128, Volume 2021 (2021)

Predicting Joint Effects on CubeSats to Enhance Internet of Things in GCC Region Using Artificial Neural Network
Faris A. Almalki and Soufiene Ben Othman
Research Article (16 pages), Article ID 1827155, Volume 2021 (2021)

A Novel Sparse Array Configuration with Low Coarray Redundancy for DOA Estimation in Mobile Wireless Sensor Network
Pin-Jiao Zhao, Guo-Bing Hu, and Liang-Tian Wan
Research Article (8 pages), Article ID 1362640, Volume 2021 (2021)

Prediction of Enterprise Economic Activity Behavior Based on Neural Network and ARIMA Hybrid Model
Zeyu Lin and Shuai Li
Research Article (12 pages), Article ID 4571840, Volume 2021 (2021)

Community Detection Algorithm Based on Intelligent Calculation of Complex Network Nodes
Yanjia Tian and Xiang Feng
Research Article (10 pages), Article ID 2931801, Volume 2021 (2021)

Signal Modulation Identification Based on Deep Learning in FBMC/OQAM Systems
Jing Chen, Jianzhong Guo, Xin Shan, and Dejin Kong
Research Article (6 pages), Article ID 4809699, Volume 2021 (2021)

Edge-Cloud-Assisted Multiuser Forward Secure Searchable Encryption (EMFSSE) Scheme in the P2P Networking Environment
Yongli Tang, Jingran Li, Xixi Yan, and Qiang Zhao
Research Article (14 pages), Article ID 7003252, Volume 2021 (2021)

Application of Adaptive Virtual Reality with AI-Enabled Techniques in Modern Sports Training
Yajun Zhang and Sang-Bing Tsai
Research Article (10 pages), Article ID 6067678, Volume 2021 (2021)

Elman-Based Secure Data Transmission Quality Prediction for Complex IoT Networks
Fagen Yin
Research Article (8 pages), Article ID 4132622, Volume 2021 (2021)

Hindawi
Mobile Information Systems
Volume 2022, Article ID 2051773, 9 pages
https://doi.org/10.1155/2022/2051773

Research Article

Multiobjective Optimal Dispatching of Smart Grid Based on PSO and SVM

Man Bao, Hongqi Zhang, Hao Wu, Chao Zhang, Zixu Wang, and Xiaohui Zhang

College of Mechanical and Electrical Engineering, Inner Mongolia Agricultural University, Hohhot 010018, China

Correspondence should be addressed to Hongqi Zhang; 050037@imau.edu.cn

Received 9 September 2021; Revised 21 October 2021; Accepted 27 October 2021; Published 18 January 2022

Academic Editor: Han Wang

The optimization of microgrid is an important part of smart grid. The global energy consumption is seriously greater than the energy it has, and the environmental pollution brought by it should not be underestimated. If we want to reduce their impact, introducing the optimization of microgrid is a good solution. Short-term load forecasting is a very important prerequisite for microgrid optimization, which lays a solid foundation for the realization of the development goal of environmental protection and the improvement of the economic benefits of microgrid. In this paper, a Multi-PSO-SVM forecasting model is proposed to forecast the actual load. By simulating four prediction models with three different samples, we can see that the average predicted value and actual load value of Multi-PSO-SVM algorithm in the three different samples are almost less than 10 MV. Compared with the other three algorithms, Multi-PSO-SVM is superior in accurately predicting the load value at each time point, which provides important conditions for the success of microgrid optimization.

1. Introduction

Traditional power generation is mostly achieved through the consumption of various energy sources, which will not only bring environmental problems, but also consume more nonrenewable resources, so the optimal scheduling of microgrid is very necessary. Microgrid mainly relies on renewable energy as the source to complete distributed generation, so it can be known that the pressure of environmental pollution is greatly reduced. Compared with other power generation modes, microgrid technology is independent, which can control and feed back the power quality and performance of power grids in different regions in time and make timely response. Its flexible installation point can make it popular all over the world. It can be installed all over the world, but nonrenewable resources such as wind and photovoltaic power have some disadvantages. They have a lot of randomness, uncertainty, and intermittence, which leads to uncontrollable output power, in addition to the need to ensure the balance between load supply and demand. In a sense, load forecasting is an essential step in the research of microgrid. In this paper, Multi-PSO-SVM forecasting model is proposed, which is unexpected in terms of forecasting accuracy and effect. We also carry out experiments with other three prediction models through experimental simulation. Through four line charts, we can see that our proposed prediction models have a very high degree of coincidence.

Literature [1] points out that it is relatively difficult to integrate photovoltaic solar power into the power grid because of the uncertainty brought by photovoltaic solar power generation, so we need to make accurate short-term prediction. At present, there are three kinds of hybrid prediction models based on intelligent technology. Based on three technologies, the wavelet transform, the particle swarm optimization to optimize some parameters, and the excellent prediction accuracy of support vector machine, we can accurately predict the power generation. Then, we also find that the simulation results are consistent with the conjecture, and this model has a high degree of coincidence with the accuracy obtained by our imagined model. Literature [2] suggests that distributed generation has many uncertainties and intermittence due to various shortcomings, so it proposes an SVM technology to predict wind power and

optimize the whole model by combining two algorithms, PSO and SA. The model parameters are self-adaptive, and then it is found that the learning and adjusting model can make the prediction model track the fluctuation and change of wind power effectively, which makes an important contribution to improving the accuracy of predicting the total power supply capacity of microgrid. Literature [3] considers the advantages of chaos theory, so it combines it with prediction model, then improves particle optimization algorithm, and then complements chaos theory. Finally, a result of optimizing LSSVM parameters can be obtained. Through layered training, the prediction results show that this method not only reduces the possibility of algorithm falling into local extremum, but also improves the learning ability and prediction accuracy. Literature [4] clearly points out that the accuracy of forecasting makes it an important influence index of short-term load forecasting, Therefore, this paper proposes a technique to combine support vector machine and dragonfly algorithm, which uses two parameters of support vector machine as the searching position of dragonfly, while the current fitness value and the optimal position of dragonfly correspond to its prediction accuracy and optimal parameters, respectively. From the results of the test, we can clearly know that, compared with this algorithm, the steps are simple and the global search ability is strong with high prediction accuracy and fast calculation familiarity. Literature [5] shows that support vector machine (SVM) has excellent experimental performance, and its performance depends on the sufficient setting of its parameters. Therefore, choosing appropriate initial parameters of SVM plays an important role in obtaining good performance and classification effect. Through several experiments, it is proved that this experiment includes two vector parameter optimization techniques, network search and particle swarm optimization. It can be intuitively concluded that this method has higher accuracy than other methods. Literature [6] proposes a parallel time-varying particle swarm optimization algorithm, which can optimize the parameters and feature selection of support vector machines. In fact, in the parallel environment, the initial convergence problem of PSO algorithm mainly depends on virtual machines. On a composite set consisting of 30 benchmark data sets, compared with other methods, this method can not only get more suitable model parameters, discriminant feature subsets, and smaller SVs sets, but also significantly reduce the calculation time and have higher prediction accuracy. Literature [7] proposes that charging state is an important condition for microgrid in energy management system. Based on the regression principle of least square support vector machine (LSSVM), the LSSVM model is able to solve the problem of poor generalization performance of general battery models under different health conditions. When PSO is selected, the efficiency and precision of the model and the initial anticipation training are improved well. Through constant discharge experiment, AC discharge experiment, and different examples, it is proved that this new model provides a very good solution for microgrid energy management system. Literature [8], in order to evaluate the level of enterprise asset quality management in power grid, proposes a mitigation model combining least square support vector machine and particle swarm optimization algorithm, which can provide an important measurement reference for improving their level to a great extent. Literature [9] evaluates the optimal foraging algorithm (OFA) to optimize the key parameters of support vector machines and reduce classification errors by using six common benchmark data sets and using five optimization algorithms. From the result of experiment, we can clearly know that OFA-SVM algorithm has good effects and ability to find the optimal values of SVM parameters. In literature [10], aiming at solving the bottleneck problem of high energy storage life loss cost in microgrid, a two-stage coordinated scheduling model and control strategy based on two groups of batteries are proposed. The established microgrid model aims to achieve the goal of minimizing its total operating cost, and there is also an energy storage system, which uses a topological structure. Later, it is verified by demonstration projects that it can prolong the life of batteries and achieve its desired economic benefits. Literature [11] describes a complete photovoltaic-vanadium redox cell (VRB) microgrid as a whole and puts forward the problem of the optimal power and energy rating of VRB system in isolated microgrid and grid-connected microgrid considering the charging and discharging efficiency and working characteristics of VRB. A new battery operating cost model considering the charge-discharge efficiency and life cycle of the battery is proposed. Literature [12] proposes a multiobjective optimization method for optimal dispatching and operation of microgrid based on dynamic dispatching of battery energy storage system, which aims at providing better economic benefits than others, maximum reliability, minimum fluctuation of renewable energy power, and best compliance with renewable energy power generation plan. Two methods, fuzzy theory and binary factors, are used to transform it into a single-objective optimization problem, to achieve the optimal scheduling strategy of BESS successfully. It has been verified that the conjecture with good optimization effect has been confirmed. Literature [13] presents a new optimization algorithm as the previous multiobjective particle swarm optimization, in which fuzzy clustering in the middle is a helper to find the optimal solution from generation to generation. This can not only improve the relative stability of the algorithm, but also greatly improve the global search ability, so that the optimization results are more evenly distributed. A typical microgrid example in Europe is used to verify its reliable results, and it can achieve better results in the process of implementation The multiobjective optimization model in literature [14] greatly reduces the economic cost, realizes the unification of environmental pollution and external power supply, and clearly constructs the weight coefficient of each subobjective by analyzing the stratification process. Therefore, for realizing the optimization and adjustment of MG, quantum particle swarm optimization algorithm is adopted. In the actual case analysis, we can clearly find that the feasibility of the model is very high, showing very effective results. A nonlinear constraint model is proposed in literature [15], which is a method to solve the problem of optimizing microgrid dispatching. This method

reduces the cost. Harmful emissions and load requirements can be reduced. The method proposed in the above literature is based on PSO method applied to power grid. However, the current training time is long, and the complexity is high. In the optimization process, the parameters of SVM training process are optimized to accelerate the convergence of the algorithm. Section 2 describes the basic algorithm and the improved algorithm of PSO and SVM. In Section 3, the combination of PSO and SVM algorithm is discussed, and the optimal parameters are given. In Section 4, PSO-SVM algorithm is simulated, and the optimal values of various algorithms are compared. Section 5 concludes this study.

2. PSO Algorithm

2.1. Basic Ideas. PSO is realized by simulating birds and other particles of birds. Particles have only two properties, velocity and position. The position vector is a good idea to denote the direction of motion and the solution. The velocity vector represents the velocity of the position and indicates changes in direction and size. Each and every particle is isolated in the searchable space to look for the optimum solution, it will remember the current individual extreme value and share the previously recorded extremum with other particles, and the optimal individual extremum is found to become the overall situation optimal solution. Each particle in the particle swarm adjusts its speed and position through the current individual extreme value. PSO algorithm is simple and easy, and its parameters are not adjusted much.

2.2. Update Rule. PSO is initialized to a random bunch of particles (random solutions) and then passes the iteration to find out the optimal solution. In each iteration, the particle usually changes the velocity vector and position vector of the two "extremes" (p, g) to update itself.

The speed and position update formula of PSO algorithm is expressed as

$$v_i = v_i + C_1 \times \text{rand}(\,) \times (p - x_i) + C_2 \times \text{rand}(\,) \times (g - x_i), \tag{1}$$

where rand () is random numbers within (0, 1); X_i is the current position of the particle; and C_1 and C_2 are learning factor, usually 2. The formula is divided into three different parts by the operation symbol plus sign. The first part represents the influence of the previous velocity vector, and the second part is a vector from the current point to the best point of the particle obtained by experience and calculation. P and G represent the velocity vector and the position vector. Rand is a random number; its function is to change the corresponding speed and position indefinitely in the optimization process, so that PSO algorithm can get the global optimization more easily in the optimization process. It shows that the particle swarm algorithm is a summary of the historical records and proposes the best solution. The third part is a vector that points from the current point to the population and obtains the best point from experience and work, which reflects the cooperation of each particle and the sharing of its own information between individuals and populations.

$$x_i = x_i + v_i, \tag{2}$$

where $i = 1, 2, 3, \ldots, N$, where N is the total number of particles. V_i is velocity of particles. PSO has the problem of premature convergence and is easy to fall into local optimum. In formula (1), the next velocity variable value is updated to get the particle velocity. Equation (2) updates the position, thus updating the position information. Repeat the above process until an optimal value is reached. In this paper, an optimal combination method of PSO-SVM is proposed, which can effectively transform local optimum to global optimum.

2.3. Improved PSO Algorithm. The inertia factor w is introduced, and its value is nonnegative. When the optimization ability of each part is strong, the global and local values need to be larger. On the contrary, if the global and local optimization ability are weak, the value needs to be smaller, and the optimization ability has a good linear relationship with the value of inertia factor. Dynamic inertia factor can obtain better optimization results than fixed values because dynamic inertia factor can change linearly in PSO search [16] and can also change dynamically according to a measure function of PSO performance.

$$v_i = w \times v_i + C_1 \times \text{rand}(\,) \times (p - x_i) + C_2 \times \text{rand}(\,) \times (g - x_i). \tag{3}$$

Linear decreasing weight is as follows:

$$w^t = \frac{(w_{\text{ini}} - w_{\text{end}})(G_k - g)}{G_k + w_{\text{end}}}, \tag{4}$$

where G_k is the maximum number of iterations; W_{ini} the is initial inertia weight, typically 0.9; and w_{end} is the inertia weight of iteration to maximum evolutionary algebra, with the typical value being $w_{\text{end}} = 0.4$. Global PSO and local PSO: The former can get a relatively fast convergence rate in the global range, but its disadvantage is that it sometimes falls into local optimization. The latter has slow convergence speed, but it is difficult to optimize locally. In common practical applications, we generally use particle swarm optimization algorithm to find a similar result, then determine that it is the most suitable in the global range, and then use local optimization particle swarm optimization algorithm to search for specific local areas.

2.4. Algorithm Instance Steps. Step 1: assume that the population size is N. The velocity vector and position vector of each connection are initialized arbitrarily in the specified search space, the adaptive function value is calculated, and then the forward optimal position of particles in the population and the global optimal position of the middle population are obtained. Step 2: update the velocity and position of particles. The two parameters are updated according to the historical optimal position and the global left and right position. Step 3: select the fitness function of

particles, and then evaluate its value. Then, it is necessary to update the historical optimal position and the global optimal position, mainly depending on whether they are different from the previous comparison, and if they are different, they will be updated. Step 4: this step belongs to judgment. If the end condition is met, it will be output. If it is not met, it will continue to be executed in Step 2.

The PSO algorithm flow chart is as shown in Figure 1.

2.4.1. Basic Concepts. SVM is a clear two-class classification model [17, 18], and at the same time, it is an excellent linear classifier, which can solve many linear problems. We can clearly see that it has the largest interval in feature space. The learning goal of an excellent linear classifier is to find a hyperplane in n-dimensional data space, and its equation is $w^T x + b = 0$. Its learning strategy is to maximize the interval. If you want to solve it simply, the interval can finally be transformed into a convex quadratic programming problem.

2.4.2. Solution of Regression Linear Equation. Convert $f(x) = w^T x + b$ to optimization:

$$\begin{aligned} &\max \frac{1}{\|w\|} \\ &\text{s.t.} \quad y_i\left(w^T x_i + b\right) \geq 1, \quad i = 1, \ldots, n. \end{aligned} \tag{5}$$

Since obtaining the maximum of $1/\|w\|$ can be equivalent to obtaining the minimum of $0.5\|w\|^2$, the corresponding objective function can be equivalent to

$$\begin{aligned} &\min \frac{1}{2}\|W\|^2 \\ &\text{s.t.} \quad y_i\left(w^T x + b\right) \geq 1, \quad i = 1, \ldots, n. \end{aligned} \tag{6}$$

By transforming the Lagrange duality into a problem where the dual variables are as large as the optimization problem, the optimal solution of the original problem is obtained by solving a problem equivalent to the original problem. Define Lagrange functions:

$$\begin{aligned} L(w, b, a) &= \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} a_i\left(y_i\left(w^T x + b\right) - 1\right), \\ L(w, b, a) &= \frac{1}{2}\sum_{i,j=1}^{n} a_i a_j y_i y_j x_i^T x_j - b\sum_{i=1}^{n} a_i y_j + \sum_{i=1}^{n} a_i, \\ &= \sum_{i=1}^{n} a_i - \frac{1}{2}\sum_{i,j=1}^{n} a_i a_j y_i y_j x_i^T x_j. \end{aligned} \tag{7}$$

From the last formula, we can see that the function only contains one variable, a_i, at this time, and if we find its value, we can find w and b; then, we can classify the function:

$$f(x) = w^T x + bs. \tag{8}$$

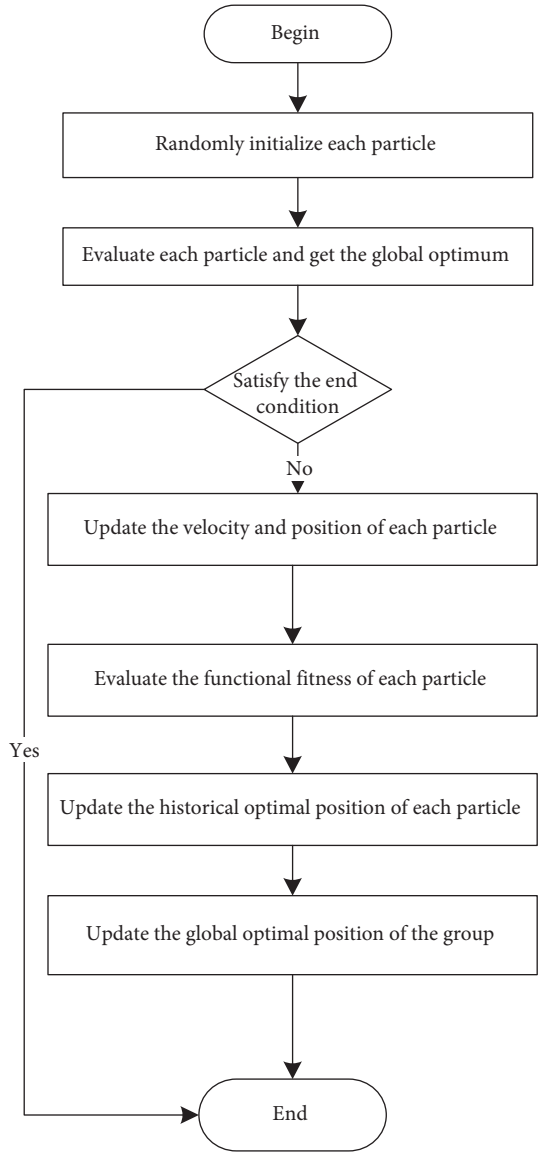


Figure 1: Algorithm flowchart.

By introducing kernel function, it can simplify the inner product operation in mapping space, which is equivalent to the original classification function:

$$f(x) = \sum_{i=1}^{n} a_i y_i \langle x_i, x \rangle + b, \tag{9}$$

mapped to

$$f(x) = \sum_{i=1}^{n} a_i y_i \langle \phi(x_i), \phi(y_i) \rangle + b, \tag{10}$$

where the kernel function is $k(x_1, x_2) = (\langle x_1, x_2 \rangle + 1)^2$.

By introducing the relaxation variable ξ_i, the objective function becomes

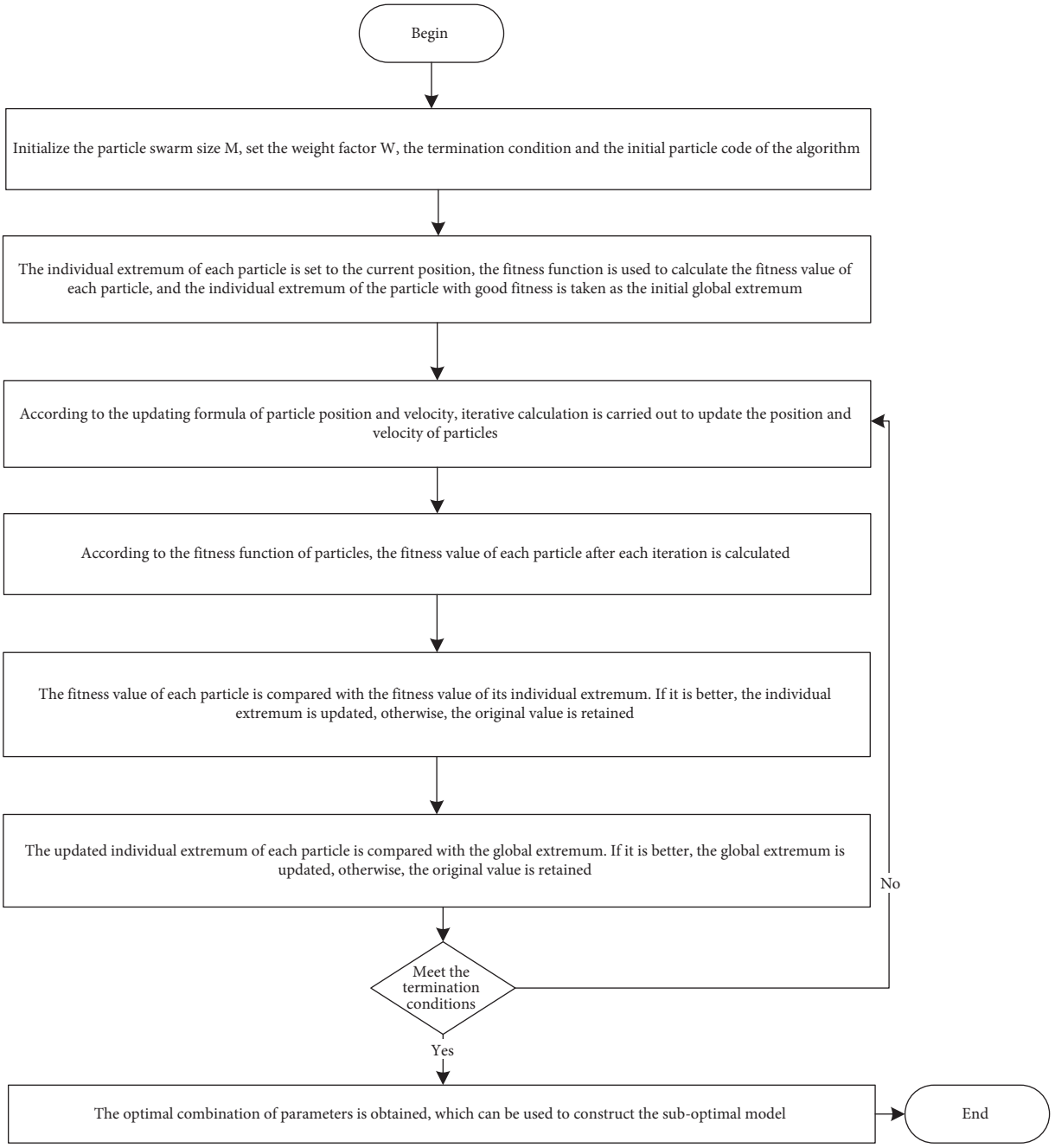


Figure 2: Flowchart of parameter optimization steps.

$$\min \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{n}\xi_i \quad \text{s.t.} \quad y_i\left(w^T x_i + b\right) \geq 1 - \xi_i, \quad i = 1, \ldots, n, \tag{11}$$

where C is the learning factor.

3. PSO-SVM

3.1. Basic Overview and Formula. The basic formula of the algorithm is

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{n} a_i^* y_i k\left(x_i, x_j\right) + b^*\right). \tag{12}$$

PSO-SVM algorithm [19, 20] needs to use particle thermal optimization algorithm to project low-dimensional plane points onto high-dimensional plane, optimize SVM kernel parameters and penalty coefficient C, and introduce linear separable data kernel function K into low-dimensional plane. The classification decision function is as follows:

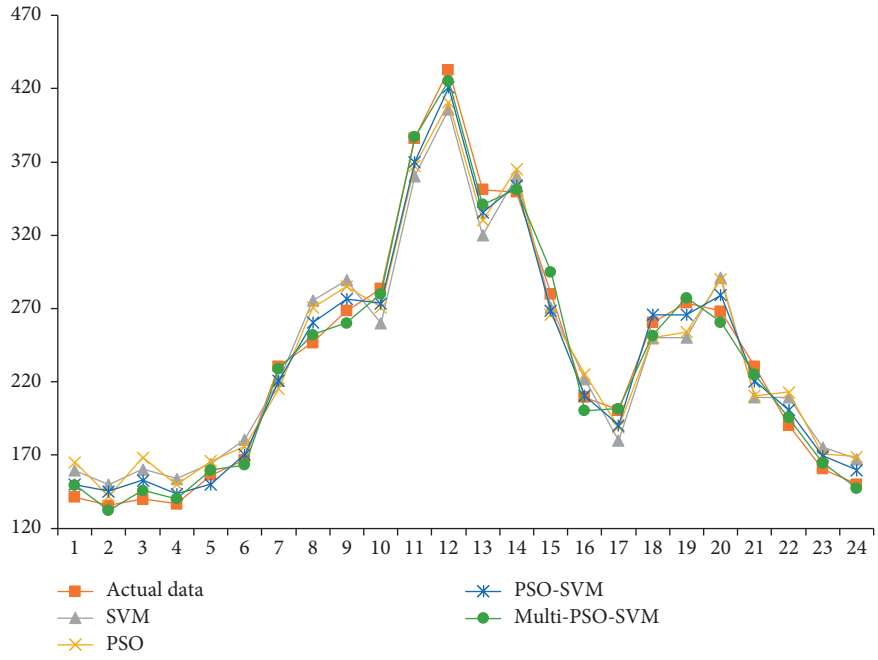


Figure 3: Comparison chart between several algorithm models and actual load on February 1, 2020.

$$f(x) = \text{sign}\left(\sum_{i=1}^{n} a_i^* y_i \exp\left(-\frac{\|x - z\|_2^2}{2\sigma^2} + b^*\right)\right). \quad (13)$$

The kernel function is $k(x, z) = \exp(-\|x - z\|_2^2/2\sigma^2)$, C is the penalty factor. The selection of kernel function parameters will affect the performance of the model. PSO and SVM, which are simple to operate and fast to converge, are combined to find the optimal combination of parameters. In the SVM classification, because of the training, the parameters in the optimization process are more complex, resulting in low efficiency in the classification process, and combined with PSO, the optimized SVM in the training can optimize the parameters and achieve better classification results.

3.2. Parameter Optimization. The flow chart of parameter optimization steps is as follows: the termination condition refers to the maximum number of iterations, the convergence of the obtained solution, or whether the obtained solution has achieved the expected effect as shown in Figure 2.

3.3. Fitness Value. This fitness value can measure the quality of a solution [21]. There is a noun called maximum fitness evaluation, which is updated according to different population variations, which can be described in relation to the maximum iteration times:

$$\text{maximum fitness evaluation} = \text{maximum iteration times} * \text{population number}. \quad (14)$$

In the iterative process, it is calculated once, that is, called once, and evaluated once. The maximization and minimization of this value are related to the product, which is the product between the weight and fitness values to handle the update. By comparing fitness values of different sizes, this can be simply described as a linear relationship. If the fitness of each element is equal, the higher the fitness, the better.

4. Experimental Simulation

The data selected in this paper is the power load values of a specific area in Zhejiang Province. The following three time points in 2020 are taken as three sampling points, and the SVM algorithm [22], PSO algorithm [23], PSO-SVM algorithm [24], and Multi-PSO-SVM forecasting model on February 1, March 1, and April 1 are compared with the actual load values, as shown in Figures 3–5.

Through the above three charts, we can find that PSO-SVM prediction is better than Multi-PSO-SVM prediction in these three sampling points, which shows that they have the best prediction effect.

The average values of load sampling data at the above three sampling points are shown in Table 1. Through specific data, the prediction accuracy and effect of the above models can be analyzed more accurately.

As can be seen from Table 1, Multi-PSO-SVM forecasting model is used to forecast the microgrid load, and the forecasting accuracy completely meets the standard. By comparing the results of the forecasting model with the other three forecasting models, we can know that the forecasting error of Multi-PSO-SVM forecasting model is the smallest; in addition, the average load predicted by it, along with the actual load, is almost less than 10 MV, so the

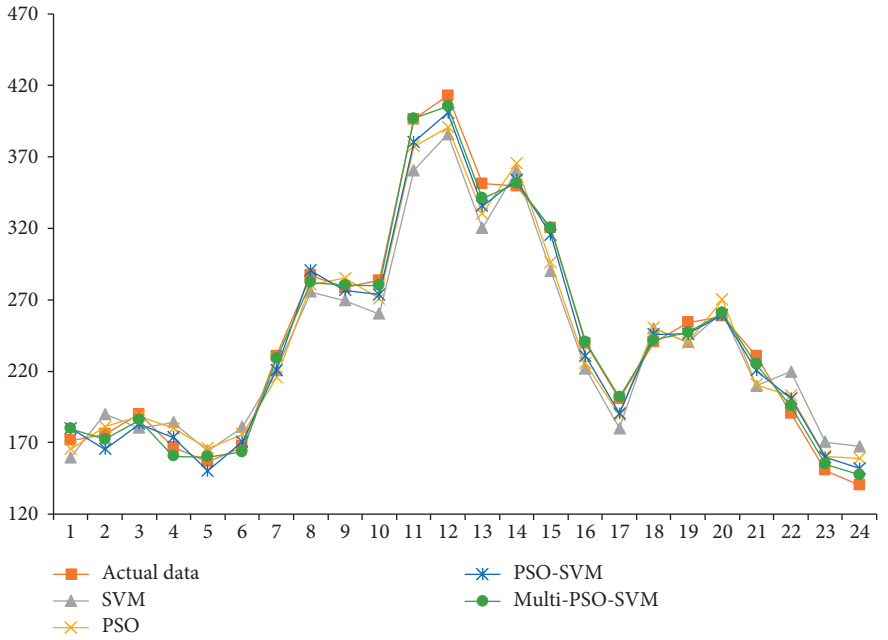


Figure 4: Comparison chart between several algorithm models and actual load on March 1, 2020.

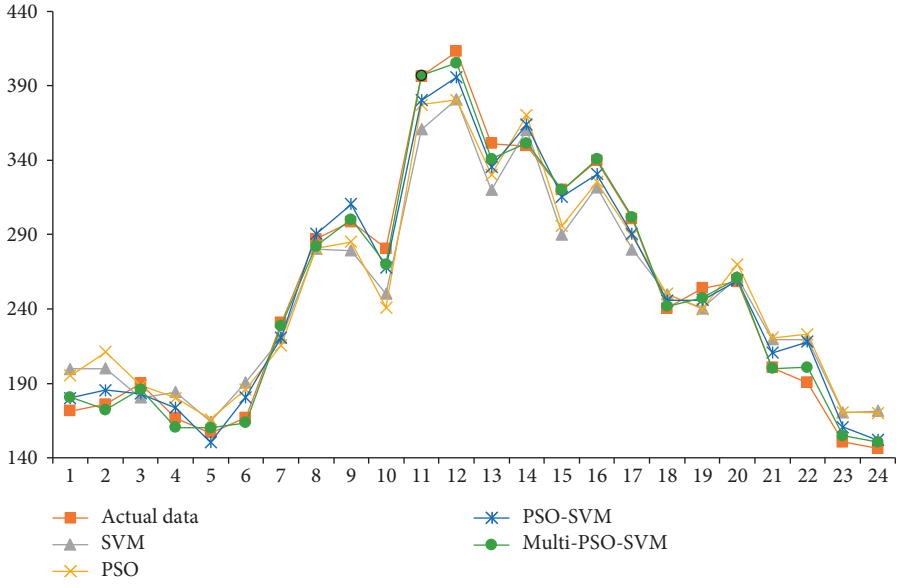


Figure 5: Comparison chart between several algorithm models and actual load on April 1, 2020.

Table 1: The average value of load sampling data.

Time	Actual	SVM	PSO	PSO-SVM	Multi-PSO-SVM
1	181.37	179.78	179.32	180.076	180.78
2	175.89	179.98	181.09	180.55	172.32
3	190	180.34	188.56	182.94	185.9
4	166.78	184.09	180.089	173.66	160.45
5	156.55	164.33	165.99	150.34	160.03
6	167	160.77	175.98	170.78	163.54
7	230.65	220.72	215.56	220.77	228.83
8	286.95	280.43	280.59	290.45	282.24
9	298.62	289.41	290.11	300.51	295.12
10	300.55	310.08	310	307.71	300
11	296.34	320.67	317.45	300.32	297.01
12	312.98	321.99	320.56	315.87	315.46
13	321.23	310	310.19	315.55	320.97
14	389.56	401.34	400.23	394	391.42
15	420	410	410	415.45	420
16	339.46	321.78	325.38	330.57	340.56
17	360.67	350	351.27	357.45	360
18	240.38	250	250.32	245.77	241.78
19	247.28	240.36	240	245.99	247.28
20	258.36	261.43	270	259.32	260.76
21	200.55	219.65	220.67	210.54	200
22	190.56	219.65	222.98	217.87	200.7
23	150.65	170.45	170.54	160.72	154.87
24	146.3	171.43	170	151.98	150.56

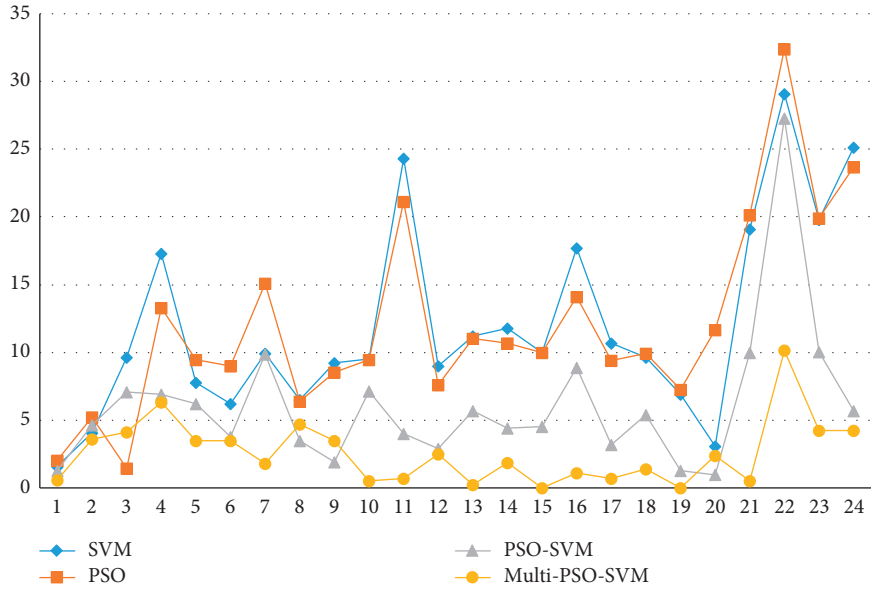


Figure 6: Error comparison of several algorithms.

Multi-PSO-SVM proposed in this paper has more advantages. The specific prediction error is shown in Figure 6.

The maximum error of the proposed method is 10, the minimum error is 0, and the overall performance is good. The time complexity of the proposed method is $O(n \log n)$.

5. Conclusion

This paper presents a superior short-term load forecasting model for microgrid, Multi-PSO-SVM prediction model; after analyzing the great importance of charge prediction to

microgrid, considering the influence of different time periods and temperatures on the microgrid, four algorithms are selected to forecast the load value. Through the combination of different algorithms, the results show that the prediction accuracy of the Multi-PSO-SVM model proposed in this paper is obviously better than that of the other three algorithms, and the model has strong accuracy and strain capacity.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] A. T. Eseye, J. Zhang, and D. Zheng, "Short-term photovoltaic solar power forecasting using a hybrid wavelet-PSO-SVM model based on SCADA and meteorological information," *Renewable Energy*, vol. 118, no. APR, pp. 357–367, 2017.

[2] L. Ning and L. Ying, "On predicted research methods of supply capacity of micro-grid based on improved particle swarm optimization," *Journal of Computers*, vol. 8, no. 10, pp. 2789–2795, 2013.

[3] H. Ma and J. M. Tang, "Short-term load forecasting of microgrid based on chaotic particle swarm optimization," *Procedia Computer Science*, vol. 166, pp. 546–550, 2020.

[4] Y. Feng, P. Zhang, M. Yang, Q. Li, and A. Zhang, "Short term load forecasting of offshore oil field microgrids based on DA-SVM," *Energy Procedia*, vol. 158, pp. 2448–2455, 2019.

[5] A. V. Phan, M. L. Nguyen, and L. T. Bui, "Feature weighting and SVM parameters optimization based on genetic algorithms for classification problems," *Applied Intelligence*, vol. 46, no. 2, pp. 455–469, 2017.

[6] H. L. Chen, B. Yang, S. J. Wang et al., "Towards an optimal support vector machine classifier using a parallel particle swarm optimization strategy," *Applied Mathematics and Computation*, vol. 239, no. 8, pp. 180–197, 2014.

[7] H. Sheng, J. Xiao, J. Jia, and X. Deng, "Estimation method for state of charge based on least square support vector machine," *Taiyangneng Xuebao/Acta Energiae Solaris Sinica*, vol. 36, no. 6, pp. 1453–1458, 2015.

[8] L. Ke, L. Zhang, W. Xu, G. Liang, and J. Zhang, "Asset quality evaluation of power grid enterprise based on LSSVM and PSO," *International Journal of Science*, vol. 3, no. 10, pp. 168–171, 2016.

[9] G. I. Sayed, M. Soliman, and A. E. Hassanien, "Modified optimal foraging algorithm for parameters optimization of support vector machine," *The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2018)*, vol. 723, pp. 23–32, 2018.

[10] Y. Zou, S. Li, Y. Dong, and Y. Niu, "A two-stage economic optimization and predictive control for EV microgrid," in *Proceedings of the IECON 2018—44th Annual Conference of the IEEE Industrial Electronics Society*, pp. 5951–5956, Washington, DC, USA, October 2018.

[11] V. Indragandhi, R. Logesh, V. Subramaniyaswamy, V. Vijayakumar, P. Siarry, and L. Uden, "Multi-objective optimization and energy management in renewable based AC/DC microgrid," *Computers & Electrical Engineering*, vol. 70, pp. 179–198, 2018.

[12] Y. Zhong, M. Huang, and C. Ye, "Multi-objective optimization of microgrid operation based on dynamic dispatch of battery energy storage system," *Dianli Zidonghua Shebei/Electric Power Automation Equipment*, vol. 34, no. 6, pp. 114–121, 2014.

[13] Y. Tian, L. Wan, and B. Ye, "Multi-objective optimization of high altitude sector operation based on environmental protection," *Cluster Computing*, vol. 22, no. S6, pp. 15429–15437, 2019.

[14] Y. Zhao, X. Song, F. Wang, and D. Cui, "Multiobjective optimal dispatch of microgrid based on analytic hierarchy process and quantum particle swarm optimization," *Global Energy Interconnection*, vol. 3, no. 6, pp. 562–570, 2020.

[15] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.

[16] X. Chen, H. G. Yang, and J. Q. Peng, "Energy optimization of high-rise commercial buildings integrated with photovoltaic facades in urban context," *Energy*, vol. 172, pp. 1–17, 2019.

[17] G. Chen and S. Li, "Research on location fusion of spatial geological disaster based on fuzzy SVM," *Computer Communications*, vol. 153, pp. 538–544, 2020.

[18] K. Shankar, S. K. Lakshmanaprabu, and D. Gupta, "Optimal feature-based multi-kernel SVM approach for thyroid disease classification," *The Journal of Supercomputing*, vol. 76, no. 28, pp. 1–16, 2020.

[19] K. Shankar, S. K. Lakshmanaprabu, D. Gupta, A. Maseleno, and V. H. Albuquerque, "Regression analysis of friction resistance coefficient under different support methods of roadway based on PSO-SVM," *Journal of Physics: Conference Series*, vol. 1941, no. 1, Article ID 012046, 2021.

[20] Y. Wang, D. Wang, and Y. Tang, "Clustered hybrid wind power prediction model based on ARMA, PSO-SVM and clustering methods," *IEEE Access*, vol. 8, p. 1, 2020.

[21] B. Zou, Y. Qiu, and L. Zhang, "Docked ships detection using PolSAR image based on GOPSO-SVM," in *Proceedings of the 2019 IEEE Radar Conference (RadarConf)*, pp. 1–6, Boston, MA, USA, April 2019.

[22] S. K. Shevade, S. S. Keerthi, C. Bhattacharyya, and K. R. K. Murthy, "Improvements to the SMO algorithm for SVM regression," *IEEE Transactions on Neural Networks*, vol. 11, no. 5, pp. 1188–1193, 2000.

[23] G. Cui, Y. Niu, W. Yanfeng, Z. Xuncai, and P. Linqiang, "A new approach based on PSO algorithm to find good computational encoding sequences," *Progress in Natural Science*, vol. 06, pp. 90–94, 2007.

[24] M. Wang and X. W. Li, "Power cable fault recognition using the improved PSO-SVM algorithm," *Applied Mechanics and Materials*, vol. 427–429, pp. 830–833, 2013.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 3168464, 10 pages
https://doi.org/10.1155/2021/3168464

Research Article

A Shadow Capture Deep Neural Network for Underwater Forward-Looking Sonar Image Detection

Taowen Xiao,[1] Zijian Cai,[1] Cong Lin,[1] and Qiong Chen[2]

[1]*College of Electronics and Information Engineering, Guangdong Ocean University, Zhangjiang 524025, China*
[2]*Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing 100084, China*

Correspondence should be addressed to Cong Lin; lincong@hainanu.edu.cn and Qiong Chen; qiongchen@mail.tsinghua.edu.cn

Received 25 October 2021; Accepted 7 December 2021; Published 30 December 2021

Academic Editor: Han Wang

Image sonar is a widely used wireless communication technology for detecting underwater objects, but the detection process often leads to increased difficulty in object identification due to the lack of equipment resolution. In view of the remarkable results achieved by artificial intelligence techniques in the field of underwater wireless communication research, we propose an object detection method based on convolutional neural network (CNN) and shadow information capture to improve the object recognition and localization effect of underwater sonar images by making full use of the shadow information of the object. We design a Shadow Capture Module (SCM) that can capture the shadow information in the feature map and utilize them. SCM is compatible with CNN models that have a small increase in parameters and a certain degree of portability, and it can effectively alleviate the recognition difficulties caused by the lack of device resolution through referencing shadow features. Through extensive experiments on the underwater sonar data set provided by Pengcheng Lab, the proposed method can effectively improve the feature representation of the CNN model and enhance the difference between class and class features. Under the main evaluation standard of PASCAL VOC 2012, the proposed method improved from an average accuracy (mAP) of 69.61% to 75.73% at an IOU threshold of 0.7, which exceeds many existing conventional deep learning models, while the lightweight design of our proposed module is more helpful for the implementation of artificial intelligence technology in the field of underwater wireless communication.

1. Introduction

The electromagnetic wave, light wave, and sound wave are commonly used in the world as the wireless communication carrier. In the field of wireless communication, seawater contains a variety of minerals and becomes a conductive medium. The electromagnetic wave propagating to the ocean will be blocked by seawater, resulting in the rapid waste of most energy. Therefore, the electromagnetic wave propagating in seawater will be greatly suppressed. Based on the above analysis, the electromagnetic wave transmitting in the seawater will be greatly refrained, because most of its energies will be wasted quickly and spreading to ocean will be blocked by the sea water. Light is also an electromagnetic wave in essence, so its transmission in seawater will also be limited [1–3]. Since the reason that the sound wave is a mechanical wave that can travel in the elastic medium of seawater without much interference, it is widely used in the field of underwater wireless communication. Inspired by sound waves, many sonar devices have been developed to measure depth of water or detect underwater objects. For example, forward-looking sonar equipment can obtain the reflection information of sound waves and generate high-resolution sonar images, which is usually used to obtain underwater information in the form of images.

Sonar image object recognition methods are mainly divided into traditional mathematical modelling methods and detection methods based on convolutional neural networks (CNN). Traditional mathematical modelling methods can handle more sonar object recognition tasks,

and it uses image process methods such as Scale Invariant Feature Transform (SIFT), Directional Gradient Histogram (HOG) [4], and Fisher vectors to extract object features and then pass machine learning or pattern matching classifies object features [5–7]. The features extracted by this method can perform well for specific data sets and tasks, but the generalization ability of most features is limited, and feature extraction still needs professional knowledge and a lot of experiments. In contrast, the convolutional neural network (CNN) optimizes its parameters with back gradient propagation, which makes it possible to combine feature extraction and model prediction into the same pipeline.

As a powerful image classification and object detection model, the detection method based on CNN is one of the most popular deep learning structures [8–12]. Compared with traditional computer vision methods, the CNN method is more widely used in the field of sonar image recognition since the advantages of automatic feature extraction and multilevel feature extraction [13–20]. There are two types of application of CNN model. One is to combine the convolution module or the fully connected layer module to build a sonar image recognition model. Matias [13] and Valdenegro-Toro [14] both designed a CNN model for object recognition in sonar images, which achieved better results than pattern matching methods. Williams [15] designed a deep convolutional neural network and applied it to multiple binary classification tasks to distinguish different categories in sonar data sets, which is better than the original classifier based on manual features. The other type of CNN model is to use advanced theory and experience in computer vision (CV) to modify some excellent models from CV to meet the needs of sonar image recognition. Galusha et al. [16] use deep convolution neural network for object detection such as object recognition and location, and the pixels of SAS image are reduced to RoIs (region of interests). This detection form resembles the standard detection stream of the two-stage model of object detection. Breistein [17] and Neves et al. [18] introduced a one-stage object detection model yolov2 for sonar image recognition, and the model achieved expected results in their respective data sets. Zacchini et al. [19] introduced Mask RCNN for sonar image recognition and object localization and tested the functions of the model at LaSpezia Naval Support and Laboratory Center. Fan et al. [20] reduced network parameters by modifying the network structure of the model without affecting accuracy.

The above CNN methods have different emphasis on using deep learning models. Although good results have been achieved in their respective data sets, they can only locate and recognize one sonar object due to the small number of data sets in most cases. How to effectively increase the feature gap between different categories and narrow the feature gap between the same category is particularly important. In addition, we noticed that the sound waves emitted by the forward-looking sonar device will bounce off the objects it touches and cannot reach the area behind the object, thus forming a shadow related to the object shape on the generated sonar image. It is difficult for traditional image processing methods to interpret the highlight area representing the object due to the above-mentioned imaging defects of sonar images, so shadow information is often introduced as an additional auxiliary identification feature [21–24]. Although shadow feature has been given considerable attention in previous recognition tasks, the methods proposed by them are exclusive, which means that it is impossible to separate the extraction steps of shadow features and transplant them to the deep learning method. Therefore, researchers have not paid enough attention to the shadow information of sonar images in the research of existing deep learning methods.

In order to make full use of the shadow information in the forward-looking sonar image and improve the accuracy of image object recognition and positioning, a neural network detection method based on shadow capture module (SCM) is designed. SCM module can capture the shadow features in the image according to the characteristic that the shadow appears directly above the object, which is very lightweight and portable. In addition, we made a data set with shadow as the object to train and test the ability of the original model to distinguish shadows. Through a large number of experiments on underwater sonar images, it is found that the detection accuracy of the network model added with SCM module is 6.12% higher than the original model, which exceeds the existing common CNN model. The main contributions of this paper can be summarized as follows:

(1) In order to utilize the shadow information in the forward-looking sonar image, we propose a structure based on CNN model that can capture the shadow information and integrate the shadow information into the feature map.

(2) We made a data set with the object shadow to train the proposed model and then used them to identify and locate the shadow of the object. The final result showed that the detection accuracy reached 94%, which proved that the CNN model can also adapt to sonar image shadow recognition and provides a priori basis for our method.

(3) The proposed module SCM in this paper can effectively increase the difference degree of features between categories of deep learning model in multicategory detection tasks. After the addition of the designed module, the accuracy of CNN model is improved by 6.12%, which is higher than the mainstream object detection model in CV field.

2. Shadow Capture Network

There will be a shadow associated with its shape directly above the target object in the forward-looking sonar picture. It is an undetected area formed by the object blocking the sound waves emitted by the sonar device. The shape of this area varies with the shape of the object. Although the angle at which the sonar device emits sound waves will affect the shape of the shadow to a certain extent, the only thing that plays a dominant role in the shape of the shadow is the shape of the object when the emission angle is basically constant.

Therefore, in the sonar picture generated by a forward-looking sonar device emitting sound waves at a stable angle, the shadow can be regarded as another characteristic expression of the object.

In this section, we describe the rationale for proposing this structure and the details of its construction. As with CenterNet [25], our model eventually outputs a heat map that contains both the classification prediction score of the object and the coordinates of the object's position in the heat map. Since there is a one-to-one mapping relationship between the object points in the heat map and the original image, the position of the object points in the heat map relative to the heat map represents the position of the object in the original image relative to the original image. The use of shadow features in the forward-looking sonar image can serve to increase the feature differences between categories and improve the response values of object points in the heat map, thus improving the network's ability to recognize objects. Since the classification scores of hotspots are generated simultaneously with the hotspot coordinates, a higher hotspot response value also means that the network is more certain about the location of the object, so the optimization of the classification effect also represents an enhancement of the localization effect. Therefore, capturing shaded features and fusing them into the feature maps in the network can fully improve the prediction of heat maps by our model.

Therefore, we replace the module in CenterNet that is used to predict the object with a feature capture module that captures the shadow features in the feature map, and also fuse the captured shadow features into the feature map of the backbone network. Then, a module for predicting the heat map alone is added after the new module, which uses the fused feature map for predicting the heat map, serving to enhance its own prediction effect.

2.1. Overview. As shown in Figure 1, we design a structure that can utilize shadow features. This structure is used to capture shadow features from the feature map and fuse them into the feature map containing the object, which serves to add features to the object. In our designed model, we use Hourglass [26] as the feature extraction network for this model, unlike CenterNet, we use nonstacked Hourglass. There are two reasons for this choice. First, the stacked Hourglass network is too deep, which will lead to extreme abstraction of the feature layer that will eventually be used as a prediction, and its large cut-off from the feature map in the front part of the network, which will cause difficulties in the design and interpretation of the model. Second, each stacked Hourglass has a different region of interest on the whole picture. Multilevel stacking of Hourglass networks will make each Stack Hourglass network has its own distinct region of interest, i.e., each different Hourglass network needs to add intermediate supervision. And this increases the uncertainty of module design. For these two reasons, we decided to use a nonstacked Hourglass network. The feature map output from this network will be used as the input of the designed module. After the feature map enters the shadow capture module (SCM), it passes through two branches, the first one is the shadow capture branch, we first use three parallel convolutional layers to obtain the initial position parameters of the object in the feature map, and then we capture the features based on the position parameters and our manually designed capture method, using the RoIs (Region of Interests) Align pool to extract the region of interest of the model for shadows, which is stripped from the spatial dimension to the channel dimension. In the second branch, the feature fusion branch, we concatenate the shadow region of interest with the feature map output by hourglass and pass through a convolution layer to complete the fusion of features. Once the fusion is complete, we feed this enhanced feature map into the heat map prediction module added at the end of the model to obtain more accurate object locations and classification scores.

2.2. Shadow Semantic Feature Capture Module. In this section, we will introduce the shadow semantic feature capture module in detail. First, in Figure 2, three parallel predictive convolution layers are set up to obtain the location parameters $\theta(x_c, y_c, w_{\text{obj}}, h_{\text{obj}})$ of the object, where x_c, y_c represents the central point coordinates of the object in the feature map and $w_{\text{obj}}, h_{\text{obj}}$ represents the width and height of the object in the feature map. In view of the imaging characteristics of the experimental data set (see the section Data sets and labels for detail), it can be determined that the shadow of the object in the sonar image generally exists directly above the object. Therefore, we design a capture method for shadow features so that no additional supervision information is required. The rule will eventually capture the shadow features in a selected region on the feature map, and the rule is as follows:

(1) The coordinates of the upper-left corner of the object (x_l, y_l) are obtained from the position parameters of the object, then the height H of the region is $y_l - y_L$, y_L is the vertical coordinate 0 of the upper-left pixel point of the feature map, and the width W is the width w_{obj} of the object.

(2) By observing the shadow images in the data sets, we found that the width of some shadows is slightly larger than the object, and there is also a small skewing of the shadows to the sides in the data sets, so we introduced a width parameter α to adjust the width of this region. The final width of this region should be $\alpha * w_{\text{obj}}$, as shown in Figure 3.

Finally, after the predetermined area is obtained, RoI-Align Pool is used to cut it. For detailed operation and gradient back propagation form, please refer to [27, 28]. The pool can be used to obtain the high response value in the predetermined area and output a feature map with the same size as the original feature map and containing the shaded high response value. Then, the newly generated feature map will be input into the subsequent semantic feature fusion module for feature fusion.

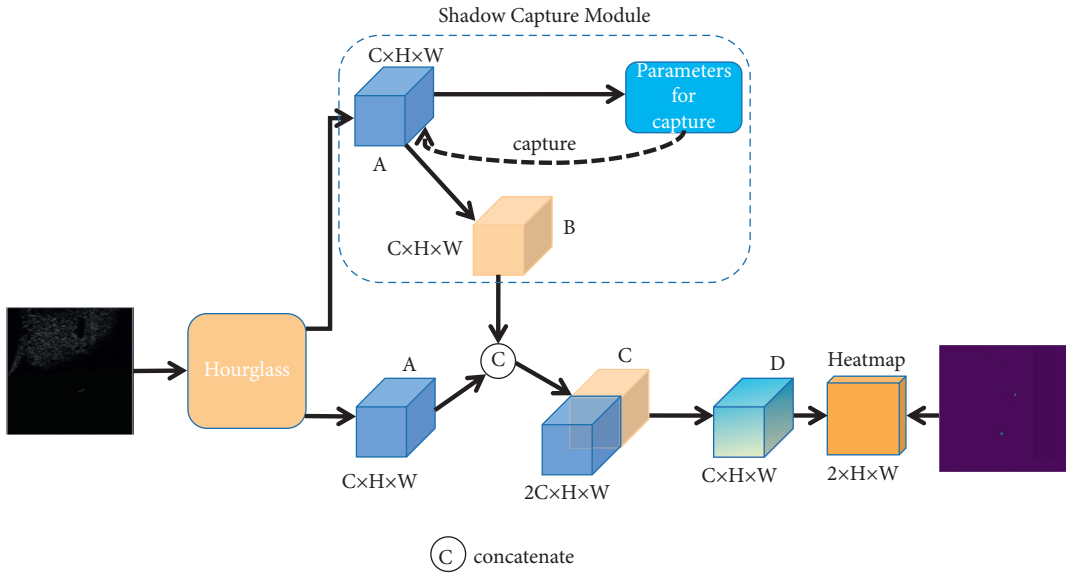


Figure 1: Feature map A generated by Hourglass network is used to predict capture parameters. The obtained parameters are used to capture shadow features in feature map A and then shadow feature map B is obtained. Then, feature figure A and shadow feature figure B are spliced to obtain spliced feature figure C Then, the spliced feature figure C is sent to the fusion module for feature fusion to obtain fusion feature figure D. Finally, the fusion feature figure D is used to predict the final Heatmap.

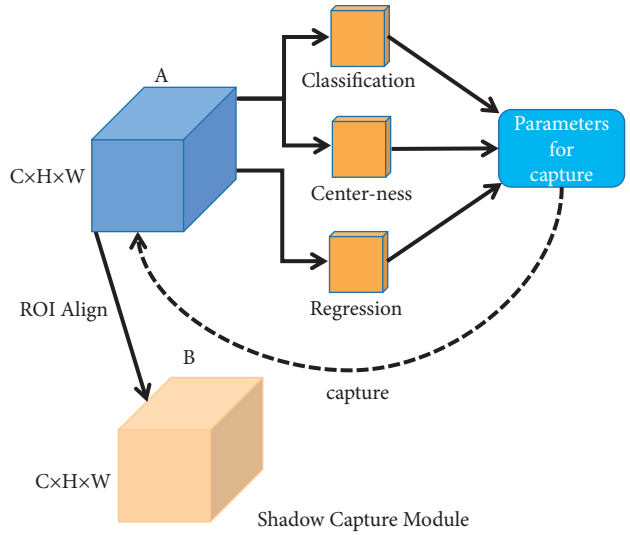


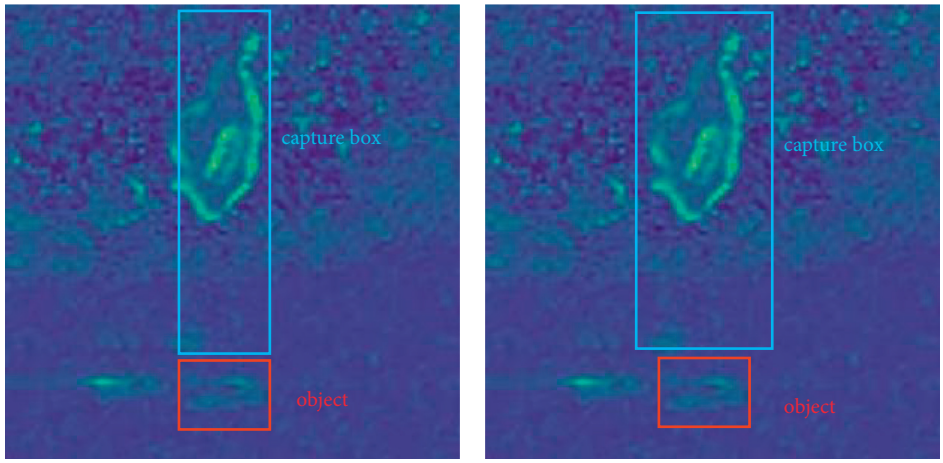
Figure 2: The structure of shadow capture module.

2.3. Semantic Feature Fusion Module. This module takes the shared convolutional feature map output from Hourglass network and the feature map output from the shadow semantic feature capture module containing the high response value of the shadow as input and fuses the two. Finally, the enhanced feature map integrated with the shadow semantic feature is used for the prediction of subsequent heat maps.

As shown in Figure 1, after the shared convolutional feature map is input into the module, it waits for the shadow semantic feature capture module to output the feature map containing the high response value of the shadow and then uses concatenate operation for both and obtains the final feature map C through a fusion function $H(\cdot)$:

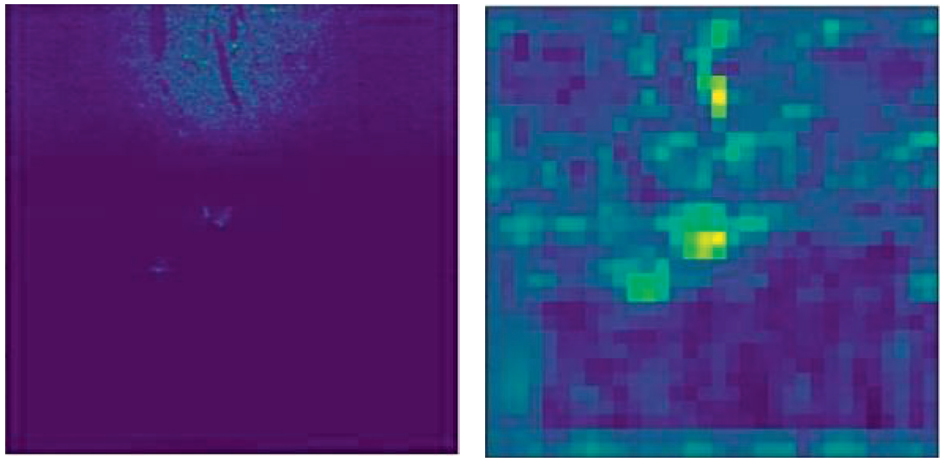
$$C = H(A, B), \tag{1}$$

where $H(\cdot)$ is a mixture of three consecutive operations: 3×3 convolutional layer, Batch Norm [29], and nonlinear activation function ReLU. 3×3 convolutional layer is used to fuse feature map B, containing the shaded high response values, with the feature map of the Hourglass network output A. The Batch Norm of [29] normalizes the values of the output of the convolutional fusion, which can alleviate the Internal Covariate Shift phenomenon (i.e. after each parameter iteration update, the output data of the previous layer will change in data distribution after being computed by this layer of the network, making it



(a) (b)

Figure 3: Capture schematic diagram of shadow features. (a) The capture schematic diagram without adjusting the cutting width and (b) capture schematic diagram after adjusting the cutting width.



(a) (b)

Figure 4: The above two figures are (a) the original figure (processed for convenient observation) and (b) the feature figure in the neural network.The highlighted area in the neural network feature map is the area of interest of the network, as shown in Figure (b). Both shadows and objects are highlighted in the feature map.

difficult for the learning of the next layer), that is inherently present in neural networks [29] and the changes to the model data distribution brought by our use of the concatenate fusion form. The nonlinear activation function ReLU can suppress and activate the fused features with smoothness.

2.4. Interpretability of Modules. In this section, we take a mathematical intuition at the model to better understand why shadow capture module (SCM) is able to incorporate the semantic features of shadows. At the beginning of Section 2, we explained that the effect of model positioning and recognition depends on the response value of pixels in the final feature image. That is, we hope that the final output of the model can have a higher response value to the actual location of the sonar object, and the pixel value of other locations is close to 0 as far as possible. In Figure 4, we found that the feature map also has highlight in shadow areas after visualizing and analyzing. But there are only two goals in the sonar image, if the shadow area still has a high response value, it will increase the fitting pressure of final prediction model and make the model more likely to cause the error detection. Therefore, we hope to use the information of the high-response area in the shadow to improve the response effect of the object area, so as to help the model classification. In order to achieve this goal, it is necessary to strip the high-concern area of the shadow feature from the spatial dimension and then splice the high-concern area to the channel dimension of the feature. After the high-concern areas of shadow features are stripped from the spatial dimension, the values of each pixel in the final feature map should be constituted as follows (taking 1×1 convolution as an example):

$$F_{ji} = W_j * X_{ji}, \tag{2}$$

where j represents different channels and i represents pixel values at different positions on the same channel. W_j is the weight parameter of the filter in the j channel of the input feature map, X_{ji} represents the pixel value of i position in the j channel of the input feature map, and F_{ji} represents the pixel value of i position in the j channel of the output feature map. When the high-response area of the shadow is stripped to the channel dimension, the model can utilize this high-concern area by itself according to the final detection requirements. Since the filters of convolutional neural network in each channel are different, it also means that the parameters of filters are different and are selected by the model, so the filter parameters can be regarded as the model's emphasis on this part of the highlighted region. With the learning of the reverse gradient propagation method, the convolution can learn the situation that is most suitable for the fusion of the feature pixels of each channel.

3. Experiment

In order to evaluate the proposed method, we conducted a comprehensive experiment on the underwater forward-looking sonar data set provided by Pengcheng Laboratory. The experimental results show that our model achieves good performance when the IOU threshold is 0.5–0.8, using PASCAL VOC 2012 as the evaluation standard. In the following sections, we will first describe the data set and the details of the experimental implementation, and then we will present the experimental results of a series of ablation experiments performed on the forward-looking underwater sonar data sets.

3.1. Data Sets and Labels. This data set is the largest and most extensive acoustic image data set in the current industry launched by Pengcheng Laboratory. The data sets have a total of 5000 images, including 3200 training sets, 800 validation sets, and 1000 test sets. The object types include cube, ball, cylinder, human body, tire, circle cage, square cage, metal bucket, and so on. Each image only marks the relevant object, and the shadow does not have any marking information. The data acquisition equipment used for the sonar images was the Tritech Gemini 1200I multibeam forward-looking sonar. The detection beam emitted by the equipment is horizontally divided into several fan-shaped beams with vertical opening angle φ, each beam irradiates the object as shown in Figure 5 and forms a set of distance intensity information, and the echo intensity information of all the beams is arranged according to position relation to form sonar image, as shown in Figure 6. The number of beams corresponds to the number of horizontal pixels of the image. When the horizontal angle θ is constant, the more the beams, the higher the angular resolution, the amount of echo data collected by each beam corresponds to the number of vertical pixels of the image, the larger the amount of data, the higher the distance resolution. The sonar data set is Cartesian (after a rectangular coordinate system) so that the image appears as a rectangle rather than a fan. As shown in Figure 6, the pixel (x, y) in the sonar image represents the acoustic reflection intensity information at direction $\theta = (W/2 - x) \cdot \varphi/W$ and range $r = (H - y) \cdot R/H$ in the polar coordinate system. φ and R, respectively, represent the horizontal opening angle and slant range of the forward-looking sonar, W and H, respectively, represent the horizontal and vertical dimensions of the image.

3.2. Experimental Details. We implemented our method based on Pytorch. After many experiments and comprehensive consideration, we adopted the following settings to train the sonar data set from scratch. The size of input resolution is fixed at 512×512, the optimizer uses Adam, the basic learning rate is set at 0.001, the number of training rounds is 300, the training strategy of learning rate fixed step size reduction is adopted, each time the reduction is 1/10, and the number of decreased rounds is 120,200,260, respectively. In order to enhance the diversity of data, we used conventional data processing methods (vertical flip, horizontal flip, etc.) for the images.

3.3. Loss Function. The training loss function of our model consists of two parts. The first part is the training loss in the stage of shadow location prediction. In this stage, not only the heatmap (heatmap1) was predicted but also the length and width of the object and the offset of the object center point. Therefore, the loss values in this stage are as follows: loss of preliminary heatmap (heatmap1), object length and width prediction loss, and object center offset loss. In the shadow prediction stage, the loss value of the predicted heatmap1 L_{k1} was calculated by focal loss [30], both the loss value of the predicted length and width L_{size} and the loss value of the object center offset L_{off} were calculated by L1 loss, and λ_{size} and λ_{off} were both set as 0.1.

$$L_{\text{det1}} = L_{k1} + \lambda_{\text{size}} L_{\text{size}} + \lambda_{\text{off}} L_{\text{off}}. \tag{3}$$

The other part is the loss function that predicts the final heatmap, where L_{k2} is the loss of heatmap2.

$$L_{\text{det2}} = L_{k2}. \tag{4}$$

The loss function of the whole model can be obtained by adding the losses of the above two parts, and the specific calculation formula is as follows:

$$L_{sum} = L_{\text{det1}} + L_{\text{det2}}. \tag{5}$$

3.4. Comparison Experiments

3.4.1. The Parameter α of Cutting Width. Since the transverse position of the shadow in this data sets is offset to a certain extent relative to the object, we design a parameter α for the multiples of the cutting width and take different values of α for experimental comparison. The higher the average accuracy (mAP) is, the more consistent the cut width is with the shadow width of the data sets, and the more

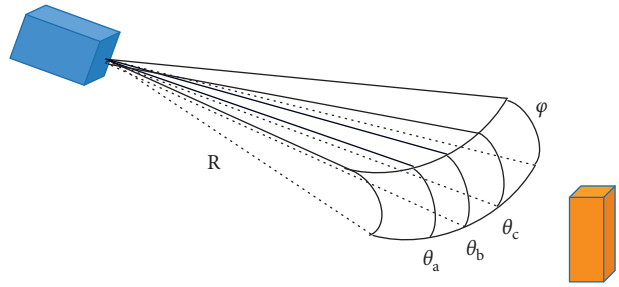


Figure 5: The schematic diagram of the detection beam.

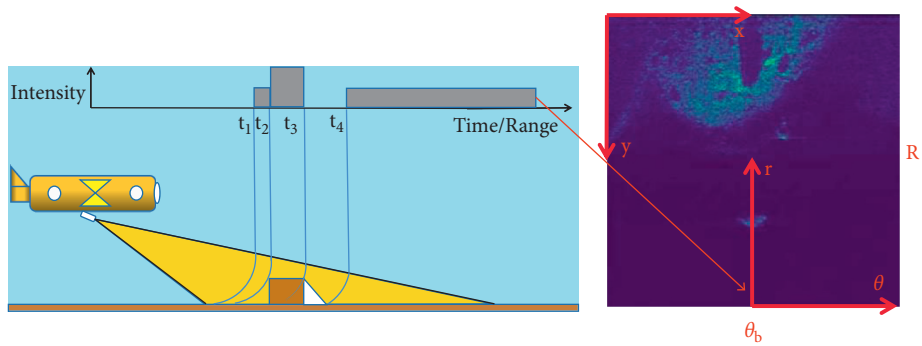


Figure 6: The schematic diagram of underwater forward-looking sonar image.

Table 1: In PASCAL VOC2012 standard, the average accuracy of each width multiple parameter α at different IOU.

α	mAP@0.5 (%)	mAP@0.6 (%)	mAP@0.7 (%)	mAP@0.8 (%)
1	98.12	93.19	71.28	21.62
1.25	97.09	**94.23**	73.95	22.33
1.5	97.51	93.39	**75.73**	**25.47**
1.75	97.11	91.76	70.13	19.83
2	**98.12**	92.24	73.75	21.01

Table 2: Under PASCAL VOC2012 standard, the average accuracy of each model under different IOU.

Method	mAP@0.5 (%)	mAP@0.6 (%)	mAP@0.7 (%)	mAP@0.8 (%)
YOLOv3 [31]	93.80	81.50	56.40	14.40
SSD [32]	91.31	82.75	58.61	19.62
CenterNet (hourglass) [25]	97.42	92.16	69.61	24.97
CenterNet (ResNet-50)	96.74	90.07	65.16	16.24
RefineDet [33]	94.64	88.38	66.62	18.76
FCOS [34]	90.72	87.95	70.70	28.63
Ours	**97.51**	**93.39**	**75.73**	**25.47**

complete the cut shadow features are. The experimental results are shown in Table 1.

3.4.2. Model Comparison. Comparison is made between our mothed and all classical object detection motheds on the same underwater forward-looking sonar data sets. Table 2 shows the comparison of different IOU under PASCAL VOC2012 evaluation standard. We can find that our model has obvious advantages in detection accuracy, by comparing different models. According to the comparison of visual images in Figure 7, it can also be seen that the detection performance of this model is improved compared with CenterNet.

3.4.3. Accuracy of each category. Under Pascal voc2012 evaluation standard, the accuracy performance of our model and classical object detection models on eight classes of objects in sonar data set is compared when IOU is 0.7. Because our structure integrates shadow information into the feature map and enhances the performance of the features of the object on the feature map, it can be seen from

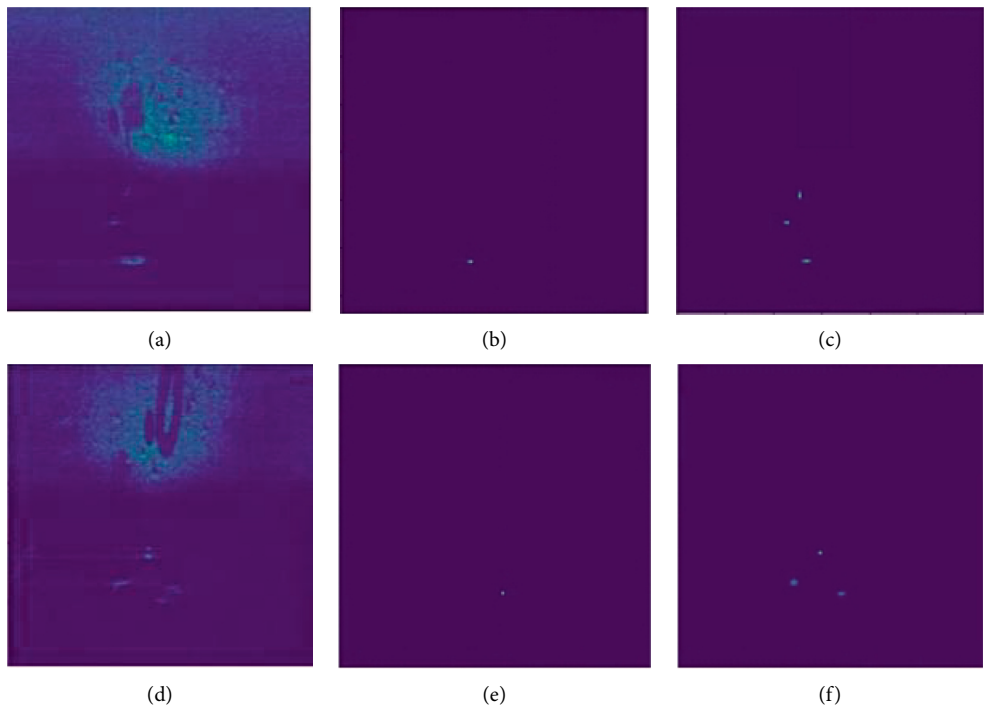


Figure 7: We visualize the heat map predicted by CenterNet and the heat map predicted after adding SCM. A row is a group of comparison pictures, (a, d) are the original pictures (processed for convenience of observation), (b, e) are the heat maps predicted by CenterNet, and (c, f) are the heat maps predicted after adding SCM. It is obvious that the object predicted by the model with SCM is more comprehensive and more accurate.

Table 3: In PASCAL VOC2012 standard, IOU is 0.7, and the accuracy of each model is different.

Methods	mAP@0.7 (%)	Ball (%)	Cylinder (%)	Square cage (%)	Cube (%)	Circle cage (%)	Human body (%)	Metal bucket (%)	Tire (%)
SSD	58.61	71.74	25.13	43.09	74.07	68.54	64.39	63.05	58.91
YOLOv3	56.40	55.20	48.80	49.00	66.40	58.30	59.80	60.80	52.50
RefineDet	66.62	72.56	61.94	54.66	77.88	78.28	57.61	70.89	59.12
CenterNet (hourglass)	69.61	78.50	67.94	64.86	77.89	74.70	67.88	64.60	60.47
CenterNet (ResNet-50)	65.16	73.79	51.30	67.10	75.55	69.28	56.30	71.52	56.24
FCOS	70.70	73.84	62.30	68.59	74.95	78.43	63.17	**76.64**	67.66
Our	**75.73**	**83.12**	**69.47**	**70.28**	**84.20**	**80.82**	**72.73**	71.47	**73.78**

Table 4: Comparison the number of parameters between CenterNet and our model.

Model	Num of pram
CenterNet	95.436 M
Ours	98.390 M
Growth rates	3.09%

Table 3 that the detection performance of this model is almost the best in all categories.

3.4.4. Analyzing the Number of Parameters of SCM. As shown in Table 4, after comparing the number of parameters between CenterNet and our model, we found that SCM only adds 2.954 M of parameters, which is only 3.09% of CenterNet, so this module is very lightweight.

4. Conclusions

In this paper, we propose a structure for capturing shadow features and fusing them into feature maps. It makes use of the correlation between shadow features and object features, as well as the difference of shadow features between different categories of objects, so as to increase the difference of features between categories. For objects with more obvious features, the recognition effect of the network is often better. Combined with our experimental results, it can be shown that the fusion of shadow and object features can indeed enhance the role of network recognition to a certain extent. In addition, there is still something worth exploring about the fusion mode of shadow and object features. How to better combine shadow features and object features will be studied in the future.

Data Availability

The data sets used and analysed during the current study are available from the corresponding author upon reasonable request.

Disclosure

Taowen Xiao and Zijian Cai are the co-first authors of the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest associated with the manuscript.

Acknowledgments

Taowen Xiao, Zijian Cai, Cong Lin, and Qiong Chen contributed equally to this work. This work was supported by the National Natural Science Foundation of China under Grant 62072121 and Natural Science Foundation of Guangdong Province 2021A1515011847.

References

[1] Q. Chen, M. Huang, and H. Wang, "A feature discretization method for classification of high-resolution remote sensing images in coastal areas," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 10, pp. 8584–8598, 2021.

[2] Q. Chen, M. Huang, H. Wang, and G. Xu, "A feature discretization method based on fuzzy rough sets for high-resolution remote sensing big data under linear spectral model," *IEEE Transactions on Fuzzy Systems*, vol. 59, no. 10, pp. 8584–8598, 2021.

[3] H. Wang, X. Li, R. H. Jhaveri et al., "Sparse bayesian learning based channel estimation in FBMC/OQAM industrial IoT networks," *Computer Communications*, vol. 176, pp. 40–45, 2021.

[4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR)*, pp. 886–893, San Diego, CA, USA, June 2005.

[5] M. Zhu, Y. Song, J. Guo et al., "PCA and kernel-based extreme learning machine for side-scan sonar image classifification," in *Proceedings of the IEEE Underwater Technology (UT)*, pp. 21–24, Busan, South Korea, February 2017.

[6] V. Myers and J. Fawcett, "A template matching procedure for automatic target recognition in synthetic aperture sonar imagery," *IEEE Signal Processing Letters*, vol. 17, no. 7, pp. 683–686, 2010.

[7] R. Fandos, A. M. Zoubir, and K. Siantidis, "Unified design of a feature based ADAC system for mine hunting using synthetic aperture sonar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2413–2426, 2014.

[8] Y. LeCun, B. Boser, J. S. Denker et al., "Backprop agation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.

[9] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classifification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.

[11] S. Wang, B. Kang, J. Ma et al., "A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19)," *European Radiology*, vol. 31, 2021.

[12] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine*, vol. 121, Article ID 103792, 2020.

[13] M. Valdenegro-Toro, "Object recognition in forward-looking sonar images with convolutional neural networks," in *Proceedings of the OCEANS 2016 MTS/IEEE Monterey*, September 2016.

[14] M. Valdenegro-Toro, "End-to-end object detection and recognition in forward-looking sonar images with convolutional neural networks," in *Proceedings of the 2016 IEEE/OES Autonomous Underwater Vehicles (AUV)*, November 2016.

[15] D. P. Williams, "Underwater target classification in synthetic aperture sonar imagery using deep convolutional neural networks," in *Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR)*, December 2016.

[16] A. Galusha, J. Dale, J. M. Keller, and A. Zare, "Deep convolutional neural network target classification for underwater synthetic aperture sonar imagery," in *Proceedings of the Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIV*, vol. 11012, April 2019.

[17] K. Breistein, "*Applying machine learning using custom trained convolutional neural networks on subsea object detection and classification*," M.S. thesis, NTNU, Trondheim, Norway, 2019.

[18] G. Neves, M. Ruiz, J. Fontinele, and L. Oliveira, "Rotated object detection with forward-looking sonar in underwater applications," *Expert Systems with Applications*, vol. 140, Article ID 112870, 2020.

[19] L. Zacchini, M. Franchi, V. Manzari et al., "Forward-looking sonar CNN-based automatic target recognition: an experimental campaign with FeelHippo AUV," in *Proceedings of the 2020 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV) (50043)*IEEE, St. Johns, Canada, 2020.

[20] Z. Fan, W. Xia, X. Liu, and H. Li, "Detection and segmentation of underwater objects from forward-looking sonar based on a modified Mask RCNN," *Signal, Image and Video Processing*, vol. 15, pp. 1–9, 2021.

[21] S. Reed, Y. Petillot, and J. Bell, "An automatic approach to the detection and extraction of mine features in sidescan sonar," *IEEE Journal of Oceanic Engineering*, vol. 28, pp. 90–105, 2003.

[22] E. Sang, Z. Shen, C. Fan, and Y. Li, "Sonar image segmentation based on implicit active contours," in *Proceedings of the 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*, vol. 4, November 2009.

[23] V. Myers and J. Fawcett, "A template matching procedure for automatic target recognition in synthetic aperture sonar imagery," *IEEE Signal Processing Letters*, vol. 17, no. 7, pp. 683–686, 2010.

[24] V. Myers and D. P. Williams, "Adaptive multiview target classification in synthetic aperture sonar images using a partially observable Markov decision process," *IEEE Journal of Oceanic Engineering*, vol. 37, pp. 45–55, 2011.

[25] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," 2019, https://arxiv.org/abs/1904.07850.

[26] A. Newell, K. Yang, and D. Jia, "Stacked hourglass networks for human pose estimation," in *Proceedings of the European Conference on Computer Vision*, October 2016.

[27] X. Wang, A. Shrivastava, and A. Gupta, "A-fast-rcnn: hard positive generation via adversary for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, July 2017.

[28] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, October 2017.

[29] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the International Conference on Machine Learning. PMLR*, Lille, France, July 2015.

[30] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, October 2017.

[31] J. Redmon and F. Ali, "Yolov3: an incremental improvement," 2018, https://arxiv.org/abs/1804.02767.

[32] W. Liu, D. Anguelov, D. Erhan et al., "Ssd: single shot multibox detector," in *Proceedings of the European Conference on Computer Vision*, October 2016.

[33] S. Zhang, L. Wen, X. Bian, Z. Lei, and S. Z. Li, "Single-shot refinement neural network for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, July 2018.

[34] Z. Tian, C. Shen, H. Chen, and T. He, "Fcos: fully convolutional one-stage object detection," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, South Korea, October 2019.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 1430512, 9 pages
https://doi.org/10.1155/2021/1430512

Research Article

Energy-Efficient UAV Trajectory Design with Information Freshness Constraint via Deep Reinforcement Learning

Xinmin Li,[1] **Jiahui Li**,[1] **and Dandan Liu**[2]

[1]*School of Information Engineering, Southwest University of Science and Technology, Mianyang 621000, China*
[2]*College of Weapon Engineering, Naval University of Engineering, Wuhan 430030, China*

Correspondence should be addressed to Dandan Liu; liudandan_nue@126.com

Received 27 October 2021; Accepted 30 November 2021; Published 23 December 2021

Academic Editor: Han Wang

Unmanned aerial vehicle (UAV) technique with flexible deployment has enabled the development of Internet of Things (IoT) applications. However, it is difficult to guarantee the freshness of information delivery for the energy-limited UAV. Thus, we study the trajectory design in the multiple-UAV communication system, in which the massive ground devices send the individual information to mobile UAV base stations under the demand of information freshness. First, an energy-efficiency (EE) maximization optimization problem is formulated under the rest energy, safety distance, and age of information (AoI) constraints. However, it is difficult to solve the optimization problem due to the nonconvex objective function and unknown dynamic environment. Second, a trajectory design based on the deep Q-network method is proposed, in which the state space considering energy efficiency, rest energy, and AoI and the efficient reward function related with EE performance are constructed, respectively. Furthermore, to avoid the dependency of training data for the neural network, the experience replay and random sampling for batch are adopted. Finally, we validate the system performance of the proposed scheme. Simulation results show that the proposed scheme can achieve a better EE performance compared with the benchmark scheme.

1. Introduction

With the explosive increasing of global mobile devices and connections in the future wireless network, Cisco forecasts that the global mobile data traffic will reach 77 exabytes per month by 2022 [1], which is almost two times over the data traffic in 2020. To meet the needs of high volume of data traffic and massive connections, the sixth generation (6G) wireless communication system enables some promising technologies to improve the communication rate, enhance the wide coverage, access the massive devices, and strengthen the intelligence and security [2, 3]. The unmanned aerial vehicle (UAV) communication working as one of the promising technologies, which has the advantages of flexible deployment, controllable maneuver, and low cost, becomes an interesting topic in the industry and academia to drive the development of Internet of Things (IoT) applications [4–7].

Due to the influence of the UAV's trajectory on the rate performance and energy consumption directly, how to design the UAV's trajectory is vital in the various types of communication scenarios. Although there exists some literatures investigating the trajectory design [8–10] for single-UAV communication systems under different settings, multiple-UAV may serve the specific area to provide better communication rate and coverage performance, which can increase the interference level of the UAV communication network. In [11], the authors analyzed the influence of UAVs' positions on the rate performance and obtained the optimal positions in the two-UAV interference channel. The authors in [12] designed the trajectory of the multiple-UAV based on the successive convex approximation (SCA) method considering the backhaul. However, exploring the energy-efficient trajectory design is vital for the energy-limited UAV to enhance the sustainable communication capability. In [13], the comprehensive energy consumption

model including the communication energy and propulsion energy was proposed for rotary-wing UAV. The SCA-based technique was employed to optimize the UAV trajectory. Furthermore, a joint of scheduling the backscatter devices, trajectory design, and transmit power was proposed in [14] to maximize energy-efficiency (EE) performance. In addition, the authors in [15] optimized a constructed trajectory to minimize the propulsion energy of fixed-wing UAV. Above literatures focus on the UAV communication to provide the high performance communication link and enable the information delivery.

With the increase in the real-time and computation-intensive applications, a new metric is required to satisfy the demand of information freshness beyond the scope of the delay time performance metric. To characterize the freshness for information delivery accurately, age of information (AoI) has been proposed in [16], which precisely describes the timeless of information updates from the original generation at the perspective of the receiver. In general, AoI is defined as the time gap between the observed time and the recent update time. Taking the generation time and information update into consideration, it is different from the conventional performance metrics, such as delay. In order to meet the demands of the various communications, AoI performance metric has been introduced to optimize the generation policies and user scheduling [17–20]. To guarantee the information freshness of the UAV communication system, the authors in [21] proposed a dynamic programming-based path planning to update the collected data to minimize the AoI value. Recently, some learning methods were proposed to make the decisions in various dynamic scenarios [22–24]. Under the UAV's energy constraint, the AoI optimization scheme based on reinforcement learning (RL) was proposed in [23] by optimizing the UAV's trajectory. In [24], the authors designed a joint trajectory and packet scheduling scheme based on deep reinforcement learning (DRL) approach to minimize the weighted AoI performance of the single-UAV system. However, how to design the trajectory that not only improves the energy efficiency but also guarantees the information update on time in the UAV communication system remains an open problem.

In this paper, we consider the UAVs' trajectory design in the multiple-UAV-enabled communication system to maximize energy-efficiency performance, in which each ground device sends the individual information to the corresponding UAV. First, we formulate an energy-efficient trajectory design optimization problem in multiple-UAV communication systems under the practical constraints, such as rest energy, safety distance, and AoI metric. However, it is difficult to solve this optimization problem due to the nonconvex objective function and the unknown spaces for UAV's trajectory. Second, a deep Q-network (DQN) method is proposed to optimize the UAV's trajectory and reduce the computational complexity. We design the state space considering the UAV's position, energy efficiency, rest energy, and AoI and construct the efficient reward function related with the objective function and the constraints. Furthermore, to avoid the dependency of training data for the neural network, the experience replay and random sampling for batch are adopted, which can grantee the stability of the proposed scheme in the dynamic environment. Finally, we verify the system performance of the proposed scheme. Simulation results show that the proposed scheme outperforms the benchmark scheme.

The rest of paper is organized as follows. In Section 2, the system model and energy-efficiency optimization problem in the multiple-UAV communication system are presented. In Section 3, we present the proposed DQN scheme to solve the optimization problem in detail. The simulation results of the proposed scheme are shown in Section 4. Finally, we conclude the work in Section 5.

2. System Model

We consider a UAV communication system consisting of M single-antenna UAV base stations denoted as $\mathcal{M} = \{1, \ldots, M\}$ and K single-antenna IoT device denoted as $\mathcal{K} = \{1, \ldots, K\}$ (e.g., smart grid, agricultural, safety, or geographic information). All the UVAs working as the aerial base stations serve the ground devices depicted in Figure 1. The UAVs service range is within a specific radius, and the flight height is fixed as H.

For simplicity, $\mathbf{q}_m^u(t) = (x_m^u(t), y_m^u(t), H)$ and $\mathbf{q}_k^d = (x_k^d(t), y_k^d(t), 0)$ denote the three-dimensional (3D) position of the m-th UAV at time t and the k-th device's position, respectively. Thus, the 3D distance between the k-th device and the m-th UAV is written as

$$\begin{aligned} d_{k,m}(t) &= \left\| \mathbf{q}_m^u(t) - \mathbf{q}_k^d \right\| \\ &= \sqrt{\left(x_m^u(t) - x_k^d\right)^2 + \left(y_m^u(t) - y_k^d\right)^2 + H^2}. \end{aligned} \tag{1}$$

In order to model the channel information of the UAV communication link more practically, we adopt the probabilistic line-of-sight (LoS) channel model proposed in [25]. The probability of LoS link for the communication link between the k-th device and the m-th UAV, which is related with the device's elevation, can be expressed as

$$P_{k,m}^{LoS}\left(\alpha_{k,m}(t)\right) = \frac{1}{1 + a_1 \exp\left(-a_2\left[\alpha_{k,m}(t) - ta_2\right]\right)}, \tag{2}$$

where a_1 and a_2 are the channel parameters related with the environment of the communication link and $\alpha_{k,m}(t)$ denotes the elevation angle between the k-th device and the m-th UAV at time t. Thus, the probability $P_{k,m}^{NLoS}$ for non-line-of-sight (NLoS) link between the k-th device and the m-th UAV is $P_{k,m}^{NLoS}(\alpha_{n,i}(t)) = 1 - P_{k,m}^{LoS}(\alpha_{k,m}(t))$. Therefore, the average channel power gain between the k-th devices and the m-th UAV is defined as follows [25]:

$$h_{k,m} = P_{k,m}^{LoS}\left(\alpha_{k,m}\right) PL_{k,m}^{LoS} + P_{k,m}^{NLoS}\left(\alpha_{k,m}\right) PL_{k,m}^{NLoS}. \tag{3}$$

In this work, we adopt the time-division multiple access to serve the ground devices, in which the interference is originated from the devices using the same time resource. Thus, the signal-to-interference-plus-noise-ratio of the k-th user at the m-th UAV at the time t is expressed as

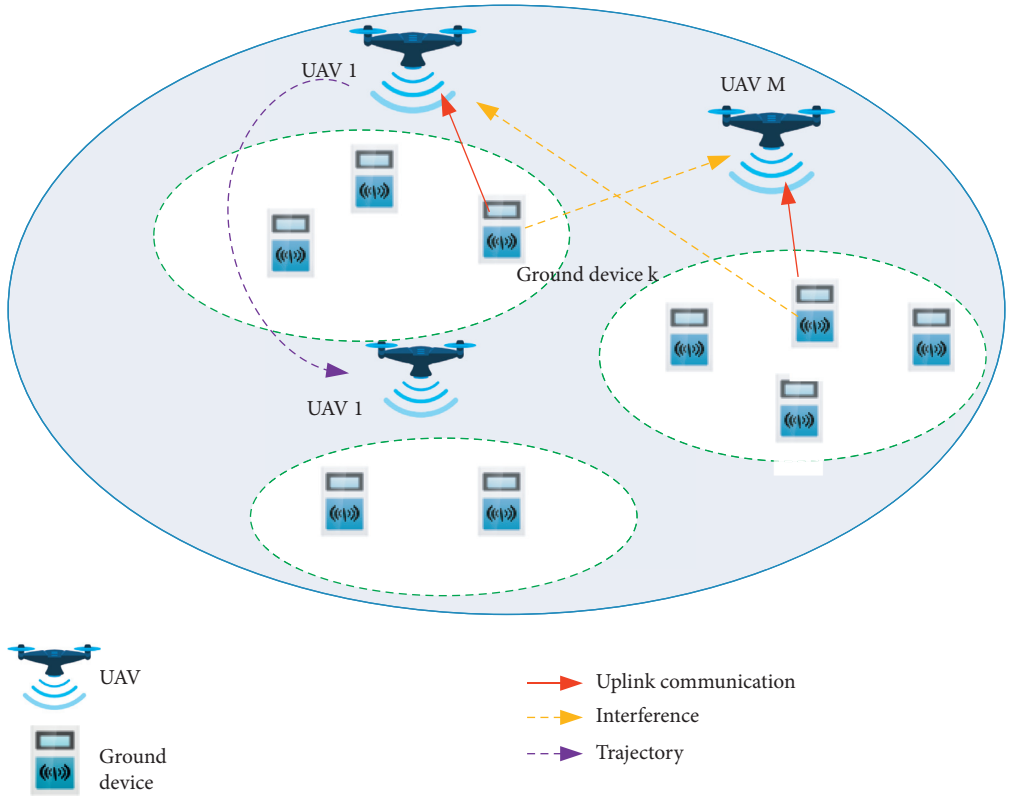


Figure 1: System model for UAV communication network serving massive IoT ground devices.

$$\gamma_{k,m}^{t} = \frac{p_k h_{k,m}}{\sum_{i \neq k} p_i h_{i,m} + \sigma^2}, \tag{4}$$

where p_k and σ^2 are the transmit power of the k-th ground device and the received noise power at each UAV, respectively. According to (4), the communication rate between the k-th device and the m-th UAV is written as

$$R_{k,m}^{t} = B \log_2\left(1 + \gamma_{k,m}^{t}\right), \tag{5}$$

where B denotes the transmission bandwidth. Since the UAV working as the aerial BS serves IoT devices with small data packets, UAV only requires comparably short time to receive the data, and we neglect the hover energy consumption. Thus, the energy consumption of UAV communication mainly consists of the communication energy consumption E_{com} and the propulsion energy consumption for flight E_{fly}. We assume that each UAV has a constant receiving power p^u, and the energy consumption consumed by the communication link from starting time to the time t is expressed as

$$E_{\text{com}}(t) = p^u t. \tag{6}$$

The energy consumption to support the UAV flight is the propulsion consumption. In general, the required power consumption can be modelled as follows [13]:

$$p_m^{\text{fly}}(t) = \left(c_1 \left\| v_m(t) \right\|^3 + \frac{c_2}{v_m(t)}\left(1 + \frac{\beta_m(t)}{g^2}\right)\right), \tag{7}$$

where $v_m(t)$ and $\beta_m(t)$ denote the flight speed of the m-th UAV and the flight acceleration of the UAV, respectively. The parameters c_1 and c_2 depend on the weight, the wing length of the UAV, and the air density in the flight environment, and g is the acceleration of gravity. Thus, the propulsion energy is $E_{m,\text{fly}}(t) = \int_0^t p_m^{\text{fly}}(\tilde{t})\mathrm{d}\tilde{t}$, and the energy consumption at the time t is written as

$$E_{m,\text{cons}}(t) = E_{\text{com}}(t) + E_{m,\text{fly}}(t). \tag{8}$$

Since the UAV has the limited energy for flight and communication, it is necessary to remain the enough rest energy for safe flight and return.

Let $E_{m,\max}$ denote the maximal energy of the m-th UAV; thus, the rest energy of the m-th UAV is defined as

$$E_m^r(t) = E_{m,\max} - E_{m,\text{cons}}(t). \tag{9}$$

In order to meet the essential safety for UAV flight, the rest energy E_m^r of UAV should not be less than the minimum rest energy $E_m^{r,th} = \chi E_{\max}$ with $\chi \in (0, 1)$ denoting the coefficient of rest energy.

The energy efficiency for UAV m to serve IoT devices at the time t is expressed as

$$\eta_m(t) = \frac{\sum_{k=1}^{K} R_{km}(t)}{p_m^{\text{fly}}(t) + p^u}. \tag{10}$$

AoI describes the age of the received packet at the destination to characterize the freshness of information collected by UAVs and becomes a new metric for the future

communication system [16]. For example, the device k sends different packets at the times $\tau_1, \tau_2, \ldots, \tau_n$ and the UAV m receives these packets at the times $\widetilde{\tau}_1, \widetilde{\tau}_2, \ldots, \widetilde{\tau}_n$. At time t, the immediate vicinity time for receiving the packets between UAV m and the device k is $\tau'_{km} = \max\{\widetilde{\tau}_1, \widetilde{\tau}_2, \ldots, \widetilde{\tau}_n\}$. Thus, AoI of UAV m is defined as the gap between the observed time and the maximal received time, which is written as

$$\Delta^t_{km} = t - \tau'_{km}. \tag{11}$$

It is noted that the smaller AoI, the higher freshness of the information while the bigger AoI, the less freshness of the information. For simplicity, $\Delta^0_{km} = 0, \forall k, m$, i.e., the AoI of the system at the time $t = 0$ sets zero. The average AoI is long-term metric to measure the total freshness of information over the duration T, which is expressed as

$$\overline{\Delta}_m = \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K}\Delta^t_{km}. \tag{12}$$

The target is to maximize the energy efficiency of the UAV system by optimizing the trajectory under the AoI and energy constraints. Therefore, the optimization problem is expressed as follows:

$$\begin{aligned}(P1):\ &\max_{\mathbf{q}^u_m(t)}\sum_{m=1}^{M}\eta_m \\ &\text{s.t. } \left\|\mathbf{q}^u_m(t) - \mathbf{q}^u_{\overline{m}}(t)\right\| \geq d^{\text{th}}, \quad \overline{m} \in \mathscr{M}\backslash m,\end{aligned} \tag{13}$$

$$E^r_m \geq E^{r,\text{th}}, \quad m \in \mathscr{M}, \tag{14}$$

$$\overline{\Delta}_m \leq \overline{\Delta}^{\text{th}}, \quad m \in \mathscr{M}. \tag{15}$$

Constraints (13) and (14) originate from the safety requirements that UAVs' distance and rest energy all should have the minimum thresholds d^{th} and $E^{r,\text{th}}$ to avoid the flight conflict, while constraint (15) comes from the demand of the data freshness with the threshold $\overline{\Delta}^{\text{th}}$. The objective function is nonconvex, and there exists the mass of the computational complexity to search the serve sequence of the devices under AoI constraints and design the UAV trajectory with unknown spaces in the system; thus, it is challenging to handle the problem (P1). To solve the nonconvex problem, a reinforcement learning method is adopted by combining the deep neural network and reinforcement learning to design the UAVs' trajectories intelligently.

3. Proposed Solution Based on Deep Reinforcement Learning

RL is one of the potential machine learning since it can make decisions by choosing the beneficial action from the action space based on its past experiences in dynamic environments, in which the agent interacts with the environments and updates the rewards on the current state [26]. There are three fundamental parts in each RL algorithm: state of environment, action of agent, and the reward from the environment.

3.1. State, Action, and Reward Function. In this work, UAVs are regarded as the agents to decide the trajectory to satisfy the requirements of energy and distance, and the state space of the environment for UAV trajectory design can be defined as a five-dimensional state space including the UAV's position, energy efficiency, rest energy of UAV, and AoI of current state, respectively, i.e., the state of the m-th UAV is expressed as follows:

$$\mathbf{s}_m = \left[x^u_m(t), y^u_m(t), E_r(t), \eta_m(t), \Delta_m(t)\right]^T. \tag{16}$$

In the initial state, each UAV equips with the maximal energy and does not build the communication link with any devices. Thus, the initial state $\mathbf{s}_m = [x^u_m(0), y^u_m(0), E_{\max}, 0, 0]^T, m \in \mathscr{M}$, where AoI in the first time is set as zero for all UAVs. The rest energy, energy efficiency, and AoI in each state can be updated according to (9)–(11), respectively. If the rest energy is smaller than the minimal threshold, UAV makes a decision to the initial position to guarantee the safety of UAV and stops the update of the state.

At any state of environment, the UAV can select the flying direction to serve the ground devices or reduce the interference in the UAV communication system. For simplicity, the action space is set as $\mathscr{A} = \{0, \ldots, 2i/L * \pi, \ldots, 2(L-1)/L * \pi\}$, with L uniform directions in $[0, 2\pi)$. If the $L = 4$, the setting is typical with four orthogonal directions {left, right, frontward, backward}. After selecting the action a_t in the action space $\mathscr{A}$, the UAV can transit the current state $\mathbf{s}_t$ to the next state $\mathbf{s}_{t+1}$, e.g., the next position of the m-th UAV in the flying time duration δ_t:

$$\begin{aligned}x^u_m(t+1) &= x^u_m(t) + v(t)\delta_t\cos(a_t), \\ y^u_m(t+1) &= y^u_m(t) + v(t)\delta_t\sin(a_t).\end{aligned} \tag{17}$$

Based on the new UAV position $\mathbf{q}^u_m(t+1) = (x^u_m(t+1), y^u_m(t+1), H)$, the communication rate is obtained while the energy efficiency and AoI values can also be updated to show the performance of the new state. Since the reward function is significant to obtain a optimal policy, the system adopts the energy efficiency to construct the reward function, i.e.,

$$r_m(s_t, a_t) = \begin{cases}\phi\eta_m(t), & \text{if (13)} - \text{(15)are satisfied,} \\ 0, & \text{otherwise,}\end{cases} \tag{18}$$

where $0 \leq \phi \leq 1$ denotes the normalized parameter for energy efficiency (since the value of energy efficiency is comparatively high, the normalized parameter can balance the reward between negative and positive reward; in general, it is almost the order of magnitude of 10^{-3}, which is based on the reward value in the specific communication system). It is observed that the UAV will obtain the positive reward when the action a_t satisfies the corresponding constraints, while it will obtain any reward or penalty if the UAV violates the constraints. Since the reward function is monotonically increasing with respect to (w.r.t) the objective function in P1, the UAV makes the decision toward the energy-efficiency maximization, and the system can obtain the optimal solution.

Since the size of the state space is nonlinear with the size of UAV's position, action space, rest energy, and AoI, it exists the intensive computation complexity to obtain the optimal action by exhaustive search exists, which is impractical for the energy-limited UAV system. Moreover, the Q-learning method based on Q-table requires the large memory to store the Q-table, and it cannot be suitable for this optimization problem.

3.2. Control Policy and Deep Reinforcement Learning Algorithm. Recently, the deep RL algorithm has become a promising technique to tackle the resource allocation and performance optimization in the wireless communication systems. However, the continuous action space in the UAV communication system needs to be quantized into discrete and formulate the state-action function to characterize the influence of the selected action on the performance with a specific state. The deep RL method is applied to solve the problem because it can use the neural network to learn the policy to reduce the high dimensionality of the state space instead of storing the value. The detailed design framework based on DRL in the UAV communication system is shown in Figure 2. The Q-function originated from the Q-learning is adopted to maximize the long-term cumulated reward. Given the control policy ξ for the m-th UAV, the Q-function is defined as

$$Q^{\xi}(\mathbf{s}_t, a_t) = E\left[r_m(\mathbf{s}_t, a_t) + \sum_{j=1}^{t-1} \gamma^{t-j} r(\mathbf{s}_j, a_j)\right], \tag{19}$$

where $\gamma \in [0, 1]$ is the discount factor. If the discount factor $\gamma = 0$, the Q-function is only related with the current reward, i.e., the selected action only maximizes the current reward without considering the future reward. Thus, the optimal action to maximize the objective function in P1 can be written as follows:

$$a_t^* = \arg\max_{a_j \in \mathscr{A}} Q^{\xi}(\mathbf{s}_t, a_j). \tag{20}$$

To obtain the optimal control policy ξ^*, the Q-function is updated based on the Bellman equation:

$$Q(\mathbf{s}_t, a_t) = Q(\mathbf{s}_t, a_t) + \nu\left(r(\mathbf{s}_t, a_t) + \gamma \max_{a_j \in \mathscr{A}} Q(\mathbf{s}_{t+1}, a_j) - Q(\mathbf{s}_t, a_t)\right), \tag{21}$$

where ν is the learning rate. According to (21), the UAV can update the Q-function and learn the control policy based on the stored Q-values by the selecting the action to maximize the reward. However, there exists one puzzle during the processing of the learning: how to select the action in the limited state-action values. At the starting of learning, the UAV only has the some partial Q-values and cannot choose an appropriate action. Thus, the UAV should explore the environment sufficiently to obtain Q-values of all state-action pairs. To tackle this issue, an ϵ-greedy strategy is applied to explore the environment with the probability ϵ, which is written as

$$a = \begin{cases} \text{random}(\mathscr{A}), & \text{with probability } \epsilon, \\ \arg\max_{a_j \in \mathscr{A}} Q^{\xi}(\mathbf{s}_t, a_j), & \text{with probability } 1 - \epsilon. \end{cases} \tag{22}$$

At each state, the UAV can take a random action with the probability ϵ to explore the environment. As the number of exploration increases, the probability can decrease to guarantee the system performance with the UAV selecting the optimal action. Since the unknown state space for the UAV's trajectory may lead to a large memory size and a slow convergence rate, thus deep neural network is an effective method to extract features from the existing data sets intelligently and reduce the computational complexity by predicting the output in parallel. According to the framework in Figure 2, the tuple consists of the state, action, reward, and next state working as the input of deep neural network to output Q-value as $Q(s_{t+1}, a_t|\theta_i)$ in the estimate and target neural networks, where θ_i denotes the parameters of the neural network during i-th training. The target neural network is the replica of the estimate neural network every N_{rep} steps to make the two neural networks as close as possible to guarantee the stability. Therefore, it is important to optimize the parameters of neural network θ_i based on the suitable loss function to obtain the optimal Q-function. The loss function based on the error is defined as follows:

$$\mathscr{L}(\boldsymbol{\theta}_{i+1}) = \left|r(\mathbf{s}_t, a_t) + \gamma \max_{a_j \in \mathscr{A}} Q(\mathbf{s}_{t+1}, a_j|\boldsymbol{\theta}_i) - Q(\mathbf{s}_t, a_t|\boldsymbol{\theta}_i)\right|^2. \tag{23}$$

Based on the loss function and the training data set, some optimizers can be used to obtain the optimal parameters of neural network, such as gradient descent algorithm and Adam algorithm.

The training data are vital for training the deep neural network. However, there exist the following challenges: first, the UAV communication system is time-varying and the objective function is related with the UAV's trajectory and AoI. How to obtain the sufficient number of training data in the dynamic environment is crucial for optimizing the neural network. Second, empirical evidence demonstrated that independent training data can enhance the stability and improve the convergence for the neural network. Thus, obtaining independent training data is another challenge to optimize the neural network. To address these aforementioned challenges, the experience replay and random sampling method are adopted. For the fixed experience replay memory with the size N_{mem}, the training data will be updated every N_{tr} steps to replace the history data, which can automatically fresh the memory to obtain the fresh training data. To avoid the dependency of the stored data to train the neural network, the random sampling method is used to form the batch by choosing the experience from the replay memory randomly, which can smooth the changes between the history data and the new observation. The proposed DQN-based trajectory design scheme for energy-efficiency maximization is shown in Algorithm 1.

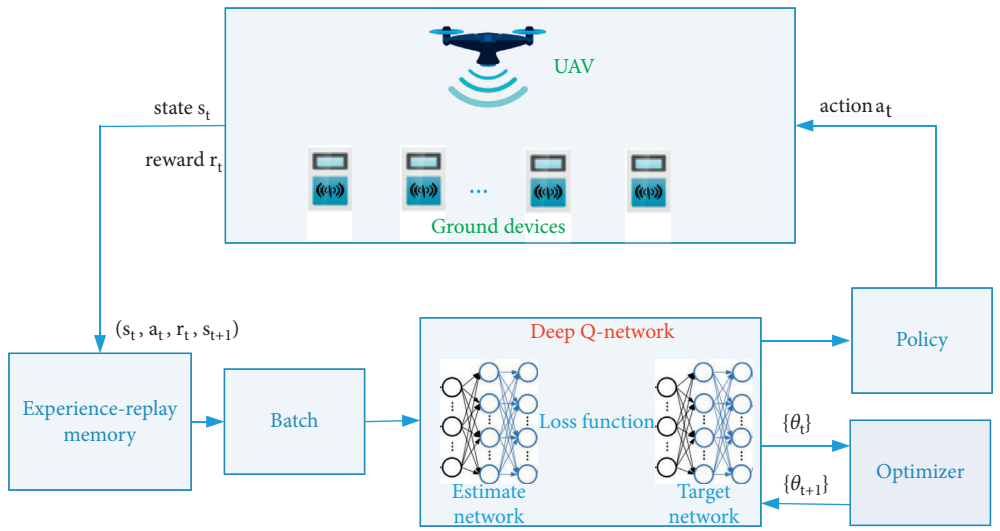


Figure 2: The design framework based on DRL for UAV's trajectory.

(1) Initialize learning parameters $\{\phi, \gamma, \epsilon\}$, memory size, batch size, the maximal episode $N_{\max}$, observed time T, and $N_{\text{step}} = 0$;
(2) **for** $n = 1$: $N_{\max}$ **do**
(3) Initialize the environment and state $\mathbf{s}_0$.
(4) **for** $t = 1$: T **do**
(5) If random() $> \epsilon$, select an action a_t using (20). Otherwise, select an random action from the action set $\mathcal{A}$.
(6) Execute the action a_t, compute EE, rest energy, and AoI, and obtain the next position $\mathbf{q}_m^u(t+1)$ to form the next state $\mathbf{s}_{t+1}$. According to (18), compute the reward. Store $\langle \mathbf{s}_t, a_t, r_t, \mathbf{s}_{t+1} \rangle$ into the experience-reply memory.
(7) If $N_{\text{step}} \% N_{\text{rep}} == 0$, duplicate the estimate neural network to target neural network.
(8) Train the neural network based on loss function in (23) to optimize the parameter θ, $N_{\text{step}} \leftarrow N_{\text{step}} + 1$.
(9) **end for**
(10) **end for**

Algorithm 1: Deep Q-network-based trajectory design scheme for energy-efficiency maximization.

4. Simulation Results

In this section, we consider three UAVs flying in the square area with $300 \times 300\,\text{m}^2$ to verify the proposed scheme. All ground devices are uniformly distributed in this area to transmit the independent information to UAVs during the observed time $T = 200$. For performance comparison, DQN without experience-reply is adopted as the benchmark scheme. Unless otherwise stated, the simulation parameters are shown in Table 1.

In order to train the deep neural network, three fully connected hidden layers with 100 hidden nodes are adopted. The size of experience-reply memory is 400 and is randomly selected from the memory to construct the batch. The gradient descent method is used to optimize the parameter of neural network. For the experience-reply memory, the new training data always update the oldest history data. To guarantee the size of efficient batch data, the training starts after 100 steps. Other parameters of the deep neural network are shown in Table 2. The simulation environment is Intel i7 CPU, Python 3.7, and TensorFlow 1.14 to train the UAV's deep neural network. All results are averaged over 500 episodes.

Figure 3 shows the effect of the flight speed v on the AoI performance. It is observed that AoI value decreases with increase in the flight speed, i.e., the information freshness can improve by increasing the UAV's flight speed, which comes from the fact that the frequency of the information update increases as the speed increases in the limited flying environment. Compared with the benchmark scheme, the AoI value decreases by 3.5% and 9.3% with $v = 20$ m/s and $v = 30$ m/s. However, since the propulsion power increases cubically w.r.t the flight speed, it is unwise only to increase the UAV's speed to maximize EE performance under the AoI constraint.

Figure 4 demonstrates the influence of the learning rate ν on the EE performance. We can find that the optimal EE performance of the proposed scheme can be achieved when the learning rate ν equals 0.4. Since the learning rate directly affects the convergence of the proposed scheme, the proposed scheme has a slow convergence rate to obtain the optimal EE performance. The proposed scheme can achieve

Table 1: Simulation parameters for the UAV communication system.

Symbol	Description	Value
B	Bandwidth	1 MHz
p	Transmit power of users	23 dBm
σ^2	Noise power	−100 dBm
K	Number of ground devices	30
M	Number of UAVs	3
L	Number of UAV's directions	10
H	UAV flight height	80 m
v	Flight speed	20 m/s
$E_{\max}$	Maximum energy of UAVs	1.5696×10^5 J
a_1, a_2	Channel parameters	[10, 0.6]
d^{th}	Distance threshold	1 m
χ	Coefficient of rest energy	0.1
Δ^{th}	AoI threshold	500

Table 2: The settings of the deep neural network.

Symbol	Description	Value
ν	Learning rate	0.4
ϵ	Greedy coefficient	0.05
γ	Discount factor	0.9
ϕ	Normalized parameter for EE	10^{-3}
N_{bat}	Batch size	64
N_{tr}	Interval of training data	3

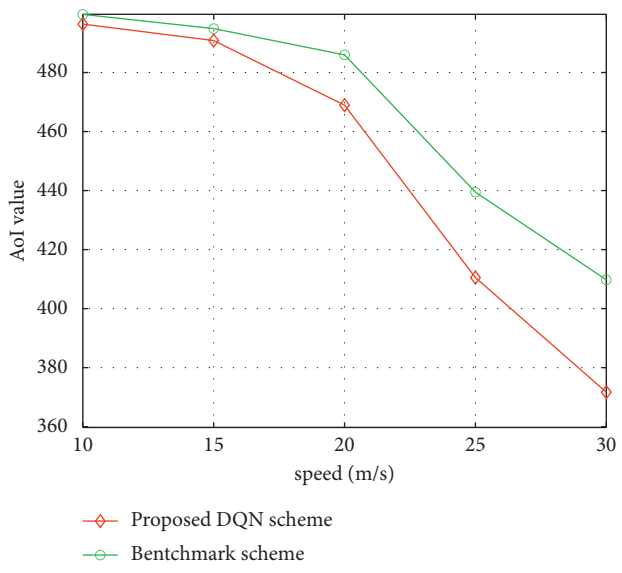


Figure 3: The effect of the flight speed on the AoI performance with $v = 0.4$.

22.2 Kbit/J at least compared with 21.7 Kbit/J when the learning rate becomes large.

Figure 5 shows the stability of the proposed scheme with the training number. It is noted that the loss fluctuates when the number of the training is small, and it converges to a small value as the number of training becomes large, which is result of the increasing number of the training batch. It is also found that the UAVs may have different convergence rates. There are two reasons as follows:

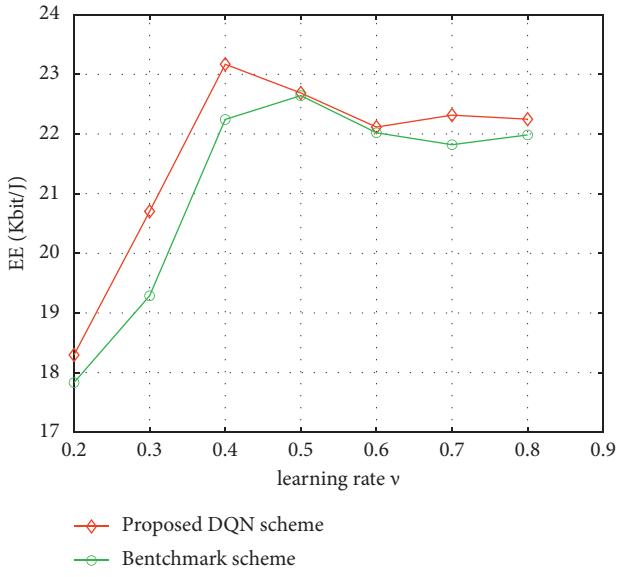


Figure 4: EE performance versus the learning rate with $v = 20$ m/s.

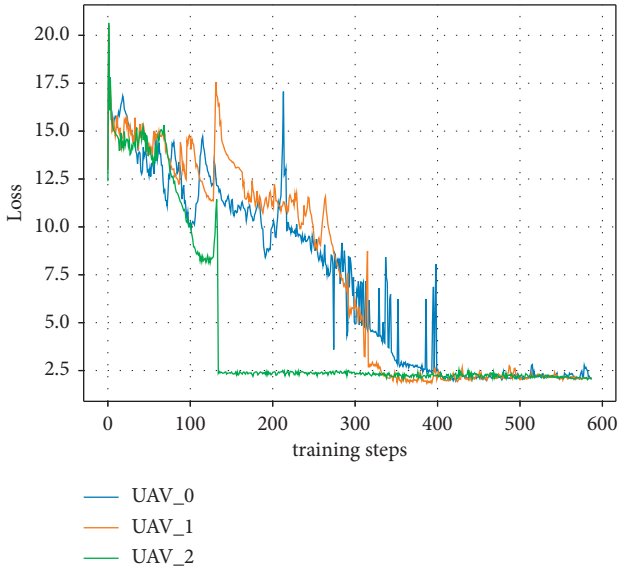


Figure 5: The effect of the training number on the loss function.

(1) UAVs have different initialized actions to result in the different estimate and target neural networks

(2) Each UAV has independent channel power gain related to the distance between UAV and served ground users to form the different state space and reward value

5. Conclusion

In this letter, we considered the UAV communication system to serve the IoT ground devices, which send the fresh information to the corresponding UAVs. Taking the rest energy and AoI into account, the DQN-based trajectory design was proposed to maximize the energy-efficiency performance. The state space related with the rest energy and AoI and the reward function related with EE performance are constructed, respectively. Under the experience replay and random sampling for batch, the simulation results show that the proposed DQN scheme achieves better performance compared with the benchmark scheme. In the future, considering the hover energy consumption for the massive IoT devices is a significant work while it is also interesting to design the collaborative UAV communication by exchanging the information for neural network optimization.

Data Availability

The data used to support the findings of this study are included within this article.

Conflicts of Interest

The authors declare there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Key R&D Program of China (2019YFB1705100), Foundation of Southwest University of Science and Technology (18zx7142 and 19xn0086), the Sichuan Science and Technology Program (2019JDTD0019), and China Scholarship Council Program (202008515123).

References

[1] Cisco, "Cisco visual networking index: Global Mobile Data traffic forecast update 2017–2022," Cisco, 2019, https://s3.amazonaws.com/media.mediapost.com/uploads/CiscoForecast.pdf.

[2] G. Flagship, *Key Drivers and Research Challenges for 6G Ubiquitous Wireless Intelligence*, G. Flagship, University of Oulu, Oulu, Finland, 2019, http://jultika.oulu.fi/files/isbn9789526223544.pdf.

[3] X. You, C. Wang, J. Huang et al., "Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Science China*, vol. 64, pp. 1–74, 2021.

[4] S. Aggarwal, N. Kumar, and S. Tanwar, "Blockchain-envisioned UAV communication using 6G networks: open issues, use cases, and future directions," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5416–5441, 2021.

[5] J. Xu, Y. Zeng, and R. Zhang, "UAV-enabled wireless power transfer: trajectory design and energy optimization," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5092–5106, 2018.

[6] H. Wang, X. Li, R. H. Jhaveri, and T. R. Gadekallu, "Sparse Bayesian earning ased hannel stimation in FBMC/OQAM industrial IoT networks," *Computer Communications*, vol. 176, pp. 40–45, 2021.

[7] Y. Huang, W. Mei, J. Xu, L. Qiu, and R. Zhang, "Cognitive UAV communication via joint maneuver and power control," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7872–7888, 2019.

[8] J. Chen and D. Gesbert, "Optimal positioning of flying relays for wireless networks: a LOS map approach," in *Proceedings of the IEEE International Conference on Communications*, pp. 1–6, (ICC), Paris, France, May 2017.

[9] X. Li, Q. Li, D. Kong, X. Zhang, and X. Wang, "Learning based trajectory design for low-latency communication in UAV-enabled smart grid networks," in *Proceedings of the IEEE Vehicular Technology Conference*, pp. 1–5, VTC2020-Fall, Victoria, B.C., Canada, November 2020.

[10] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: a tutorial on UAV communications for 5G and beyond," *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, 2019.

[11] X. Li and J. Xu, "Positioning optimization for sum-rate maximization in UAV-enabled interference channel," *IEEE Signal Processing Letters*, vol. 26, no. 10, pp. 1466–1470, 2019.

[12] P. Li and J. Xu, "Placement optimization for UAV-enabled wireless networks with multi-hop backhauls," *Journal of Communications and Information Networks*, vol. 3, no. 4, pp. 64–73, 2018.

[13] Y. Zeng, J. Xu, and R. Zhang, "Energy minimization for wireless communication with rotary-wing UAV," *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2329–2345, 2019.

[14] G. Yang, R. Dai, and Y.-C. Liang, "Energy-efficient UAV backscatter communication with joint trajectory design and resource optimization," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 926–941, 2021.

[15] F. Dong, L. Li, Z. Lu, Q. Pan, and W. Zheng, "Energy-efficiency for fixed-wing uav-enabled data collection and forwarding," in *Proceedings of the 2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Shanghai, China, May 2019.

[16] R. Y. S. Kaul and M. Gruteser, "Real-time status: how often should one update?" in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, pp. 2731–2735, Orlando, FL, USA, March 2012.

[17] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: how to keep your data fresh," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[18] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, pp. 1844–1852, Honolulu, HI,USA, April 2018.

[19] B. Zhou and W. Saad, "Joint status sampling and updating for minimizing age of information in the internet of things," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7468–7482, 2019.

[20] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," *IEEE/ACM Transactions on Networking*, vol. 28, no. 1, pp. 15–28, 2019.

[21] Z. Jia, X. Qin, Z. Wang, and B. Liu, "Age-based path planning and data acquisition in uav-assisted iot networks," in *Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Shanghai, China, May 2019.

[22] H. Wang, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2020.

[23] B. Yin, X. Li, X. Zhang, and L. Wei, "Trajectory optimization for age of information minimization in uav communication systems," in *Proceedings of the Accteped by IEEE Vehicular Technology Conference (VTC2021-Fall)*, Helsinki, Finland, April 2021.

[24] M. A. Abd-Elmagid, A. Ferdowsi, H. S. Dhillon, and W. Saad, "Deep reinforcement learning for minimizing age-of-information in uav-assisted networks," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Waikoloa, HI, USA, December 2019.

[25] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 569–572, 2014.

[26] R. S. Sutton and A. G. Barto, *Reinforcement Earning: An Introduction*, MIT Press, Cambridge,MA, USA,, 2018.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 2292128, 18 pages
https://doi.org/10.1155/2021/2292128

Research Article

Adaptive Polymorphic Fusion-Based Fast-Tracking Algorithm in Substations

Wenbin Shi,[1] **Jingsheng Lei**,[1] **Xingli Gan**,[1] **and Zhongguang Yang**[2]

[1]*Zhejiang University of Science and Technology, Hangzhou 310000, China*
[2]*Electric Power Research Institute of State Grid Shanghai Electric Power Company, Shanghai 200051, China*

Correspondence should be addressed to Zhongguang Yang; yzg_sgcc@qq.com

Received 1 October 2021; Accepted 13 November 2021; Published 29 November 2021

Academic Editor: Han Wang

Tracking multiple objects in a substation remains a challenging problem since pedestrians often overlap together and are occluded by infrastructures such as high-tension poles. In this paper, we propose an adaptive polymorphic fusion-based fast-tracking algorithm to address the problem. We first leverage the fast segmentation algorithm to obtain the fine masks of pedestrians and then combine the motion and performance information of pedestrians to realize the fast-tracking in substations. Our model is evaluated on the widely used MOT19 dataset and real-substation scenarios. Experimental results demonstrate that our model outperforms state-of-the-art models with a significant improvement in the MOT19 dataset and occlusion cases in substations.

1. Introduction

We have witnessed a significant change in the intelligent management of grid corporations promoted by the high demand for power grid construction. Substations are the bridges of the power transmission. However, due to the remote construction location and steep terrain, the maintenance and the safety precautions in substations become inevitably difficult. The rapid development of intelligent software and hardware technology achieves the extensive application of intelligent power equipment in grid construction [1]. Unattended substations based on video surveillance technology significantly improve the efficiency of the substation management and gradually replace the manned mode, becoming the main way forward in smart substations. Recently, most positioning systems in substations exploit hardware technologies such as ultrawide waves to localize and track targets [2] whereas the method cannot automatically track targets and can only collect videos in a fixed area. The appearance of machine learning introduces the traditional machine learning algorithms, such as support vector machines (SVMs), Kalman filter, and adaptive enhancement to locate and track targets [3]. With the emergence of the fifth-generation wireless communication technology [4], data communication with ultrahigh reliability and ultralow time delay [5, 6] becomes possible. It can achieve good results in areas with high requirements for time delay and reliability, such as real-time target detection, target tracking, target positioning, and automatic driving [7, 8]. Nevertheless, traditional methods cannot detect targets efficiently and deal with the scenarios such as occlusion and blurry shooting.

Lately, the introduction of deep learning models has promoted the robustness and accuracy of tracking in substations. The SORT (Simple Online and Real-time Tracking) algorithm [9] is an online real-time multiobject tracking algorithm, which uses the Kalman filter and the Hungary algorithm to achieve data association and determine whether the targets detected in different frames are the same object or not. Chen et al. [10] proposed the MOTDT algorithm, which formulated a standard pedestrian detection trajectory scoring mechanism through fusing the object classifiers and the tracker credibility, utilizing the generated standard credibility as nonmaximum suppression (NMS) input to obtain more accurate candidates. Wang et al. [11] proposed the JDE model, which uses the YOLOv3 algorithm for the purpose of combining the detection and the embedded model to improve time efficiency. Besides, the JDE model makes use of triplet

loss to train the network, which automatically learns the loss weight strategies so as to achieve the weighting of nonuniform loss. Zhang et al. [12] proposed a one-shot MOT-based fair algorithm, which takes advantage of an anchor-free tactic to reduce the impact of the detection frame on pedestrian recognition. By estimating the center of the object on a high-resolution feature map, the features of the pedestrian can be better aligned with the center of the object. Therefore, the mainstream tracking algorithms usually identify and track the target by combining the motion information with the performance information of the target. Based on the aforementioned methods, this paper realizes the rapid and accurate detection and tracking of pedestrians in the substation.

In the substation, the tracking targets are generally made up of working pedestrians and vehicles. Working pedestrians are uniformly dressed, which frequently causes the occlusion and loss of pedestrians during tracking. Meanwhile, the complex infrastructures coupled with the hindrance when obtaining video surveillance data make pedestrian detection and tracking in the substation become difficult. Consequently, we first enlarge the tracking dataset with a data expansion method based on existing substation surveillance video data. And we proposed an adaptive segmentation strategy to train the model so that the model can automatically update when different pedestrian information is input. Besides, aiming at the pedestrian loss and identity change caused by obstructions in the substation, we adopted the correlation algorithm and priority allocation strategy to address the problem. We verified the effectiveness of our model on the established test dataset, which consists of real-substation monitoring scenarios, and we compared our model with other state-of-the-art models on widely used datasets. As a result, our method is able to detect and track pedestrians quickly and robustly.

The contributions of the proposed model are mainly in the following four aspects:

(1) We proposed a lightweight real-time pedestrian detection network Light-YOLOv4. We took advantage of the segmentation algorithm to obtain the fine mask of pedestrians and the segmentation results. Then, we matched the cosine similarity of the segmentation results in multiple frames to obtain the results of tracking.

(2) We proposed a multimodal fusion substation pedestrian tracking method. Combining with the motion and performance information of pedestrians, our method tracks pedestrians in the substation on the basis of the detection results of the Light-YOLOv4 algorithm.

(3) We proposed a weight distribution method based on integrated metric learning, which adopts the results of segmentation branches to accurately express the performance information of pedestrians. Besides, aiming at solving the problems of pedestrian loss and identification change caused by dense obstructions in the substation, we adopted the correlation algorithm and priority allocation strategy.

(4) We proposed the algorithm to be in accord with the real scenarios in the substation. We expanded existing datasets using the data expansion method and adopted an adaptive segmentation training method to automatically update when various pedestrians enter the substation.

1.1. Methodology. Existing deep learning-based target detection methods are mainly divided into two categories: one is the candidate frame-based two-stage target detection algorithm represented by Fast RCNN network and Faster RCNN network, and the other is the end-to-end one-stage target detection algorithm represented by SSD and YOLO network. Because of the excellent real-time performance and flexibility of the detection algorithm, the detection-based tracking method has become one of the compelling research topics. In the field of tracking, how to confirm the trajectories of pedestrians has turned into a huge challenge. Nowadays, the classic strategies to solve the problem consist of the fluid network conception and the probabilistic graphic model. However, the abovementioned methods are unsuitable for online situations where a target exists in every frame. We combine the motion and the performance information to solve the aforementioned issues in the substation. We use the Kalman filter to update the trajectories of pedestrians so that tracking multiple objects in the substation can be realized.

1.2. YOLOv4. The YOLO series is an end-to-end one-stage target detection algorithm, which has a faster detection speed. The YOLOv4 algorithm is on the basis of the original YOLO detection architecture and adopts the best optimization strategies in five aspects, including data processing, backbone network, training, activation function, and loss function. The architecture of YOLOv4 is shown in Figure 1.

YOLOv4 is mainly composed of three components, comprising the feature extraction network, feature fusion network, and YOLO detection head. The feature extraction network is based on the CSPDarknet53 network, which contains 29 convolutional layers, $725 * 725$ receptive fields, and 27.6 M parameters. The CSPDarknet53 feature extraction network divides the feature into two parts. The first part is directly constructed to generate residual edges, and the second part ensures the accuracy of the model while reducing model calculation complexity by first convolving with the main branch and then concatenating with the input of the first part. In addition, YOLOv4 transforms feature maps of any size into fixed-size feature vectors by fusing with the SPP module. Meanwhile, we merge features in different levels together with the PANet.

1.3. Kalman Filter. Kalman filter is one of the classic recursive filters. The main idea of the Kalman filter comes from Bayesian Estimation Theory. Kalman filter is based on the optimal estimation value at the current moment and uses the error covariance matrix to calculate the predicted value of

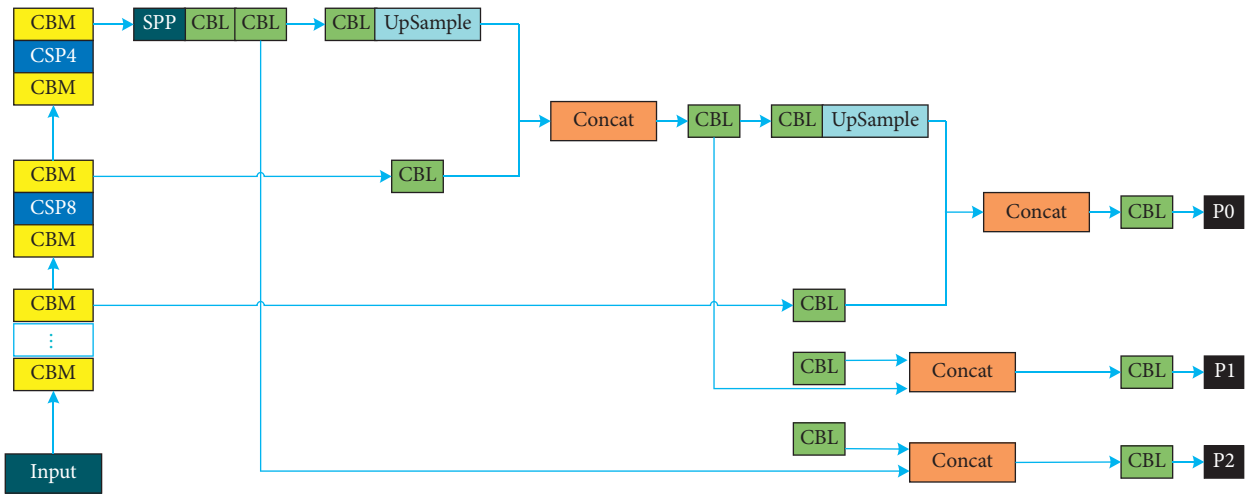


Figure 1: The architecture of the YOLOv4 network.

the state variable at the next moment. Meanwhile, the Kalman filter observes the state to obtain the observed variables and then uses the observed variables to correct the forecasts. Finally, we obtain the optimal state estimator at the next moment.

The Kalman filter prediction algorithm tracks the changes of the linear system quickly and accurately. The prediction model is mainly divided into two parts, which consist of the time update and the measurement update. The update equation assumes that the state at the time t evolves from the state at the time $(t-1)$, and the measurement equation calculates time based on the estimated state and noise at the time t.

1.4. Cosine Distance. The proposed model reidentifies pedestrians based on the performance information of the cosine distance. Specifically, the cosine similarity is used as a measurement function to calculate the cosine distance between all detected targets in different frames. The smaller the cosine distance is, the greater the probability that the two targets are the same target is. The proposed model reduces tracking errors caused by motion information and improves the accuracy of tracking.

2. The Framework of the Proposed Model

We proposed a lightweight Light-YOLOv4 real-time detection network, which combines motion and performance information of pedestrians, and proposed a pedestrian tracking method in substations using integrated metric learning. Inspired by Mask Region Proposal Network [13], we first utilize the detection algorithm to obtain the candidate box containing the pedestrian, then segment the pedestrian in the candidate box, and calculate the similarity of the candidate boxes in adjacent frames to confirm the identity of the pedestrian. Finally, we weigh the similarity of candidate boxes with the tracking boxes generated from the motion information of the pedestrian to obtain a more refined tracking box.

The framework of our model is shown in Figure 2. Concretely, (1) we preprocess the input real-time video frame signal and then detect pedestrians through the proposed lightweight Light-YOLOv4 algorithm and obtain the coordinates of the pedestrian's circumscribed rectangular boxes. (2) We crop the pedestrians according to the obtained coordinates, and the cropped pedestrian images are input to the segmentation branch, in order to obtain the fine mask of the pedestrians. In addition, input the segmentation results in the pretrained convolutional neural network. Moreover, we calculate the cosine similarity between the segmentation results of adjacent frames and determine the identity of the pedestrian through the performance information. (3) We use the Kalman filter algorithm to predict the trajectory of the detected pedestrian and make use of the Mahalanobis distance to indicate the predicted state and current state of the pedestrian. (4) Finally, we use the fusion metric learning method to calculate the degree of association between two independent objects in the front and rear frames in combination with motion and performance information. Furthermore, we introduce the weight coefficients and set a threshold to obtain a more refined tracking box. If the difference between the current frame time and the frame time that the pedestrian successfully matched last time is greater than the threshold, the pedestrian's trajectory is considered to be terminated and would be deleted in the subsequent tracking. Otherwise, it is considered that the trajectory is not lost. The above process is used to solve the problems of occlusion and crossing scenarios in the substation. In summary, we proposed a real-time pedestrian detection network called Light-YOLOv4. Light-YOLOv4 has the ability to track targets accurately and rapidly.

3. Light-YOLOv4 Detection Network and Segmentation Branch

In the substation, the high installation height of cameras and real-time requirement in surveillance demand tracking algorithms have the ability of real-time tracking and detecting accurately. The YOLO algorithm is a detection algorithm

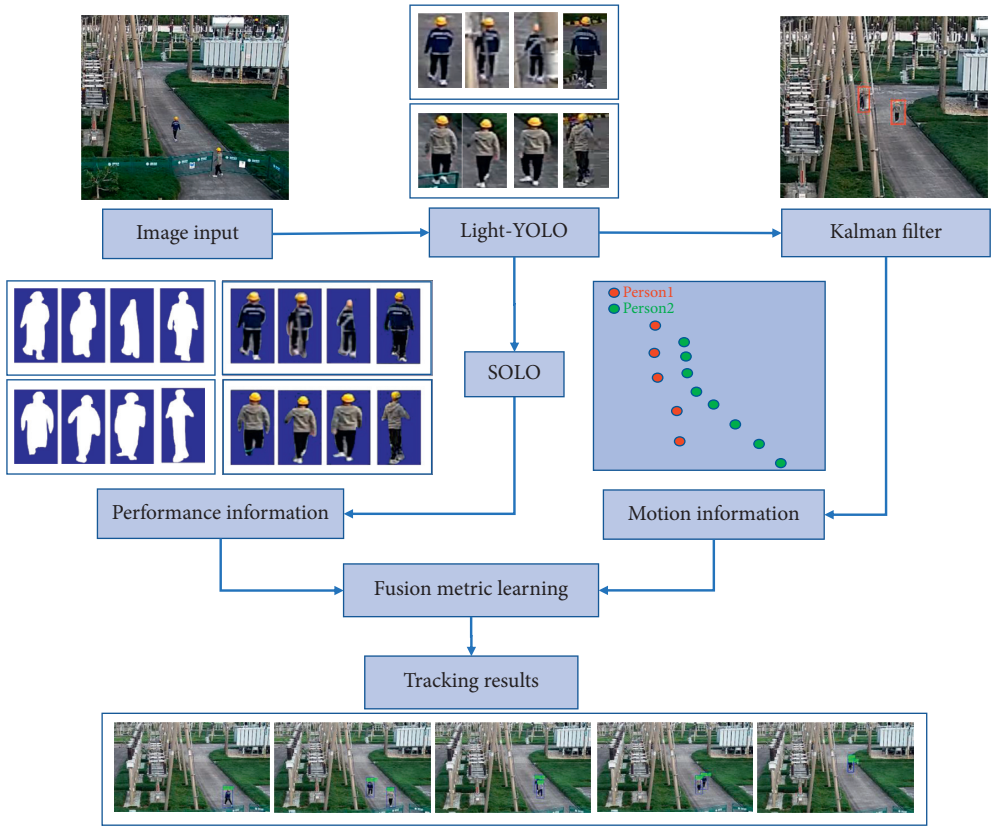


Figure 2: The framework of the proposed algorithm.

based on regression. The entire network is only composed of convolutional layers. And YOLO can complete the classification and location only once through the network. YOLOv4-Tiny is a lightweight version of YOLO, the running speed of which is faster than the YOLO network. However, YOLOv4-Tiny pursues the real-time requirement too much, thus giving up the accuracy.

Therefore, we proposed the Light-YOLOv4 real-time detection network, and the structure of Light-YOLOv4 is shown in Figure 3. The light-YOLOv4 network mainly contains three modules: multistream feature fusion module (MSFBlock), adaptive module (AdBlock), and hierarchical multiscale prediction module. The MSFBlock introduces an attention mechanism to obtain the different spatial and channel features and then merge the features of multistream branches to enhance the depth of features. In addition, we make use of the AdBlock to mine the connection between the points on the spatial and the channel feature maps. AdBlock can not only obtain rich global information but also improve detection accuracy and real-time performance. Finally, we optimize the multiscale prediction in the YOLO algorithm and combine multistage feature maps in the detection network to gain fine features and elevate the accuracy of the detection.

3.1. Multistream Feature Fusion Module. We adopt the MSFBlock to replace three CBL modules in the YOLOv4-Tiny network. CBL refers to the components made of conv, Batch Normalization, and Leaky ReLU function. The attention mechanism in the MSFBlock improves the ability of extraction, accuracy, and efficiency of detection.

As shown in Figure 4, the MSFBlock is made of two branches. The first branch is composed of a 3×3 convolutional layer connected with the output of the complete attention module (CAM) to extract features with different receptive fields. The second branch is shortcut by a 1×1 convolutional layer. The full convolution module in the first branch extracts semantic features. The first 1×1 convolution layer of the full convolution module reduces the dimensionality and the 3×3 convolution layer extracts features. The CAM in the first branch suppresses irrelevant noise. CAM uses a 1×1 convolutional layer to reduce dimensionality, then learns the weight of each feature group, and extracts important features through weights. Hence, the CAM filters out irrelevant features and extracts important features.

The structure of the CAM is depicted in Figure 5. By adding residual connections, the CAM makes better use of the previously extracted feature information. While introducing spatial feature information, as well as adopting the LeakyReLU activation function and a 1×1 convolutional layer, the CAM acquires the feature response.

3.2. Adaptive Module. The AdBlock is divided into spatial adaptation (SAd) and channel adaptation (CAd). The former module mainly obtains the spatial connections of features.

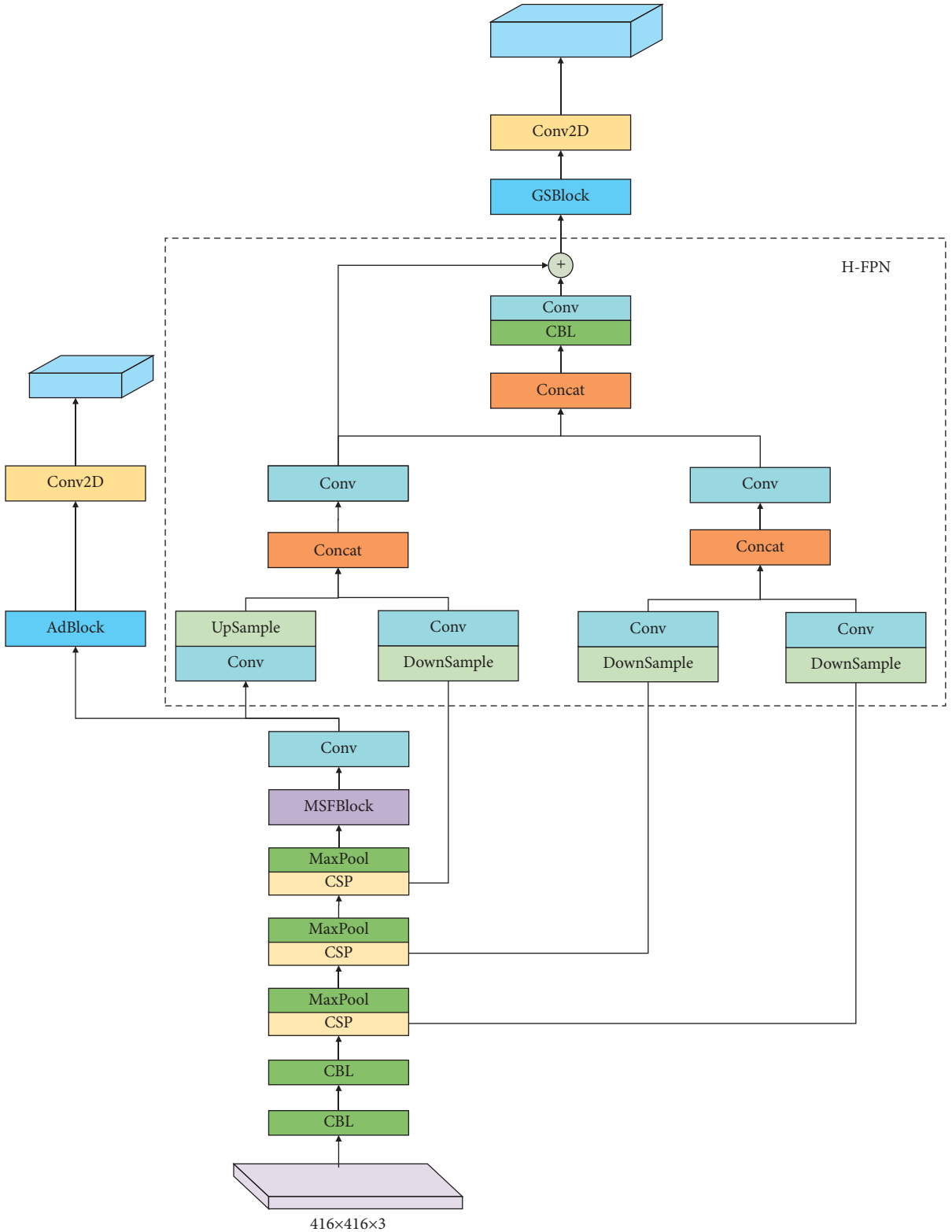


Figure 3: The structure of Light-YOLOv4.

Since high-level features in the channel are related to and share similarities with each other, the CAd is responsible for gaining feature information in the channel. Since the last MSFBlock removes redundant information and emphasizes important information but cannot obtain the global information of the target, the SAd is used to obtain the rich global

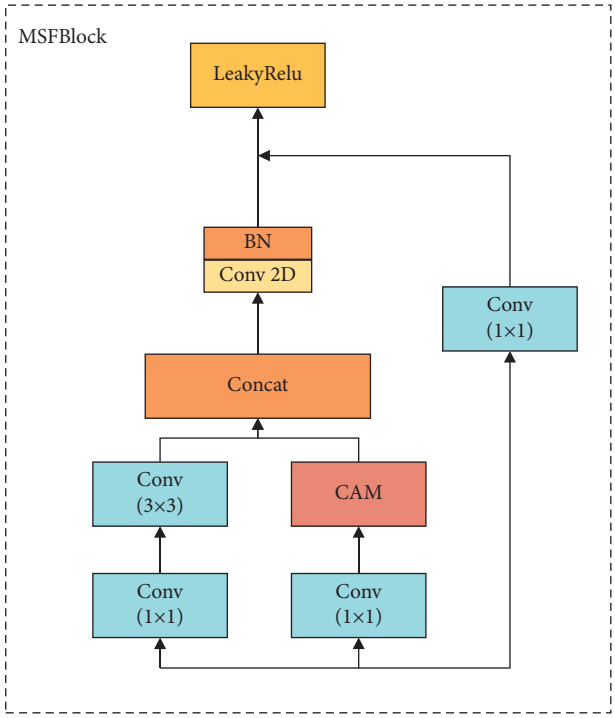


Figure 4: The structure of MSFBlock.

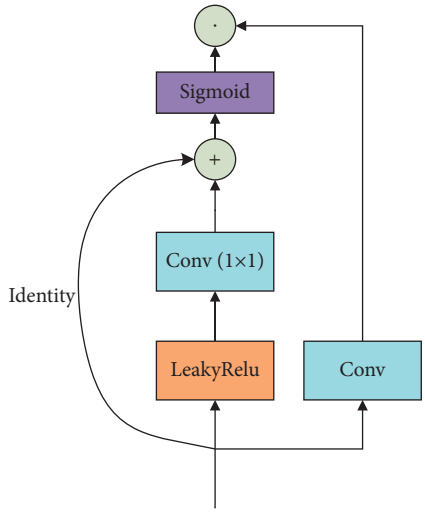


Figure 5: The structure of CAM.

feature information. The AdBlock mines the feature maps from the two dimensions of space and channel, which express the global information of targets comprehensively.

3.3. Hierarchical Multiscale Prediction. Feature Pyramid Network (FPN) improves the accuracy of pedestrian detection algorithms by extracting multiscale features. As shown in Figure 6, YOLOv4-Tiny first extracts and generates features at each level from the bottom up, which are denoted as {C2, C3, C4, C5}. Then, FPN obtains P5 from C5 with the convolution operation and uses top-down upsampling and horizontal connection operations to generate fusion feature P4. However, P4 fails to effectively leverage low-level feature

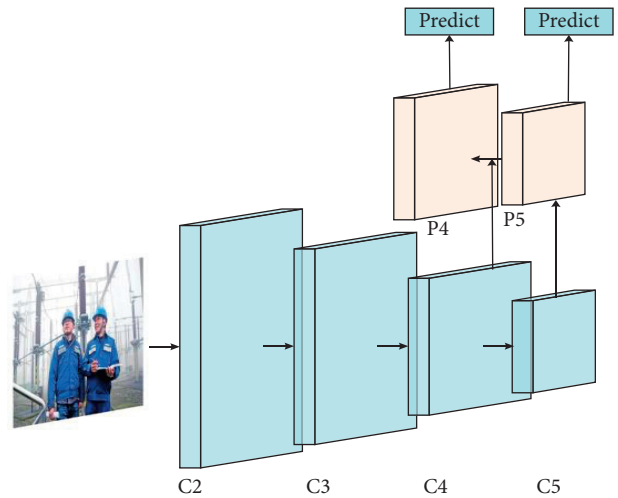


Figure 6: The structure of the FPN network in the YOLOv4-Tiny network.

maps; hence, FPN cannot accurately detect distant pedestrians. We remedy the multiscale prediction method. Specifically, we hierarchically fuse high-level semantic features with low-level features to make full use of all-scale feature maps, which improves the detection accuracy of distant pedestrians.

In order to fully utilize all-scale feature maps, we conduct the downsampling operation on C2 and C3, which are on a large scale, and upsampling C4 and C5, fourfold downsampling C2, and twofold downsampling C3. Then, we concatenate the results horizontally and reduce the dimensionality of the result to feature map N1 with the 1×1 convolutional layer. Similarly, we concatenate the results of C4 and C5 and reduce the dimensionality of the result to feature map N2 with the 1×1 convolutional layer. Finally, we concatenate N1 and N2 horizontally and reduce the dimensionality with a 1×1 convolutional layer to obtain a feature map {P4, P5} with more detailed information, which was shown in Figure 7. The structure in Figure 7 combines the low-level with the high-level semantic features, which can enrich the fine-grained information without increasing the amount of calculation and improve the detection of distant pedestrians without losing real-time performance.

3.4. Pedestrian Segmentation. Considering the requirements of the real-time performance in segmentation, we choose the SOLO [14] as the segmentation algorithm. SOLO introduces the concept of "instance category" to distinguish object instances. Specifically, the "instance category" refers to the quantized center position and the size of objects, which makes it possible to segment objects using their positions. SOLO is in an end-to-end manner and transforms coordinates into the problems of classification with discrete quantization. Hence, it has good real-time performance and is usually used with YOLO detectors.

SOLO directly distinguishes instances through the center position and object size, which are denoted as location and sizes, respectively. SOLO uses the location to allocate which

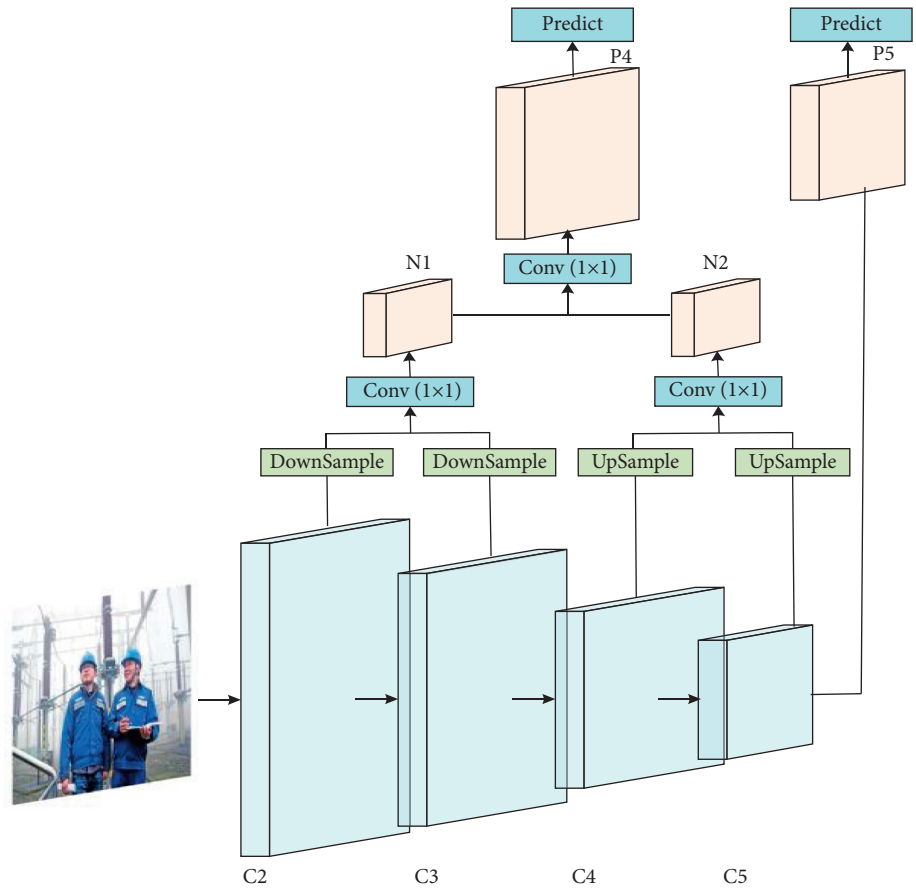


Figure 7: The network of Hierarchical FPN.

channel the instance should fall into and uses the FPN to solve the size problem. The specific process is shown in Figure 8.

SOLO is similar to the method in YOLO, which first separates the input image into $S \times S$ grids. The output is divided into two branches: category branch and mask branch. The size of the category branch is SSC, and C is the number of semantic categories. The mask branch size is HWS^2, and S^2 is the maximum number of predicted instances, which is corresponding to the position of the original image from top to bottom and from left to right. When the center of the pedestrian object falls into a grid, the corresponding position of the category branch and the corresponding channel of the mask branch are responsible for the prediction. We adopt the FPN to solve the problems of sizes. The large feature maps of FPN predict small pedestrians, and the small feature maps predict large pedestrians. FPN can also alleviate the problem of overlapping, and instances with different sizes will be assigned to different FPN output layers for prediction.

The SOLO algorithm also uses CoordConv to enhance the processing of location information. CoordConv concatenates two channels directly on the original tensor, storing the x and y coordinates, respectively, and normalizing them to $[-1, 1]$, which explicitly bring the position information into the next convolution operation. Meanwhile, CoordConv improves Decoupled head to solve the problem of too large output channels caused by too many grid cell settings. The final prediction process is that we first extract the predicted probability values of all category branches and use a threshold (such as 0.1) to filter the predicted values. Then, we obtain the i, j index corresponding to the remaining classification positions and divide the i channel in the X branch and the j channel in the Y branch using elementwise multiplication to obtain the mask of this category. Moreover, we use a threshold (such as 0.5) to filter the mask and conduct NMS on all masks. Finally, the mask is scaled to the size as original images to obtain the instance segmentation of the pedestrian, which is shown in Figure 9.

4. Pedestrian Tracking

In this section, we adopt the Kalman filter to predict the trajectory of the pedestrian based on the coordinate information detected by Light-YOLOv4 and the performance information (cosine similarity). Besides, we leverage the metric learning to match the predicted trajectory and the movement of the pedestrian, combining the performance information to realize the tracking.

4.1. Kalman Filter. The Kalman filter selects the pedestrian as the tracking object. The midpoint $(dx\ (k), dy\ (k))$ on the bottom edge of the detection box is taken as the tracking

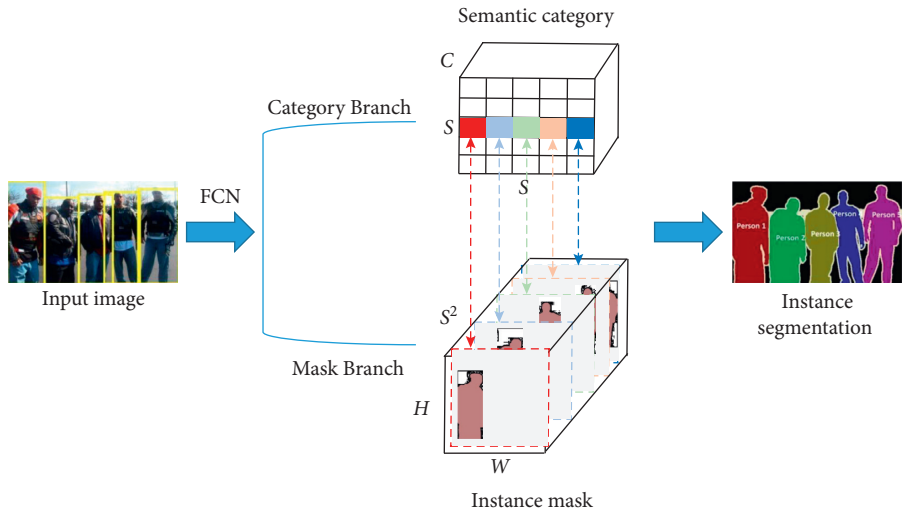


Figure 8: The processing of the SOLO algorithm.

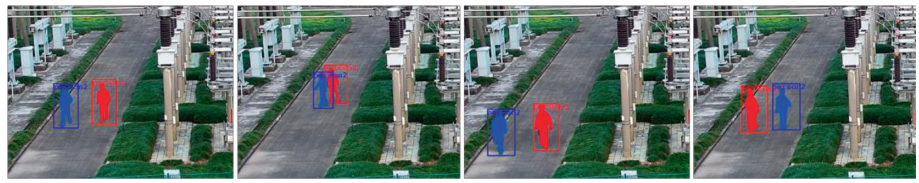


Figure 9: Visually results of the segmentation.

feature point, and the length w and height h of the detection box are selected as the other two feature variables. The above four variables form a four-dimensional state variable. Then, the workflow of the Kalman filter algorithm is used to predict the four-dimensional state variables of the detection box.

First, from the state at time k, the state function at time $k-1$ can be derived, as shown in

$$x_k = Fx_{k-1} + Bu_k + w_k, \tag{1}$$

where x_k refers to the state variable at time k, and F is a $n \times n$ gain matrix of the state variable. U_k is composed of c-dimensional vectors of input control. B is the $n \times c$ matrix related to input control and state change. The variable w_k represents the process noise.

Secondly, since the measured z_k may be equal to the state variable x_k, or may not to, the calculation of measured z_k is shown in

$$z_k = H_k x_k + v_k, \tag{2}$$

where H_k is an $m \times n$ observation matrix, and v_k represents the measurement noise.

Assume that the elements in w_k follow the Gaussian distribution $N(0, Q_k)$, where Q_K refers to the $n \times n$ covariance matrix. The elements in v_k follow the Gaussian distribution $N(0, R_k)$ and R_k represents the $m \times m$ covariance matrix.

First, we calculate the prior probability estimate of the current state $\hat{x}_k^-$ and use the superscript "–" to indicate "before the new measurement." The calculation of $\hat{x}_k^-$ is shown in

$$\hat{x}_k^- = F\hat{x}_{k-1} + Bu_{k-1} = w_k. \tag{3}$$

P_k^- denotes the error covariance, and P_k^-'s prior probability estimate at time k is obtained from its value at time $k-1$, the calculation of which is shown in

$$P_k^- = FP_{k-1}F^T + Q_{k-1}. \tag{4}$$

Equations (3) and (4) constitute the prediction part of the filter, from which the Kalman gain coefficient can be obtained, as shown in the following formulation:

$$K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k\right)^{-1}. \tag{5}$$

From equation (5), the optimal observation value and the corrected error covariance can be obtained, which is shown in equations (6) and (7).

$$\hat{x}_k = \hat{x}_k^- + K_k\left(z_k - H_k\hat{x}_k^-\right), \tag{6}$$

$$P_k = P_k^- - P_k^- K_k H_k. \tag{7}$$

The workflow of the Kalman filter is generally sorted into two stages: one is the prediction stage and the other is the update stage of prediction results. The two stages are illustrated in Figure 10. In the prediction stage, the system state $\hat{x}_{k-1}$ at time $k-1$ is input into equations (2) and (3) to calculate a prior probability estimate $\hat{x}_k^-$ and measured value z_k. In the update stage of the prediction result, the Q_{k-1} at

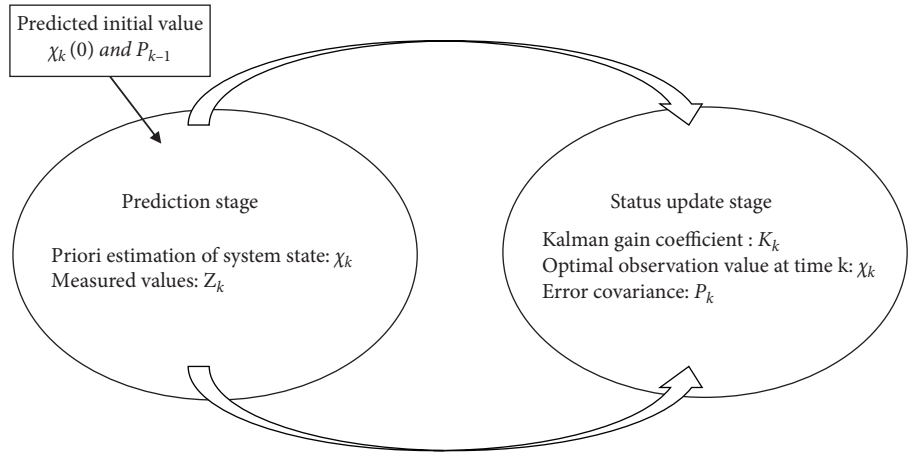


Figure 10: The workflow of the Kalman filter.

time $k-1$, the covariance P_{k-1}, and equation (4) are input to obtain a prior estimate P_k^- of the covariance at time k and then input P_k^- and RK at time k into equation (5) to obtain the Kalman gain coefficient K_k. Next, according to equations (6) and (7), the optimal observation value $\hat{x}_k$ and the corrected error covariance P_k are calculated. Then, we return to equation (2) and start the next pedestrian detection. Since the midpoint of the bottom edge $d_x(k)$ and $d_y(k)$ of the pedestrian detection frame and the length w and height h of the detection frame update every time, the initial position of $\hat{x}_k$ is the coordinates $(d_x(k), d_y(k))$, and the initial values of the length w and height h of the detection frame are the length and height of the pedestrian detection frame at the beginning, respectively. And the initial values of $\hat{x}_k(0)$, F, and H_k are shown in

$$\begin{aligned} x_k(0) &= \begin{bmatrix} d_x(k) \\ d_y(k) \\ w \\ h \end{bmatrix}, \\ F &= \begin{bmatrix} 1 & 0 & \mathrm{d}t & 0 \\ 0 & 1 & 0 & \mathrm{d}t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\ H_k &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{8}$$

4.2. Motion Information. Considering that the detection target in the current frame is composed of a four-dimensional vector, we adopt the Mahalanobis distance to measure the similarity between the current and historical trajectories of the target pedestrian. Mahalanobis distance represents the covariance distance of data, and it effectively calculates the similarity between two multidimensional pedestrians by taking into account the correlation between various features of the target.

Given a multivariable $t = (t_1, t_2, t_3, \ldots, t_p)^T$, the mean value and the covariant matrix are $\mu = (\mu_1, \mu_2, \mu_3, \ldots, \mu_p)^T$ and S, respectively. The formulation of t's Mahalanobis distance is shown as

$$D_M(t) = \sqrt{(t-\mu)^T S^{-1}(t-\mu)}. \tag{9}$$

In this paper, we calculate the Mahalanobis distance $M(i-1, i)$ of target personnel in i-th frame and in $(i-1)$-th frame, which is represented as

$$M(i-1, i) = \sqrt{(t_i - g_{i-1})^T S_{i-1}^{-1}(t_i - g_{i-1})}, \tag{10}$$

where t_i represents the state $(d_x(k), d_y(k), w, h)$ in #i frame. g_{i-1} denotes the prediction of i-th frame in $(i-1)$-th frame. S_{i-1} is the covariance matrix of target trajectory predicted by the Kalman filter algorithm in i-th frame. Since the motion in the video frame is continuous, we use $M(i-1, i)$ to filter the target and set 3.08 as the threshold of filtering. Equation (11) shows the filtering rule; filter represents the filtering function.

$$a(i-1, i) = \begin{cases} \text{filter}\{M(i-1, i) \le 3.08\}, & i \ge 2, \\ 0, & 0 < i \le 0. \end{cases} \tag{11}$$

4.3. Performance Information. When pedestrians are occluded or crossed, the identities of the pedestrians change frequently. Therefore, only using the motion information to track pedestrians will greatly reduce the accuracy. And thus, it is necessary to combine the segmentation to reidentify the performance information of the pedestrian, which remedies tracking errors caused by motion information and improves the accuracy of tracking.

In this paper, we adopt a deep convolutional neural network to extract the pedestrian masks segmented in each frame and utilize the cosine similarity as a metric function. The cosine distance is calculated for the feature vector of the object b extracted from the segmentation result of the i-th frame and the average value of the feature vector corresponding to the pedestrian segmentation result after f times

of successful tracking before the i-th frame. The cosine distance calculation is in

$$\cos \text{dis} = 1 - \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \cdot \sqrt{\sum_{i=1}^{n} B_i^2}}. \tag{12}$$

Different from other distance algorithms that directly measure between two points, the cosine distance algorithm calculates the cosine value of the angle of two vectors and normalizes the result to [0, 1], which is not related to absolute value and can measure the similarity between vectors. Therefore, we use the cosine distance algorithm as the measurement function, in order to calculate the matching degree of the appearance of the pedestrians in the current frame and previous frame.

The feature extraction network used in this article is a deep convolutional neural network, consisting of three convolutional layers, a maximum pooling layer, and four residual layers. Since online training takes a long time, we pretrained the network on a large-scale pedestrian dataset in advance to obtain an appearance model. The features are suitable for tracking. Then, we use the pedestrian segmentation result as the input of the network. And the output feature is a 256-dimensional vector.

Let $r_i = \{r_1^{(b)}, r_2^{(b)}, \ldots, r_k^{(b)}\}$ denote feature vectors $r_k^{(b)}$ of object b before i-th frame. In continuous tracking, the performance information of pedestrians may change due to the long tracking duration. Therefore, we keep the average value of the feature vector corresponding to the pedestrian mask after f times of successful tracking before the i-th frame of the object b, which is shown in equation (13). And the cosine distance calculation is shown in equation (14):

$$\overline{r}_k^{(b)} = \frac{\sum_{k=1}^{f} r_k^{(b)}}{f}, \tag{13}$$

$$\cos \text{dis}\,(k, i) = 1 - r_i^T \overline{r}_k^{(b)}. \tag{14}$$

Similarly, the measurement of cosine distance needs a threshold th, which is obtained through training. When the cosine distance is less than a specific threshold th, two objects are associated. The calculation of the measurement is shown in

$$c\,(k, i) = \text{filter}\{\cos \text{dis}\,(k, i) \le \text{th}\}. \tag{15}$$

4.4. Integrated Metric Learning. We set a weight coefficient K and obtain a weighted average of Mahalanobis distance and cosine distance through:

$$u = Ka\,(i - 1, i) + (1 - K)c\,(k, i). \tag{16}$$

4.5. Solving the Occlusion. The occlusion of the trajectory for a long duration will lead to the uncertainty of Kalman filter prediction. If there are multiple trackers matching the detection results of the same target at the same time, then the position of the trajectory will not update for a long time, which will cause a relatively large deviation in the position predicted by the Kalman filter. It can be calculated from equations (5)–(7) that the covariance will become larger. Since the Mahalanobis distance is proportional to the reciprocal of the covariance, the Mahalanobis distance will be small, which makes the detection result more likely associated with the longer trajectory, causing a decrease in the performance of the tracking algorithm. We set a frame time threshold, which means that if the difference between the current frame time and the frame time that the target successfully matched last time is greater than the threshold, the target trajectory is considered as terminated. Furthermore, the target trajectory will be deleted in the subsequent tracking. Otherwise, it is considered that the target is not lost.

5. Datasets and Experimental Results

In order to verify the effectiveness and generalization of the algorithm, we train and test the proposed network on a widely used dataset. The training of this article includes 3 stages: first is pretraining the Light-YOLOv4 network in the widely used tracking dataset, second is training the segmentation branch on the video target segmentation dataset, and third is regulating the weight of the segmentation branch network by pretraining the appearance model on a large-scale pedestrian dataset.

5.1. The Training of Light-YOLOv4 Network. The images in the dataset are selected from large target tracking datasets such as OBT50, COCO [15], and ImageNet [16]. Since constantly updating the training samples is time-consuming, we choose a graphics processing server with high performance. The proposed method is implemented on a PC with Intel (R) Core (TM) i7-8086K CPU@4.00 GHz, x64-based processor, 64 G memory, RTX3090 ∗ 4 GPU, Linux operating system. The whole experiments are implemented on the Pytorch framework, using the stochastic gradient descent optimization algorithm, totally training 60 epochs, and the batch_size is set to 8, every epoch of which input 8000 photos for training. The training samples are shown in Figure 11.

5.2. The Training of Segmentation Branch. The segmentation branch only extracts the mask of the pedestrian. We distinguish the background and the foreground in the picture and set them to 0 and 1. In order to better distinguish people and objects, we train the network on a large number of different types of datasets, such as Open Images V4 and ImageNet. The Open Images V4 dataset contains 1.9 million images, 600 categories, and 15.4 million bounding-box annotations. It is currently the largest dataset with object location annotation information. The ImageNet dataset has more than 14 million pictures, covering more than 20,000 categories, which has more than one million pictures with clear category annotations and the location of objects. We train the network with the location of bounding boxes generated by the Light-YOLOv4. We uniformly adjust the image size to 64×64. Some training samples are shown in Figure 12.

Figure 11: The training samples of Light-YOLOv4.

Figure 12: Samples of the segmentation.

5.3. The Training of the Appearance Model. We select the Market-1501 [17] and MSMT17 [18] as the pedestrian reidentification datasets to guarantee the generalization capabilities of the proposed appearance model. The feature extraction network is a deep convolutional neural network. The Market-1501 dataset contains a large number of identities, each of which is taken by six disjoint cameras, and separated into a training set with 625 people and a test set with 636 people. The Market-1502 also includes 2793 false alarms from DPM, which are used as interference factors in simulated real scenes. Samples of Market-1502 are shown in Figure 13.

Figure 13: Samples of MSMT-1501 dataset.

MSMT17 is a large-scale reidentification dataset collected on a campus with 12 outdoor cameras and 3 indoor cameras. MSMT17 selects data from 4 days with different weather conditions in a month, and 3 hours of video is collected every day, covering three time periods: morning, noon, and afternoon. MSMT17 has similar points of view with Market-1501, but the scenarios are much more complicated. Sample images of MSMT17 are shown in Figure 14.

5.4. Experimental Results. We test the proposed model on MOT19, and the experimental results are shown in Figure 15, which demonstrates the tracking results of multiple moving objects. As we can see from the experimental results, the proposed algorithm has better accuracy when tracking multiple objects, and the accuracy will not be affected in the case of small objects. Besides, the proposed algorithm has satisfied the high requirement in terms of real-time. In complex situations such as occlusion and pedestrian crossing, the algorithm still maintains a high tracking accuracy without losing the target. We also compare the proposed model with several state-of-the-art tracking algorithms (such as: KCF [19], STRN [20], INARLA [21], Tracktor [22], STAM [23], and AMIR [24]).

5.4.1. Evaluation Criteria. We adopt several evaluation metrics to verify the effectiveness and availability of tracking algorithms. The calculation of the multiple objects tracking accuracy (MOTA) is shown in

$$\mathrm{MOTA} = 1 - \frac{\mathrm{FN} + \mathrm{FP} + \mathrm{IDSW}}{\mathrm{GT}} \in (-\infty, 1], \tag{17}$$

where FN represents false negatives, which is the sum of the number of false negatives in the entire video. FP refers to false positives, which is the sum of the number of false positives in the whole video. IDSW means the ID Switch. The smaller the value is, the greater the effectiveness is. GT is the sum of the number of ground truth objects in the video. The closer the MOTA is to 1, the better the performance of the tracker is. Due to the existence of the number of hops, there is a possibility that MOTA is less than 0. MOTA prevailingly indicates all matching errors of objects in tracking. MOTA intuitively measures the performance of the detection algorithm in detecting objects and keeping trajectories. Besides, MOTA is unrelated to detection accuracy.

The calculation of multiple objects tracking precision (MOTP) is shown in

$$\mathrm{MOTP} = \frac{\sum_{t,i} d_{t,i}}{\sum_t c_t}, \tag{18}$$

where c_t represents the number of matching in the t-th frame, and the matching error $d_{t,i}$ is the IOU of the detection frame in the t-th frame and the GT.

The calculations of Recall and Precision are formulated as $P = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP})$, $R = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$, where TP represents true positives, which is defined as the number of true answers in positive samples. FP represents false positives, which is defined as the number of false answers in positive samples. TN represents true negatives, which is defined as the number of true answers in negative samples. FN represents false negatives, which is defined as the number of false answers in negative samples.

The formulation of Identification Precision (IDP) is shown in equation (19), which refers to the accuracy of ID identification in each bounding box of the pedestrian.

$$\mathrm{IDP} = \frac{\mathrm{IDTP}}{\mathrm{IDTP} + \mathrm{IDFP}}, \tag{19}$$

where IDTP and IDFP represent the amount of ID's TP and FP, respectively.

Equation (20) illustrates the identification recall, which refers to the recall of ID identification in each bounding box of the pedestrian.

$$\mathrm{IDR} = \frac{\mathrm{IDTP}}{\mathrm{IDTP} + \mathrm{IDFN}}, \tag{20}$$

where IDFN represents the number of ID's FN.

IDF_1 means the identification F-score, which refers to the F-score of ID identification in each bounding box of the pedestrian.

$$\mathrm{IDF}_1 = \frac{2\mathrm{IDTP}}{2\mathrm{IDTP} + \mathrm{IDFP} + \mathrm{IDFN}}. \tag{21}$$

MT is the proportion of trajectories where most of the targets are tracked. A trajectory can be considered as MT if more than 80% of it is tracked. ML is the proportion of trajectories where most of the objects are lost. Only the trajectory tracked less than 20% can be regarded as ML.

IDS is the total number of identity switches. Frag, which is also called FM, refers to fragmentation. And the FPS means frames per second.

5.4.2. Comparison. The comparison of evaluation metrics on MOT19 is shown in Table 1. "↑" means the higher the score is, the better the performance is; "↓" means the opposite.

It can be seen that the MOTA of the proposed algorithm on MOT19 is 0.7% higher than the FFT algorithm, which has the best performance in 6 state-of-the-art methods. The

Figure 14: Samples of MSMT17 dataset.

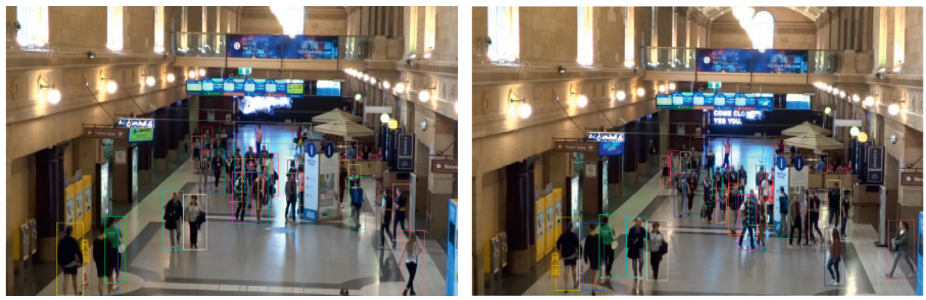


Figure 15: Experimental results of the algorithm on mot19 data set.

Table 1: Comparison effect of experimental model on MOT19 dataset.

Method	MOTA↑	MOTP↑	IDF_1↑	IDP↑	IDR↑	MT↑ (%)	ML↓ (%)	FP↓	FN↓	IDS↓	Frag↓	FPS↑
KCF [19]	48.8	75.7	47.2	65.9	36.7	15.8	38.1	5875	86567	906	**1116**	106
STRN [20]	48.5	73.7	**53.9**	72.8	**42.8**	17.0	34.9	9038	84178	747	2919	90
FFT [21]	56.5	78.1	50.1	64.4	41.1	23.6	29.4	5831	71825	1635	1607	54
Tracktor [22]	54.4	78.2	52.5	71.3	41.6	19.0	36.9	3280	79149	682	1480	132
STAM [23]	46.0	74.9	48.1	62.7	38.2	14.6	43.6	6895	91117	473	1422	73
Sadeghian et al. [24]	47.2	75.8	42.6	52.1	36.6	14.0	41.6	**2681**	92856	774	1675	36
Ours	**57.2**	**78.3**	51.7	**75.2**	38.8	**26.1**	**23.6**	6913	**25753**	**384**	1431	152

results in Table 1 show that the proposed method improves the accuracy of object tracking. Meanwhile, we optimize the structure of the detection and segmentation algorithm. While elevating the accuracy, the improved structure can also reduce the amount of calculation. Therefore, the proposed tracking algorithm can not only work at a fast speed but meet the requirements of practical projects. The speed of the proposed model is 152 frames per second on a single graphics card, which can fully achieve real-time detection and tracking. Under the same conditions, the fastest tracking speed of the FFT algorithm is only 5 frames per second. The Tracktor algorithm has only 11 frames per second. Our model has the ability to accurately detect, segment, and accurately track objects. Moreover, our model has the basic characteristic of universality.

6. The Application in the Substation

Considering the requirements of real-substation situations, we expand the dataset. Test videos come from the real-world scenarios in the substation. Taking pedestrians in the substation as an example, the collected substation monitoring videos are divided into two groups with the same lighting but different occlusion conditions to evaluate the performance of the moving object tracking method. The first group of videos is used to evaluate the tracking performance when there are fewer obstructions. The second group of videos is used to test the robustness and accuracy of the tracking method in the case of serious obstructions. The detailed parameters are shown in Table 2. The entire experiment is implemented on the Pytorch framework, using the stochastic gradient descent optimization algorithm, a total of 60 epochs are trained, batch_size is set to 8, the initial learning rate is set to 0.001, and the number of images trained for each epoch is 15,073.

6.1. Dataset. Due to the complicated environment in the substation, we convert the surveillance video of the pedestrian in the substation into a picture frame by frame in order to more accurately detect the position of the pedestrian in the image and ensure the tracking effect. In addition,

Table 2: Statistics of substation scene data set.

Video	Video time (s)	Number of frames	Average number of staff members	Illumination, camera downward tilt angle	Occlusion
Group 1	41	1640	5	Strong illumination, 30° downward	There are few obstructions and the environment is relatively open
Group 2	52.8	2112	2	Strong illumination, 30° downward	There are few obstructions and the environment is relatively open

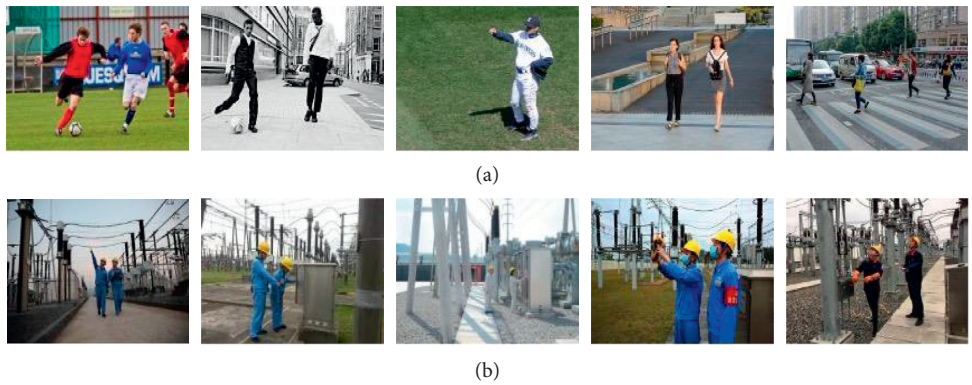


Figure 16: Data set example. (a) Common dataset example. (b) Example of the substation dataset.

in order to ensure the generalization of the model, we select the same number of pictures containing pedestrians from the public dataset and then mix the pictures to form the detection dataset, and the number of datasets is about 15,073. Next, we divide the dataset into a training set, validation set, and test set, the number of images in the training set is 9,011, the validation set is 3,041, and the test set is 3,021. We annotate the images according to the image annotation standard of the Pascal VOC dataset and output only the person category, which is also counted at the same time (for example, person1, person2, person3, person4, ...) as the pedestrian ID for the initial frame tracking. Examples of the dataset and the labeled dataset are shown in Figures 16 and 17.

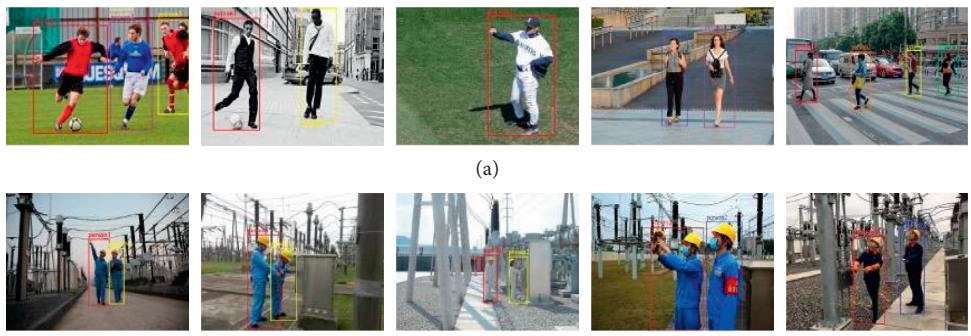
6.2. Results and Analysis. We test the proposed tracking model on two datasets containing substation surveillance videos, as shown in Figures 18 and 19. Figures 18 and 19 show the tracking results on the first and second groups of videos (when the occlusion is not serious), and the detection and tracking results are output every 2.5 s (100 frames). Experimental results prove that the proposed pedestrian tracking method in substations using integrated metric learning can effectively detect and track multiple pedestrians in the substation and has better robustness in the severely occluded environment (the second group of substation surveillance videos).

We can see from Figure 18 that under the circumstance that the occlusion is not serious, the proposed tracking method can accurately track multiple working pedestrians in the substation. When a small amount of pedestrians is crossed, overlapped, and occluded, the proposed tracking method still accurately reidentifies multiple working pedestrians.

There are more insulators, poles, and towers in the second group of surveillance videos of the substation, which causes dense obstruction. And the proposed algorithm still accurately tracks pedestrians in the case of severe occlusion. The visually tracking results are shown in Figure 19. Since our model can involve both the motion information and the performance information of pedestrians, it boosts the reidentification ability for the problem of pedestrians' loss. Tables 3 and 4 represent the extensive experiments under occlusion in two degrees.

It can be seen from the two tables that no matter there are light or heavy occlusions in the video, introducing more training data, adaptive segmentation training methods, and fusion measurement methods can significantly improve the tracking effect of pedestrians. The MOTA, MOTP, IDP, ML, and IDS results of our model are better than other algorithms. Moreover, in the light obstruction scenes, the MOTA of our model improves 1.6% compared to the FFT. In a scene with severe obstructions, the MOTA of our model increases to 4.7%, and the IDS of our model keeps the lowest.

In addition, since we adopt the lightweight Light-YOLOv4 real-time detection network framework and fast segmentation algorithm, the proposed algorithm guarantees a fast running speed. We can see from the experimental results that the running speed of the adaptive polymorphic fusion tracking algorithm is 134 FPS in a scene with light obstructions and 127 FPS in a scene with dense obstructions. Compared with other algorithms, the real-time performance of our method has been greatly improved. Experiments show that the adaptive polymorphic fusion tracking algorithm proposed in this paper can not only ensure tracking accuracy



(a)

(b)

Figure 17: Example of the annotated data set. (a) Example of the annotated common dataset. (b) Example of the annotated substation dataset.

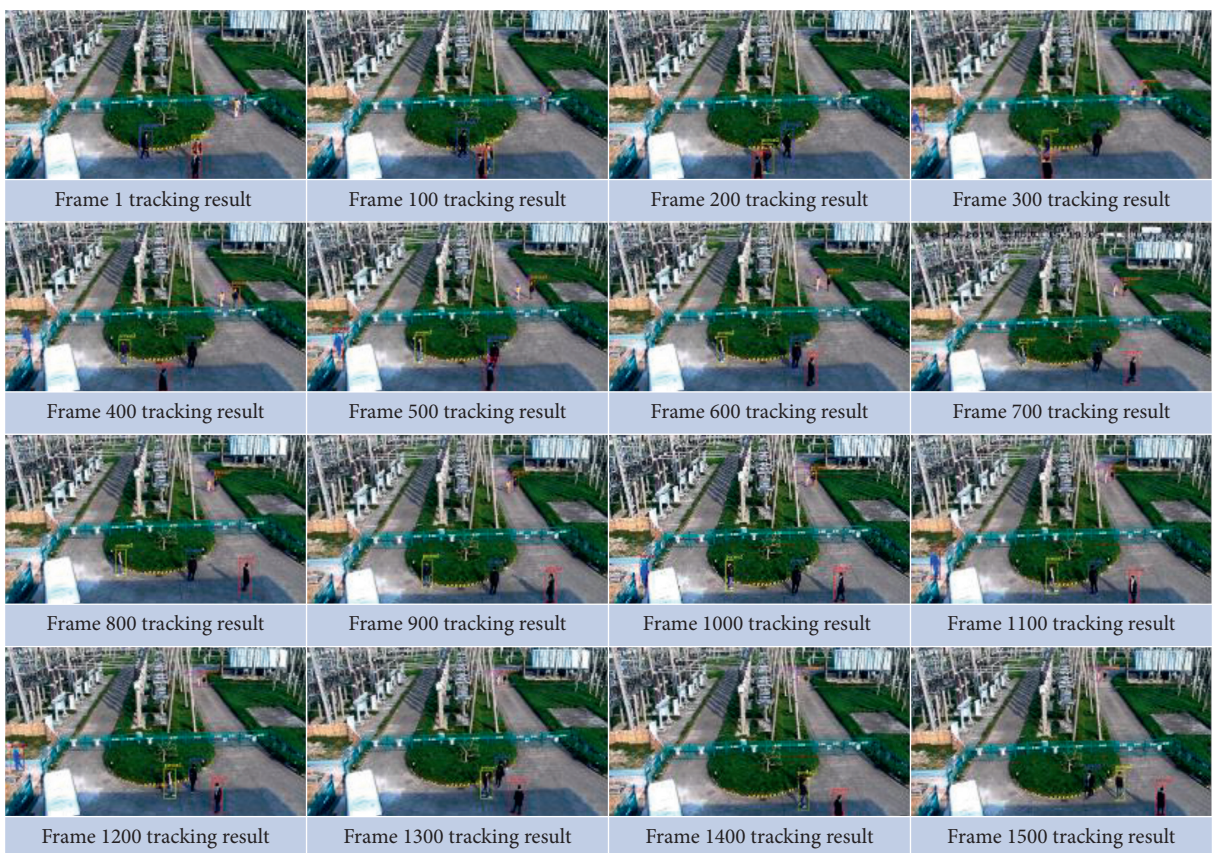


Figure 18: Tracking effect of this algorithm in the first group of videos (occlusion is not serious).

but also has extremely fast running speed. Our algorithms also solve the problems of dense occlusion and pedestrian loss by fusing pedestrian motion and performance information. Considering the actual scenarios and demands of substations, the proposed algorithm is more practical.

We also tested the proposed model on MOT19. The experimental results are shown in Figure 15 in Section 5.4. Under complex conditions such as occlusion and pedestrian crossing, the algorithm can still maintain high tracking accuracy without losing the target. Therefore, our algorithm

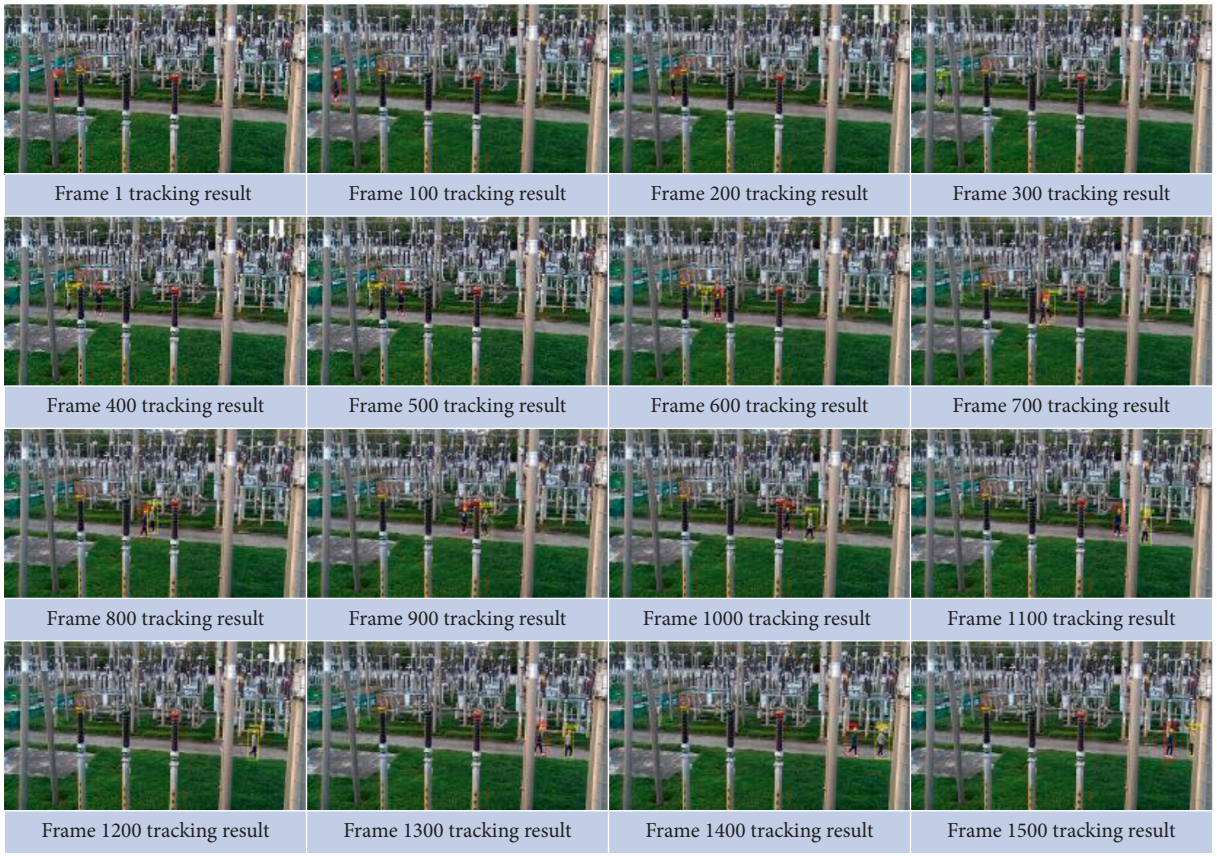


Figure 19: Tracking effect of this algorithm in the second group of videos (serious occlusion).

Table 3: Tracking performance of tracker under the first group of videos (less occlusion).

Method	MOTA↑	MOTP↑	IDF$_1$↑	IDP↑	IDR↑	MT↑ (%)	ML↓ (%)	FP↓	FN↓	IDS↓	Frag↓	FPS↑
KCF [19]	48.9	75.8	47.3	66.1	36.9	16.9	39.1	5865	86567	906	**1116**	97
STRN [20]	48.6	73.8	**54.1**	72.9	**43.9**	17.0	35.9	9038	84178	747	2919	74
FFT [21]	56.6	78.2	50.2	64.5	42.1	23.6	30.4	5931	69825	1635	1607	42
Tracktor [22]	54.5	78.3	52.6	71.4	42.6	20.0	37.9	3365	79249	682	1480	114
STAM [23]	46.1	74.4	48.2	62.9	39.2	15.6	45.6	6796	91137	473	1422	68
Sadeghian et al. [24]	47.3	75.9	42.7	52.2	37.6	15.0	42.6	**2693**	92856	774	1675	33
Ours	**57.2**	**78.3**	51.7	**75.3**	38.8	**26.1**	**25.6**	5913	**25753**	**384**	1431	**134**

Note. The data marked in bold indicates the optimal algorithm of each index corresponding to each video.

Table 4: Tracking performance of the tracker under the second group of videos (more obstructions).

Method	MOTA↑	MOTP↑	IDF$_1$↑	IDP↑	IDR↑	MT↑ (%)	ML↓ (%)	FP↓	FN↓	IDS↓	Frag↓	FPS↑
KCF [19]	43.7	64.7	42.2	48.9	34.7	14.8	38.1	5972	92352	1023	**1237**	92
STRN [20]	42.4	62.5	**50.9**	50.8	**42.5**	15.7	34.9	9934	87230	843	3326	73
FFT [21]	49.5	67.1	48.1	46.4	40.3	21.2	29.4	6027	75292	1843	1849	41
Tracktor [22]	50.4	63.2	49.5	49.3	40.3	17.4	36.9	4980	80422	704	1593	112
STAM [23]	41.0	63.9	45.1	41.7	35.5	13.4	43.6	6895	97739	536	1744	67
Sadeghian et al. [24]	42.2	62.8	39.6	39.1	32.4	14.2	41.6	**4581**	10745	832	1763	31
Ours	**51.2**	**68.3**	49.7	**54.2**	36.7	**22.3**	**28.6**	7613	**47762**	**523**	1631	**127**

Note. The data marked in bold indicates the optimal algorithm of each index corresponding to each video.

has general promotion potential, but the current research is based on a single camera. Cross-camera and multicamera scenes are the directions of our follow-up research.

7. Conclusion

In this paper, we have carefully designed an adaptive polymorphic fusion-based fast tracking algorithm, which integrates motion and performance information of pedestrians to achieve real-time tracking in substations. First, we proposed a lightweight real-time target detection method called Light-YOLOv4, the backbone network of which is based on CSPDarknet-Tiny network to enhance the performance of detection from three aspects, containing multibranch feature fusion, grouped self-attention, and hierarchical multiscale prediction. Second, we exploited the SOLO segmentation algorithm to gain the performance information by inputting the fine masks of a pedestrian into the deep convolutional neural network. Third, we not only adopted the Kalman filter to predict the trajectory of the detected pedestrian but also used the Mahalanobis distance to represent the matching degree between the current state and the previous state of pedestrians and denoted the motion information with the motion matching degree. Finally, we obtained the tracking results using the integrated metric learning method. Experimental results prove that the proposed algorithm has better real-time tracking performance, and it can solve the severe occlusions in the substation. However, there are still some problems in real-world scenarios, such as the cross-camera tracking scene, the night, and other scenes. For future work, we expect to investigate more in the above challenging scenarios [10].

Data Availability

Figures 8, 11–14, 16(a), and 17(a) were taken from public datasets: (1) Figure 8 was taken from the public dataset PASCAL VOC2012, which you can download from this website: http://host.robots.ox.ac.uk/pascal/VOC/voc2012/. (2) Figure 11 was taken from the public dataset INRIA-Person, which you can download from this website: https://pascal.inrialpes.fr/data/human/. (3) Figure 12 was taken from the public dataset DAVIS 2016, which you can download from this website: https://davischallenge.org/. (4) Figure 13 was taken from the public dataset Market-1501, which you can download from this website: https://drive.google.com/file/d/0B8-rUzbwVRk0c054eEozWG9COHM/view?usp=sharingttps://www.pkuvmc.com/publications/msmt17.html. (5) Figure 14 was taken from the public dataset MSMT17, which you can download from this website: https://www.pkuvmc.com/publications/msmt17.html. (6) Figure 15 was taken from the public dataset MOT19, which you can download from this website: https://motchallenge.net. (7) Figure 16(a) was taken from the public dataset PASCAL VOC2012, which you can download from this website: http://host.robots.ox.ac.uk/pascal/VOC/voc2012/. (8) Figure 17(a) was taken from the public dataset PASCAL VOC2012, which you can download from this website: http://host.robots.ox.ac.uk/pascal/VOC/voc2012/.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Fang, X. Shu, and L. I. Dewu, "A method based on machine vision for opening-closing status recognition of substation disconnecting switches," *Journal of Hunan University of Technology*, vol. 31, 2017.

[2] J. Gabela, G. Retscher, S. Goel et al., "Experimental evaluation of a UWB-based cooperative positioning system for pedestrians in GNSS-denied environment," *Sensors*, vol. 19, no. 23, p. 5274, 2019.

[3] S. S. Blackman and S. Samuel, "Abstracts of previous tutorials in this series: multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, no. 3, pp. 90–96, 2016.

[4] L. Wan, K. Liu, Y.-C. Liang, and T. Zhu, "DOA and polarization estimation for non-circular signals in 3-D millimeter wave polarized massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3152–3167, 2021.

[5] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2021.

[6] H. Wang, X. Li, R. H. Jhaveri et al., "Sparse Bayesian learning based channel estimation in FBMC/OQAM industrial IoT networks," *Computer Communications*, vol. 176, pp. 40–45, 2021.

[7] L. Wan, L. Sun, K. Liu, X. Wang, Q. Lin, and T. Zhu, "Autonomous vehicle source enumeration exploiting non-cooperative UAV in software defined internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3603–3615, 2021.

[8] L. Wan, Y. Sun, L. Sun, Z. Ning, and J. J. P. C. Rodrigues, "Deep learning based autonomous vehicle super resolution DOA estimation for safety driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4301–4315, 2021.

[9] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, 2016.

[10] L. Chen, H. Ai, Z. Zhuang, and C. Shang, "Real-time multiple people tracking with deeply learned candidate selection and person re-identification," in *Proceedings of the IEEE International Conference on Multimedia & Expo*, San Diego, CA, USA, 2018.

[11] Z. Wang, L. Zheng, Y. Liu, Y. Li, and S. Wang, *Towards Real-Time Multi-Object Tracking*, Springer, Cham, Switzerland, 2020.

[12] Y. Zhang, C. Wang, X. Wang, W. Zeng, and W. Liu, "FairMOT: on the fairness of detection and re-identification in multiple object tracking," *International Journal of Computer Vision*, vol. 129, pp. 3069–3087, 2021.

[13] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017.

[14] X. Wang, T. Kong, C. Shen, Y. Jiang, and L. Li, "Solo: segmenting objects by locations," in *Proceedings of the European Conference on Computer Vision*, pp. 649–665, Glasgow, UK, 2020.

[15] T.-Y. Lin, M. Maire, S. Belongie et al., "Microsoft coco: common objects in context," in *Proceedings of the European Conference on Computer Vision*, pp. 740–755, Zurich, Switzerland, 2014.

[16] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: a large-scale hierarchical image database," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, 2009.

[17] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, "Scalable person re-identification: a benchmark," in *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 2015.

[18] L. Wei, S. Zhang, W. Gao, and Q. Tian, "Person transfer gan to bridge domain gap for person re-identification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018.

[19] P. Chu, H. Fan, C. C. Tan, and H. Ling, "Online multi-object tracking with instance-aware tracker and dynamic model refreshment," in *Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, 2019.

[20] J. Xu, Y. Cao, Z. Zhang, and H. Hu, "Spatial-temporal relation networks for multi-object tracking," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, Republic of Korea, 2019.

[21] J. Zhang, S. Zhou, X. Chang et al., "Multiple object tracking by flowing and fusing," 2020, https://arxiv.org/abs/2001.11180.

[22] P. Bergmann, T. Meinhardt, and L. Leal-Taixé, "Tracking without bells and whistles," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, Republic of Korea, 2019.

[23] Q. Chu, W. Ouyang, H. Li, X. Wang, B. Liu, and N. Yu, "Online multi-object tracking using CNN-based single object tracker with spatial-temporal attention mechanism," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017.

[24] A. Sadeghian, A. Alahi, and S. Savarese, "Tracking the untrackable: learning to track multiple cues with long-term dependencies," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 1827155, 16 pages
https://doi.org/10.1155/2021/1827155

Research Article

Predicting Joint Effects on CubeSats to Enhance Internet of Things in GCC Region Using Artificial Neural Network

Faris A. Almalki[1] **and Soufiene Ben Othman**[2]

[1]*Department of Computer Engineering, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*
[2]*Prince Laboratory Research, IsitCom, University of Sousse, Hammam Sousse, Tunisia*

Correspondence should be addressed to Soufiene Ben Othman; ben_oth_soufiene@yahoo.fr

Received 4 September 2021; Revised 28 October 2021; Accepted 5 November 2021; Published 25 November 2021

Academic Editor: Han Wang

Satellite telecommunication systems promise to bridge digital gaps and deliver wireless communication services to any corner of the world. However, despite satellites' global connectivity and wide footprint, still atmospheric and dust impairments are open challenges that face satellite systems, especially at high-frequency bands in arid and semiarid regions. Therefore, this paper aims to predict joint effects of atmospheric and dust attenuations in Gulf Cooperation Council (GCC) countries on CubeSat communications using Artificial Neural Network (ANN). The prediction model has been carried out using a massive Multiple-Input Multiple-Output (MIMO) antenna payload at K-frequency Bands. Consider these joint effects have positive relations in calculating satellites link margin, which leads to obtaining efficient communication system, delivering better quality of service (QoS), and enhancing Internet of Things (IoT) connectivity, or even Internet of Space Things (IoST). Predicated results infer that the ANN attenuation predictions, along with the 5G MIMO antenna on-board the CubeSat, offer much promise channel model for satellite communications, which in turn leads to not only supporting IoT connectivity but also reducing power consumption, thus enhancing lifetime of CubeSat. Also, this study can provide a reference for CubeSat engineers to guarantee large-capacity communication.

1. Introduction

Traditional wireless communication systems provide services with a good level of data rates, reconfigurable provision with various dynamic coverage demands. However, the deployment of these enabling technologies has led to a huge rise in the demand for mobile communications, partly due to the exponential growth in multimedia traffic and the emergence of a new type of technology such as the Internet of Things (IoT) or Big Data. However, space-based wireless communication systems such as satellites and unmanned aerial vehicles (UAVs) promise to provide global connectivity, because they contribute effectively to linking trillions of objects and sensors, all of which generate real-time data. This also helps in boosting IoT and Internet of Everything (IoE) technologies, which are considered fuel for the Fourth Industrial Revolution (4IR) [1–3]. Figure 1 shows the main technological pillars of 4IR include IoT and space-based vehicles (e.g., Drones and Satellites). The 4IR and its enormous processing power represent creative digitization based on a combination of symbiotically interacting technological breakthroughs through innovative algorithms. Further, the combination between IoT and Drones and Satellites is vital, which would massively pave the way to hyperconnected societies, thus leading to global smart connectivity [4–6].

Figure 1 displays that Artificial Intelligent (AI) is one of the key pillars of 4IR, where such an innovative technology can be considered a great booster of the 4IR and digital transformation in any modern society. One of those AI techniques is an Artificial Neural Network (ANN), which is a computing system inspired by the biological neural networks that constitute biological brains, where all neurons contain individual weights and bias values that are parallel and organized nonlinear components. Further, the ANN technique

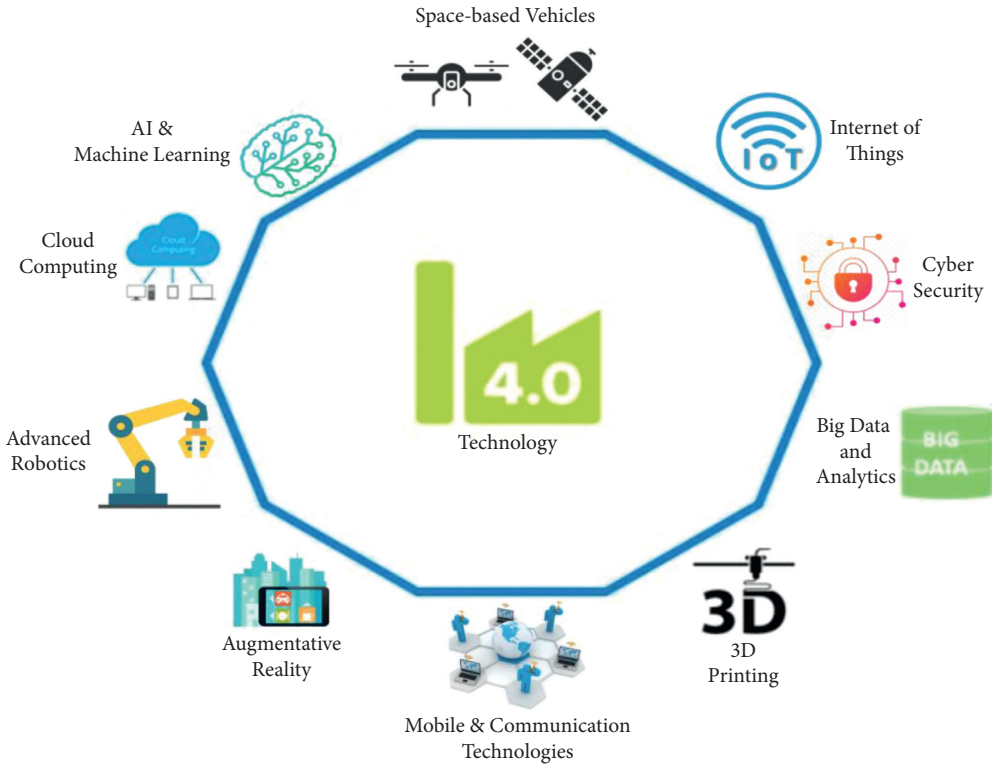


Figure 1: The main 4IR pillars.

is widely used for highly accurate predictions and/or optimization due to its simplicity and efficiency [7]. The authors in [8–11] are representative examples of related work that considered ANN for optimization and enhancement.

Researchers in space-based systems confirm that reliability, flexibility, efficiency, wide applicability, Line of Sight (LoS) connectivity, low latency, rapid deployment, and cost-effectiveness are the main advantages of using such a technology. The authors have emphasised that space is stepping up to the connectivity challenge posed by the 4IR, where one of the driving forces of this change has been the introduction of next-generation high-technology cube satellite systems [12–14]. Further, space-based communication systems include satellites enticing due to their various applications and low design-and-deployment cost [15–17]. Furthermore, in the last few years, there has been a tremendous interest in cube satellites, which are also known as "CubeSats" or "Nanosatellites," built to standard dimensions of few tens on centimetres and weigh few kilograms. These small satellites can fly on a rocket to space to provide opportunities for small satellite payloads to serve some applications as follows [18, 19]:

(i) Technology demonstration: in the harsh environment of space to test new instruments or materials

(ii) Academia and educational projects: which provide a unique hands-on experience in developing space missions from design to launch and operations

(iii) Commercial applications: telecommunications, remote sensing, high-resolution aerial imagery, ship tracking, and formation flying

The nanosatellites that operate at Low Earth Orbit (LEO) based on Van Allen radiation belts at altitude start from around 160 km to 600 km from the Earth facing some challenges, for instance, propagation impairments in last-mile communications, especially when using high-frequency bands often assigned to space-based communication systems (e.g., above 10 GHz). Although high-frequency bands offer wide bandwidths, they are more vulnerable to signal degradation, as radio signals get absorbed by atmospheric rain, snow, ice, or even dust. The attenuations' factors vary from one region to another worldwide [20–22]. The authors in [23, 24] introduced a holistic overview of various aspects of CubeSats missions, services, recent advances, and future challenges for further research. The authors stressed that the research on cube satellites for communications is still in its early phase. IoT, Internet of Space Things (IoST), low-power long-range networks, Iridium communications, earth-remote sensing, and machine learning are opened research areas for cube satellites. The authors in [25] have introduced CubeSats as "Eyes in the Sky" because it monitors and offers various remote sensing with capabilities of IoST. According to a forecast by SpaceWorks Enterprises, Inc. (SEI), more than 2,800 nano-/microsatellites will require launch over the next five years. Figure 2 shows a comparison between small satellites application trends that weigh (1–50 kg), where the left pie chart shows application trends from 2014 to 2018, while the right one from 2019 to 2023 [26]. Figure 3 illustrates Cube satellite wide applications.

Woellert et al. [27] shed light on CubeSats capabilities in science education and spur technology for developing nations. Aslan et al. [28] constructed a CubeSat project in Sharjah–Emirates, one of the Gulf Cooperation Council

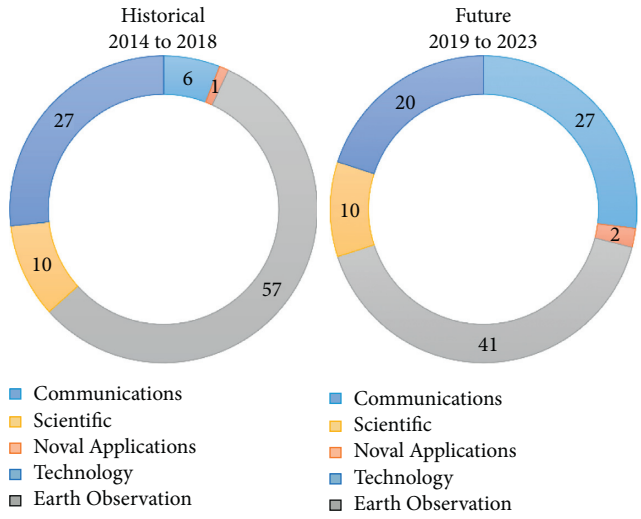


Figure 2: Cube satellite application trends (1–50 kg).

Figure 3: Cube satellite wide applications.

(GCC) countries. The project aims to educate students in cutting edge technologies including space technology, which puts human capacity development in its centre and support United Nations Sustainable Development Goals (UN SDG) for 2030. Akyildiz et al. [29] highlighted the importance of atmospheric attenuation when designing link budget analysis at an altitude of 400 km with above 10 GHz frequency.

Qu et al. [30] focused on rain attenuation caused by satellite communication operating on frequency bands above 10 GHz, as they are sensitive to atmosphere physical phenomena. The work predicted real-time satellite channel fading caused by rain using stochastic dynamic modelling (SDM). Luini and Capsoni [31] addressed the joint effects of rain and clouds on Ka-Band for LEO satellite using

stochastic model. The results give better understanding of link budget; however, more atmospheric effects (e.g., gasses, water vapor, and oxygen) are recommended for further work.

Mpoporo and Owolawi [32] inferred that rain attenuation is the most unfavourable impact on signal transmission at high-frequency bands, especially in places with high rainfall. This paper aims to estimate rain attenuation on satellite links using the ITU-R model in South Africa. Further, Wang et al. [33] gave rain attenuation analysis of the Ka-Band satellite communication in the Indian and Pacific Oceans. Reasonable results of the ITU-R model in predicting rain attenuation have been presented. Averly and Suryana [34] also considered the ITU-R model and simulated it in MATLAB at below 3 GHz frequencies. However, Regonesi et al. [35] emphasised that the ITU-R model with some limitations and inaccuracies, especially in high frequencies elevation angles and rainfall cases.

An experimental work of radio link design for a CubeSat communication system is presented in [36]. The CubeSat is placed at 400 km altitude using a dipole antenna with an omnidirectional pattern and frequency band of 401 MHz. Experimental results of link budget are 137 dB of free space path loss (FSPL), 5 Kb/S of effective throughput, and signal-to-noise ratio (SNR) floats between 10 dB and 24.54 dB at bit error rate (BER) of 1×10^{-5}. Atmospheric effects have been neglected.

Charbit et al. [37] showed a Narrow-Band IoT (NB-IoT) for satellite-to-ground communication with atmospheric effects. Carrier-to-noise ratio (CNR) of satellite links are evaluated at altitude of 600 km and 2 Ghz frequency band with wide consideration of atmospheric effects (gasses, rain, scintillation). Results are adequate, but the work does not include high-frequency bands, where it truly shows the predictions of the atmospheric effects.

Harb et al. [38] confirmed that dust is a vital cause of wireless channel impairments, which are regularly observed in arid and semiarid areas worldwide. References [38, 39] presented a dust storm model that affects high-frequency satellite-to-ground communications in Saudi Arabia and other countries in the Middle East. The dust model includes mathematical variables such as visibility, average dust particles size, and permittivity indices. SNR of various elevation angles, altitude, and frequency bands have been analysed in case of using a directional-hemispherical antenna. Predications of dust attenuations help in optimally designing satellite link budgets.

Motivated by the below observations through the related studies, major portions of satellite link budgets have considered either part of atmospheric impairments mainly rains attenuation or dust attenuation separately. Therefore, this paper aims to provide a comprehensive prediction model that encompasses joint effects of atmospheric and dust impairments from CubeSat communication aspect in the GCC region using ANN. The prediction model has been carried out using a massive Multiple-Input Multiple-Output (MIMO) antenna payload at K-frequency Bands. The massive MIMO has combined with Orthogonal Frequency Division Multiplexing (OFDM), which can divide a radio channel into many close-spaced subchannels to deliver extra reliable communications over multiple antennas at high speeds. The atmospheric effects comprise rain, gasses, clouds, scintillation, and dust. Figure 4 shows a bird's eye view of a cube satellite network communication architecture, which consists of space segment and ground segment. The former has the cube satellite with communication payloads at the LEO layer. The latter shows a ground station and many wireless communication links within a footprint that can serve fixed/mobile users. Clearly, signals get attenuated as they pass by atmospheric effects, affecting the received signals at the ground users. Services such as Internet connectivity and active or passive sensing can benefit from the strong received signal from the CubeSat side.

The rest of this paper is organized as follows: Section 2 presents the proposed predictions model followed by Section 3 that details the performance analysis and discusses results. Section 4 concludes this paper.

2. Proposed Predictions Model

Predicating weather impairments is an essential component when evaluating optimal link budgets for satellite communication links. This section sheds light on the proposed prediction model that comprehensively includes joint effects of atmospheric and dust impairments from CubeSat to GCC countries using an ANN technique. Since researchers and designers need to consider a wide range of parameters that can impact communication links. This work considers a CubeSat simulation library for Aerospace Blockset in MATLAB. It simulates, analyses, and visualizes the motion and dynamics of CubeSats and nanosatellites, as Figure 5 shows the Cube satellite network communication architecture. The CubeSat simulation toolbox provides standards-based tools and enormous parameters and network configurations for designing, simulating, and verifying satellite communications systems and links. This section contains three subsections: (a) CubeSat Link Margin, (b) Atmospheric Effects Calculation, and (c) ANN Framework.

2.1. CubeSat Link Margin. Calculating atmospheric and dust impairments is affecting the total carrier-to-noise ratio (C/N), as per equations (1) and (2), because it is one of the performance quality indicators of a satellite communication channel. Further, the most critical performance parameter is bit energy per noise density ratio (E_b/N_0) as per equation (3), which is also influenced over a satellite network due to its characteristics such as long propagation, frequency, and bandwidth [40, 41]:

$$\left(\frac{C}{N}\right)_T = P_t + G_C + G_S - \left(\frac{4\pi\ d}{\lambda^2}\right) - F - 10\ \log_{10}\left(KT_0\ B\right) - \text{Losses}, \tag{1}$$

$$\text{d[km]} = 2\,E_r\left[\cos^{-1}\left(\frac{E_r}{E_r + h_t}\ *\ \cos\left(\theta\right)\right) - \theta\right], \tag{2}$$

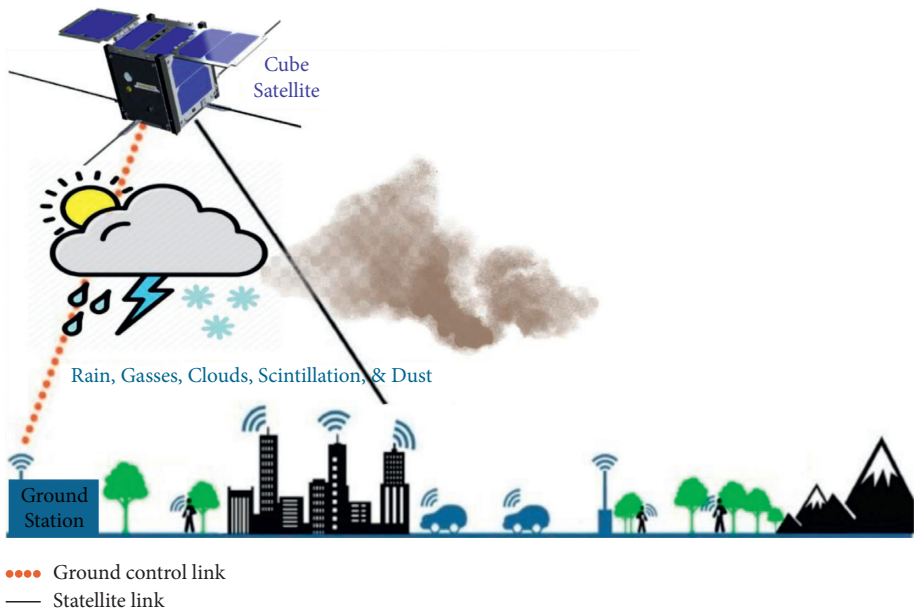


Figure 4: Cube satellite network communication architecture.

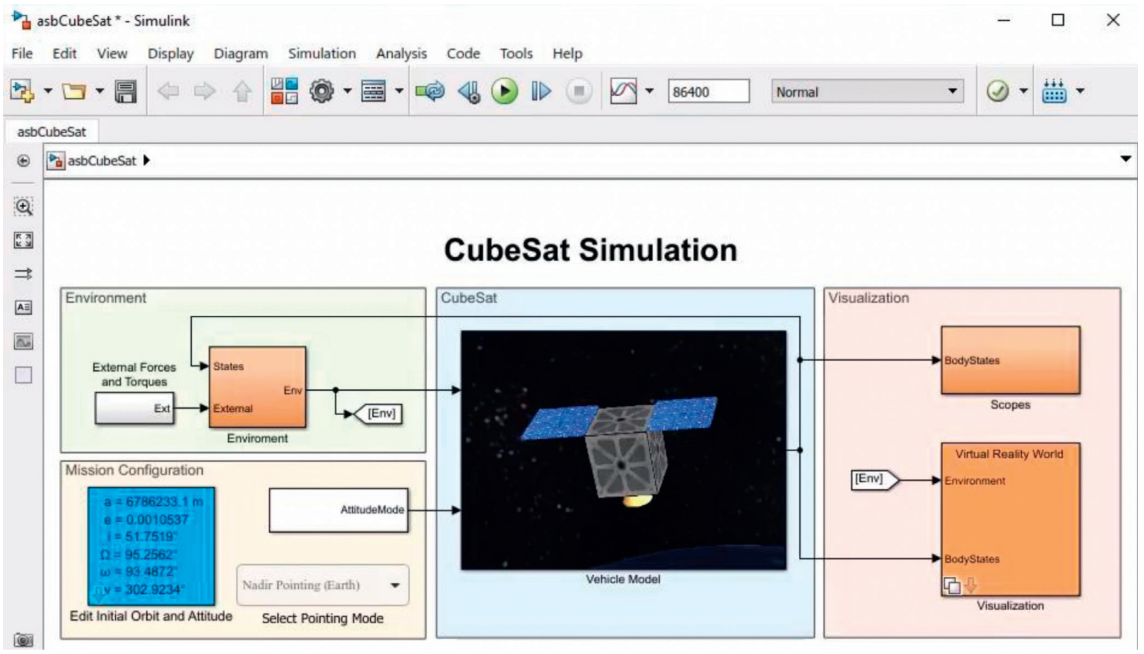


Figure 5: CubeSat network communication architecture.

$$\frac{E_b}{N_0} = \left(\frac{C}{N}\right)_T \cdot \left(\frac{Bw}{R_b}\right), \tag{3}$$

where P_t denotes transmitter power; G_C denotes CubeSat antenna gain; G_S denotes ground station antenna gain; d denotes the distance from CubeSat to Earth station/terrestrial users; λ denotes wavelength, which is the ratio of the speed of light (c) to the frequency of transmission f; d is calculated with consideration of elevation angle θ, which is a vital aspect in space-based wireless communication systems to achieve better LoS connectivity; E_r denotes the Earth's radius (6,371 km); h_t denotes CubeSat's altitude; F denotes the receiver noise figure; K denotes Boltzmann's constant (1.38.10–23 Ws/K); T_0 denotes the absolute temperature (290K); B denotes the receiver noise bandwidth; *Losses* represent atmospheric (rain, gasses, clouds, and scintillation) and dust attenuations; Bw denotes the required bandwidth; and R_b denotes data bit rate.

A massive 5G MIMO antenna is used in the proposed model design, where it consists of a planar phased array of 64 patches designed for operation at up to tens of GHz, along with an optimum feed location to produce optimal return

loss. Further, the considered massive 5G MIMO antenna uses adaptive beamforming with spatial multiplexing to increase data rates and decrease interference. To propagate signals well from CubeSat to Earth stations/users, each patch of massive 5G MIMO antenna uses an adaptable phase offset to steer beamforming at various angles (°) towards the desired direction, where the main beams can be calculated as follows [42–44]:

$$W_n = \exp\left[-j\left(\frac{2\pi}{\lambda}\right)\sin\ (\theta_d)\,[x_n\ \ \cos(\phi_d) + y_n \sin(\phi_d)]\right], \tag{4}$$

where θ_d, ϕ_d denote phases at $x_n$and y_n.

2.2. Atmospheric Effects Calculation. This section presents the mathematical modelling used to evaluate the performance of CubeSat links and determine different types of atmospheric losses. These atmospheric and dust impairments can seriously affect satellite-to-Earth links at high frequencies, which are often above 10 GHz. Equations (5)–(9) represent losses (attenuations) of rain, gasses, clouds, scintillation, and dust, respectively [33, 38, 39, 45, 46]:

$$R_A = \alpha R^{K\prime}.\, h_t, \tag{5}$$

$$G_A = \frac{A_0 + A_w}{\sin\theta}, \tag{6}$$

$$C_A = \gamma_{\text{cloud}}\left(\frac{\text{LWC}}{\sin\theta}\right), \tag{7}$$

$$S_A = a(p).\,\sigma, \tag{8}$$

$$D_A = \frac{\left(5.67 \times 10^2\right)}{V(\theta).\, r_{e0}^2 .\, \lambda} * \frac{\varepsilon''}{(\varepsilon\prime + 2)^2 + \varepsilon''^2} \sum_i^n p_i * r_i^3, \tag{9}$$

where R_A denotes rain attenuation, α and $K\prime$ are empirical coefficients that depend on frequency band and temperature, R denotes rain rate in mm/hr, G_A denotes gasses attenuation, denotes oxygen absorption, A_w denotes water vapor, C_A denotes cloud attenuation, LWC denotes liquid water content, S_A denotes scintillation attenuation, $a(p)$ denotes the time percentage factor for time percentage, σ denotes standard deviation of the signal amplitude, D_A denotes dust attenuation as a function of propagation angle and frequency, V denotes visibility in Km, r_{e0}^2 denotes equivalent particle radius, $\varepsilon\prime$and $\varepsilon''$are real and imaginary parts of the dielectric constant of dust, and $\sum_i^n p_i * r_i^3$ denote summation of the probability particle size between r_i series of particle volume at the lowest and highest layer level.

2.3. ANN Framework. There are various types of machine learning techniques, which are widely used for high accurate predictions/optimization. One of those well-known techniques is ANN that mainly comprises three layers: input layer, hidden layer, and output layer. Figure 6 shows universal illustration of an ANN, where all neurons contain individual weights and bias values that are parallel and organized nonlinear components, which are repeatedly appropriate for improving performance during training [47–51].

This work uses Multilayer Feedforward Network Architecture (MLP), which is one of the most common type of ANNs used in the literature. MLP is used to predict the link margin of CubeSat as per the following equations:

$$\alpha = b + \sum_{n=1}^{N} y_n t_n, \quad n = 1, 2, \ldots, N, \tag{10}$$

$$\chi(\alpha) = \frac{1}{\left(1 + e^{-\alpha}\right)}, \tag{11}$$

$$E = \frac{1}{2}\sum_{h-1}^{H} (A_0 - D_0)^2\left(\frac{mm}{h}\right), \tag{12}$$

$$\frac{\partial E_{\text{total}}}{\partial A_0} = -(D_0 - A_0), \tag{13}$$

$$\Delta w_i = \eta \frac{\partial E}{\partial A_i}, \quad i = 1, 2, \ldots, I, \tag{14}$$

where α denotes the additive and output function, which is a sigmoid activation function of χ that denotes neuron prediction, b denotes bias input, y_n denotes nth weight, t_n denotes nth input, and N denotes the number of inputs. E estimates the error when training the ANN occurs, D_0 denotes expected output (target), mm denotes multilayer of the feedforward network, A_0 denotes real output, H denotes the data point number, w denotes optimal weight vector, Δw_i denotes the change in weight on the ith input, η denotes learning rate, and i denotes the input, and I denotes the last input of the neural network. The training algorithm stops when the size of the error function is minimal enough.

Figure 7 demonstrates the proposed ANN framework to predict the link margin of CubeSat with consideration of atmospheric and dust impairments. Such an ANN framework seems massive parallel architectures, yet it operates with interconnected mechanisms efficiently to solve difficult problems. A Levenberg–Marquardt backpropagation algorithm is used in the feedback forward fitting tool to evaluate the performance of the ANN predictions, where the Levenberg–Marquardt algorithm uses the Hessian matrix approximation, as per equation (15), where it is considered faster and more accurate near an error minimum. Hence, the scalar μ decreases after each drop-in performance function, which indicates the performance function is always reduced at each iteration of the algorithm [52–56]:

$$xk + 1 = xk - [JTJ + \mu J] - 1JT_e, \tag{15}$$

where J denotes a Jacobian matrix, which consists of the first values of the network errors considering the assigned weights and biases and e denotes a vector of network errors. Input vectors and target vectors are arbitrarily divided into three sets:

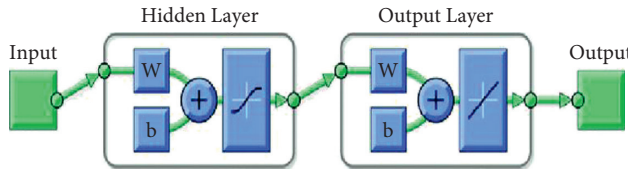


Figure 6: ANN's three universal illustration layers.

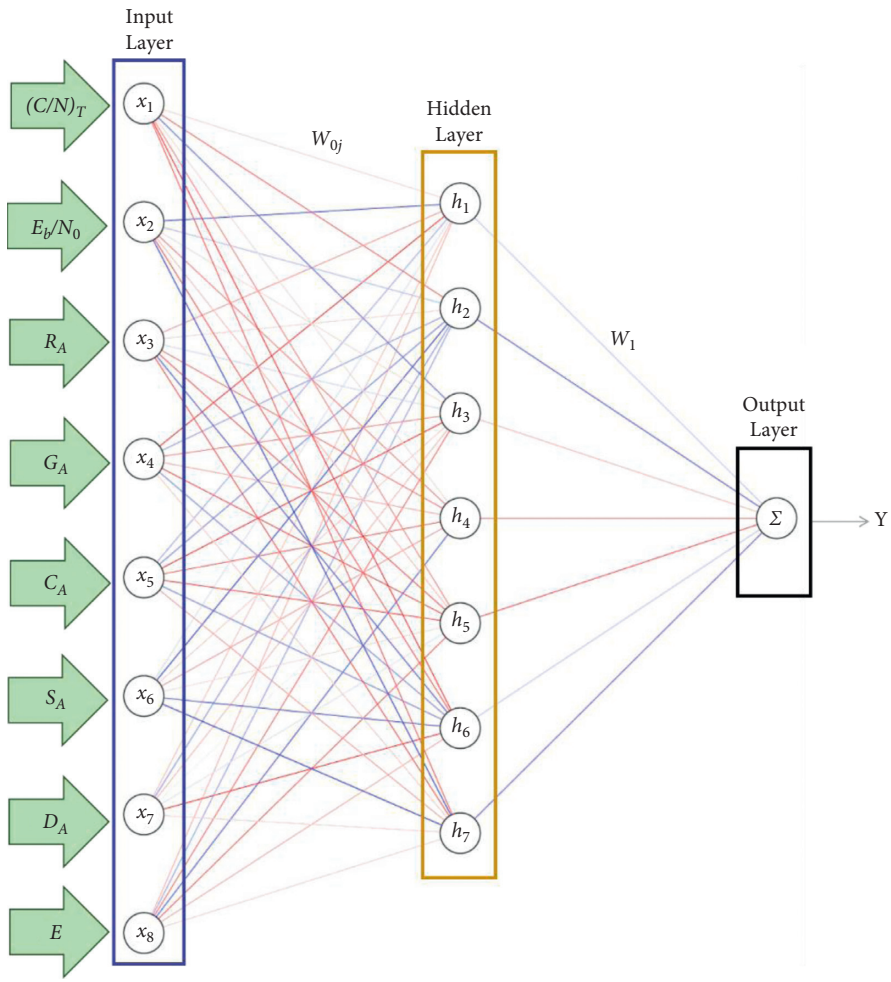


Figure 7: The proposed ANN framework.

(1) Training makes offerings to the ANN while training, and the network is tuned in response to its error, consequently calculating the gradient and updating the weights and biases.

(2) Validation measures network generalization and stops training when generalization stops improving.

(3) Testing delivers an autonomous performance measure during and after training, hence with no effect on training.

3. Performance Analysis and Discussions

Predictions of joint effects of atmospheric and dust attenuations for CubeSat communications associated with a massive MIMO antenna in GCC countries are presented in this section, where the predictions take into consideration the proposed ANN framework, using MATLAB. K-frequency bands include Ku-Band 13–14 GHz, K-Band 18 GHz, and Ka-Band 26.6 GHz, at CubeSat altitude of 400 km above the ground. These three types of variables in the framework of the proposed prediction can be classified into three groups as follows:

(i) Fixed parameters: these parameters are fixed input to the ANN, which are P_t, G_C, G_S, d, λ, E_r, h_t, F, K, T_0, B, Bw, and R_b

(ii) Variable parameters: these parameters are variable input to the ANN, which are θ, f, longitude and latitude of the specific area in the GCC countries

(iii) Predicted parameters: these parameters are expected output from the ANN, which are $(C/N)_T$, E_b/N_0, R_A, G_A, C_A, S_A, D_A, and E

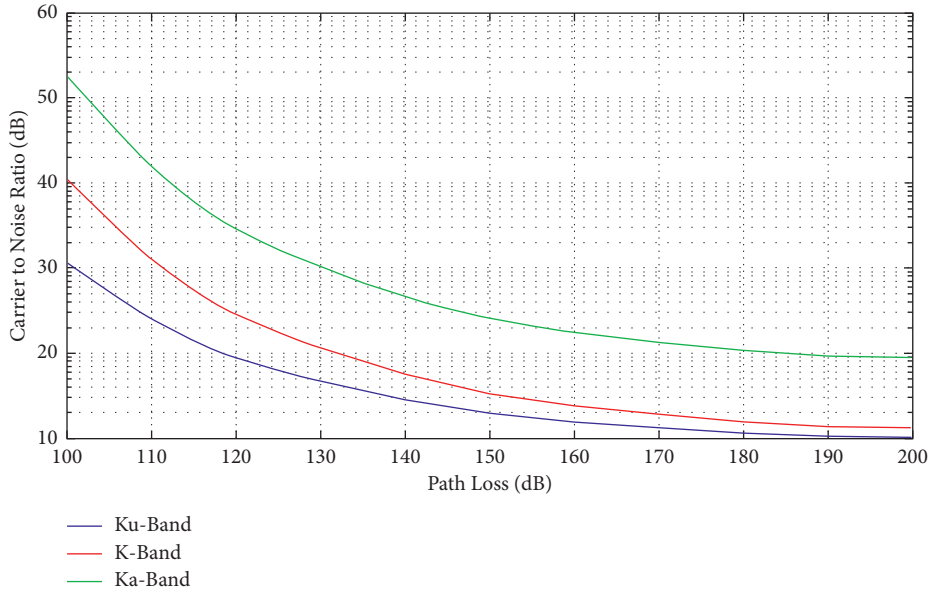


Figure 8: C/N as functions of path loss at different K-frequency bands.

Figures 8 and 9 visualize the link margin of the proposed CubeSat framework. Figure 8 shows C/N as functions of path loss at different K-frequency bands. When the frequency band increases, path loss increases too. Figure 9 reveals the BER of the signal as a function of E_b/N_o at different K-frequency bands. It is observed that an increase of the frequency offset increases the BER value at a fixed E_b/N_o ratio. The best link performance is the one that allows for the lowest possible BER with the lowest possible E_b/N_o. At the lowest BER achieved of $1x10^{-6}$, the E_b/N_o with range floats between 13.5 dB and 16.5 dB. It is observed that, as PL/frequency band increases, both BER and E_b/N_o increase, and system performance drops gradually. Overall, these two performance quality indicators of the proposed CubeSat communication channel framework seem reasonable.

Figures 10–13 show the performance of the simulated MIMO antenna at K-frequency Bands. Figure 10 displays the specification and geometry of the 8×8 patch MIMO antenna configuration. Figure 11 presents an antenna gain where all patches are in-phase and beamforming has been steered at various angles (°) towards the desired direction. Figure 12 displays the return loss result S11 of the proposed feed geometry with respect to port. S11, which is also known as "reflection coefficient" represents how much power is reflected from the antenna. The S parameter measures the resonant frequency of the antenna in relation to the resonance of the TM12 mode. Herein, it shows an accepted radiation power with low loss at 14, 18, and 26 GHz, respectively. Figure 13 demonstrates the Effective Isotropic Radiated Power (EIRP) of the proposed MIMO antenna to measure the coverage of the arrays, where it plots a Cumulative Distribution Function (CDF). Figure 13 shows that coverage just above 56% of the sphere has a positive gain at 23 dBmW input power, representing an acceptable performance level.

Figures 14 and 15 show attenuations of rain, gasses, clouds, scintillation, and dust across wide range of frequency bands and elevation angles. Also, total attenuation, which includes rain, gasses, clouds, scintillation, and dust, of the CubeSat communication channel framework has been calculated. The average of the total attenuation at different frequency bands floats between 5 and 7 dB. Clearly, dust attenuation registers the highest attenuation value compared to other attenuations. This is due to the physical and chemical characteristics of dust and sand particles (e.g., permittivity, size, concentration) that have an extra amount of radio wave energy dissipated by means of scattering and absorption. Thus, dust is the most vital cause of wireless satellite channel impairments, often observed in arid and semiarid areas such as GCC countries.

Another observation is that the attenuations increase as frequency increases along with a decrease in elevation angle. Unsurprisingly, high frequencies are more vulnerable to signal degradation, as radio signals get absorbed by atmospheric rain, snow, ice, or even dust, whereas the reason why attenuations increase positively with elevation angle is that distance increases accordingly at lower elevation angles. Path loss increases due to distance increases, whereas at high elevation angles, more LoS connectivity is achieved but less coverage footprint. Thus it is a trade-off.

Figures 16–19 depict performance results of the proposed ANN prediction framework. Batch training has been considered to train input sets via a learning algorithm in one epoch, which relates to the maximum iterations before weights get updated. Then, the process regulates the optimal number of iterations, during which validation produces a minimal value that can be visualized in regression plotting. Figure 16 shows the process of training the proposed ANN

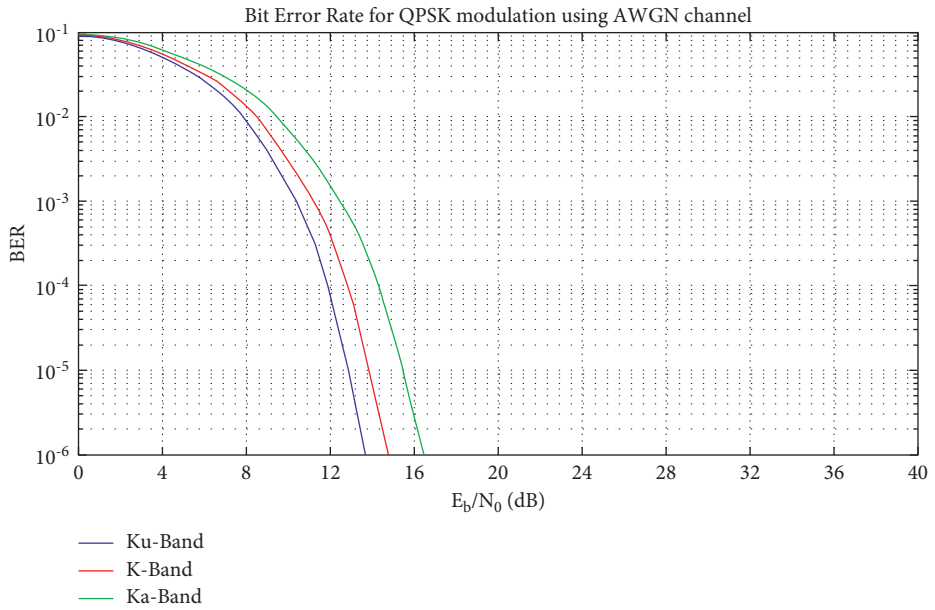


Figure 9: BER of the signal as a function of E_B/N_O at different K-frequency bands.

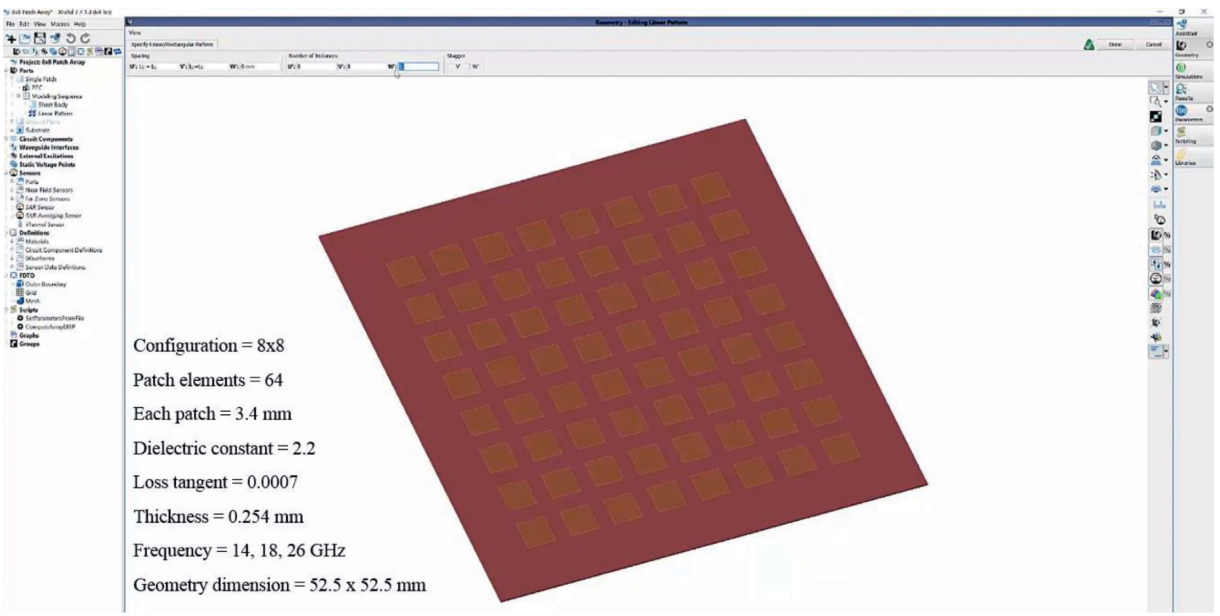


Figure 10: The antenna geometry and specification.

continues for 6 more iterations, during which error rates do not drop lower. However, in the 7^{th} iteration, training stops as the error rate increases. The result is reasonable due to three reasons: first, the concluding mean square error (MSE) is small; second, no substantial overfitting has occurred by iteration 7, before which the best validation performance occurs; and third, the test set error and the validation set error have similar characteristics.

Figure 17 shows a 3-layer training performance of the proposed ANN prediction framework using Stochastic gradient descent (SGD), which refers to the value of the backpropagation gradient on each iteration in a logarithmic scale. So, at each iteration, the weights and biases get updated. Observing both values of gradient coefficient and Mu at their lowest minimum results indicate better training and testing of the ANN. However, validation checks represent the number of successive epochs in which the validation performance fails to decrease. Training stops when the validation parameter reaches the supreme number of validations of 6 at epoch 13 with the lowest gradient coefficient and Mu values. Overall performance shows that precise predictions may be obtained, where no

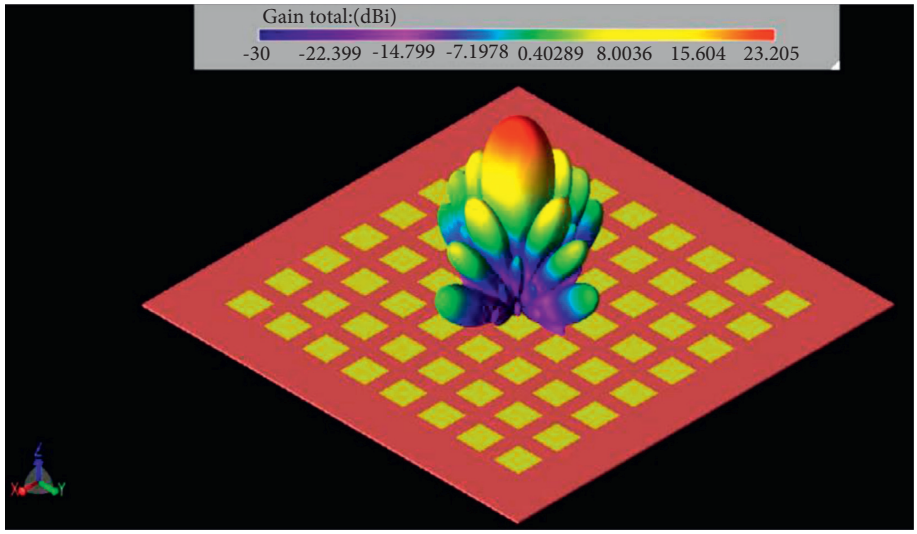


Figure 11: Antenna gain at various angles.

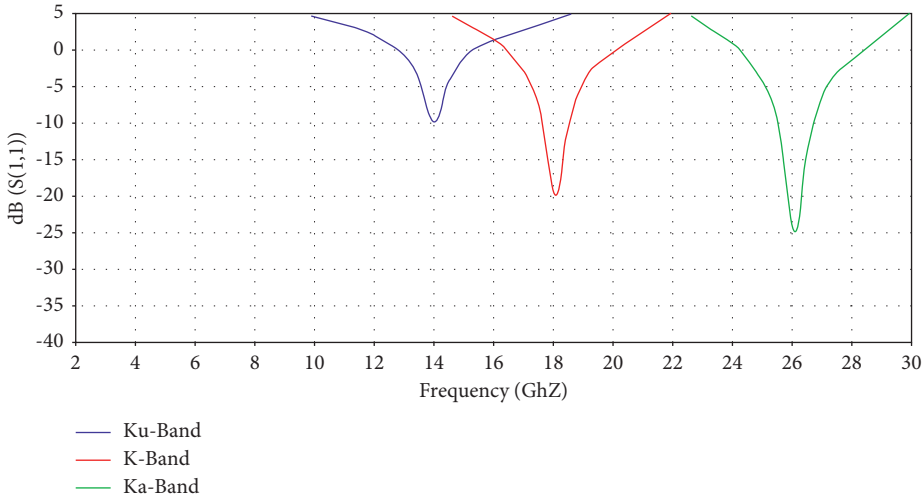


Figure 12: 5G MIMO antenna S11 results at K-frequency bands.

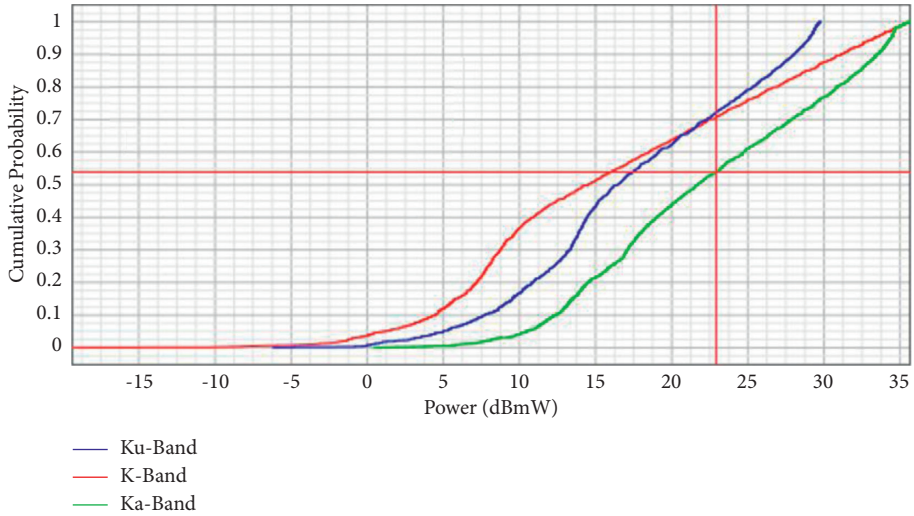


Figure 13: 5G MIMO antenna CDF of EIRP.

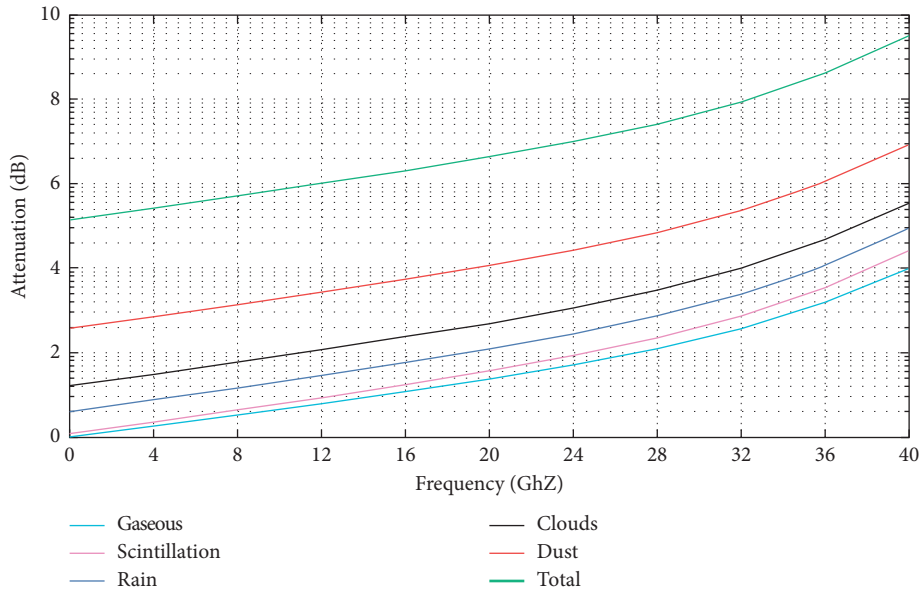


Figure 14: Different attenuations across a wide range of frequency bands.

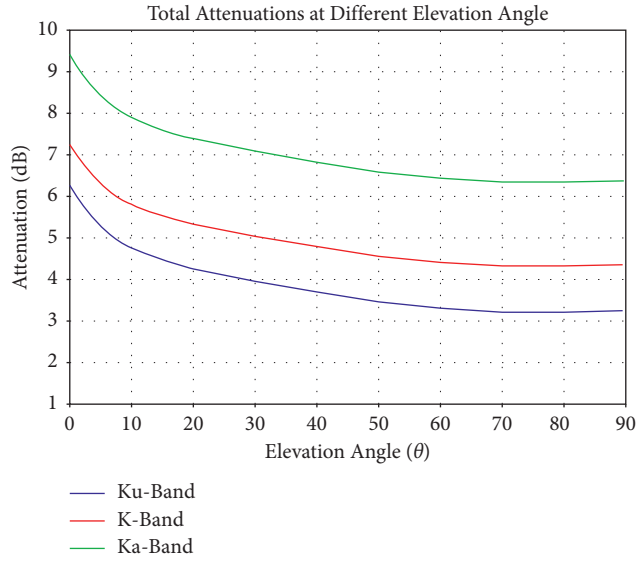


Figure 15: Total attenuations across K-frequency bands.

overfitting occurs before its best validation performance occurs.

Figure 18 demonstrates that regression plots target against training, validation, and test sets. Data must fall along a 45-degree line, where the outputs are equal to the targets to achieve a perfect fit. Targets mean the difference between the perfect result and the outputs. The dashed lines refer to the targets, while the solid lines refer to the best fit between targets and outputs. The R value denotes the relation among the outputs and targets, where R is equal or close to 1. Then, there is a linear relation. Overall performance shows that R values in the regression plots are satisfactory and represent the best fitness levels. Figure 19 displays the E performance of attenuations prediction against the attenuations ANN prediction framework, where the proposed ANN approach is aimed at offering robust channel model planning, with full-range consideration of different attenuation impairments. The error performance of the proposed ANN approach shows lower value in comparison to nonoptimized predictions that describe a robust channel, where you can achieve low error rate without requiring a lot of transmission power.

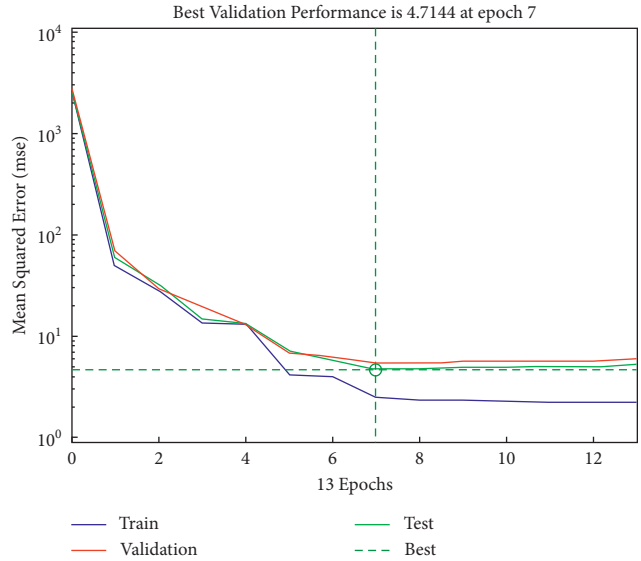


Figure 16: Training performance of the proposed ANN prediction framework.

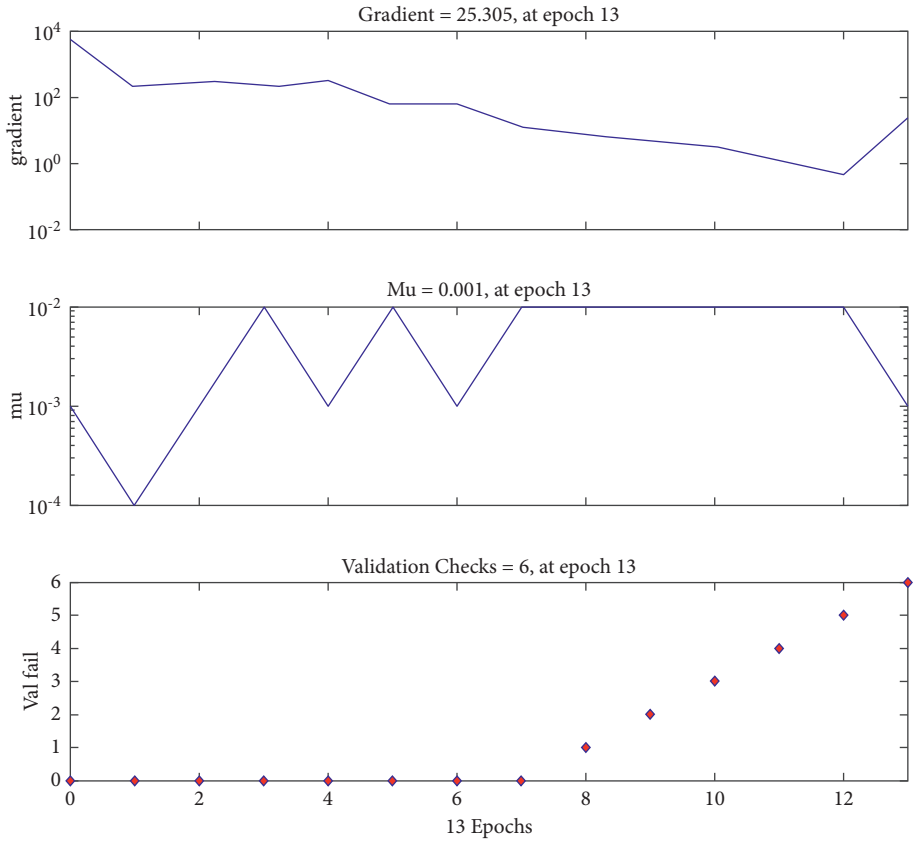


Figure 17: Training data of proposed ANN prediction framework using GSD.

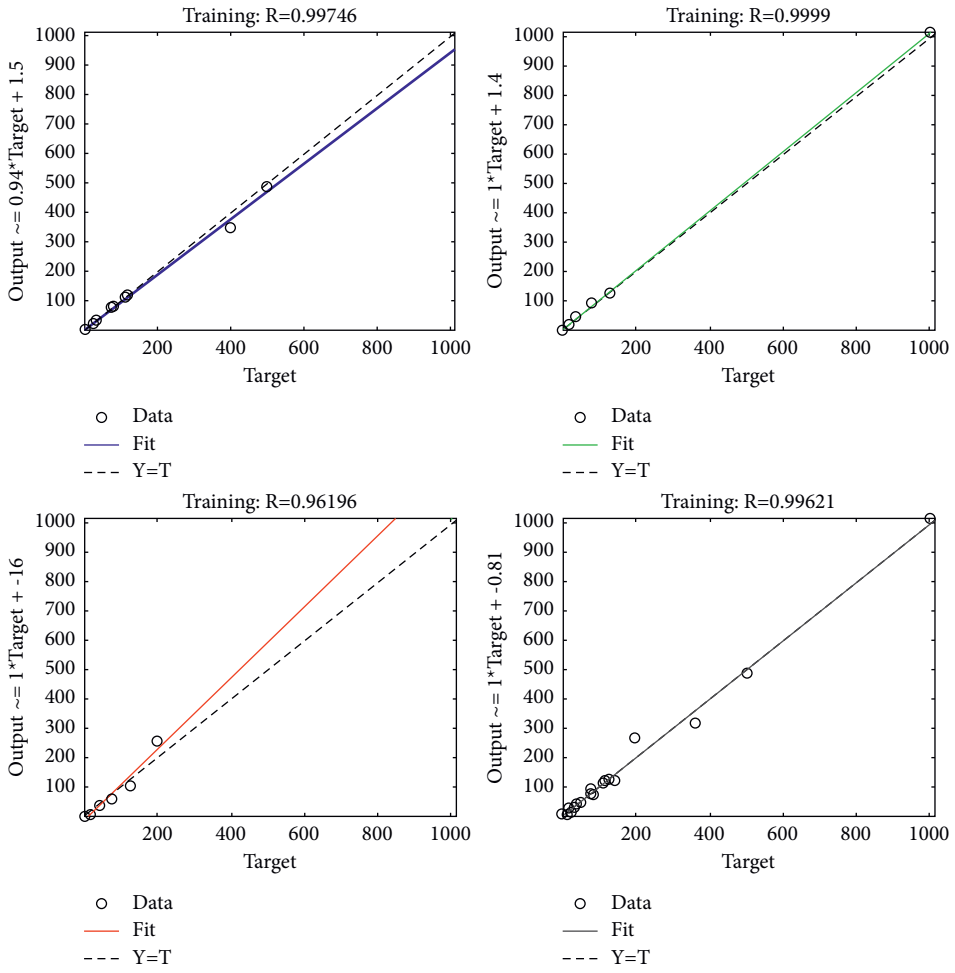


Figure 18: Regression plots of the ANN prediction framework.

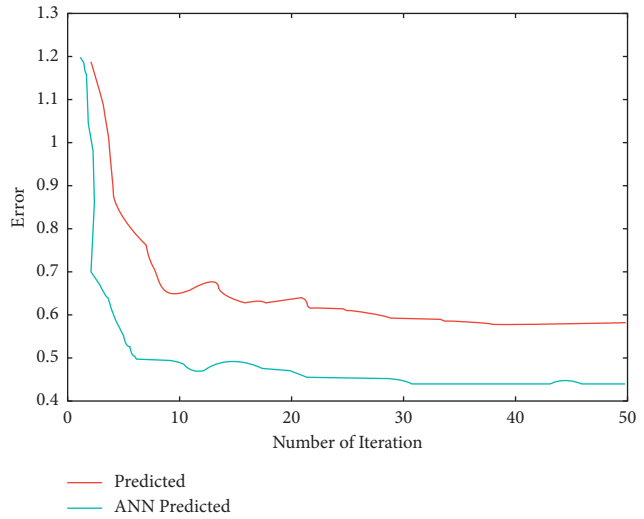


Figure 19: E performance of attenuations prediction against the attenuations ANN prediction framework.

4. Conclusion

Providing seamless communication services using space-based systems, such as micro-/nanosatellites (e.g., CubeSat), would massively pave the way to hyperconnected societies and thus more global smart connectivity. Still, atmospheric and dust impairments are open challenges that face satellite systems, especially at high-frequency bands in arid and

semiarid regions. Motivated by this, this paper aims to provide a comprehensive prediction model that encompasses joint effects of atmospheric and dust impairments from CubeSat communication aspect in GCC) region using ANN framework for CubeSat communication. The prediction model also considers a MIMO antenna to obtain better QoS and capacity. The predicted results confirm that the comprehensive prediction model for CubeSat using ANN offers much promise channel model for satellite communications, which in turn leads to not only supporting IoT connectivity but also reducing power consumption, thus enhancing lifetime of CubeSat. Further, the overall ANN model outperforms the nonoptimized model. As future work, generating predictions with fine-tune environmental considerations in any location using a model that integrates Google Maps is gaining momentum. Further, this work has considered OFDM technology with the proposed massive MIMO antenna. Nevertheless, Filter Bank Multicarrier (FBMC) can be investigated further as proposed in [57–59] but from spec-based systems such as CubeSat perspectives because of its efficiency of high date rate and low latency.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was funded by the Deanship of Scientific Research at Taif University, Kingdom of Saudi Arabia, Taif University Researchers Supporting Project (no. TURSP-2020/265).

References

[1] N. Saeed, A. Elzanaty, H. Almorad, H. Dahrouj, T. Y. Al-Naffouri, and M.-S. Alouini, "CubeSat communications: recent advances and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1839–1862, 2020, thirdquarter 2020.

[2] C. Nieto-Peroy and M. R. Emami, "CubeSat mission: from design to operation," *Applied Sciences*, vol. 9, no. 15, p. 3110, 2019.

[3] F. A. Almalki, "Utilizing drone for food quality and safety detection using wireless sensors," in *Proceedings of the 2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP)*, pp. 405–412, Shanghai, China, 2020.

[4] F. Almalki, "Developing an adaptive channel modelling using a genetic algorithm technique to enhance aerial vehicle-to-everything wireless communications," *International Journal of Computer Networks & Communications*, vol. 13, no. 2, pp. 37–56, 2021, Available:.

[5] J. W. Cheong, B. J. Southwell, W. Andrew et al., "A robust framework for low-cost CubeSat scientific missions," *Space Science Reviews*, vol. 216, no. 1, 2020.

[6] X. Fang, W. Feng, T. Wei, Y. Chen, N. Ge, and C.-X. Wang, "5G embraces satellites for 6G ubiquitous IoT: basic models for integrated satellite terrestrial networks," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14399–14417, 2021.

[7] S. H. Alsamhi, F. A. Almalki, H. Al-Dois et al., "Machine learning for smart environments in B5G networks: connectivity and QoS," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6805151, 23 pages, 2021.

[8] J. Junior, A. Ambrosio, and F. Sousa, "Real-time cubesat thermal simulation using artificial neural networks," *Journal of Computational Interdisciplinary Sciences*, vol. 8, no. 2, 2017.

[9] F. A. Almalki, A. A. Alotaibi, and M. C. Angelides, "Coupling multifunction drones with AI in the fight against the coronavirus pandemic," *Computing*, vol. 103, no. 11, 2021.

[10] G. Villarrubia, J. F. De Paz, P. Chamoso, and F. D. la Prieta, "Artificial neural networks used in optimization problems," *Neurocomputing*, vol. 272, pp. 10–16, 2018.

[11] J. E. Rayas-Sanchez, "EM-based optimization of microwave circuits using artificial neural networks: the state-of-the-art," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 1, pp. 420–435, 2004.

[12] S. H. Alsamhi, O. Ma, M. S. Ansari, and F. A. Almalki, "Survey on collaborative smart drones and internet of Things for improving smartness of smart cities," *IEEE Access*, vol. 7, pp. 128125–128152, 2019.

[13] F. A. Almalki, B. O. Soufiene, S. H. Alsamhi, and H. Sakli, "A low-cost platform for environmental smart farming monitoring system based on IoT and UAVs," *Sustainability*, vol. 13, no. 11, p. 5908, 2021.

[14] F. A. Almalki, S. H. Alsamhi, R. Sahal et al., "Green IoT for eco-friendly and sustainable smart cities: future directions and opportunities," *Mobile Networks and Applications*, vol. 26, no. 4, 2021.

[15] F. A. Almalki, S. Ben Othman, F. A. Almalki, H. Sakli, and EERP-DPM, "Energy efficient routing protocol using dual prediction model for healthcare using IoT," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9988038, 15 pages, 2021.

[16] A. Camps, *"Nanosatellites and Applications to Commercial and Scientific Missions Satellites Missions And Technologies For Geosciences*, pp. 1–25, Intechopen, London, UK, 2019.

[17] F. A. Almalki and M. C. Angelides, "Considering near space platforms to close the coverage gap in wireless communications: the case of the Kingdom of Saudi Arabia," in *Proceedings of the 2016 Future Technologies Conference (FTC)*, pp. 224–230, San Francisco, USA, 2016.

[18] Canadian Space Agency, "What is a CubeSat," 2017, http://www.asc-csa.gc.ca.

[19] A. Kak and I. F. Akyildiz, "Designing large-scale constellations for the internet of space Things with CubeSats," *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 1749–1768, 2021.

[20] F. Almalki, "Comparative and QoS performance analysis of terrestrial-aerial platforms-satellites systems for temporary events," *International journal of Computer Networks & Communications*, vol. 11, no. 6, pp. 111–133, 2019.

[21] R. Sinha, A. Verma, and S. Choudhary, "General analysis of cube satellite technology: an overview," *International Journal of Computer Application*, vol. 178, no. 21, pp. 19–22, 2019.

[22] F. A. Almalki and M. C. Angelides, "Empirical evolution of a propagation model for low altitude platforms," in *Proceedings of the 2017 Computing Conference*, pp. 1297–1304, London, UK, 2017.

[23] M. M. Nair, A. K. Tyagi, and N. Sreenath, "The future with industry 4.0 at the core of society 5.0: open issues, future

opportunities and challenges," in *Proceedings of the 2021 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–7, Coimbatore, India, 2021.

[24] P. Machuca, J. P. Sánchez, and S. Greenland, "Asteroid flyby opportunities using semi-autonomous CubeSats: mission design and science opportunities," *Planetary and Space Science*, vol. 165, pp. 179–193, 2019.

[25] I. F. Akyildiz and A. Kak, "The internet of space things/CubeSats," *IEEE Network*, vol. 33, no. 5, pp. 212–218, 2019.

[26] Spaceworks Enterprises Inc, "Nano/microsatellite market forecast," 2019, https://www.spaceworks.aero/wp-content/uploads/Nano-Microsatellite-Market-Forecast-9th-Edition-2019.pdf.

[27] K. Woellert, P. Ehrenfreund, A. J. Ricco, and H. Hertzfeld, "CubeSats: cost-effective science and technology platforms for emerging and developing nations," *Advances in Space Research*, vol. 47, no. 4, pp. 663–684, 2011.

[28] A. R. Aslan, H. M. K. AlNaimiy, I. Fernini et al., "Space technology capacity building in support of SDG 2030 through CubeSat SharjahSat-l," in *Proceedings of the 2019 9th International Conference on Recent Advances in Space Technologies (RAST)*, pp. 955–958, Istanbul, Turkiye, 2019.

[29] I. F. Akyildiz, J. M. Jornet, and S. Nie, "A new CubeSat design with reconfigurable multi-band radios for dynamic spectrum satellite communication networks," *Ad Hoc Networks*, vol. 86, pp. 166–178, 2019.

[30] Z. Qu, G. Zhang, H. Cao, and J. Xie, "Stochastic dynamic modeling of rain attenuation: a survey," *China Communications*, vol. 15, no. 3, pp. 220–235, 2018.

[31] L. Luini and C. Capsoni, "Joint effects of clouds and rain on ka-band earth observation data downlink systems," in *Proceedings of the 9th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, Lisbon, Portugal, 2015.

[32] L. J. Mpoporo and P. A. Owolawi, "Comparison of measured and estimated rain attenuation on earth-satellite links in South Africa," in *Proceedings of the 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*, pp. 1–5, Kimberley, South Africa, 2019.

[33] J. Wang, L. Wang, and M. Xu, "Rain attenuation analysis of Ka band ship-borne satellite communication station in Indian Ocean and pacific ocean," in *Proceedings of the 2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP)*, pp. 385–388, Shanghai, China, 2020.

[34] M. R. Averly and J. Suryana, "CubeSat communication system for maritime needs," in *Proceedings of the 2020 27th International Conference on Telecommunications (ICT)*, pp. 1–5, Shanghai, China, 2020.

[35] E. Regonesi, L. Luini, and C. Riva, "Limitations of the ITU-R P.838-3 model for rain specific attenuation," in *Proceedings of the 13th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–4, Krakow, Poland, 2019.

[36] A. G. Cappiello, D. C. Popescu, J. S. Harris, and O. Popescu, "Radio link design for CubeSat-to-ground station communications using an experimental license," in *Proceedings of the 2019 International Symposium on Signals, Circuits and Systems (ISSCS)*, pp. 1–4, Iasi, Romania, 2019.

[37] G. Charbit, D. Lin, K. Medles, L. Li, and I.-K. Fu, "Space-terrestrial radio network integration for IoT," in *Proceedings of the 2020 2nd 6G Wireless Summit (6G SUMMIT)*, pp. 1–5, Levi, Finland, 2020.

[38] K. Harb, O. Butt, A. A. Al-Yami, and S. Abdul-Jauwad, "Probabilistic dust storm layers impacting satellite communications," in *Proceedings of the 2013 IEEE International Conference on Space Science and Communication (IconSpace)*, pp. 407–411, Melaka, Malaysia, 2013.

[39] K. Harb, B. Omair, S. Abdul-Jauwad, A. Al-Yami, and A. Al-Yami, *"A Proposed Method for Dust and Sand Storms Effect on Satellite Communication Networks*, Innovations On Communication Theory, Istanbul, Turkey, 2012.

[40] F. A. Almalki and M. C. Angelides, "Propagation modelling and performance assessment of aerial platforms deployed during emergencies," in *Proceedings of the 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*, pp. 238–243, Cambridge, UK, 2017.

[41] D. Barbaric, J. Vukovic, and D. Babic, "Link budget analysis for a proposed CubeSat Earth observation mission," in *Proceedings of the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 0133–0138, Cambridge, UK, 2018.

[42] S. Chen, J. Zhang, Y. Jin, and B. Ai, "Wireless powered IoE for 6G: massive access meets scalable cell-free massive MIMO," *China Communications*, vol. 17, no. 12, pp. 92–109, 2020.

[43] F. A. Almalki and M. C. Angelides, "An enhanced design of a 5G MIMO antenna for fixed wireless aerial access," *Cluster Computing*, vol. 24, no. 3, 2021.

[44] S. Abulgasem, F. Tubbal, R. Raad, P. I. Theoharis, S. Lu, and S. Iranmanesh, "Antenna designs for CubeSats: a review," *IEEE Access*, vol. 9, pp. 45289–45324, 2021.

[45] D. Mansukhani, "Satellite communication: signal attenuation by rain and clouds," *International Journal of Electrical, Electronics And Data Communication*, vol. 5, no. 11, pp. 36–38, 2017.

[46] N. Aida M.Nor, M. R. Islam, W. Al-Khateeb, and S. A.Z, "Atmospheric effects on free space earth-to-satellite optical link in tropical climate," *International Journal of Computer Science, Engineering and Applications*, vol. 3, no. 1, pp. 17–36, 2013.

[47] A. Mansourkhaki, M. Berangi, and M. Haghiri, "Comparative application of radial basis function and multilayer perceptron neural networks to predict traffic noise pollution in tehran roads," *Journal of Ecological Engineering*, vol. 19, no. 1, pp. 113–121, 2018.

[48] F. A. Almalki and M. C. Angelides, "Deployment of an aerial platform system for rapid restoration of communications links after a disaster: a machine learning approach," *Computing*, vol. 102, no. 4, pp. 829–864, 2019.

[49] W. Jiang and H. D. Schotten, "Neural network-based fading channel prediction: a comprehensive overview," *IEEE Access*, vol. 7, pp. 118112–118124, 2019.

[50] Y. Yanqin Zuo, F. Chengrui Jia, and C. Jia, "Prediction for satellite communication based on data mining," in *Proceedings of the 15th International Conference on Optical Communications and Networks (ICOCN)*, pp. 1–3, 2016.

[51] F. A. Almalki and Ben Othman Soufiene, "EPPDA: an efficient and privacy-preserving data aggregation scheme with authentication and authorization for IoT-based healthcare applications," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5594159, 18 pages, 2021.

[52] A. O. Ayo, P. A. Owolawi, J. S. Ojo, and L. J. Mpoporo, "Rain impairment model for satellite communication link design in South Africa using neural network," in *Proceedings of the 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*, pp. 1–8, Kimberley, South Africa, 2020.

[53] F. A. Almalki and M. C. Angelides, "A machine learning approach to evolving an optimal propagation model for last

mile connectivity using low altitude platforms," *Computer Communications*, vol. 143, pp. 9–33, 2019.

[54] P. Thiennviboon and S. Wisutimateekorn, "Rain attenuation prediction modeling for earth-space links using artificial neural networks," in *Proceedings of the 2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pp. 29–32, 2019.

[55] S. H. Alsamhi, F. Almalki, O. Ma, M. S. Ansari, and B. Lee, "Predictive estimation of optimal signal strength from drones over IoT frameworks in smart cities," *IEEE Transactions on Mobile Computing*, vol. 2021, Article ID 6805151, 1 page, 2021.

[56] F. Fourati and M.-S. Alouini, "Artificial intelligence for satellite communication: a review," 2021, https://arxiv.org/abs/2101.10899.

[57] D. Kong, X. Zheng, Y. Zhang, and T. Jiang, "Frame repetition: a solution to imaginary interference cancellation in FBMC/OQAM systems," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1259–1273, 2020.

[58] H. Wang, X. Li, R. H. Jhaveri et al., "Sparse Bayesian learning based channel estimation in FBMC/OQAM industrial IoT networks," *Computer Communications*, vol. 176, pp. 40–45, 2021.

[59] D. Kong, J. Li, K. Luo, and T. Jiang, "Reducing pilot overhead: channel estimation with symbol repetition in MIMO-FBMC systems," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7634–7646, 2020.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 1362640, 8 pages
https://doi.org/10.1155/2021/1362640

Research Article

A Novel Sparse Array Configuration with Low Coarray Redundancy for DOA Estimation in Mobile Wireless Sensor Network

Pin-Jiao Zhao,[1] **Guo-Bing Hu**,[1] **and Liang-Tian Wan**[2]

[1]*Department of Electronic and Information Engineering, Jinling Institute of Technology, Nanjing 211169, China*
[2]*Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, School of Software, Dalian University of Technology, Dalian 116620, China*

Correspondence should be addressed to Pin-Jiao Zhao; zhaopinjiao@hrbeu.edu.cn

Received 24 September 2021; Accepted 25 October 2021; Published 22 November 2021

Academic Editor: Lingwei Xu

For tacking and localizing sources in the mobile wireless sensor network, underdetermined direction of arrival (DOA) estimation with high-accuracy is a crucial issue. In this paper, a novel sparse array configuration is developed for accurate DOA estimation from the perspective of sum-difference coarray (SDCA). As compared with most of the existing sparse array configurations, the proposed array can effectively reduce the overlap between difference coarray (DCA) and sum coarray (SCA) and can achieve more consecutive degrees of freedom (DOF), more sources can be resolved accordingly. Additionally, the proposed array has hole-free DCA and SDCA. Then, the concept of coarray redundancy ratio (CRR) is introduced for evaluating the coarray overlap quantitatively and the closed-form CRR expressions of the proposed array are derived in detail. Based on the good properties of the proposed array, vectorized conjugate augmented MUSIC (VCAM) is adopted for underdetermined DOA estimation. The theoretical propositions and numerical simulations demonstrate the superior performance of the proposed array in terms of CRR, consecutive DOF, and DOA estimation accuracy.

1. Introduction

In the large-scale mobile wireless communication network [1, 2], underdetermined direction of arrival (DOA) estimation [3–6] with high-accuracy is a key issue to solve for tacking and localizing sources. The optimized array configuration design is a prerequisite for accurate DOA estimation. Recently, the emerging sparse arrays break through the constraint of spatial sampling theorem by sparsely arranging the array elements and have remarkable advantages in array element layout flexibility, degrees of freedom (DOF), and virtual array aperture [7, 8]. Benefiting from the sparse arrays, more sources than the number of the physical sensors can be resolved with high-accuracy.

From the view of array configuration design, coprime array (CPA) [9] and nested array (NA) [10] are the two most typical sparse arrays. In comparison to the minimum redundancy array and the minimum hole array, CPA and NA are easy to construct since their sensor locations are analytically tractable. Specifically, CPA is obtained from a pair of coprime uniform linear arrays (ULA), which can offer $O(MN)$ DOF with $O(M+N)$ sensors, but has holes in the virtual coarray. By contrast, NA contains two "nesting" ULAs with increasing element spacing, which has a consecutive virtual coarray with $O(N^2)$ DOF from $O(N)$ sensors. On the basis of these prototypes, several modifications are developed, such as generalized CPA [11], thinned CPA [12], super NA [13], and augmented NA [14], aiming at increasing consecutive DOF and expanding effective array aperture. Nevertheless, note that these arrays construct synthetic virtual arrays from the view of difference coarray (DCA), the number of consecutive DOF in the virtual array cannot be more than twice of the physical aperture. Accordingly, the number of the resolvable sources and the DOA estimation accuracy are also affected.

Alternatively, several array configurations based on sum-difference coarray (SDCA) have attracted considerable interests, wherein the introduction of sum coarray (SCA) originating from multiple-input multiple-output (MIMO) system [15, 16] can bring more DOF. In [17], a novel CPA based on SDCA is presented, from which the consecutive DOF more than twice of array aperture can be obtained. Following this, a modified nested array configuration named as sum-diff NA (SdNA) is proposed in [18]. To further improve the available DOF, two kinds of improved NAs [19], termed as INAwSDCA-I and INAwSDCA-II, are developed by rearranging the translational NA. In [20], the unfold CPA configuration with SDCA is proposed. However, for the above array configurations based on SDCA, lots of overlapping virtual sensors between their DCA and SCA lead to unsatisfactory coarray utilization. In addition, there also exist holes in the SDCA. Toward this end, transformed NA (TNA) is designed in [21], which can reduce the overlap between DCA and SCA to some extent, but the degree of overlap under different number of array elements is only evaluated via simulations, lacking quantificational analysis.

To tackle these problems, a novel sparse array configuration with low coarray overlap is proposed in this paper. The proposed array is constructed with two concatenated subarrays with different interelement spacings. From the perspective of virtual array, both the DCA and SDCA of the proposed array are hole-free and can provide more consecutive DOF accordingly. Moreover, the coarray redundancy ratio (CRR) of the proposed array is derived in detail to evaluate the coarray overlap quantitatively. Based on the proposed array, the vectorized conjugate augmented MUSIC (VCAM) approach can be employed for DOA estimation. Now, we briefly summarize the contributions of this work as follows:

(1) A novel array configuration is proposed to reduce the overlap between DCA and SCA

(2) The proposed array has hole-free DCA and SDCA, which has been proved via theoretical propositions

(3) More sources than twice of the number of sensors can be resolved

The mathematical notations used throughout this paper are denoted as follows. Vectors and matrices are denoted by lowercase and uppercase bold-face letters, respectively. $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^*$ denote transpose, conjugate transpose, and conjugate. $E[\cdot]$ and $|\cdot|$ represent the statistical expectation and modulus of the internal entity, $\partial f(a)/\partial a$ denotes the derivation of $f(a)$ with respect to a, $\mathbf{Z}$ denotes the non-negative integer set, $\mathbf{Z}^+$ denotes the positive integer set, $\mathbf{Z}(n)$ denotes a set of consecutive positive integers ranging from 1 to n in the case of $n > 0$, and $\complement(n) = \{0\}$ holds in the case of $n = 0$; otherwise, $\complement(n)$ denotes a set of consecutive negative integers ranging from $-n$ to -1. The symbol $\otimes$ denotes the Kronecker product and $\odot$ denotes the Khatri–Rao product.

The remainder of this paper is organized as follows. The DOA estimation model based on SDCA is formulated in Section 2. Section 3 presents the proposed low coarray redundancy array configuration in detail. Section 4 exhibits the simulation results in terms of CRR, consecutive DOF, and DOA estimation accuracy. Conclusions are drawn in Section 5. The proofs of Propositions 1 and 2 are derived in Appendix A and Appendix B.

2. Problem Formulation

Assume that K far-field narrowband sources from directions $\{\theta_k | k = 1, 2, \ldots, K\}$ impinge on an N-element nonuniform sparse antenna array. Then, the array output at time t can be represented as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) = \sum_{k=1}^{K} \mathbf{a}(\theta_k)\mathbf{s}_k(t) + \mathbf{n}(t), \tag{1}$$

where the value of K is known in advance which can be estimated via source number estimation methods [22]. $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_K(t)]^T$ is the source vector with the kth element being $s_k(t) = G_k e^{jw_k t}$, G_k is the deterministic complex amplitude, and w_k is the frequency offset. $\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_K)]$ denotes the array manifold matrix. After normalizing by the unit element spacing d, the steering vector can be expressed as $\mathbf{a}(\theta_k) = e^{-j\mathbf{g}\pi \sin\theta_k}$ with $\mathbf{g} = [g_1, g_2, \ldots, g_N]^T$ being the sensor location vector. $\mathbf{n}(t)$ is the Gaussian white noise vector with zero mean and variance being σ_n^2. By collecting T_s samples, the time average function can be expressed as

$$R_{x_i^* x_j}(\tau) = \frac{1}{T_s}\sum_{t=1}^{T_s} x_i^*(t)x_j(t+\tau) \approx \sum_{k=1}^{K} e^{j\pi(g_j - g_i)\sin\theta_k} R_{s_k^* s_k}(\tau) + R_{n_i^* n_j}(\tau), \tag{2}$$

where $\tau \neq 0$ is the time lag, $x_i(t)$ and $x_j(t)$ are respectively the i^{th} and the j^{th} $(1 \le i, j \le N)$ row of $\mathbf{x}(t)$, $R_{s_i^* s_j}(\tau) = |E_k|^2 e^{jw_k t}$, and $R_{n_i^* n_j}(\tau) = 0$. For the T_p pseudosnapshots, the pseudodata matrix is defined as

$$\mathbf{G} = \left\{ \begin{bmatrix} \mathbf{g}_x(\tau) \\ \mathbf{g}_x^*(-\tau) \end{bmatrix} | \tau = n_p P_s,\ n_p = 1, \ldots, T_P \right\}, \quad (3)$$

where P_s is the pseudosampling period that satisfies the Nyquist sampling theorem. For convenience, let the 1st sensor be the reference, $\mathbf{g}_x(\tau) = [R_{x_1^* x_1}(\tau), R_{x_1^* x_2}(\tau), \ldots, R_{x_1^* x_N}(\tau)]^T$. Then, the covariance matrix of $\mathbf{G}$ is calculated as

$$\mathbf{R}_G = E\left[\mathbf{G}\mathbf{G}^H\right] \approx \frac{1}{T_p} \sum_{n_p=1}^{T_p} \mathbf{g}\left(n_p P_s\right)\mathbf{g}^H\left(n_p P_s\right). \quad (4)$$

Vectorizing $\mathbf{R}_G$ yields

$$\mathbf{r}_G = \mathrm{vec}\left(\mathbf{R}_G\right) = \mathbf{B}\mathbf{z}, \quad (5)$$

where $\mathbf{z} = [|E_1|^4, |E_2|^4, \ldots, |E_K|^4]^T$ and the k^{th} column of $\mathbf{B}$ can be denoted as

$$\tilde{\mathbf{a}}(\theta_k) = \begin{bmatrix} \mathbf{a}(\theta_k) \otimes \mathbf{a}^*(\theta_k) \\ \mathbf{a}(\theta_k) \otimes \mathbf{a}(\theta_k) \\ \mathbf{a}^*(\theta_k) \otimes \mathbf{a}^*(\theta_k) \\ \mathbf{a}^*(\theta_k) \otimes \mathbf{a}(\theta_k) \end{bmatrix}, \quad (6)$$

where the term $\tilde{\mathbf{a}}(\theta_k)$ in (6) behaves like a virtual steering vector of a longer virtual ULA. More specifically, the union of $\mathbf{a}(\theta_k) \otimes \mathbf{a}^*(\theta_k)$ and $\mathbf{a}^*(\theta_k) \otimes \mathbf{a}(\theta_k)$ is respectively corresponding to DCA and its mirror version; $\mathbf{a}(\theta_k) \otimes \mathbf{a}(\theta_k)$ and $\mathbf{a}^*(\theta_k) \otimes \mathbf{a}^*(\theta_k)$ are respectively corresponding to SCA and its mirror version. Then, the spatial smoothing MUSIC or sparse construction approach can be performed for underdetermined DOA estimation, and the above procedure is termed as the vectorized conjugate augmented MUSIC (VCAM) [17].

3. Low Coarray Redundancy Array Configuration

Definition 1. Assume $2 < N_1 \le N_2$, $N_1 \in \mathbf{Z}^+$, and $N_2 \in \mathbf{Z}^+$, after normalizing by d, the configuration of the proposed array with $N = N_1 + N_2$ antenna elements is defined by

$$\begin{cases} \mathbb{L} = \mathbb{L}_1 \cup \mathbb{L}_2, \\ \mathbb{L}_1 = 1 + (l-1)N_1, \quad l \in \mathbb{Z}(N_1), \\ \mathbb{L}_2 = N_1^2 - N_1 + N_2 + l, \quad l \in \mathbb{Z}(N_2). \end{cases} \quad (7)$$

Intuitively, an example of $N_1 = N_2 = 3$ is given in Figure 1, where white circles and black circles denote the sensor locations of subarray $\mathbb{L}_1$ and $\mathbb{L}_2$, respectively.

From the perspective of coarray equivalence, the proposed array, NA, and TNA are compared in Figure 2 with the number of antenna elements being $N = 6$, where the red circles and blue triangles indicate the virtual sensors in DCA and SCA, respectively, and black crosses indicate holes. Note that the DCA, SCA, and SDCA are origin-symmetric, here the nonnegative parts of the abovementioned coarrays are selected for performance comparison, which are respectively abbreviated as n-DCA, n-SCA, and n-SDCA. The sensors in the rectangular dashed box indicate the overlapping virtual sensors between the n-DCA and the consecutive n-SCA. We can see from Figure 2 that the n-SDCA is continuous in the range of (0, 24) for the proposed array, while (0, 16) for NA and (0, 22) for TNA. Moreover, there exists only 1 overlapping virtual sensor for the proposed array, while 10 for NA and 4 for TNA.

Proposition 1. *Some properties of the proposed array configuration in the virtual coarray are listed as follows:*

(i) The DCA of the proposed array is continuous in set of $\mathbb{Z}(-S_1) \cup \mathbb{Z}(0) \cup \mathbb{Z}(S_1)$ *with* $S_1 = N_1^2 - N_1 + 2N_2 - 1$

(ii) The SCA of the proposed array is continuous in set of $\{¢(S_3) - ¢(S_2)\} \cup \{¢(-S_3) - ¢(-S_2)\}$ *with* $S_3 = 2(N_1^2 - N_1 + 2N_2)$, *let* $\varepsilon = N_2 - N_1 \in ¢$, *and the value of* S_2 *is given as follows:*

(a) if $0 \le \varepsilon < 2$, $S_2 = N_1^2 + 2$

(b) if $\varepsilon = 2$ *or* 3, $S_2 = N_1^2 - N_1 + N_2 + 2$

(c) if $\varepsilon \ge 4$, $S_2 = \begin{cases} N_1^2 + N_1 + 2, & N_1 = \varepsilon - 1, \\ N_1^2 - N_1 + N_2 + 2, & \text{else} \end{cases}$

Proof. See Appendix A. □

Proposition 2. *The proposed array has hole-free DCA and SDCA, and the maximum consecutive DOF is* $4(N_1^2 - N_1 + 2N_2) + 1$.

Proof. See Appendix B. Note that for a sparse array configuration with the fixed number of antenna elements, the higher consecutive DOF means larger virtual array aperture and more accurate DOA estimation. Therefore, it is essential to look for appropriate N_1 and N_2 to maximize the consecutive DOF. □

Corollary 1. *After optimizing* N_1 *and* N_2, *the optimal consecutive DOF and the corresponding solutions are given as*

$$\mathrm{DOF}_{\max} = \begin{cases} (N+1)^2, & \text{if } N_1 = N_2 = \dfrac{N}{2}, \\ N^2 + 8, & \text{if } N_1 = \dfrac{N-1}{2}, N_2 = \dfrac{N+1}{2}. \end{cases} \quad (8)$$

Proof. The optimal DOF under the constraint of $N = N_1 + N_2$ can be cast as the following optimization problem:

$$\max_{N_1, N_2} \mathrm{DOF} = 4\left(N_1^2 - N_1 + 2N_2\right) + 1 \quad \text{s.t.} N = N_1 + N_2. \quad (9)$$

Using arithmetic mean-geometric mean (AM-GM) inequalities, the solutions are obtained from (8). According to the above analysis, the proposed array exhibits the advantages in terms of consecutive DOF and virtual array aperture. From the view of coarray efficiency, the overlapping

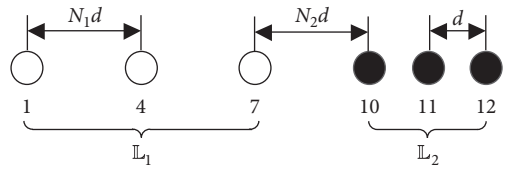


FIGURE 1: An example of the proposed array configuration with $N_1 = N_2 = 3$.

FIGURE 2: Virtual coarrays of the proposed array and NA with $N = 6$, (a) NA, (b) TNA, and (c) the proposed array.

virtual sensors between the DCA and SCA lead to coarray redundancy, which, to some extent, reduce the number of available virtual coarrays. In view of this, the concept of CRR is introduced to evaluate the coarray overlap. □

Definition 2. CRR denotes the ratio of the overlapping virtual sensors between the DCA and SCA, which is given by

$$\kappa = \frac{2(S_1 - S_2 + 1)}{(2S_3 + 1)}. \tag{10}$$

More specifically, CRR of the proposed array configuration can be calculated as

$$\text{if } 0 \le \varepsilon < 2, \quad \kappa = \frac{4N_2 - 2N_1 - 4}{4\left(N_1^2 - N_1 + 2N_2\right) + 1},$$

$$\text{if } \varepsilon = 2 \text{ or } 3, \quad \kappa = \frac{2N_2 - 4}{4\left(N_1^2 - N_1 + 2N_2\right) + 1},$$

$$\text{if } \varepsilon \ge 4, \quad \kappa = \begin{cases} \dfrac{4N_1}{4\left(N_1^2 - N_1 + 2N_2\right) + 1}, & N_1 = \varepsilon - 1, \\ \dfrac{2N_2 - 4}{4\left(N_1^2 - N_1 + 2N_2\right) + 1}, & \text{else.} \end{cases} \tag{11}$$

Proposition 3. *For the optimal DOF given in Corollary 1, if $N_1 = N_2 = (N/2) > 2$, the corresponding CRR is no more than 0.0496; if $N_2 = (N + 1/2) > N_1 = (N - 1/2) > 2$, the corresponding CRR is no more than 0.1053.*

Proof. If $N_1 = N_2 = (N/2) > 2$, the corresponding CRR can be simplified to $\kappa = (4N_2 - 2N_1 - 4/4(N_1^2 - N_1 + 2N_2) + 1) = (N - 4/(N + 1)^2)$. Then, the first-order derivative of κ is taken as $(\partial\kappa/\partial N) = (9 - N/(N + 1)^3)$. Since N is even, the maximum of κ is calculated as $\kappa_{\max} = \max(\kappa_{N=8}, \kappa_{N=10}) = 0.0496$. If $N_2 = N + 1/2$, $N_1 = N - 1/2$, the corresponding CRR can be simplified to $\kappa = (4N_2 - 2N_1 - 4/4(N_1^2 - N_1 + 2N_2) + 1) = (N - 1/N^2 + 8)$. Similarly, the first-order derivative of κ is taken as $\partial\kappa/\partial N = (-N^2 + 2N + 8/(N^2 + 8)^2)$. Note that N is odd; the maximum of κ is calculated as $\kappa_{\max} = \max(\kappa_{N=7}) = 0.1053$. □

4. Numerical Simulations

In this section, numerical simulations have been carried out to evaluate the performance of the proposed array, where SdNA [18] and TNA [21] are selected for performance comparison. In the first simulation, the CRR scatters versus $\varepsilon = N_2 - N_1$ are plotted in Figure 3, and the blue asterisks indicate the case where N is even and the red squares indicate the case where N is odd. It can be observed that the CRR of the proposed array tends to increase with the increase of ε, thus the optimal CRR can be obtained from $\varepsilon = 0$ when N is even and $\varepsilon = 1$ when N is odd.

In the second simulation, the CRR of three array configurations with $\varepsilon = 0$ and $\varepsilon = 1$ are respectively compared in Figures 4 and 5. The results show that the proposed array can achieve the smallest CRR compared to SdNA and TNA with the same number of antenna elements. Also, it has been concluded that the CRR of the proposed array is no more than 0.0496 when N is even and no more than 0.1053 when N is odd, which has been verified in Proposition 3.

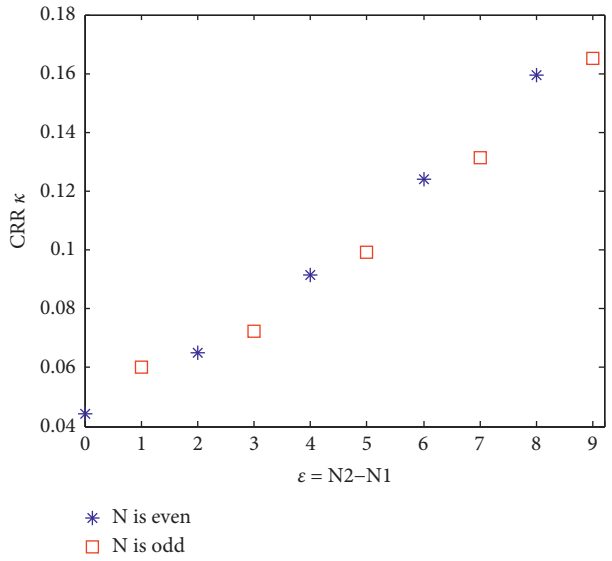


Figure 3: The scatter plot of CRR.

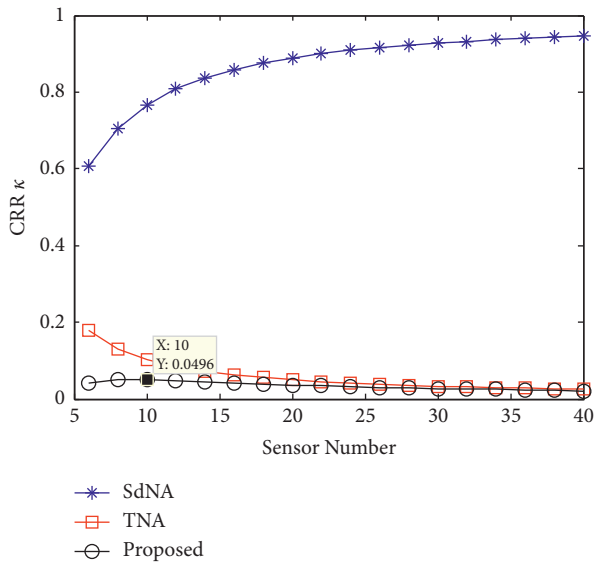


Figure 4: Comparison of CRR with different sensor number ($\varepsilon = 0$).

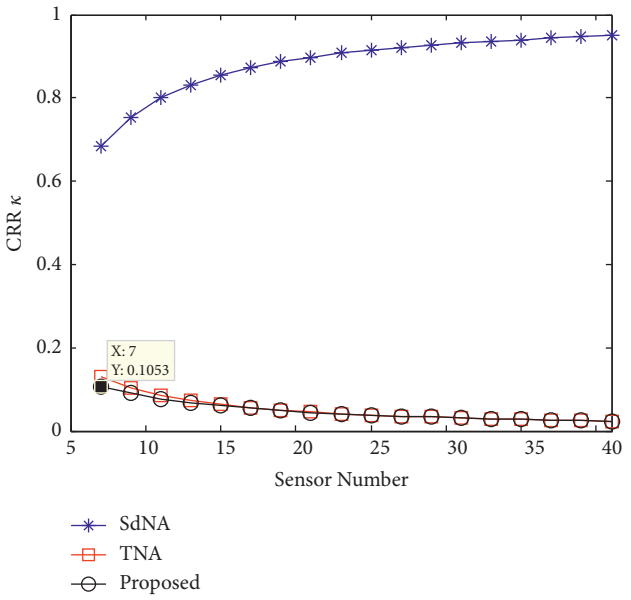


Figure 5: Comparison of CRR with different sensor number ($\varepsilon = 1$).

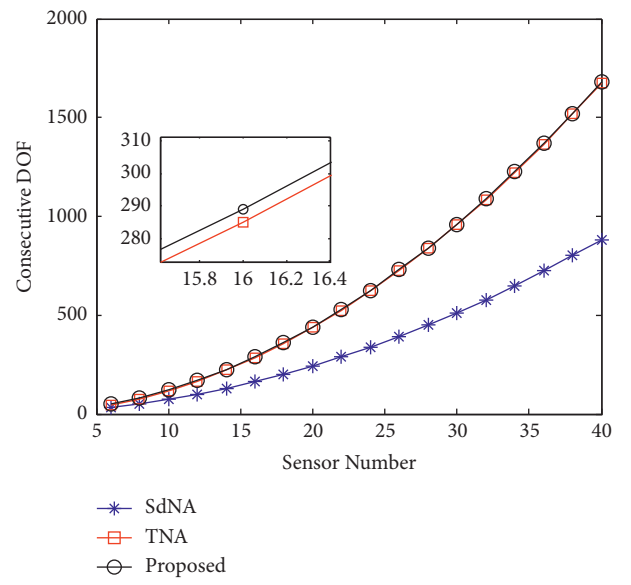


Figure 6: Comparison of consecutive DOF.

In the third simulation, we compare the consecutive DOF of three array configurations and the simulation results are depicted in Figure 6. As can be seen from Figure 6, the proposed array can obtain more consecutive DOF than the other two arrays benefitting from the lower CRR. This implies that there exist few overlapping sensors between the DCA and SCA of the proposed array, thus more effective virtual sensors can be utilized for the construction of SDCA.

Then, the DOA estimation performance is investigated in the last simulation. We consider $K = 27$ narrowband sources uniformly distributed between -60° and 60° impinging on the proposed sparse array with 8 sensors, where SNR being 0 dB and $T_s = T_p = 200$. Figure 7 depicts the MUSIC spectrum of the proposed array where the blue solid lines denote the angle estimates and the red dotted lines denote the incident sources. It can be seen that the proposed array can detect 27 impinging sources with only 8 sensors, and the estimated spectrum has sharp and clearly discernible peaks in the vicinity of the true impinging sources.

Figures 8 and 9 respectively depict the root mean square error (RMSE) of DOA estimation versus SNR and snapshots via 200 Monte Carlo trials, where RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{200K}\sum_{n=1}^{200}\sum_{k=1}^{K}\left(\hat{\theta}_{n,k} - \theta_{n,k}\right)^2}, \tag{12}$$

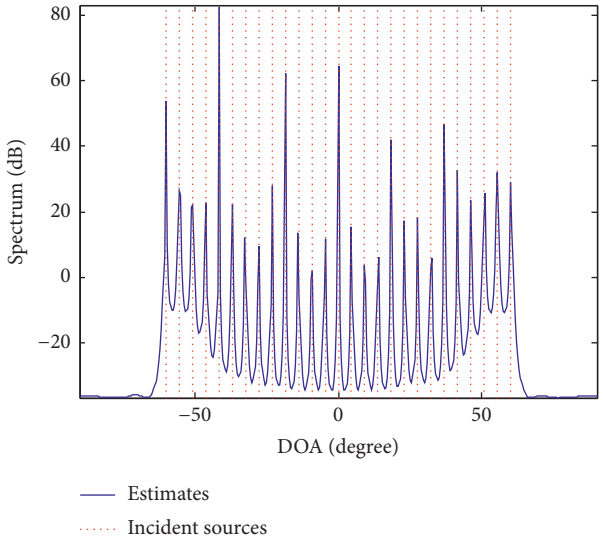


Figure 7: The MUSIC spectrum of the proposed array with $N = 8$.

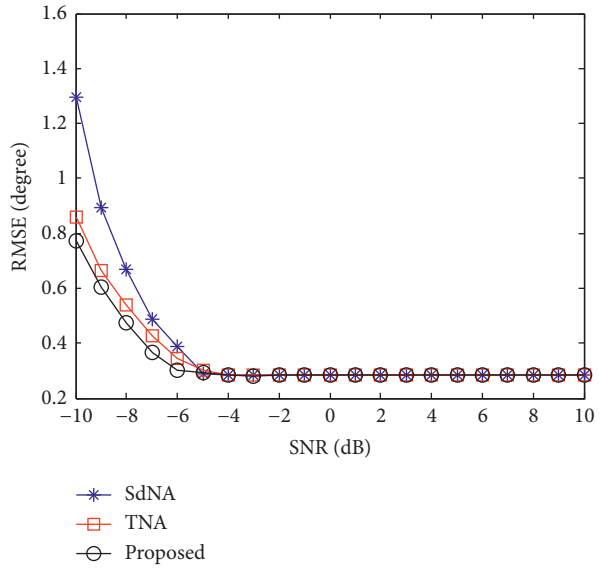


Figure 8: Comparison of DOA estimation performance (RMSE versus SNR).

with $\hat{\theta}_{n,k}$ being the estimate of the kth impinging source for the nth Monte Carlo trial. All the simulation conditions are the same as Figure 7, except that the snapshot number is fixed at 200 in Figure 8 and the SNR is fixed at −6 dB in Figure 9. The results show that the proposed array is superior to the other two array configurations in terms of DOA estimation accuracy, which mainly attributed to the larger consecutive DOF and the lower CRR of the proposed array.

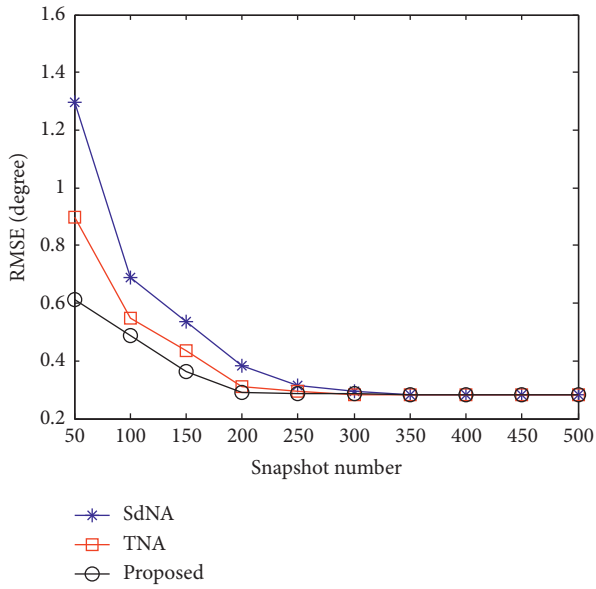


Figure 9: Comparison of DOA estimation performance (RMSE versus snapshots).

5. Conclusions

This paper presents a novel sparse array configuration for accurate DOA estimation, which has fewer coarray overlapping sensors and more consecutive DOF as compared with most of the existing sparse array configurations. From the perspective of coarray equivalence, the proposed array has hole-free DCA and SDCA. The closed-form CRR expressions of the proposed array are derived in order to evaluate the coarray overlap quantitatively. With the utilization of the proposed array, the VCAM approach is adopted for DOA estimation. The results of the theoretical analysis and the numerical simulations demonstrate the effectiveness and favorable performance of the proposed array configuration and DOA estimation performance.

Appendix

A. Proof of Proposition 1

(i) The positive part of the difference set between $\mathbf{L}_1$ and $\mathbf{L}_2$ is $\{\mathbf{L}_2 - \mathbf{L}_1\}$, which distributes continuously in the set of $\{\mathbb{Z}(S_1) - \mathbb{Z}(N_2)\}$ with $S_1 = N_1^2 - N_1 + 2N_2 - 1$. Similarly, the positive part of $\{\mathbf{L}_2 - \mathbf{L}_2\}$ is continuous in the set of $\{\mathbb{Z}(0) \cup \mathbb{Z}(N_2 - 1)\}$. Thus, the whole DCA (containing both the positive part and the negative part) of the proposed array is hole-free, which is continuous in the set of $\{\mathbb{Z}(-S_1) \cup \mathbb{Z}(0) \cup \mathbb{Z}(S_1)\}$.

(ii) The positive part of the sum set $\{\mathbf{L}_1 + \mathbf{L}_2\}$ is continuous in the set of $\{\mathbb{Z}(2N_1^2 - 2N_1 + 2N_2 + 1) - \mathbb{Z}(N_1^2 - N_1 + N_2 + 2)\}$ and $\{\mathbf{L}_2 + \mathbf{L}_2\}$ is also continuous in the set $\{\mathbb{Z}(S_3) - \mathbb{Z}(2N_1^2 - 2N_1 + 2N_2 + 2)\}$ with $S_3 = 2(N_1^2 - N_1 + 2N_2)$. Thus, the positive set of $\{(\mathbf{L}_1 + \mathbf{L}_2) \cup (\mathbf{L}_2 + \mathbf{L}_2)\}$ is continuous in the set $\{\mathbb{Z}(S_3) - \mathbb{Z}(N_1^2 - N_1 + N_2 + 2)\}$. By contrast, the set $\{\mathbf{L}_1 + \mathbf{L}_1\}$ is a discontinuous set composed of discrete points which are directly determined by N_1 and ε. More specifically, if $0 \le \varepsilon < 2$, the largest element in $\{\mathbf{L}_1 + \mathbf{L}_1\}$ is $N_1^2 + 2$, where $N_1^2 - N_1 + N_2 + 2 = N_1^2 + 2$ when $\varepsilon = 0$ and $N_1^2 - N_1 + N_2 + 2 - (N_1^2 + 2) = 1$ when $\varepsilon = 1$. Then, the positive SCA can be written as $\{\mathbb{Z}(S_3) - \mathbb{Z}(S_2)\}$ with $S_2 = N_1^2 + 2$ in the case of $0 \le \varepsilon < 2$. If $\varepsilon = 2$ or 3, the largest element in $\{\mathbf{L}_1 + \mathbf{L}_1\}$ is $N_1^2 - N_1 + N_2 + 2$, which means the positive SCA is $\{\mathbb{Z}(S_3) - \mathbb{Z}(S_2)\}$ with $S_2 = N_1^2 - N_1 + N_2 + 2$. If $\varepsilon \ge 4$ and $N_1 = \varepsilon - 1$, the largest element in $\{\mathbf{L}_1 + \mathbf{L}_1\}$ is $N_1^2 + N_1 + 2$ and $N_1^2 - N_1 + N_2 + 2 - (N_1^2 + N_1 + 2) = 1$, then the positive SCA is $\{\mathbb{Z}(S_3) - \mathbb{Z}(S_2)\}$ with $S_2 = N_1^2 + N_1 + 2$. If $\varepsilon \ge 4$ and $N_1 \ne \varepsilon - 1$, the largest element in $\{\mathbf{L}_1 + \mathbf{L}_1\}$ is $N_1^2 - N_1 + N_2 + 2$, the corresponding positive SCA is $\{\mathbb{Z}(S_3) - \mathbb{Z}(S_2)\}$ with $S_2 = N_1^2 - N_1 + N_2 + 2$. Since the negative SCA and the positive SCA are symmetrical about the zero, the whole SCA of the proposed array is continuous in set of $\{\mathbb{Z}(S_3) - \mathbb{Z}(S_2)\} \cup \{\mathbb{Z}(-S_3) - \mathbb{Z}(-S_2)\}$.

B. Proof of Proposition 2

Based on Proposition 1, the whole DCA is continuous in the set of $\{\mathbb{Z}(-S_1) \cup \mathbb{Z}(0) \cup \mathbb{Z}(S_1)\}$ with $S_1 = N_1^2 - N_1 + 2N_2 - 1$, which means the DCA of the proposed array is hole-free. Also, $S_2 = N_1^2 + 2$ holds in the case of $0 \le \varepsilon < 2$, then we have $S_1 - S_2 = 2N_2 - N_1 - 3 \ge 0$ since $N_2 \ge N_1 > 2$ ($N_1 \in \mathbb{Z}^+$, $N_2 \in \mathbb{Z}^+$). This implies that S_2 falls within the continuous interval of DCA; as a result, the SDCA is hole-free. Similar analyses can be utilized for the cases of $\varepsilon = 2$ or 3, $\varepsilon \ge 4$, i.e., if $\varepsilon = 2$ or $\varepsilon = 3$, $S_1 - S_2 = N_2 - 3 \ge 0$; if $\varepsilon \ge 4$ and $N_1 = \varepsilon - 1$, $S_1 - S_2 = 2N_2 - 2N_1 - 3 \ge 5$; if $\varepsilon \ge 4$ and $N_1 \ne \varepsilon - 1$, $S_1 - S_2 = N_2 - 3 \ge 4$. Consequently, the SDCA of the proposed array is continuous in the set of $\{\mathbb{Z}(-S_1) \cup \mathbb{Z}(0) \cup \mathbb{Z}(S_1)\} \cup \{\mathbb{Z}(S_3) - \mathbb{Z}(S_2)\} \cup \{\mathbb{Z}(-S_3) - \mathbb{Z}(-S_2)\}$ and can be simplified to $\mathbb{Z}(-S_3) \cup \mathbb{Z}(0) \cup \mathbb{Z}(S_3)$, which means the proposed array has hole-free SDCA. Accordingly, the maximum consecutive DOF is $2S_3 + 1 = 4(N_1^2 - N_1 + 2N_2) + 1$.

Data Availability

No data were used in this study.

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 the National Natural Science Foundation of China under Grant no. 62101223 and Major Basic Research Project of the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant nos. 20KJB510027 and 20KJA510008.

References

[1] H. Wang, X. Li, R. H. Jhaveri et al., "Sparse bayesian learning based channel estimation in FBMC/OQAM industrial IoT networks," *Computer Communications*, vol. 176, pp. 40–45, 2021.

[2] M. Agatonović, Z. Stanković, and N. Dončov, "Application of artificial neural networks for efficient high-resolution 2D DOA estimation," *Radio Engineering*, vol. 21, no. 4, pp. 1178–1186, 2012.

[3] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[4] L. Wan, Y. Sun, L. Sun, Z. Ning, and J. J. P. C. Rodrigues, "Deep learning based autonomous vehicle super resolution DOA estimation for safety driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4301–4315, 2021.

[5] S. Qin, Y. D. Zhang, and M. G. Amin, "Improved two-dimensional DOA estimation using parallel coprime arrays," *Signal Processing*, vol. 172, Article ID 107428, 2020.

[6] W. Zheng, X. Zhang, Y. Wang, J. Shen, and B. Champagne, "Padded coprime arrays for improved DOA estimation: exploiting hole representation and filling strategies," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4597–4611, 2020.

[7] R. Cohen and Y. C. Eldar, "Sparse array design via fractal geometries," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4797–4812, 2020.

[8] G. Qin, M. G. Amin, and Y. D. Zhang, "DOA estimation exploiting sparse array motions," *IEEE Transactions on Signal Processing*, vol. 67, no. 11, pp. 3013–3027, 2019.

[9] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 573–586, 2011.

[10] P. Pal and P. P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4167–4181, 2010.

[11] S. Qin, Y. D. Zhang, and M. G. Amin, "Generalized coprime array configurations for direction-of-arrival estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 6, pp. 1377–1390, 2015.

[12] A. Raza, W. Liu, and Q. Shen, "Thinned coprime array for second-order difference co-array generation with reduced mutual coupling," *IEEE Transactions on Signal Processing*, vol. 67, no. 8, pp. 2052–2065, 2019.

[13] C.-L. Liu and P. P. Vaidyanathan, "Super nested arrays: linear sparse arrays with reduced mutual coupling-part I: Fundamentals," *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 3997–4012, 2016.

[14] J. Liu, Y. Zhang, Y. Lu, S. Ren, and S. Cao, "Augmented nested arrays with enhanced DOF and reduced mutual coupling," *IEEE Transactions on Signal Processing*, vol. 65, no. 21, pp. 5549–5563, 2017.

[15] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2021.

[16] L. Wan, K. Liu, Y.-C. Liang, and T. Zhu, "DOA and polarization estimation for non-circular signals in 3-D millimeter wave polarized massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3152–3167, 2021.

[17] X. Wang, Z. Chen, S. Ren, and S. Cao, "DOA estimation based on the difference and sum coarray for coprime arrays," *Digital Signal Processing*, vol. 69, pp. 22–31, 2017.

[18] Z. Chen, Y. Ding, S. Ren, and Z. Chen, "A novel nested configuration based on the difference and sum co-array concept," *Sensors*, vol. 18, no. 9, pp. 1–17, 2018.

[19] W. Si, Z. Peng, C. Hou, and F. Zeng, "Improved nested arrays with sum-difference coarray for DOA estimation," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6986–6997, 2019.

[20] P. Zhao, K. Wang, G. Hu, and L. Wan, "Underdetermined DOA estimation using unfold coprime array from the perspective of sum-difference co-array," *IEEE Access*, vol. 7, pp. 168557–168564, 2019.

[21] Y. Wang, W. Wu, X. Zhang, and W. Zheng, "Transformed nested array designed for DOA estimation of non-circular signals: reduced sum-difference co-array redundancy perspective," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1262–1265, 2020.

[22] F. Izedi, M. Karimi, and M. Derakhtian, "Joint DOA estimation and source number detection for arrays with arbitrary geometry," *Signal Processing*, vol. 140, pp. 149–160, 2017.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 4571840, 12 pages
https://doi.org/10.1155/2021/4571840

Research Article

Prediction of Enterprise Economic Activity Behavior Based on Neural Network and ARIMA Hybrid Model

Zeyu Lin[1] and Shuai Li[2]

[1]*Hong Kong Baptist University, Kowloon Tong, Hong Kong 999077, China*
[2]*Central University of Finance and Economics, Beijing 100081, China*

Correspondence should be addressed to Shuai Li; cufejwc_ls@cufe.edu.cn

Received 4 September 2021; Revised 21 October 2021; Accepted 25 October 2021; Published 15 November 2021

Academic Editor: Han Wang

Enterprise economy refers to the comprehensive situation reflected in the gross product, production scale, total production and efficiency, technological content, marketing means, and so on; under certain social conditions, enterprises use resources obtained by law to engage in economic activities. Under the guidance of consciousness or culture, enterprises use "legally obtained resources" to promote economic development. Enterprise economy is affected by manpower, capital, management, operation, policy, and other aspects. In the context of the rapid development of big data in the current era, this paper proposes a prediction model of enterprise economic activity behavior based on neural network and ARIMA by investigating a variety of artificial intelligence models and verifies its feasibility. Commodity circulation enterprises have a more urgent demand for the development of business audit due to their operation characteristics. Therefore, this paper takes commodity circulation enterprises as representatives and predicts business audit in the big data environment based on the model proposed in this paper.

1. Introduction

Prediction is a study of what has not happened and what is not clear in advance, speculating and realistically studying the results of things that should happen. Since Lapebest and Farbe first applied artificial neural network technology to forecast, neural network method has been paid more and more attention. Many experts and scholars at home and abroad have done a lot of research on the application of neural network in prediction and achieved certain results [1].

In the economics field, as an important branch of forecasting field, economic forecasting has attracted much attention. Although traditional enterprise economic forecasting methods have advantages, with the complexity of economic management problems, these traditional methods cannot meet the requirements of scientific management and correct management. Neural network technology provides a new research conception and method for forecasting research. Since the previous research has achieved certain research results, it has become a valuable research topic to apply neural network to the research of enterprise economic forecasting, which forms the basis and idea of this topic selection. However, because the economic operation system is a complex system, the success of neural network technology in economic prediction and alarm is obviously a huge systematic project [2]. Therefore, the research in this field is mostly based on exploratory nature. With the deepening of financial sharing and industrial property integration, the management and supervision of commodity circulation enterprises has become a trend. At this time, enterprises can develop an embedded audit system that conforms to their own business characteristics and realize real-time monitoring and alarm such as the occurrence, collection, and analysis of business data, so as to truly monitor before and after the event at the same time [3].

Aiming at the lack of previous studies, this paper puts forward some views and ideas of its own and establishes a mixed model of enterprise economic activity action prediction based on neural network and ARIMA. On this basis, an empirical study is carried out by using actual data. It lays a good foundation for further research in the future.

CPU, hard disk utilization, and resource consumption often have a great impact on the response speed of programs, and insufficient memory may even lead to program crash. Performance models pay more attention to these contents. We propose a method based on memory measurement and derived measurement to narrow the gap and predict the impact of resources. In the past, there were many models to solve this problem. However, the accuracy is too low, so we use the performance model generator to automatically extract memory behavior and evaluate it with industry benchmarks. The results are satisfactory. It can play a role in capacity planning (internal), cost forecasting (cloud), and so forth and can also be used for continuous model-based evaluation of memory usage of enterprise architecture [4]. We study the potential factors that affect the economic cycle and turbulence and take the venture capital market as a case to study the timing in organizational behavior. It is shown that once an enterprise carries out more frequent business activities or transactions in a short time, sudden effects will occur, which is the potential factor. Similarly, when the memory effect of organization occurs, it can predict the future behavior. From this, we find the changing characteristics of organizational behavior and economic system, which makes researchers think it is necessary to extend the study of human dynamics to the study of organizational dynamics [5].

Performance appraisal is very important for enterprises. With the expansion of enterprise scale and the increase of employees, how to match the key characteristics of employees' performance and how to accurately identify potential employees are the key issues of research. The application of deep learning to analyze data sets and identify and predict employees' potential and reasons for turnover is significantly improved compared with traditional machine learning methods [6]. We build an employee topology diagram based on interactive network and then use a large amount of collected employee data to communicate and analyze with employee data from all walks of life to get its key features. Finally, compared with the traditional model, the accuracy and AUC indicators are improved. Because of its characteristics, the design of composite cap stiffening plate will involve a large number of geometric and material parameters. Therefore, we put forward an analysis tool with high analysis and calculation efficiency based on artificial neural network for high-efficiency calculation and analysis. The tool uses the training set, verification set, and test set of neural network and compresses the original characteristics and then establishes back-propagation neural network to predict buckling and finally the test is performed. Experiments show that the neural network can effectively predict the buckling and ultimate load of composite cap-ribbed plates under in-plane shear action, which further reflects that the analysis is of great significance for practical design [7]. We use 1000 small businesses as a sample to study the potential of small business default prediction. Combining industry priority payment behavior with Kohonen map, two prediction models based on financial ratio and payment behavior related variables are calculated. The results show that the prediction accuracy of Kohonen map is inversely proportional to the enterprise scale; if small businesses add variables related to payment, the accuracy can be greatly improved [8]. The previously mentioned research methods are also good, but when many enterprises predict their economic activities, they do not study the fusion of various algorithms, and there are defects in the prediction accuracy.

2. Hybrid Model of Neural Network and ARIMA

2.1. Neural Network Model. Artificial neural network is referred to as neural network. As a branch of artificial intelligence, it has been widely used in model recognition, optimization, intelligent robot, control, prediction, detection, and so on. At present, there are many ways to build neural networks, including neural network models based on mathematical methods, neural network models based on thermodynamic methods, and neural network models based on fuzzy and chaotic methods. It has become an important direction in the field of science. The basic characteristics of neural network make it have unique advantages. When predicting the data with few training samples, it can get the network threshold and network weight quickly and accurately and has strong learning and fitting ability. Neuron is the most basic unit of neural network. It is the simplest neural network, which can process information nonlinearly. It can have multiple inputs but only one output. Its structure is shown in Figure 1, which consists of three parts: input, activation function, and output [9].$x_1, x_2, \ldots, x_r$ are r input parameters, $\omega_i (i = 1, 2, \ldots, r)$ and b are the weights and biases of neural network connection, respectively, $n = \sum \omega_i x_i + b$ will be brought into an activation function $f(\cdot)$ as input, and the final output is $a = f(\sum \omega_i x_i + b)$.

In the input neuron structure proposed in Figure 1, economic indicators are used as input variables to learn and train neurons. Generally, the input economic index data needs to be preprocessed, and the relevant processing process has been given in the article.

Then, a neural network will be obtained by connecting a certain number of neurons according to the required structure, in which the activation function is the key to realize the nonlinear mapping of the neural network. The commonly used activation functions are as follows:

(1) Sigmoid function:

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{1}$$

(2) Tanh function:

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}. \tag{2}$$

(3) Threshold function:

$$\begin{aligned} (a) f(x) &= \begin{cases} 1, & x \geq 0, \\ 0, & x < 0, \end{cases} \\ (b)\ \operatorname{sgn}(x) &= \begin{cases} 1, & x \geq 0, \\ -1 & x < 0. \end{cases} \end{aligned} \tag{3}$$

(4) ReLU function:

$$f(x) = \max(0, x). \tag{4}$$

2.2. BP Neural Network. Feedforward neural network, referred to as BP neural network for short, was established in 1986. When the training error is obtained, the network propagates it backwards, and the parameters in the network are modified in this way. When the sum of squares of the error meets the requirements of network accuracy, the iteration will stop. Its signal propagation is unidirectional; that is, the signal of each layer only affects the next layer but has no influence on the upper layer and other layers [10]. It transforms the input and output of experimental samples into a nonlinear optimization problem, and the gradient descent algorithm is adopted. In fact, for this network, it is worth noting that it can effectively approximate any nonlinear function and only needs one hidden layer. It can be seen that this network has strong learning ability and strong feasibility, so it has become the most widely used artificial neural network. Its topology is shown in Figure 2.

The modeling steps of BP neural network are as follows:

Step 1. Data preprocessing:When forecasting, if the values of the data in the samples differ a lot, it is easy to have numerical problems. In order to solve these problems, it is necessary to normalize the original data. The normal way of normalization is

$$Y_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad i = 1, 2, \ldots, n. \tag{5}$$

After the original data is normalized, the data Y_i, x_i are the input values, and $x_{\max}, x_{\min}$ are the maximum and minimum values of the normalized input sequence, respectively. It is worth noting that, in order to obtain the final predicted value of the model, the predicted value of the network needs to be denormalized in the subsequent process:

$$x_i = Y_i * (x_{\max} - x_{\min}) + x_{\min}. \tag{6}$$

Step 2. Determining the structure of neural network: According to the characteristics of data, the optimal number of structural layers and the number of nodes of each structural layer needed by the built neural network model are determined [11]. For the problem of determining the number of nodes in each layer, it is important and difficult to determine the number of nodes N needed in the hidden layer. At present, there is no accurate method for this problem, which cannot be generalized. People need to try many times to finally determine the value of N which makes the prediction effect of the model optimal. It should be noted that the value of N should not be too large or too small. If the number of nodes is too large, the training process will be too complicated and the training time will be increased. If the number of hidden layer nodes is too small, the model cannot learn enough nonlinear features. The value of N can be preliminarily determined by the following empirical formula, and people need to try to compare many times near the determined number of nodes and finally determine the value of the number of nodes n.

$$\begin{cases} n = \sqrt{NM}, \\ n = \log_2 N, \\ n = \sqrt{N + M} + a, \quad a \in [1, 10], \end{cases} \tag{7}$$

where n represents the number of nodes in network training; formula (7) uses three formulas to describe the value process of n, which is difficult to determine in the training process. Different ways lead to different results and have different influences on the performance of the algorithm. In order to affect the overall algorithm effect, three formulas are used to describe n, and the formula value of n is determined according to the sample size and the corresponding algorithm requirements.

Step 3. Determining the activation function: The hidden layer and the output layer need to determine the activation function, and there are many choices for the activation function $f(x)$, such as tansig, logsig, purelin, and many other functions.

Step 4. Initializing weights, thresholds, and learning rates

Step 5. Forward-propagation calculation: The normalized data is input and propagated forward through the determined network to obtain the following:

Hidden layer output H_j:

$$H_j = f\left(\sum_{i=1}^{n} \omega_{ij} + a_j\right), \quad j = 1, 2, \ldots, L, \tag{8}$$

ω_{ij} is the weight from the input layer to the hidden layer, a_j is the hidden layer threshold, and L is the number of hidden layer nodes.

Output layer output O_k:

$$O_k = \sum_{i=1}^{n} \omega_{jk} H_j + b_k, \quad k = 1, 2, \ldots, m, \tag{9}$$

ω_{jk} is the weight from hidden layer to output layer and b_k is the threshold of output layer.

Step 6. Error calculation

$$e_k = Y_k - O_k, \quad k = 1, 2, \ldots, m. \tag{10}$$

Step 7. Weight updating: According to the prediction error e_k calculated by the above formula, the weights ω_{ij} and ω_{jk} in the established neural network are updated and adjusted.

$$\begin{aligned} \omega_{ij} &= \omega_{ij} + \eta H_j \left(\left(1 - H_j\right) x_i \sum_{k=1}^{m} \omega_{ik} e_k\right), \\ \omega_{jk} &= \omega_{jk} + \eta H_j e_k. \end{aligned} \tag{11}$$

Step 8. Threshold updating: The thresholds a_j and b_k are updated according to the prediction error e_k.

$$a_j = a_j + \eta H_j\left(1 - H_j\right)\sum_{k=1}^{m}\omega_{jk}e_k,$$
$$b_k = b_k + \eta e_k. \tag{12}$$

Step 9. Judging whether the iteration can be terminated after obtaining new weights and biases; if not, return to step 5 to iterate and update parameters repeatedly until the minimum mean square error is less than the set value, and the BP neural network modeling is finished.

2.3. ARIMA Model

2.3.1. Time Series

(1) Definition of Time Series. Time series refers to a series of numbers arranged by the values of the same statistical index according to the time sequence of their occurrence [12]. Its formation is very simple, but it is widely used, and the process and phenomenon of the development of many things can be expressed by it, like the annual precipitation in meteorological statistics, in industrial production, the number of finished products in each quarter, in the field of national economy, annual GDP, and so on; the research fields involved in time series are numerous.

In people's research on time series, there are generally two purposes: one is to establish a model to understand the mechanism of generating the time series; and the second is to predict the possible level of the time series in the next time period based on the historical data of the series and some other influencing factors.

(2) Basic Concepts of Time Series.

Step 1: Stationary process: It is a special stochastic process whose main characteristic is that its statistical characteristics are fixed with time.

Let $\{X_t\}$ be a time sequence; for all positive integers r, take $t_1, t_2, \ldots, t_r \in T$ for any integer t; there are

$$F_{t_1,t_2,t_3,\ldots,t_r}(x_1, x_2, x_3, \ldots, x_r) = F_{t_1+\tau,t_2+\tau,t_3+\tau,\ldots,t_r+\tau}(x_1, x_2, x_3, \ldots, x_r). \tag{13}$$

Then $\{X_t\}$ is called a strictly stationary time series.

In our practical application, it is very difficult to obtain the distribution of random sequences; even if we know the joint distribution, it is very difficult to apply on its basis. Therefore, the discussion of Yan Ping is generally studied at the theoretical level. In practical applications, more people will choose wide stationary sequences with relatively loose restrictions.

Let $\{X_t\}$ be a time series if the following conditions are met:

(1) Take $t \in T, EX_t^2 < \infty$.

(2) Take $t \in T$; there is $EX_t = \mu = \mu$, where μ is a constant.

(3) The time series $\{X_t\}$ is called a broadly stationary time series if $r, s, t \in T$ has $\gamma(r, s) = \gamma(t, t + s - r)$.

Generally speaking, a time series that satisfies the strict stationary condition must satisfy the wide stationary condition, and vice versa. But this is not absolutely true. There are special examples in these two aspects. When the time series satisfies the condition of normal distribution, they are consistent.

Pure random sequence (also called white noise sequence) is a special stationary time series. White noise test is very important in time series, and it is an important basis to judge the rationality of the model. The definition of pure random sequence is introduced below.

Sequence $\{X_t\}$ satisfies the following conditions:

(1) Take $t \in T, EX_t = \mu$.

(2) If $t, s \in T$ has $\begin{cases}\sigma^2, & t = s,\\ 0, & t \neq s\end{cases}$, then $\{X_t\}$ is said to be a pure random sequence.

(3) Characteristic statistics.

Some commonly used low-order and correlation functions are as follows.

Taking any time series $\{X_t, t \in T\}$, the mean value function is defined as

$$\mu_t = E\left(X_t\right) = \int_{-\infty}^{\infty} x\mathrm{d}F_t(x), \tag{14}$$

where $F_t(x)$ is the distribution function of X_t and μ_t is the mean value function of sequence $\{X_t\}$ at t.

The variance function is defined as follows when there is $\int_{-\infty}^{\infty} x\mathrm{d}F_t(x) < \infty$:

$$DX_t = E\left(X_t - \mu_t\right)^2 = \int_{-\infty}^{\infty}\left(X_t - \mu_t\right)2\mathrm{d}F_t(x). \tag{15}$$

Autocovariance is defined as

$$r(t, s) = E\left(X_t - \mu_t\right)\left(X_s - \mu_s\right). \tag{16}$$

The autocorrelation coefficient is defined as

$$\rho(t, s) = \frac{r(t, s)}{\sqrt{DX_t \cdot DX_s}}. \tag{17}$$

The closer the autocorrelation coefficient is to zero, the weaker the sequence correlation is, and the closer it is to l, the stronger the sequence correlation is.

2.3.2. Stationary Time Series Model

(1) AR Model. If the time series model satisfies the following structural conditions, it is called p-order autoregressive model, abbreviated as AR(P):

$$\begin{cases} x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + \varepsilon_t,\\ E\left(\varepsilon_t\right) = 0, \mathrm{Var}\left(\varepsilon_t\right) = \sigma_\varepsilon^2, E\left(\varepsilon_t\varepsilon_s\right) = 0,\\ E\left(\varepsilon_s\varepsilon_t\right) = 0. \end{cases} \tag{18}$$

(2) MA Model Q. -order moving average model, abbreviated as MA (Q), has the following structure:

$$\begin{cases} x_t = \mu + \varepsilon_t - \Theta_1\varepsilon_{t-1} - \Theta_2\varepsilon_{t-2} - \cdots - \Theta_q\varepsilon_{t-q}, \\ E(\varepsilon_t) = 0, \mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_s\varepsilon_t) = 0. \end{cases} \tag{19}$$

(3) ARMA Model. Autoregressive moving average model, abbreviated as ARMA (P, Q), has the following structure:

$$\begin{cases} x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \ldots + \varphi_p x_{t-p} + \varepsilon_t - \Theta_1\varepsilon_{t-1} - \Theta_2\varepsilon_{t-2} - \cdots - \Theta_q\varepsilon_{t-q}, \\ (\varepsilon_t) = 0, \mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_s\varepsilon_t) = 0, \\ E(\varepsilon_s\varepsilon_t) = 0. \end{cases} \tag{20}$$

2.3.3. Nonstationary Time Series Model. Because the sequences obtained in practical applications are usually unstable sequences and the stability of time series is required in the traditional linear model introduced earlier, the unstable sequences cannot be directly applied to the time series analysis model introduced above. This requires a method of converting an unstable sequence used to apply the ARMA model to the obtained stable sequence into a stable sequence, wherein the process of first transforming the sequence to fit the ARMA model is the ARIMA model as applied herein.

(1) ARIMA Model. Differential autoregressive integrated moving average model, abbreviated as ARIMA (P, D, Q) model, has the following structure:

$$\begin{cases} \varphi(B)\nabla^d X_t = \theta(B)\varepsilon_t, \\ (\varepsilon_t) = 0, \mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_s\varepsilon_t) = 0, \\ E(\varepsilon_s\varepsilon_t) = 0. \end{cases} \tag{21}$$

When $d = 0$, model ARIMA (p, d, q) is actually model ARIMA (p, q). Fitting ARIMA (P, Q) model with D-order difference for nonstationary sequence is equivalent to fitting ARIMA (P, D, Q) model for original nonstationary sequence. The following summarizes the transformation process between the above time series models, as shown in Figure 3.

According to Cramer decomposition theorem, for any time series, its fluctuation is not only the result of random influence but also the result of deterministic influence.

It is worth noting that although the difference operation can fully extract the deterministic trend of time series, the nonstationary sequence is transformed into a stationary sequence, but every difference will be accompanied by certain information loss. Therefore, when applying this method of difference operation, it usually follows the principle of "being low but not high"; if you can use low-order difference, you will not use high-order difference.

2.3.4. ARIMA Modeling Steps

Step 1. The common inspection methods for stationarity inspection are as follows:

(1) Sequence diagram test method:Time sequence diagram is an indispensable part of time series analysis. By making a time series diagram to observe the data distribution, if the data distribution fluctuates around a constant value, it can be roughly judged that the time series is stationary. Although this method is simple and convenient, it has great subjectivity and is not accurate enough. Therefore, it can only be used as a preliminary judgment.

(2) Unit root test:PP test is mainly suitable for the case of nonhomogeneous variance; DF test is used in the case of first-order autoregression; ADF test is mainly applicable to the case of homogeneous variance.

(3) Correlation function judgment method:The sample autocorrelation function is

$$\widehat{\rho}_k = \frac{\widehat{\gamma}_k}{\widehat{\gamma}_0} = \frac{\mathrm{cov}(X_t, X_{t-k})}{\sigma_{X_t}, \sigma_{X_{t-k}}} = \frac{\sum_{i=1}^{n=k}(X_t - \overline{X})(X_{t-k} - \overline{X})}{\sum_{i=1}^{n=k}(X_t - \overline{X})^2}. \tag{22}$$

Observe the ACF diagram. When k increases, if $\widehat{\gamma}_k$ is in a state of rapid attenuation, it is judged to be stable. If $\widehat{\gamma}_k$ shows a slow decay state, it is judged to be nonstationary.

Step 2. Smoothing:If the test result in Step 1 is nonstationary, it needs stationary treatment. Generally, there are two ways: one is direct difference treatment, and the other is logarithmic treatment before difference treatment. If the original sequence is stationary, Step 3 is directly carried out.

Step 3. Pure random test:The data processed in Step 2 is stable, and then white noise test is done. If the sequence is white noise sequence, the modeling is finished; otherwise, go to the next step.

Step 4. Identifying the model structure:Observe the sample ACF graph and PACF graph to identify the model and determine the order, and determine the range of model parameters. The basic principles of model recognition and ranking are shown in Table 1. It should be noted that this model identification and

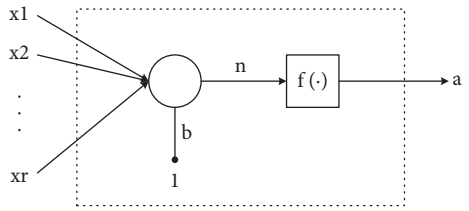


Figure 1: Neuron structure.

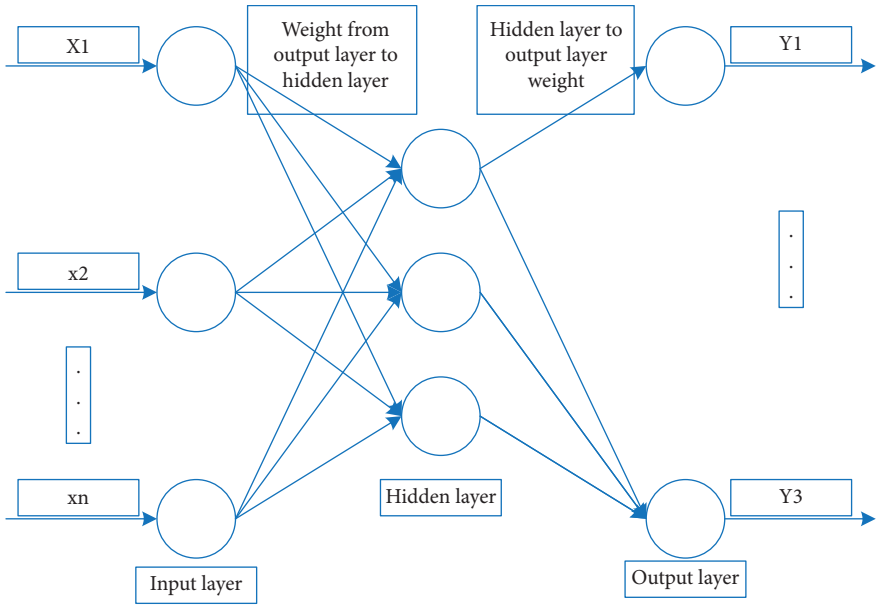


Figure 2: Topological structure of BP neural network.

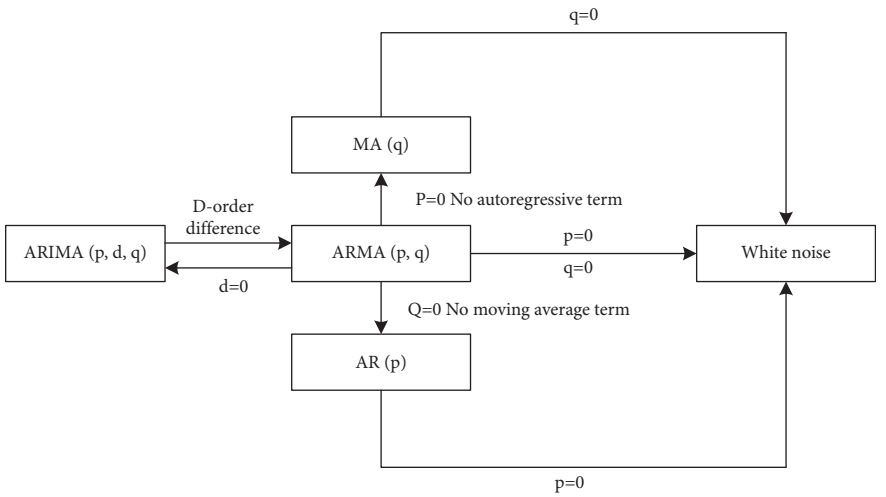


Figure 3: Model transformation diagram.

Table 1: Model identification criteria.

ACF	PACF	Selection model
Tailing	P-order truncation	AR(P)
Q-order truncation	Tailing	MA(q)
Tailing	Tailing	ARMA(p, q)

order determination method is subjective and not very accurate, so we calculate the AIC value and BIC value of the model corresponding to each group of parameters within the range of model parameters and find the optimal model according to AIC criterion and BIC criterion.

AIC criterion (Minimum Information Criterion), which is based on the prediction error of the model and is used to judge whether the lag order finally determined by the autoregressive model is appropriate, can be expressed as

$$\mathrm{AIC} = -2\ln(L) + 2k, \tag{23}$$

where L is the exponential likelihood function and K is the number of selected parameters.

Step 5. Model validation:It is divided into parameter significance and model adaptability test. The former is to test whether the parameter is significantly zero to judge whether the independent variable corresponding to a parameter has a significant effect on the model and then consider whether to eliminate the variable. The latter is to test whether the residual sequence is white noise sequence and then judge whether the model is valid.

Step 6. Prediction and analysis of the model

The model is used to predict the short-term sequence values in the future. The modeling process of ARIMA model is shown in Figure 4.

2.4. Hybrid Model of Neural Network and ARIMA

2.4.1. Advantages of Hybrid Models. A hybrid model is a new model formed by combining different models. From the perspectives of model itself and data itself, on one hand, there are many models in the field of forecasting; different models have their own advantages and disadvantages. At the same time, there are many similarities between different models, which are not opposite. In order to be compatible with the advantages of different models, different models can be combined to construct a new model. Because the new model has the advantages of two or more models at the same time, the final prediction effect is more ideal. On the other hand, in practice, most time series contain both linear and nonlinear components. In this case, the mining of information by a single model is limited, which often cannot achieve the best prediction effect. Then, combining the model suitable for linear fitting with the model suitable for nonlinear fitting makes the hybrid model have both linear and nonlinear processing abilities, thus making the prediction effect more ideal. We know that neural network is very good at data processing, and learning ability is also the advantage of neural network, which is good at mining nonlinear features of sequences. Neural network also has certain advantages in operation. First, it requires simple number of experimental samples; second, its structure is flexible and its design is convenient. However, ARIMA model, a traditional time series model, is good at mining linear features of data. Then, this section aims to combine a variety of neural network models with the ARIMA models to form a hybrid model of multiple neural network models and ARIMA models, so that the hybrid model has both linear processing ability and nonlinear processing ability. Finally, the constructed hybrid model is applied to the prediction of enterprise economic data.

2.4.2. Introduction to Mixed Modeling Methods. The hybrid model can be divided into two hybrid methods: The first method is to combine different models through the method of weight allocation. Because the prediction results of different models are different, it is necessary to allocate the weights of the models participating in the combination reasonably, so the new mixed model is formed by the method of allocating different weights to different models. For this combination, the distribution of weights should be set in advance and should be reasonable and effective. The second method is to use model x_1 to predict and get prediction result 1 and residual error and then use the residual error as the input of model x_2 to predict and get result 2 and finally combine result 1 and result 2. Superposition is used as the prediction result of the hybrid model. Compared with the first method, the advantage of this method is that it does not need to set reasonable and effective weight distribution in advance, and the disadvantage is that the data processing will be relatively complicated. This paper adopts the second combination mode. According to the ability of neural network model and ARIMA model, they are combined to form a hybrid model of neural network and ARIMA. Then, among the different data predictions applied by the hybrid model, the combination method is as follows.

Firstly, the ARIMA model is used to model the data, and the linear characteristics of the data are mined. The predicted value is recorded as L (1), and the residual e_t is calculated from the original sequence y_t and $L^{'}(t)$; namely,

$$e_t = y_t - L^{'}(t). \tag{24}$$

Sequence e_t contains the nonlinear characteristics of the original sequence y_t, so the nonlinear characteristics of the data are further mined, and the neural network model is applied to approximate this nonlinear relationship in e_t; that is, sequence e_t is used as the input of the neural network after matrix transformation and normalization, and the output result $M^{'}(t)$ is obtained. Finally, the two are added as the prediction result $N^{'}(t)$ of the mixed model:

$$N^{'}(t) = L^{'}(t) + M^{'}(t). \tag{25}$$

At this time, the prediction effect of the hybrid model is compared with that of the single neural network model in the previous paper.

2.4.3. Concrete Steps of Hybrid Prediction Model Modeling.

Step 1: Establish ARIMA time series model for original data. The ARIMA model is used to fit the rule of sample

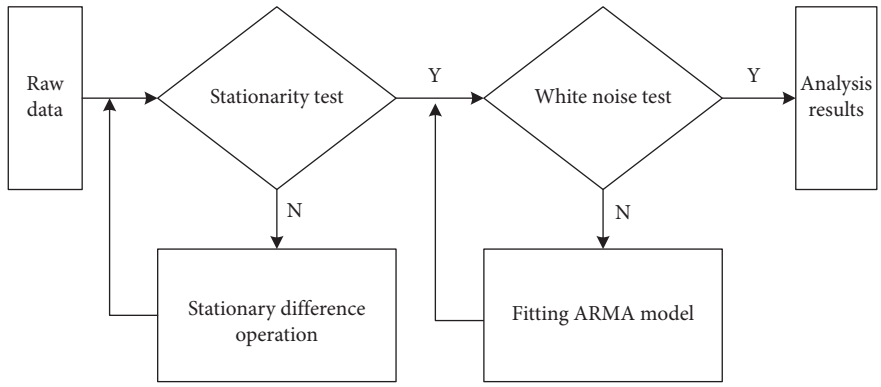


Figure 4: ARIMA model modeling process.

data, and the future data is predicted to get the predicted value and residual error.

Step 2: Establish three neural network models for the residuals obtained in Step l. Firstly, the training set and test set are generated by the sliding window method and then normalized as the input matrix of single neural network prediction model, and then the network is trained to predict and get the predicted value.

Step 3: Add the prediction results in the previous two steps, respectively, to obtain the final prediction values of the three mixed models. The hybrid model flow is shown in Figure 5.

3. Case Analysis of Economic Audit Forecast of Sales Business of Commodity Circulation Enterprises

The main business process of commodity circulation enterprises is procurement-warehouse-sales-transportation, and whether the sales data is accurate and effective directly affects the authenticity of enterprise interests. In addition, the correctness of sales forecast is very important for enterprises to make purchase decisions and sales plans. Inappropriate purchases lead to inventory backlog and other problems, which affect the capital flow of enterprises. Therefore, it is meaningful to audit sales data.

This paper validates the application effect of outlier analysis method and the mixed model of neural network and ARIMA in the business audit of commodity circulation enterprises by using the business case data of Y Enterprise.

3.1. Business Background. Y Enterprise was established in 1992, headquartered in Chongqing. It is mainly engaged in the procurement, sales, and distribution of electronic products, parts, and supporting products. As a chain enterprise, Y Enterprise has the operating characteristics of "integrating business flow, logistics and information flow" and will integrate the vertical chain management of "taking the head office as the axis, going forward to various suppliers, middle and chain stores, and going back to consumers." In terms of operation, it adopts the mode of "unified procurement of the company-unified distribution of distribution centers-sales of several chain stores." Finance adopts the method of "unified price setting, unified calculation, and unified tax payment." In terms of service, the standard of "unified service for all stores" is established and the "six unifications" of operation mode are realized as a whole. Y Enterprise's asset management focuses on various electronic products, and the capital flow is relatively simple. Buy goods with money, and then turn them into money after being sold, that is, "currency-commodity-currency." Enterprises obtain the intermediate difference in the mode of "buy low and sell high" and deduct various expenses to make profits. That is to say, buying goods from manufacturers, providing distribution to consumers, setting up services, selling various goods, and earning differences have the characteristics of commodity circulation enterprises.

Y Enterprise has developed and built a management information system with enterprise characteristics and realized management informationization and financial informationization, but the audit work still stays in the traditional manual audit stage. One is that enterprises do not have special monitoring software, so they assist in monitoring business. Inspectors can use the information management system to investigate the shared accounting and banking flows on the platform, but the management system can initially find out the suspected evidence. Second, inspectors have no information foundation, and the level of computers is limited. According to business experience, data are compared, analyzed, checked, and confirmed.

3.2. Audit Planning Phase. Due to the limitation of data acquisition, the audit objectives of this case mainly have two points: (1) Detect abnormal audit data, delete abnormal data and select them, and find out the doubtful points of audit. (2) Make sales forecast to judge whether the predicted value in the enterprise sales plan is reliable, and take it as one of the references for whether the enterprise sales plan is effective.

The analysis of this example is based on audit objectives; the original sales data is collected from the marketing management information system of Y Enterprise by SQL sentence. According to the outlier algorithm and the data analysis requirements of neural network and ARIMA mixed model, the original sales data is pretreated by Python tool, and the pretreated data is applied to the two algorithm

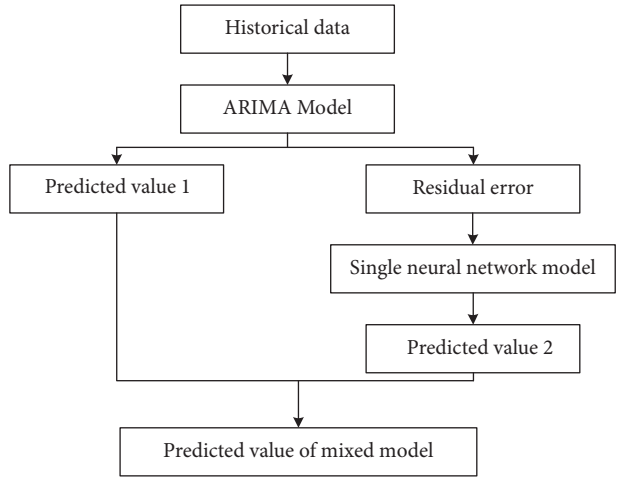


Figure 5: Flow chart of mixed prediction model.

models for data analysis. The specific audit procedure is shown in Figure 6.

3.3. Audit Data Analysis based on Neural Network and ARIMA Hybrid Model. For Y Enterprise, the reliability of sales plan directly affects the economic effect of the enterprise the next year. Therefore, doing a good job in reviewing the sales plan is conducive to improving the economic benefits of enterprises. This time, using the actual sales volume of Y Enterprise from January 2019 to June 2020, a hybrid model of neural network and ARIMA is constructed to predict the sales volume of the model from July to December 2020. The predicted results are compared with the actual sales volume and the predicted values of enterprise sales plan to verify the feasibility and effectiveness of the hybrid model of neural network and ARIMA. It provides a new method for the inspector to review the sales plan.

3.3.1. Application of Hybrid Model of Neural Network and ARIMA. (1) Stationarity Test. In this section, the stability of predicted samples is checked according to ADF test method by using time series diagram. In the historical sales data sequence diagram in Figure 7, the data is checked unilaterally using the ADF check method, and its P is judged. If the value is less than 0.05, the sequence is stable. The value of the received sample is 0.994, which is greater than 0.05, so the original sequence is unstable.

(2) Differential Conversion. The unsteady sample sequence is derived from the time series diagram of the data after differential transformation by the first-order difference in Figure 8, and the ADF check is carried out on the sequence, and the value obtained by p is $5.057e-11$, which is less than 0.05. Therefore, the sequence after primary difference is a stable sequence, which is called primary difference sequence, and the next parameter estimation can be carried out.

(3) Parameter Estimation. Autocorrelation and deflection autocorrelation analyses are carried out for the first difference sequence. According to the correlation diagram and deflection autocorrelation diagram, as shown in Figures 9 and 10, it can be seen that the autocorrelation value is selected to be of order 2 and the partial autocorrelation value is selected to be of order 3.

Using BIC matrix to test parameters, the optimum values of P and Q are 3 and 3, respectively, so the optimum mixed model of neural network and ARIMA is ARIMA (3, 1, 3).

(1) Model checking:The partial correlation coefficient is used to test the residual correlation, and the test result is 1.77, which shows that there is no autocorrelation and it can be effectively predicted.

(2) Model prediction:Use ARIMA (3, 1, 3) to forecast the sales volume from July to December 2020, and the results are shown in Figure 11.

3.4. Analysis of Application Results of Hybrid Model of Neural Network and ARIMA. Using neural network and ARIMA hybrid model to forecast and analyze the sales business data, the predicted value of the model, the actual value, and the predicted number in the sales plan are compared and analyzed to judge the effectiveness of the application of the model. A specific comparison is shown in Figure 12.

Generally speaking, the prediction results of neural network and ARIMA hybrid model are close to the actual value, and the prediction results are more reliable than those in the sales plan. According to the survey, Y Enterprise has complete purchasing and preentry management information system. When carrying out sales forecast, according to the experience of the salesperson in charge, it makes a simple linear analysis of sales data to get the forecast result, which is the basis of sales plan. The prediction results of the hybrid model of neural network and ARIMA are more accurate and reliable than the existing sales prediction methods of enterprises, so the hybrid model of neural network and ARIMA can be used to review the sales prediction of enterprises.

The reliability of sales forecast has an important influence on the purchasing arrangement, inventory

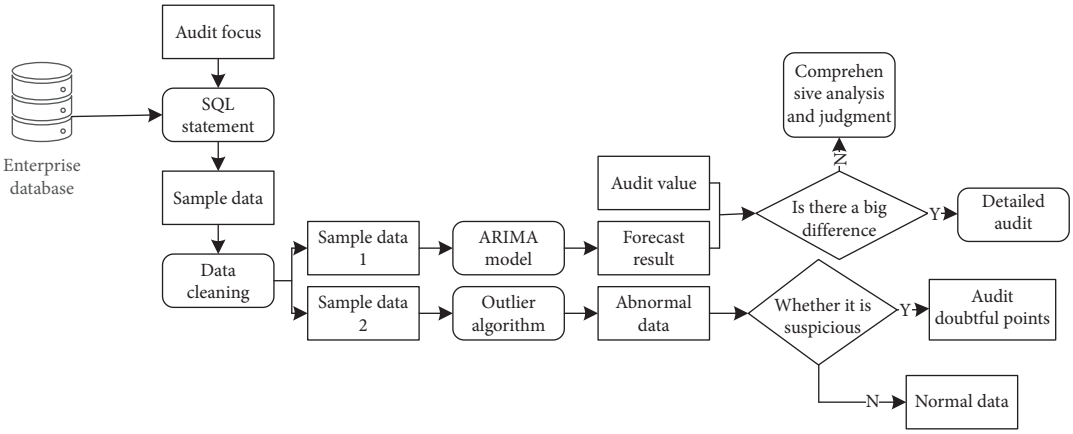


Figure 6: Audit program design.

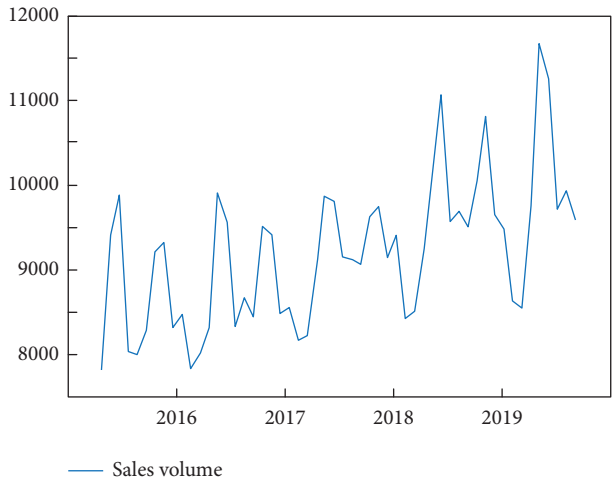


Figure 7: Sequence diagram of sample sequence.

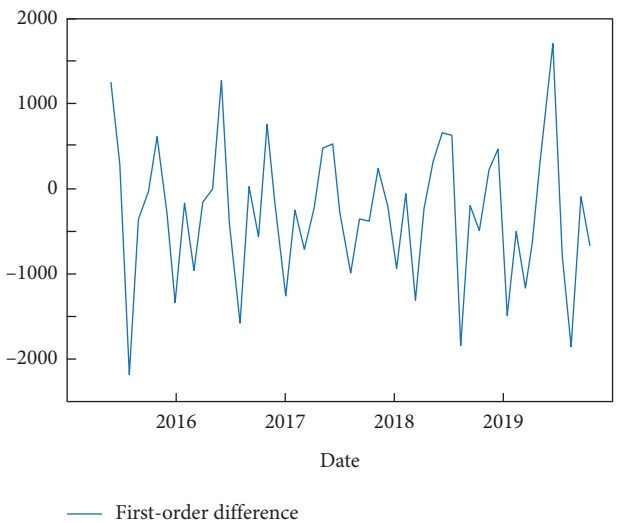


Figure 8: First-order difference sequence diagram.

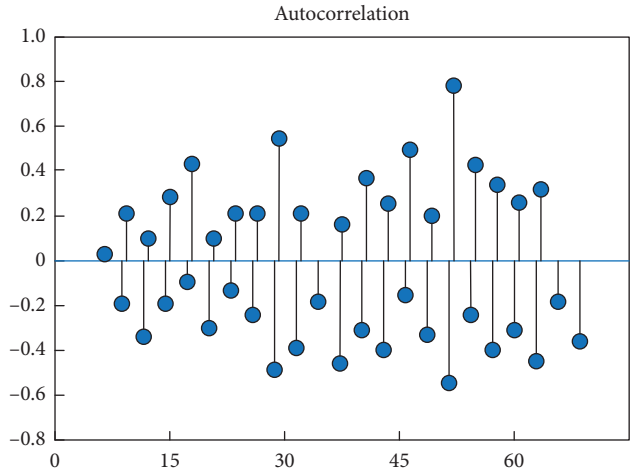


Figure 9: Autocorrelation diagram of first-order difference sequence.

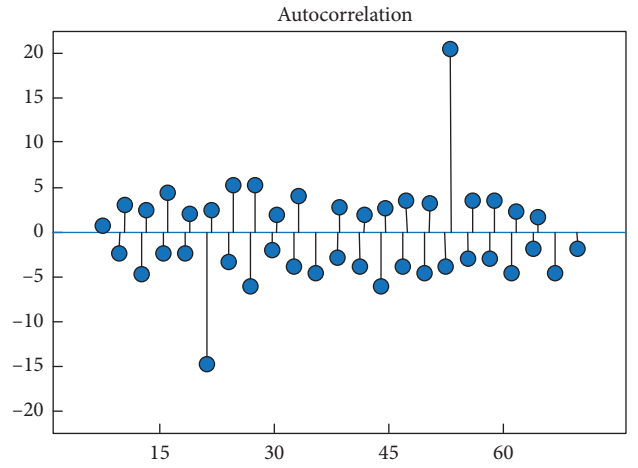


Figure 10: Partial autocorrelation diagram of first-order difference sequence.

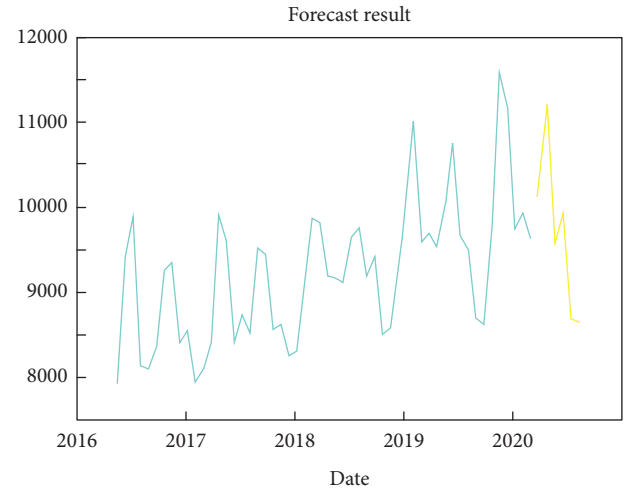


Figure 11: Sales forecast results from July to December 2020.

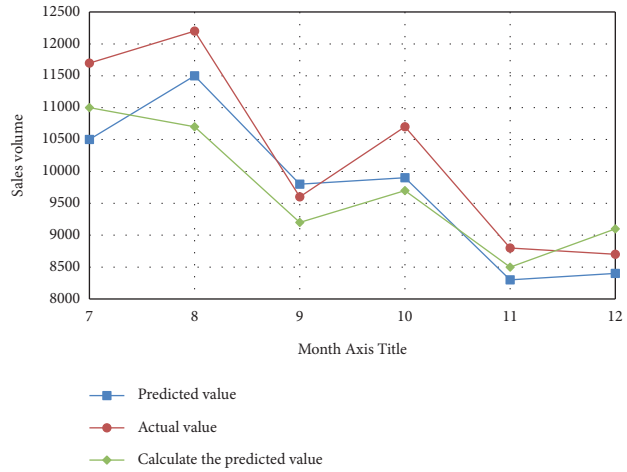


Figure 12: Comparison of predicted value, actual value, and planned predicted value.

management, and distribution of sales resources. The examination of sales forecast has an effect on improving the business performance of enterprises and can effectively prevent the possible business risks of enterprises in the future.

When the inspector reviews the forecast value of the sales plan by using the hybrid model of neural network and ARIMA, if the difference between the forecast result of the model and the forecast value of the plan is large, it is necessary to review the sales plan in detail during the forecast period. This is mainly reflected in the following aspects: (1) whether the market demand is considered; (2) whether the market competition situation is analyzed in detail, especially the sales situation of competitors; (3) whether the completion of the previous sales plan has been summarized, and, on the basis of the summary, the existing sales forecasting methods have been optimized; and (4) whether the changes in the sales force are considered. Therefore, auditors can comprehensively analyze the effectiveness of the enterprise sales plan in combination with the above situation and judge whether there are factors such as large changes in organizational personnel and changes in market demand, which lead to the discrepancy between the planned forecast value and the previous sales data rules.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] P. Shahnoori and G. P. Jenkins, "Valuation of the quality attributes of online banking services by small and medium enterprises engaged in international trade," *South African Journal of Economics*, vol. 87, no. 1, pp. 65–81, 2019.

[2] M. D. Liu, L. Ding, and Y. L. Bai, "Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction," *Energy Conversion and Management*, vol. 233, no. FEB, Article ID 113917, 2021.

[3] D. Li, L. Deng, and Z. Cai, "Statistical analysis of tourist flow in tourist spots based on big data platform and DA-HKRVM algorithms," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 87–101, 2020.

[4] F. Willnecker and H. Krcmar, "Model-based prediction of automatic memory management and garbage collection behavior," *Simulation Modelling Practice and Theory*, vol. 93, pp. 164–191, 2019.

[5] Z. Xin, X. A. Sheng, A. L. M. Vilela, and H. Eugene Stanley, "Inter-event time interval analysis of organizational-level activity: venture capital market case," *Physica A: Statistical Mechanics and Its Applications*, vol. 516, pp. 346–355, 2019.

[6] J. Yuan, "Research on behavior prediction based on deep learning–take chengdu economic innovation enterprise as an example," *E3S Web of Conferences*, vol. 275, no. 1, Article ID 03060, 2021.

[7] Z. Sun, Z. Lei, J. Zou, R. Bai, H. Jiang, and C. Yan, "Prediction of failure behavior of composite hat-stiffened panels under in-plane shear using artificial neural network," *Composite Structures*, vol. 272, p. 15, Article ID 114238, 2021.

[8] F. Ciampi, V. Cillo, and F. Fiano, "Combining Kohonen maps and prior payment behavior for small enterprise default prediction," *Small Business Economics*, vol. 54, no. 4, pp. 1007–1039, 2020.

[9] S. Y. Liu, S. Liu, Y. Tian, L. Quan, and Y. Y. Tang, "Research on forecast of rail traffic flow based on ARIMA model," *Journal of Physics: Conference Series*, vol. 1792, no. 1, Article ID 012065, 2021.

[10] Y. Zheng, K. Yuan, and L. Dai, "Research on coal price forecast based on ARIMA and SVM combination model," *E3S Web of Conferences*, vol. 257, no. 26, Article ID 02008, 2021.

[11] J. Wang, J. Cheng, F. Liu, Y. Lei, and T. Taijie, "Research on the air quality prediction model of Wuhai mining area based on deep learning," *E3S Web of Conferences*, vol. 300, no. 2, Article ID 02005, 2021.

[12] X. Wang, B. Guo, X. Ma, and B. Xuemei, "Research on emotional music reconstruction method based on DBN-GRU," *Journal of Physics: Conference Series*, vol. 1966, no. 1, Article ID 012016, 2021.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 2931801, 10 pages
https://doi.org/10.1155/2021/2931801

Research Article

Community Detection Algorithm Based on Intelligent Calculation of Complex Network Nodes

Yanjia Tian [1,2] **and Xiang Feng** [1,3]

[1]*Department of Computer Science and Engineering, East China University of Science and Technology, Shanghai 200237, China*
[2]*School of Electronics and Information, Shanghai DianJi University, Shanghai 201306, China*
[3]*Shanghai Engineering Research Center of Smart Energy, Shanghai 200237, China*

Correspondence should be addressed to Xiang Feng; xfeng_ecust@163.com

Received 30 August 2021; Revised 21 October 2021; Accepted 26 October 2021; Published 13 November 2021

Academic Editor: Han Wang

With the explosive development of big data, information data mining technology has also been developed rapidly, and complex networks have become a hot research direction in data mining. In real life, many complex systems will use network nodes for intelligent detection. When many community detection algorithms are used, many problems have arisen, so they have to face improvement. The new detection algorithm CS-Cluster proposed in this paper is derived by using the dissimilarity of node proximity. Of course, the new algorithm proposed in this article is based on the IGC-CSM algorithm. It has made certain improvements, and CS-Cluster has been implemented in the four algorithms of IGC-CSM, SA-Cluster, W-Cluster, and S-Cluster. The result of comparing the density value on the entropy value of the Political Blogs data set, the DBLP data set, the Political Blogs data set, and the entropy value of the DBLP data set is shown. Finally, it is concluded that the CS-Cluster algorithm is the best in terms of the effect and quality of clustering, and the degree of difference in the subgraph structure of clustering.

1. Introduction

Complex systems in social life and nature can be intelligently detected by network nodes. When the community algorithm can solve many problems in life, it has been widely used, and it has also promoted the improvement of community detection in complex network systems. A study about directed network module found that its value will penetrate and change [1], and the directed network module overlaps, and the overlap center includes two aspects, inward and outward. Revealing the structure in the community is one of the key issues in the study of complex networks [2]; that is, the node group formed by the module and the community will form a close unit, but the connection between the units is very weak. The current research has found that the network connection partition allows the community to overlap at the node [3], and the connection partition can be generated by the algorithm of multiple node partitions through overlap and through the line chart to show the role of degree heterogeneity. Combined with network dynamics, a new detection community algorithm is proposed [4], through two principles to explain the new algorithm, and used for the detection of strong overlapping communities. Moreover, researchers analyze the real society based on the research of overlapping network structures [5] and then propose a novel general framework to detect the entire community. According to the research, it is found that SLPA has excellent performance in identifying nodes and overlapping communities of different degrees. An algorithm is also proposed, which is very practical and can find overlapping communities in a huge network [6]. The same as the traditional algorithm is that the vertices have labels spread between the same vertices. The main contribution of this paper is to expand the labeling and propagation steps, and this algorithm is very effective in restoring the community. Researchers have found that it is necessary to verify the information when identifying the community structure of the large-scale network real world [7]; in the process of

replacement, densely connected nodes form a consensus on a unique label, and a community is generated. At present, researchers use the new community detection algorithm in the study to delete the high-different unconnected node subgroups in the splitting process [8]. To provide a general framework for the realization of this method, the application of this method to computer-generated networks and different real-world networks shows the effectiveness of this method. The identification of building blocks is very important for understanding the network structure and functional characteristics [9]. Use new technologies to find overlapping communities on a large scale, and define new features through community statistics. Based on the studied topological characteristics [10], researchers proposed a new penetration algorithm SCP, which performs rapid community detection in selected modules of weighted and unweighted networks. With recognition network being used more and more [11], it has led to the lack of feature consensus in overlapping communities, so the previous algorithm scan was modified accordingly to solve some attribute defects. Furthermore, researchers have found a completely solvable social network model [12] that provides models for simple one-party and two-party networks and also compares the prediction results of the model with social networks in the real world. The hidden relationship is revealed by the complex network [13], but the detection method at this time is unstable, and the result is affected by many factors. In the case of self-consistent combination with any method, the stability and accuracy of the partition result will be greatly improved. A study proposes a simple and intuitive network cohesion method [14], which can be implemented in a few lines of code. According to research, node separation can achieve quite obvious effects on the best graph and clustering method in terms of quality and efficiency. The network community structure algorithm proposed in this paper densely divides connected subgroups [15]. After a series of calculations, it is proved that this method also has a significant effect on network data in the real world research on the community detection algorithm based on the above-mentioned complex network intelligence.

With the increasing maturity of data mining technology [16], and the wide application of complex networks in various fields, clustering technology has been developed day by day. Analyzing the community structure in complex networks can not only make researchers have a deeper understanding of the structural and performance characteristics of nodes in complex networks, but also help find the evolution rules of complex networks.

2. Basic Community Detection Algorithm

Nowadays, with the advent of the era of big data explosion, data mining technology has made rapid progress, and complex networks have become a hot research direction in the field of data mining. A community detection method is proposed [17]; it is a hierarchical clustering method. It uses the random walk model to calculate the dissimilarity of nodes in the network to complete the division of the network. Radicchi et al. used the idea of edge-based clustering coefficient to complete the identification of community structure [18]; this method first calculates the clustering coefficient of edges, then selects lower-value edges from them, and then gradually deletes them to identify the community. The advantage of this algorithm is that it can be successfully applied to large networks. A community detection algorithm (GN algorithm) is proposed in 2002 [19, 20]; this algorithm is an algorithm based on the idea of splitting. It initially calculates the betweenness values of the edges contained in the network and keeps updating the betweenness values of the remaining edges in the network until all edges in the network are deleted, thereby obtaining community structure. A network community detection algorithm based on minimum spanning tree and modularity is proposed [21]; it realizes community detection through a hierarchical idea. In addition, a detection algorithm that recognizes communities based on the idea of tag propagation is proposed [22]; this algorithm initially defines a label for all nodes in the graph and then sequentially replaces the label of node i with the label owned by most of the neighboring nodes around it. And a new algorithm is to minimize the number of edges that exist outside the community to divide the network graph [23] and finally complete the recognition of the community, but the algorithm must know the number of communities in advance. Furthermore, a hierarchical clustering algorithm is developed [24]; according to the rule of information centrality, this method uses information centrality to detect edges in communities and gradually remove the edge with the highest value of information centrality in the graph to identify the community structure. A detection algorithm, NMF algorithm, is also proposed [25], which uses the overlap of the Laplacian matrix of a given network to realize community recognition.

In this chapter, we introduce the degree of node dissimilarity-node proximity. Based on the dissimilarity, a new clustering algorithm, CS-Cluster, is proposed. This algorithm comprehensively considers the topological structure and semantic information of the node to compare complete clustering process with high efficiency. Compared with other algorithms, this algorithm has the following characteristics:

(1) Use the structural and semantic characteristics of the node to calculate the proximity of the node;

(2) According to the different connection modes of nodes, the contribution degree and the matching degree are, respectively, added to participate in the calculation of the structural dissimilarity;

(3) The algorithm can automatically determine the initial clustering center point, which improves the accuracy of clustering and reduces the inaccuracy caused by human subjective judgment, because the selection of different initial cluster centers will have a greater impact on the final clustering results.

In the network graph to be explored in this chapter, the number of semantic types contained in nodes is constant,

and it is impossible for the number of semantic types to be inconsistent. An undirected weighted graph with semantic information can be represented $G = (V, E, W, \Gamma)$, which represents$V = \{v_1, v_2, \ldots, v_n\}$ as the collection of all nodes in the graph; E represents the set of all connected edges in the graph. Each edge in the edge set E of the network corresponds to two nodes in the node set V in the graph $E = \{e_1, e_2, \ldots, e_m\}$, that is, $\{e_1, e_2, \ldots, e_m\} \subseteq V \times V; W$, the weight set of the undirected edges in the graph, The weights of the connecting edges of nodes v_m and nodesv_n represent the set of semantic information (attributes) of the nodes $\omega_{mn} > 0; \Gamma = \{s_1, s_2, \ldots, s_n\}$ in the graph; the value set of the node v_m on the semantic information Γ can be expressed as$\{s_1(v_m), s_2(v_m), \ldots, s_k(v_m)\}$. The total number of nodes $v_k$directly connected to the node is called the degree of this node $d(v_k)$.

The goal of graph clustering algorithm is to divide a large graph into several closely connected and disjoint subgraphs (i.e., communities in the following) through the clustering process and should satisfy the following:

(1) The internal nodes of the same community are in the structure the above is closely connected, and the object connection between different communities is relatively sparse;

(2) The semantic features of the nodes in the same community are similar, but in different communities, the semantic information of the nodes is different.

3. Calculation and Replacement of Community Test Algorithm

3.1. Calculation of Node Proximity

3.1.1. Relevance and Matching. This paper proposes a community detection algorithm based on the proximity of nodes. The closeness of the relationship between nodes depends on the correlation between nodes. It shows the characteristics of information transfer between nodes. For interconnected nodes, there may be multiple paths between the initial node and the target node. The IGC-CSM algorithm is proposed and considers the selection of weighted connecting edges [26]. The path with the smallest weight participates in the calculation of structural dissimilarity instead of all paths participating in the calculation, which reduces the amount of calculation. The algorithm proposed in this chapter introduces the correlation degree and the matching degree on the basis of the IGC-CSM algorithm to complete the calculation of structural dissimilarity.

Definition 1. (degree of association). In an undirected network graph with weights, a pair of directly connected nodes isvm and$v$. The degree of connection intimacy is defined as the degree of association. Then, the calculation formula for the degree of association between the node vmandvn the node is $H(vm, vn)$ as follows:

$$T(v_m, v_n) = 1 - e^{-\left(\left(\omega_{mn}/\sum_{k-1}^{d(v_m)}\omega_{ma_k}\right) * d(v_m)\right)}, \quad v_m \leftrightarrow v_n,$$

$$H(v_m, v_n) = \frac{T(v_m, v_n) + T(v_n, v_m)}{2}, \tag{1}$$

which defined $T(v_m, v_n)$ as the correlation coefficient between v_m nodes v_n. It indicates the degree to which the node is associated with v_m the node v_n:$d(v_m)$represents the degree of the node v_m, and ω_{mn} represents the weight of the edge connecting the node v_m and the node v_n. It represents the sum of the weights of $\sum_{k=1}^{d(v_m)}\omega_{ma_k}$, all the edges connected to the nodev_m; because the degrees $v_m$of the nodes $v_n$are different, the weights on the respective connected edges are not the same, so $T(v_m, v_n) \neq T(v_n, v_m)$. For example, in the DBLP network, an author may be associated with many authors. These authors are other authors who are associated with them, and the correlation coefficients of these authors are therefore different. The larger the value $H(v_m, v_n)$, the higher the degree of association between the two nodes, and the closer the relationship between the two nodes. Suppose that there are two directly connected nodes v_m and v_n, $d(v_m)$ and $d(v_n)$ mean $v_m$sum, respectively; v_n, w_{mm} represent the node v_m and the node v_n, the weight on the connecting edge. Then, the attraction factors v_m and v_n of the sum $f(v_m, v_n)$ are calculated as follows:

$$f(v_m, v_n) = \ln\left(1 + \frac{d(v_m)}{\sum_{j=1}^{d(v_m)} w_{mj}} * w_{mn}\right), \quad v_m \leftrightarrow v_n. \tag{2}$$

Definition 2. (matching path). A pair of indirectly connected nodes v_m, from the source node through several intermediate nodes v_n to the end point, a path passed is called a matching path $OP(v_m, v_k, \ldots, v_n)$, and all nodes pass only once on the path.

$$R(v_m, v_n) = \sum_{i-\text{srcnode}}^{\text{desnode}} f(v_i, v_{i+1}), \quad v_m \Theta v_n. \tag{3}$$

which represents *sno* as the initial node, *dno* represents the end node, and $T(v_i, v_j)$ represents the correlation coefficient between the two directly connected nodes. In the undirected weighted network graph, there may be multiple matching paths in the path from the initial node v_m to the end point v_n. We choose the path with a high degree of matching as the best path from the initial node v_m to the end point v_n.

3.1.2. Structural Dissimilarity. The calculation method of the adjacency between nodes maps the relationship between the nodes. An undirected weighted network graph is composed of the structural characteristics and semantic information of the nodes, so when calculating the proximity of the nodes, it should be composed of the structural characteristics and semantic characteristics of the nodes., And add the concept of relevance and matching in the calculation process of structural dissimilarity. We decided to

use the Jaccard correlation coefficient to calculate the basic structural dissimilarity. The formula is as follows:

$$D(A,B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|},$$
$$\text{sim}(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}. \tag{4}$$

The IGC-CSM algorithm uses the weighted Jaccard correlation coefficient for calculation. Due to the differences in the connection methods between nodes, the corresponding structural dissimilarity calculation methods in the IGC-CSM algorithm are also different. The specific calculation formula is as follows:

(1) When two nodesv_m are directly connectedv_n, the calculation method is expressed as

$$D(v_m, v_n)_{\text{struct}} = \frac{\omega_{mn}}{\sum_{c=1}^{dc\ g} \omega_{mc} + \sum_{c=1}^{dc\ g(v_n)} \omega_{nc} - \omega_{mn}}, \quad v_m \leftrightarrow v_n. \tag{5}$$

(2) For the sum of two indirectly connected nodes v_m and v_n, the IGC-CSM algorithm is calculated in the form of the product of the directly connected node sequence pair. The specific calculation method is as follows:

$$D(v_m, v_n)_{\text{struct}} = \prod_{sno}^{dn\ o} D(v_j, v_k), \quad v_j \in V \text{ and } v_m \Theta v_n. \tag{6}$$

And its characteristic inspection method:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j). \tag{7}$$

Among them, (v_j, v_k) represents a pair of directly connected nodes. Since there are multiple paths in the process from the initial node v_m to the target node v_n, IGC-CSM selects the path with the least weight and does not take all paths into the calculation, which greatly simplifies the amount of calculation increased, and the efficiency of calculation is improved.

Among them, m is the total number of edges in the network, k_i and k_j the degree of the node i and j respectively, C_i and C_J the node and the community, respectively; if $C_i = C_j$, then $\delta(C_i, C_j) = 1$, and the value is 0. In a specific network structure, the result of the division of the community is not the same, so the obtained modularity values are also different; if the modularity is larger, the algorithm partition is larger.

(3) When the two nodes are relatively independent, the structural dissimilarity is 0.

After the concepts of relevance and matching are introduced, the structural dissimilarity and semantic dissimilarity of the nodes in the graph can be calculated. The calculation method is based on the IGC-CSM algorithm to add the correlation degree and matching degree of the node. The calculation of structural dissimilarity can be divided into three cases according to the connection mode of the nodes: (1) For two directly connected nodes, the dissimilarity calculation is completed by combining the weighted Jaccard correlation coefficient and the correlation between the nodes. (2) For two indirectly connected nodes, the structural dissimilarity and the matching degree between the nodes in the IGC-CSM algorithm are used to complete the calculation of the structural dissimilarity. In the path between the node and the indirect connection, there may be more matching paths; we choose a path with a high degree of matching as the best path between the initial node and the end to participate in the calculation. (3) For unconnected nodes (independent nodes), the structural dissimilarity in the IGC-CSM algorithm is 0. The calculation formula is as follows:

$$D(v_m, v_n)_{str} = \begin{cases} D(v_{mv}, v_n)_{con} + H(v_{m,}, v_n), & v_m \leftrightarrow v_n, \\ D(v_m, v_n)_{in\ di\ r} + R(v_m, v_n), & v_m \Theta v_n, \\ 0, & v_m \otimes v_n. \end{cases} \tag{8}$$

Among them, $D(v_m, v_n)_{\text{sem}}$ means the degree of structural dissimilarity, $\longrightarrow$ means that two nodes are directly connected, and Θ means that the source nodes v_m and v_n are indirectly connected to the end point, which means that the node has no connected edges and belongs to an independent node.

The structural similarity calculation of indirect nodes is realized through the product of the node sequence on the path, and the calculation method is

$$\text{sim}(v_m, v_n)_{\text{indirconn}} = \prod_{i-\text{srcnode}}^{\text{desnode}} \text{sim}(v_i, v_{i+1}) \text{connected}, \quad v_m \Theta v_n, v_i \in V. \tag{9}$$

Among them, (v_i, v_{i+1}) is the sequence pair meaning two points v_m and v_n directly connected. There may be multiple paths from the source node to the node, but IGC-CSM uses the shortest weighted path strategy to avoid a large number of calculation processes.

In the process of graph clustering according to the dissimilarity of nodes, if only the structural features between nodes are concerned, the result of the division may be inaccurate, because, in real life, a node may carry multiple semantic information. For example, in a social network, a

user can perform different roles in different places. In addition, it may also include attributes such as occupations and hobbies. Therefore, when we compare node dissimilarity, we should also consider their semantic dissimilarity. A node contains multiple semantic information, and each semantic may take different values. We use the method of calculating dissimilarity in the K-Modes algorithm as the semantic similarity between nodes. For the convenience of calculation, we set the semantic weight to 1.

$$D(v_m, v_n)_{\text{sem}} = \begin{cases} \sum_{i=1}^{j} \dfrac{\text{com}(v_m, v_n, \text{sem}) * \omega_{\text{sem}}}{j}, & v_m \leftrightarrow v_n \text{ or } v_m \otimes v_n, \\ \sum_{i=\text{sno}}^{\text{dno}} D(v_i, v_j)_{\text{sem}}, & v_m \otimes v_n, \end{cases}$$

$$\text{com}(v_m, v_n, \text{sem}_l) = \begin{cases} 1, & \text{the node } v_m \text{ and } v_n \text{ the } l \text{ semantic value are equal.} \\ 0, & \end{cases} \tag{10}$$

Table 1 shows the meaning of symbols in complex network structure.where $D(v_m, v_n)_{\text{sem}}$ represents the degree of semantic dissimilarity, and ω_{sem} is the semantic weight; $j$represents the number of semantics; when the semantic value is the same, the value is 1; when the semantic value is different, the value is 0; so the semantic value range is [0.1].

3.1.3. Node Proximity. The calculation of similarity between nodes should comprehensively consider the structural characteristics of nodes and the semantic characteristics between nodes, so we draw the proximity of nodes to complete the calculation of dissimilarity between nodes. The formula is shown in (13):

$$D(v_m, v_n) = \lambda\, D(v_{m,} v_n)_{\text{str}} + (1 - \lambda) D(v_m, v_n)_{\text{sem}}. \tag{11}$$

Among them, λ is the balance factor, which aims to adjust the proportion of structure and semantics in the calculation of dissimilarity. Its value range is [0, 1]; $D(v_m, v_n)_{\text{str}}$ represents the structural dissimilarity; $D(v_m, v_n)_{\text{sem}}$ represents the semantic dissimilarity. In the later part of the experiment, this chapter will give the most suitable value λ.

When the value $y_i$is larger, it will be more likely to become the cluster center; when the noncluster center transitions to the cluster center, the value will undergo a jump change, so we define a transfer ϕ_i function to indicate that the jump has found that the specific calculation formula of the clustering center point in the process is as follows:

$$\phi_i = \left|\gamma_{i-1} + \gamma_{i+1} - 2\gamma_i\right|. \tag{12}$$

It can be seen from the formula that the greater the value ϕ_i, the greater the change in the value of the node γ_i; then, the node ϕ_i should become the initial cluster center, so the results are sorted in descending order, and the point with the largest value z is taken as ϕ_i the initial cluster center point.

After calculating the node proximity, we use the node distance to complete the clustering process. Since the distance is inversely proportional to the proximity, that is, the greater the proximity, the closer the relationship between the nodes, and the smaller the corresponding distance value, so we use the inverse of the proximity to calculate the distance. The calculation formula is as follows:

$$\text{distance}(v_m, v_n) = \frac{1}{D(v_m, v_n)}. \tag{13}$$

According to the algorithm framework in this chapter to achieve clustering, the calculation formula is as follows:

$$f(w, s) = \sum_{l=1}^{k} \sum_{j=1}^{n} \omega_{lj}\, \text{distance}(s_l, v_j), \quad 1 \le l \le k \text{ and } 1 \le j \le n, \tag{14}$$

which represents the adjacency matrix of order. At that time, the representative node belongs to the first subcommunity; at that time, it means that the node does not belong to the community. It represents a collection of cluster center points.

The selection of the cluster center point should meet the following two criteria:

(A) The local density of the center point should be large enough

(B) The centers of different subassociations are far apart from each other

Regarding the local density definition of the node, the calculation formula can be expressed as

$$\partial_i = \sum_{v_m \in V/(v_n)} e^{-(\text{distance}(v_m, v_n)/d_c)}. \tag{15}$$

Among them, d_c is the cutoff distance, v_m is its value and is the average value of the distances from other nodes. $V/\{v_n\}$ represents other nodes except for v_n. It can be seen from the definition that the greater the number of nodes v_m whose distance to the node d_c is smaller, the larger the value ∂_i.

3.2. Iterative Update of the Algorithm. The iterative update steps of the CS-Cluster algorithm are as follows, in which $s_{(l)}$, $w_{(l)}$, respectively, represent the cluster center point and adjacency matrix at the first update.

Table 1: Common symbols in this chapter.

Symbol	Meaning
$\longrightarrow$	Represents two nodes directly connected
Θ	Indicates that the source node is indirectly connected to the end point
$\otimes$	Indicates that the node has no connected edges
ω_{sem}	Represents semantic weight
j	Represents the number of semantics
$D(v_m, v_n)_{\text{str}}$	Indicates the degree of structural dissimilarity
$D(v_m, v_n)_{\text{sem}}$	Represents the degree of semantic dissimilarity

(1) The point z with $\phi_i$the largest previous values$_{(l)}$ is calculated by formula (8) as the initial cluster center point$w_{(l)}$, and the adjacency matrix is calculated$f(w_{(l)}, s_{(l)})$ according to formula (9)

(2) When w does not change, calculates$_{(l+1)}$ and obtain the minimum value$f(w_{(l)} \cdot s_{(l+1)})$ according to formula (12); if it is$f(w_{(l)}, s_{(l+1)}) = f(w_{(l)}, s_{(l)})$, the algorithm iteration ends; otherwise, perform step (3)

(3) When s does not change, calculate $w_{(l+1)}$ and obtain $f(w_{(l+1)} \cdot s_{(l)})$, the minimum value according to formula (9). If it is$f(w_{(l+1)}.s_{(l)}) = f(w_{(l)}, s_{(l)})$, the algorithm update ends; otherwise, proceed to step (2)

When s does not change, the update method$w$ is as shown in formula (9), where the value range is $[1, z]$ and the value range is $[1, n]$.

$$w_{(lj)} = \begin{cases} 1, & \text{distance}\left(s_l, v_j\right) \leq \text{distance}\left(s_{\phi_i}, v_j\right), \\ o, & \text{Others.} \end{cases} \tag{16}$$

When w is not changed, the update method s is as follows:

$$s_l = v_l. \tag{17}$$

During the execution of the algorithm, the cluster center point will be updated after each iteration, and finally the remaining nodes will be divided into the nearest category with higher density from each cluster center.

The pseudocode of the clustering algorithm CS-Cluster mentioned in this chapter is shown in Algorithm 1:

where the balance factor$\lambda = 0.6$. In this paper, different values of balance factor are applied to the data set to observe the influence of balance factor on the density and entropy. When the balance factor is 0.6, there are higher density value and lower entropy value. See Section 4.3 for details.

4. Experimental Verification

In this experiment, five algorithms are used as comparison algorithms, namely, IGC-CSM, SA-Cluster, W-Cluster, S-Cluster, and CS-Cluster. Among them, IGC-CSM, SA-Cluster, and W-Cluster are three algorithms in the cluster. The topological structure of the node and the semantic information of the node are comprehensively considered in the class process. The IGC-CSM algorithm uses the K-Medoids framework to achieve clustering, and the SA-Cluster uses random walk technology. The iteration process is very time-consuming. The W-Cluster algorithm passes the weight function used to calculate the degree of dissimilarity between nodes. The S-Cluster algorithm also uses random walk technology, but the algorithm only considers the structural characteristics of the node. The CS-Cluster algorithm is the algorithm proposed in this chapter. The concept of correlation and matching is given in the algorithm to complete the calculation of the structural dissimilarity between the nodes in the graph and integrate the structural dissimilarity of the nodes; the degree of semantic dissimilarity calculates the proximity of nodes, redefines the selection of the initial clustering center point, and finally realizes the division of communities.

4.1. Data Set. This chapter uses two classic data sets to verify the effectiveness of the algorithm.

(1) Political blogs data set: this data set is composed of links between different blogs. For each blog contained in the data set, there is a semantic meaning to describe political orientation, with 0 for members of the Liberal Party and 1 for the Conservative Party member. The data on political orientation comes from blog directories.

(2) DBLP data set: we used one of the sub-data-sets-author collaboration network; we created a graph that reflects the cooperative relationship between authors. In addition, we added two phases to each node included in the network graph.

Related semantic information: the information is reflected by the number of evaluation indicators and themes of the papers. There are three optional values for the number of papers, which are less than 10, between 10 and 20, and greater than or equal to 20, to determine whether the author is "prolific." The other semantic is the topic. For topic attributes, we obtained the paper titles of the selected authors and organized these paper titles into documents. Finally, we extracted research topics containing 100 possible values from the documents, which were randomly assigned the keywords that represent each topic. The theme attributes of each author will correspond to these 100 themes in Table 2.

4.2. Evaluation Index. The comparison algorithm we used evaluates the results of clustering based on density and entropy, so this chapter also uses density and entropy to measure the effect of clustering.

(1) Density: It is defined as the ratio of the sum of edges in the subnetwork to the edges in the entire network.

Input: an undirected weighted graph containing attributes G, Number of clusters$z$;
Output: get a category $M_1, M_2, \ldots, M_z$;
Step 1: initialize the distance of each node (v_m, v_n) in the graph distance $[V_m][V_n] = 0$
Step 2: cluster center point centroid = 0, Semantic extremum $\omega_{\text{sem}_1}, \omega_{\text{sem}_2}, \ldots, \omega_{\text{sem}_l} = 1$, Balance factor $\lambda = 0.6$
Step 3: each node included in the set of all nodes V in the For graph
Step 4: each node v_n and $v_m \neq v_n$ contained in the set of all nodes in the For graph
Step 5: calculate the proximity of nodes according to formula (8)
Step 6: calculate the node distance according to formula (8)

Algorithm 1: The pseudocode of the clustering algorithm.

The higher the density, and the clustering of communities, the better the division effect for the clustering of communities, and the formula is as follows:

$$\text{Density}\left(\{V_c\}_{c=1}^{k}\right) = \sum_{c=1}^{k} \frac{\left|\left\{\left(v_{m,} v_n\right)\right\}\right| v_m, v_n \in V_c, \left(v_m, v_n\right) \in E}{|E|}, \tag{18}$$

where $E = \{e_1, e_2, \ldots, e_m\}$ represents the set of connected edges of all nodes in the graph, and m represents the total number of edges.

(2) Entropy: It is a measure of the semantic similarity between nodes, when the semantic similarity between nodes is a category.

The higher the value, the smaller the entropy value in the subnetwork, and the better the result of community division. Its definition is shown in formulas (19) and (20):

$$\text{Entropy}\left(\{V\}_{cc=1}^{k}\right) = \sum_{i=1}^{m} \left(\frac{w_i}{\sum_{s=1}^{m} w_s} \sum_{c=1}^{k} \frac{|V_C|}{|V|}\right) \text{entropy}, \tag{19}$$

$$\text{entropy}\left(a_i, V_c\right) = -\sum_{n=1}^{n_j} \text{prcnt}_{icn} \log\left(\text{prcnt}_{icn}\right). \tag{20}$$

Among them, $i = \{1, 2, \ldots, m\}$ and $s = \{1, 2, \ldots, m\}$ represent the number of semantics, and w represents the semantic weight; $|V_c|$ represents the percentage of nodes whose semantic values are in the divided subgraph; and V_C represents the number of semantic values $n = \{1, 2, \ldots, n\}$.

4.3. Experimental Results. Figure 1 is on the Political Blogs data set. We show the comparison results of the density values of the five algorithms when the number of clusters is 3, 5, 7, and 9, respectively. The density value of CS-Cluster changes slowly along with the increase in the number of clusters and has always been higher than the values of the other four comparison algorithms. When $Z = 7$, its density value is the highest 0.93, indicating that the CS-Cluster has reached the structural dissimilarity among nodes in the same subgraph in the cluster division and is the best. Among the five algorithms, the density value of W-Cluster is the lowest among the five algorithms, indicating that the structural dissimilarity of nodes in the same subgraph is the worst. As the number of clusters increases, the density value of IGC-CSM remains basically unchanged, and the density value is relatively high, so the clustering effect of IGC-CSM is second only to CS-Cluster.

In Figure 2, we show the comparison of the density values of the five algorithms in the DBLP data set. As the number of clusters continues to increase, the density value of the W-Cluster algorithm changes more obviously, and compared with the other four comparison algorithms, its density value reaches the lowest value. Therefore, the detection results obtained by the W-Cluster algorithm are used. The structural dissimilarity of nodes in the same subgraph is the worst; IGC-CSM and CS-Cluster have roughly the same density values at $Z = 50$, and the density values are all above 0.85; but the number of clusters is higher than 50 when the density value of CS-Cluster is slightly higher than that of IGC-CSM; the density value of SA-Cluster is in the middle of all algorithms; from the overall situation, the average density of CS-Cluster algorithm is higher than the values of the other four algorithms, which shows that, compared with other comparison algorithms, CS-Cluster has the best structural dissimilarity of the nodes in the subgraph in the clustering results obtained.

In Figure 3, we show the comparison of the entropy values of the five algorithms used on the Political Blogs data set. The number of clusters set in the experiment is $z = 3, 5, 7$, and 9. From the experimental figure, we can see that the entropy value of the CS-Cluster algorithm is always 0; as the number of clusters continues to increase, the entropy value of IGC-CSM is almost stable at about 0.1 and remains unchanged, so it can explain the two algorithms of IGC-CSM and CS-Cluster. It can accurately classify the semantically similar nodes into the same category; but from the comparison of the density values in Figure 3, it can be concluded that the average density of CS-Cluster is slightly higher than that of the IGC-CSM algorithm, so CS-clustering quality of the Cluster algorithm has priority over IGC-CSM. When the z value increases from 3 to 7, the entropy value of SA-Cluster basically remains unchanged, and the value remains at about 0.1, but when the z value increases to At 9 o'clock, the entropy value of SA-Cluster suddenly increased, indicating that when the number of clusters is set to 9, the semantic dissimilarity of nodes in the same subnetwork is poor. Among all the comparison algorithms, the

Table 2: Data subset meeting of DBLP.

Research areas	Meeting name
Database	SIGMOD, VLDB, PODS, ICDE, EDBT
Data mining	KDD, ICDE, SDM, PAKDD, PKDD
Information retrieval	SIGIR, CIKM, ECIR, WWW
Artificial intelligence	IJCAI, AAAI, VAI, NIPS

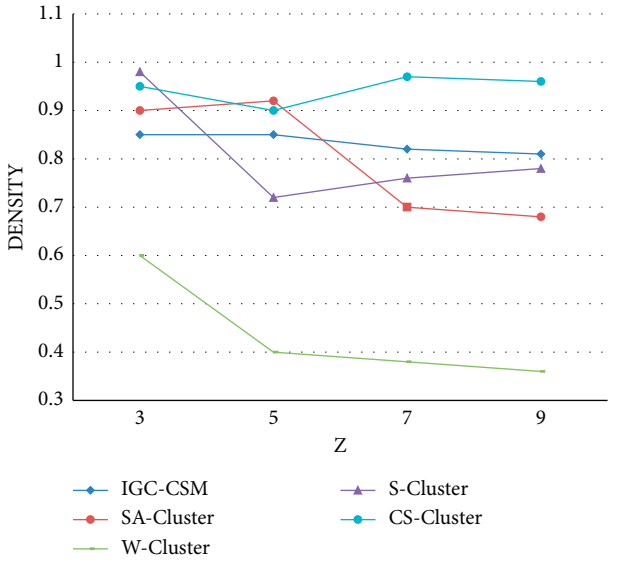


Figure 1: Density analysis on political blogs dataset.

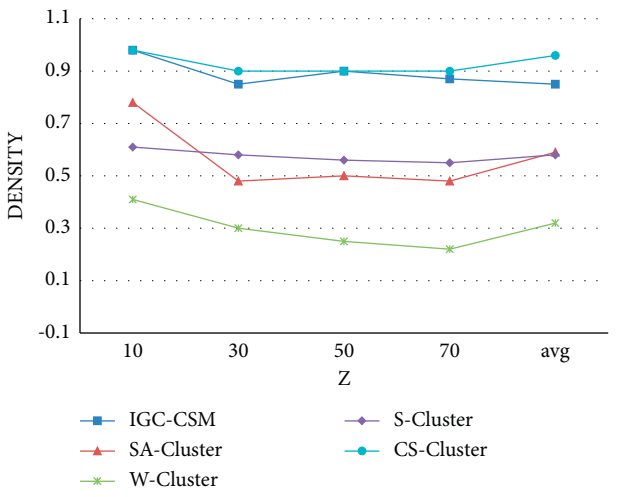


Figure 2: Density analysis on DBLP data.

S-Cluster algorithm has the highest entropy value, indicating that the subgraphs of the clustering division obtained by the S-Cluster algorithm have the worst semantic dissimilarity between nodes.

Figure 4 is the comparison result of the entropy of the four algorithms on the DBLP data set. We set the number of clusters to $z = 10, 30, 50, 70$. Among the five comparison algorithms, the W-Cluster algorithm has the lowest entropy, and the average value remains at 0.3; it means that, in the

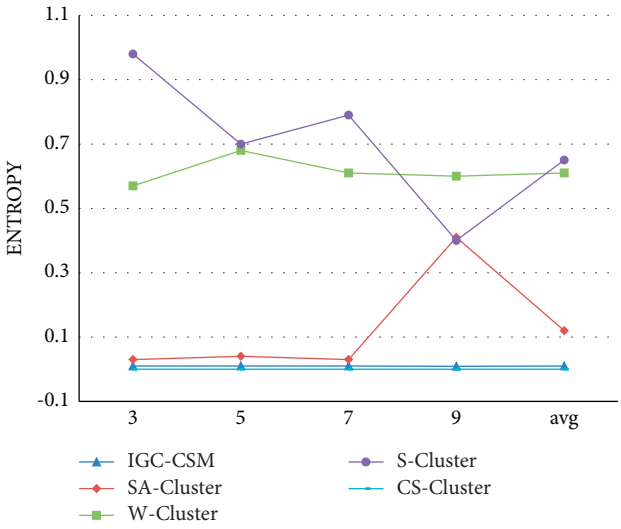


Figure 3: Entropy analysis on political blogs dataset.

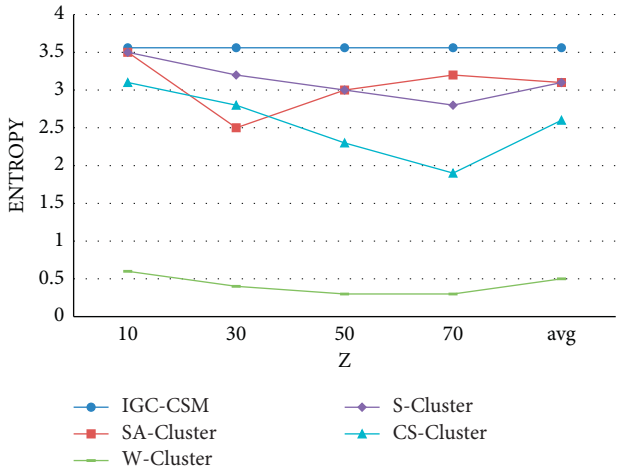


Figure 4: Entropy analysis on DBLP data set.

clustering results obtained, the semantic dissimilarity of the nodes in the subgraph is very high, but in Figure 2, we can see that the density value of W-Cluster is also relatively low. Therefore, it can be judged that the structural dissimilarity between nodes is poor; that is, the connection between nodes is relatively distant. The entropy of the IGC-CSM algorithm is the highest. As the number of clusters increases, the entropy value of the CS-Cluster algorithm shows a downward trend. When the number of clusters is higher than 30, compared with SA-Cluster and CS-Cluster, in the clustering results obtained, the semantic dissimilarity between nodes is relatively high, and the accuracy of the classification results is relatively high.

Figures 5 and 6 show the influence of the balance factor λ on the density and entropy on the Political Blogs dataset when the number of clusters z is 15. It can be seen from Figure 5 that, with the increase of x, the density as a whole shows a trend of gradual increase first, then decrease, and then increase. The density value in the range of λ from 0.5 to

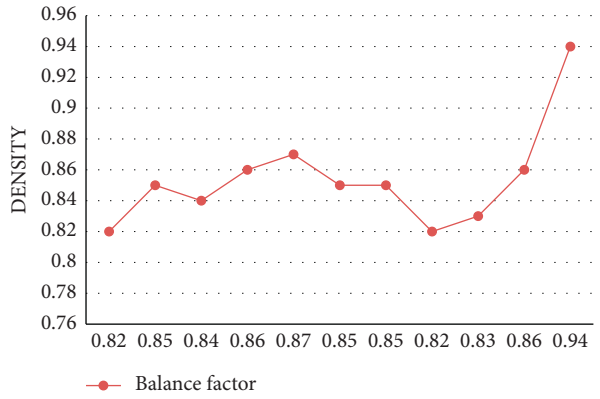


Figure 5: The influence of balance factor λ on density.

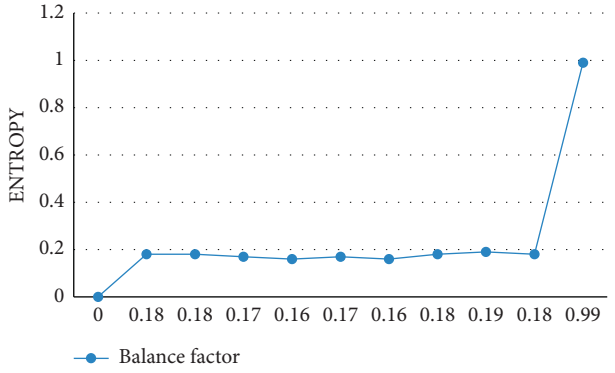


Figure 6: The influence of balance factor λ on entropy.

0.6 is almost unchanged. It can be seen from Figure 6 that the entropy value slowly decreases during the process of the value of a from 0.5 to 0.6, so the value of λ is 0.6 is more appropriate.

5. Conclusion

This article first summarizes the conventional community clustering algorithm and believes that the structural characteristics and semantic information between nodes should be considered comprehensively in the clustering process. Based on this, this article proposes node proximity to complete the calculation of dissimilarity between nodes, introduces associations of the concepts of degree and matching degree, and completes the calculation of structural dissimilarity between nodes. Afterwards, the method of selecting the initial clustering center point was redefined. This method avoids the drawbacks caused by human judgment and improves the accuracy of clustering. Finally, the CS-Cluster algorithm uses the K-Medoids framework to achieve community division. In other algorithm ideas, the experimental comparison on two practical and effective data sets shows that the algorithm proposed in this paper has achieved good clustering results. Data mining and community detection in complex network are a very meaningful research topic. This paper conducts relevant research on the characteristics of nodes in the network and the community detection in the complex network. However, there are still many problems that need to be improved. In the community detection algorithm based on node proximity, this paper only studies the undirected weighted network graph and ignores the clustering of nodes in directed network graph. Secondly, how to effectively select semantic categories in the case of different semantic weight values is also not discussed in this study, which can be further discussed in the future.

Data Availability

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

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

[1] G. Palla, I. J. Farkas, P. Pollner, I. Derényi, and T. Vicsek, "Directed network modules," *New Journal of Physics*, vol. 9, no. 6, p. 186, 2007.

[2] A. Lancichinetti, S. Fortunato, and J. Kertész, "Detection of overlapping and hierarchical community structures in complex networks," *Acta New Physics*, vol. 11, no. 3, pp. 19–44, 2009.

[3] T. S. Evans and R. Lambiotte, "Line graphs, link partitions, and overlapping communities," *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 80, no. 80, pp. 145–148, 2009.

[4] R. Cazabet, F. Amblard, and C. Hanachi, "Detection of overlapping communities in dynamic social networks," in *Proceedings of the 2010 IEEE Second International Conference on Social Computing (Social-Com)*, pp. 309–314, IEEE, Minneapolis, MN, USA, August 2010.

[5] J. Xie, B. K. Szymanski, and X. Liu, "SLPA: uncovering overlapping communities in social networks via a speaker-listener interaction dynamic process," in *Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops*, IEEE, Vancouver, Canada, December 2012.

[6] S. Gregory, "Searching for overlapping communities in the network through tag propagation," *New Journal of Physics*, vol. 12, no. 10, pp. 2011–2024, 2009.

[7] U. Raghavan, R. Albert, and S. Kumara, "Near-linear time algorithm to detect community structure in large-scale networks," *Physical Review E*, vol. 76, p. 36106, 2007.

[8] B. Saoud and A. Moussaoui, "Community detection in networks based on minimum spanning tree and modularity," *Physica A: Statistical Mechanics and Its Applications*, vol. 460, pp. 230–234, 2016.

[9] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," *Nature*, vol. 435, no. 7043, pp. 814–818, 2005.

[10] J. M. Kumpula, M. Kivela, and K. Kaski, "Sequence algorithm for fast clique penetration," *Physical Review E*, vol. 78, no. 2, pp. 1–7, 2008.

[11] M. Goldberg, S. Kelley, M. Magdon-Ismail, K. Mertsalov, and A. Wallace, "Finding communities in overlapping social networks," in *Proceedings of the Second IEEE Conference on*

Internal Society, pp. 104–113, Minneapolis, MN, USA, August 2010.

[12] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: percolation on random graphs," *Physical Review Letters*, vol. 85, no. 25, pp. 5468–5471, 2000.

[13] A. Lancichinetti and S. Fortunato, "Consensus clustering in complex networks," *Scientific Reports*, vol. 2, no. 13, p. 336, 2012.

[14] J. Kim and T. Wilhelm, "Spanning tree separation reveals community structure in networks," *Physical Review E*, vol. 87, no. 3, p. 32816, 2013.

[15] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, no. 2, p. 26113, 2004.

[16] S. Deng, C. Wang, M. Wang, and Z. Sun, "A gradient boosting decision tree approach for insider trading identification: an empirical model evaluation of China stock market," *Applied Soft Computing*, vol. 83, Article ID 105652, 2019.

[17] P. Pons and M. Latapy, "Computing communities in large networks using random walks," *Journal of Graph Algorithms and Applications*, vol. 10, no. 2, pp. 191–218, 2006.

[18] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying Communities in networks," *Proceedings of the National Academy of Sciences*, vol. 101, no. 9, pp. 2658–2663, 2004.

[19] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 6, pp. 7821–7826, 2002.

[20] M. E. J. Newman, "Goldon Fast algorithm for detecting community structure in networks," *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 69, no. 6, p. 66133, 2003.

[21] D. Saoud and A. Moussaoui, "Community detection in networks based on minimum spanning Applications: S0378437116301996.tree and modularity," *Physica A Statistical Mechanics*, vol. 460, 2016.

[22] U. Brunn, R. Albert, and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," *Physical Review E*, vol. 76, p. 36106, 2007.

[23] B. W. Kernighan and S. Lin, "An efficient heuristic procedure for partitioning graphs," *Bell System Technical Journal*, vol. 49, no. 2, pp. 291–307, 1970.

[24] S. Fortunato, V. Latora, and M. Marchiori, "Method to find community structures based on Information centrality," *Physical Review E, Statistical, nonlinear, and soft matter physics*, vol. 70, no. 5, p. 56104, 2004.

[25] M. Zarei, D. Lzadi, and K. A. Samani, "Detecting overlapping community structure of networks based on vertex-vertex correlations," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 11, p. 11013, 2009.

[26] W. Nawaz, K.-U. Khan, Y.-K. Lee, and S. Lee, "Intra graph clustering using collaborative similarity measure," *Distributed and Parallel Databases*, vol. 33, no. 4, pp. 583–603, 2015.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 4809699, 6 pages
https://doi.org/10.1155/2021/4809699

Research Article

Signal Modulation Identification Based on Deep Learning in FBMC/OQAM Systems

Jing Chen,[1] Jianzhong Guo,[2] Xin Shan,[1] and Dejin Kong[2]

[1]*Electronic Information School, Wuhan University, Wuhan 430072, China*
[2]*School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan 430200, China*

Correspondence should be addressed to Dejin Kong; djkou@wtu.edu.cn

Received 6 August 2021; Accepted 27 September 2021; Published 14 October 2021

Academic Editor: Han Wang

Signal modulation identification (SMI) has always been one of hot issues in filter-bank multicarrier with offset quadrature amplitude modulation (FBMC/OQAM), which is usually implemented by the machine learning-based feature extraction. However, it is difficult for conventional methods to extract the signal feature, resulting in a limited probability of correct classification (PCC). To tackle this problem, we put forward a novel SMI method based on deep learning to identify FBMC/OQAM signals in this paper. It is noted that the block repetition is employed in the FBMC/OQAM system to achieve the imaginary interference cancelation. In the proposed deep learning-based SMI technique, the in-phase and quadrature samples of FBMC/OQAM signals are trained by the convolutional neural network. Subsequently, the dropout layer is designed to prevent overfilling and improve the identification accuracy. To evaluate the proposed scheme, extensive experiments are conducted by employing datasets with different modulations. The results show that the proposed method can achieve better accuracy than conventional methods.

1. Introduction

Filter-bank multicarrier with offset quadrature amplitude modulation (FBMC/OQAM) has been considered as one of the potential physical-layer techniques for future wireless communications [1–4]. Owing to the use of pulse shaping filter with low spectrum sidelobe, FBMC/OQAM offers the high spectrum utilization and good ability of asynchronous transmission [1, 2]. In addition, the cyclic prefix is not required in FBMC/OQAM, leading to a high spectral efficiency. However, in contrast to the classical orthogonal frequency division multiplexing (OFDM), the orthogonality of FBMC/OQAM only holds in the real-valued field. To meet the orthogonality condition, FBMC/OQAM systems transmit real-valued symbols obtained by the real and imaginary parts of complex-valued QAM symbols, and there exist imaginary interferences among the transmitted real-valued symbols, called the intrinsic imaginary interference [3]. The imaginary interference will make a major effect on the algorithms about the parameter estimation, which has to be considered in the algorithm design.

With the rapid development of wireless communications, noncooperative communications will be very common in military and civilian areas. In order to identify different modulations of the received signals, signal modulation identification (SMI) is required in the noncooperation wireless communications [5, 6]. For instance, diverse eavesdropping risks can be identified in wireless links by the SMI technique, which ensure the system integrity [7]. In the electronic countermeasure, the signal modulation type is first required to estimate from the intercepted electromagnetic waves. On this basis, the intercepted signal could be decrypted further [8]. Therefore, SMI has been regarded as one of the most crucial techniques when we design a noncooperative communication system. Developing the accurate SMI method is necessary to identify FBMC/OQAM signals [9] since the SMI design is a big challenge to identify types of FBMC/OQAM signals.

So far, most existing modulation identification approaches are based on feature extraction and machine learning classification. A lot of SMI schemes are developed by combining different machine learning-based classifiers and feature extraction strategies [10]. However, since conventional feature extraction schemes rely on statistics, it is difficult for these methods to extract the signal features of different modulation types. As a result, the classification result is vulnerable to mix up. In addition, machine learning-based SMI schemes will suffer performance bottleneck for the problem associated with big data. That is to say, the probability of correct classification is not good enough, which indicates unpreferred in practical FBMC/OQAM systems. To solve this issue, deep learning (DL) has been deemed to be one of effective techniques to deploy SMI [11]. In [12], an enhanced identification scheme was presented based on deep neural network (DNN). In [13], the genetic programming (GP) was combined with the k-nearest neighbor (KNN) to identify four modulation types accurately.

In this paper, we propose a novel SMI technique based on the convolutional neural network (CNN) to make an identification on FBMC/OQAM signals. It is noted that the block repetition is employed in the FBMC/OQAM system to achieve the imaginary interference cancelation. In the proposed scheme, 3 fully connected layers and 2 convolutional layers are designed. In addition, the dropout layer is designed to lower the interaction neurons of the same layer. To evaluate the proposed scheme, extensive experiments are conducted by employing datasets with different modulations. The results show that the proposed method can achieve better accuracy than conventional methods.

The rest of the paper is organized as follows. The system model of FBMC/OQAM with block repetition is introduced briefly, as well as the model of the deep learning in Section 2. Subsequently, the proposed SMI method is presented in Section 3. Section 4 gives experimental results, followed by the conclusion in Section 5.

2. Model of FBMC/OQAM

Figure 1 depicts the baseband FBMC/OQAM system diagram with M subcarriers. The transmitted signal of each FBMC/OQAM subcarrier gets through a pulse shaping filter. $d_{m,n}$ is the transmitted symbol at the time frequency position (m,n). Denote $g[l]$ as the pulse shaping filter for all subcarriers in FBMC/OQAM, which has the even and symmetric coefficients and exhibits the extremely low spectrum sidelobe [2]. According to Figure 1, the transmitted signal of FBMC/OQAM can be expressed as follows [3, 14]:

$$s[l] = \sum_{m=0}^{M-1} \sum_{n\in\mathbb{Z}} d_{m,n} g\left[l - n\frac{M}{2}\right] e^{((j2\pi ml)/M)} e^{((j\pi(m+n))/2)}. \tag{1}$$

Assume $h[l]$ is the Rician fading channel. The transmitted signal $s[l]$ passes through the channel $h[l]$, and the received signal can be written as

$$r[l] = h[l] * s[l] + \eta[l], \tag{2}$$

where $*$ represents the convolution operator and $\eta[l]$ stands for the additive white Gaussian noise [15].

Then, after FBMC/OQAM demodulation at the receiver, it can obtain the following [16, 17]:

$$\begin{aligned}\widehat{d}_{m,n} &= \sum_{l=-\infty}^{\infty} r[l] g\left[l - n\frac{M}{2}\right] e^{-(j2\pi ml/M)} e^{-(j\pi(m+n)/2)} \\ &= H_{m,n}\left(d_{m,n} + \sum_{(m_0,n_0)\neq(m,n)} d_{m_0,n_0}\zeta_{m,n}^{m_0,n_0}\right) + \eta_{m,n} \\ &= H_{m,n}\left(d_{m,n} + jd_{m,n}^{c}\right) + \eta_{m,n},\end{aligned} \tag{3}$$

where $\zeta_{m,n}^{m_0,n_0}$ is the imaginary interference factor in FBMC/OQAM, which is a pure imaginary value, $d_{m,n}^{c}$ is the imaginary interference to the symbol $d_{m,n}$ [17], and $\eta_{m,n}$ is the noise to the symbol $d_{m,n}$. Note that, although $\eta[l]$ is white Gaussian noise, $\eta_{m,n}$ is not white due to the real-valued orthogonality of FBMC/OQAM systems. $H_{m,n}$ is the channel frequency response at the subcarrier m. For the time-invariant channels, the value of $H_{m,n}$ is not relevant to the time index n.

It has been demonstrated that the imaginary interference factor $\zeta_{m,n}^{m_0,n_0}$ is symmetrical [18]. On this basis of symmetric, the block repetition is designed in FBMC/OQAM to achieve the imaginary interference cancelation [18, 19]. As shown in Figure 2, the frame consists of two parts, i.e., original block and repeated block. $d_{m,n}$ $m \in [0, M-1]$, $n \in [0, N-1]$ are the symbols in the original block, which are obtained from the QAM constellation. $a_{m,n+N}$ are the symbols in the repeated block, and it is noted that $a_{m,n+N} = a_{m,N-1-n}$.

At the receiver, the demodulations of original block and repeated block are, respectively, written as

$$\widehat{d}_{m,n} = H_{m,n}\left(d_{m,n} + jd_{m,n}^{c}\right) + \eta_{m,n}, \quad n = 0, 1, \ldots, N-1, \tag{4}$$

$$\begin{aligned}\widehat{d}_{m,2N-1-n} &= H_{m,2N-1-n}\left(d_{m,2N-1-n} + jd_{m,2N-1-n}^{c}\right) \\ &\quad + \eta_{m,2N-1-n} \\ &= H_{m,n}\left(d_{m,n} + jd_{m,2N-1-n}^{c}\right) \\ &\quad + \eta_{m,2N-1-n}, \quad n = 0, 1, \ldots, N-1.\end{aligned} \tag{5}$$

Note that, the imaginary interferences of the original block and repeated block satisfy the following equation [18], i.e.,

$$jd_{m,n}^{c} + jd_{m,2N-1-n}^{c} = 0. \tag{6}$$

Subsequently, the imaginary interferences can be removed by the following linear combination:

$$\frac{\widehat{d}_{m,n} + \widehat{d}_{m,2N-1-n}}{2} = H_{m,n}d_{m,n} + \frac{\eta_{m,n} + \eta_{m,2N-1-n}}{2}. \tag{7}$$

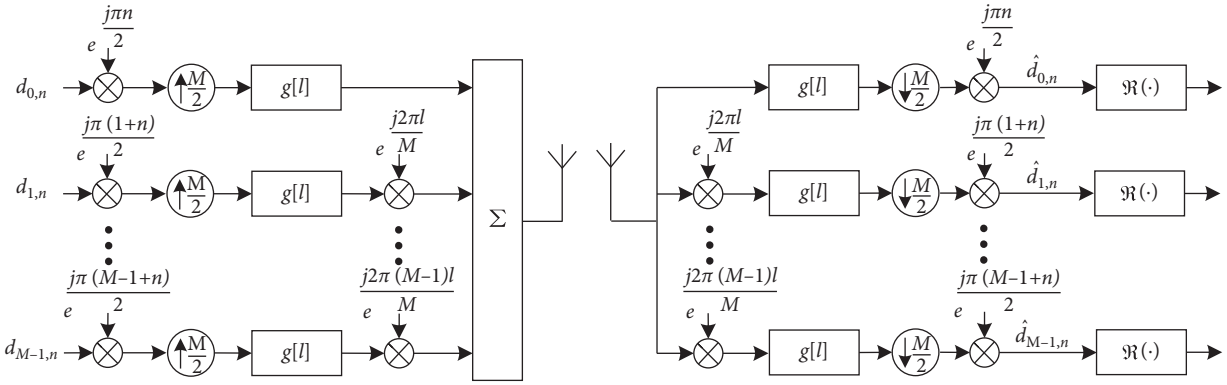


Figure 1: The FBMC/OQAM system diagram.

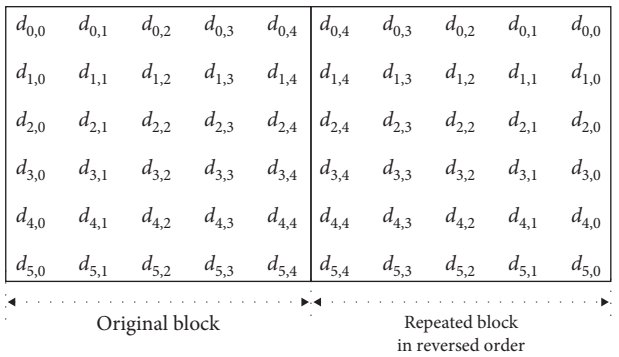


Figure 2: The reversed-order block for FBMC/OQAM.

Let $\overline{d}_{m,m} = (\hat{d}_{m,n} + \hat{d}_{m,2N-1-n})/2$ and $\overline{\eta}_{m,n} = (\eta_{m,n} + \eta_{m,2N-1-n})/2$. Then, equation (7) can be rewritten as

$$\overline{d}_{m,m} = H_{m,n} d_{m,n} + \overline{\eta}_{m,n}. \tag{8}$$

Note that, the noise $\overline{\eta}_{m,n}$ satisfies the white Gaussian distribution [18].

Suppose the Rician fading channel only consists of a line of sight (LOS) between the receiver and the transmitter. Then, the probability of correct classification (PCC) of the Rician distribution is written as

$$f[x] = \frac{x}{\sigma^2} e^{-\left(x^2 + A^2/2\sigma^2\right)} \cdot I_0\left[\frac{Ax}{\sigma^2}\right], \tag{9}$$

where σ^2 stands for the power of the multipath signal component, A represents the amplitude peak of the main signal, $I_0[\cdot]$ stands for the modified Bessel function of the 0-th order. The Rician channel model can be written as

$$h[l] = \sqrt{\frac{\kappa}{\kappa+1}}\sigma e^{j\theta} + \sqrt{\frac{1}{\kappa+1}}\mathbb{N}\left(0, \sigma^2\right), \tag{10}$$

where the first part stands for the mirror path which has the uniform phase θ. The second part represents a majority of scattering paths and reflection paths, which are independent of θ. κ is the Rician factor that indicates the Rician distribution. The Rician factor is defined as

$$\kappa = \frac{Ax}{\sigma^2}. \tag{11}$$

When the factor κ goes to zero gradually, the Rician distribution will become a Rayleigh distribution.

3. Proposed Deep Learning-Based SMI

3.1. CNN Algorithm. After converting to IQ samples, the FBMC/OQAM signal can be used to train the CNN. As a commonly used algorithm, CNN is one of hot research issues in the area of artificial intelligence. The CNN structure mainly includes input layers, convolution layer, pooling layer, fully connected layer, and output layer. The main features of CNN are summarized as follows:

(i) *Local Connection*. For this connection, each neuron is not required to connect the whole neurons from the upper layer but only a small part of neurons, which can reduce the parameters significantly.

(ii) *Weight Sharing*. It not required that each connection corresponds to one weight. Instead, one set of connections has the same weight, which is also beneficial to reduce the number of parameters.

(iii) *Down Sampling*. The number of samples for each layer can be reduced by the pooling layer, which can enhance the model robustness.

In the CNN structure, the convolution layer is the most important part, which can be compared with the convolution operation in calculus. For instance, the convolutional signal between a time-domain signal $x[k]$ and a time-domain signal $w[k]$ can be written as

$$y[k] = \sum_l x[k-l]w[l]. \tag{12}$$

For the two-dimension signals $x[p,q]$ and $w[p,q]$, the convolutional signal can be written as

$$y[p,q] = \sum_k \sum_l x[p-k, q-l]w[k,l]. \tag{13}$$

It is noteworthy that the convolution formula in the CNN algorithm has a difference slightly with the above-mentioned convolution definition. For example, the two-dimensional convolution in CNN is written as

$$y[p,q] = \sum_{k}\sum_{l} x[p+k, q+l]w[k,l]. \tag{14}$$

3.2. Deep Learning-Based SMI Method. In this subsection, the deep learning-based SMI method is proposed via the CNN algorithm, in which 2 convolutional layers and 3 fully connected layers are included, as shown in Figure 3. Specifically, the first one of convolutional layers consists of 128 convolution kernels with the 1×16 dimension matrix for every convolution kernel. As for the second one of convolutional layers, there exist 64 convolution kernels with the 2×8 dimension matrix for every convolution kernel. The neuron numbers of the 3 fully connected layers are 256, 128, and λ, respectively. Note that, λ denotes the number of the modulation modes used in FBMC/OQAM systems. In addition, except the last one of 3 fully connected layers, the activation function of all layers is the parametric rectified linear unit (PReLU), which can effectively reduce the problem of the gradient disappearance during the operation of back-propagation. As for the last one of 3 fully connected layers, softmax is taken to acquire the probability distribution matrix.

In addition, the dropout layer is added in the first 4 layers in the proposed deep learning-based SMI structure to decrease the overfitting effectively. In a circle, some neurons are randomly selected in the neural layer and are hidden temporarily. Subsequently, the training process of the CNN is carried out. In the next circle, some other neurons will be hidden until the training ends. Note that, the proposed scheme can decrease the interaction among neurons, which makes the technique more generalizable.

3.3. Dataset. To verify the proposed method, two datasets are created for the task of SMI. The modulations in dataset θ_1 consist of BPSK, QPSK, 8PSK, and 16QAM, and the modulations in dataset θ_2 include BPSK, QPSK, 8PSK, 16QAM, and 64QAM. For the training and testing of CNN, 40000 data samples are created for each modulation. For example, for a certain signal-to-noise ratio (SNR), there are 160000 data samples that are input into the neural network. Note that, the samples for training take up 70% of the whole samples and the samples for testing take up 30% of the whole samples.

(a) *IQ Samples*. Through the FBMC/OQAM modulation and the channel, the i-th samples can be obtained as

$$S_i = \left[s_0, s_1, \ldots, s_{N_s}\right], \tag{15}$$

where N_s stands for the number of samples and s_l is the value of the l-th sample. Since s_l is a complex-valued number, it can be rewritten as

$$s_l = \mathfrak{R}\{s_l\} + j\mathfrak{T}\{s_l\}, \tag{16}$$

where $\mathfrak{R}\{\cdot\}$ and $\mathfrak{T}\{\cdot\}$ represent the operations of taking real and imaginary parts, respectively. $\mathfrak{R}\{s_l\}$ and $\mathfrak{T}\{s_l\}$ stand for the in-phase component and quadrature component of the signal.

(b) *AP Samples*. The module and phase of s_l can be obtained as

$$A = \sqrt{\mathfrak{R}\{s_l\}^2 + \mathfrak{T}\{s_l\}^2}, \tag{17}$$

$$\theta = \arctan\frac{\mathfrak{T}\{s_l\}}{\mathfrak{R}\{s_l\}}. \tag{18}$$

(c) *Manmade Features*. The higher order cumulants (HOC) feature can be calculated, and the manmade feature can be constituted by combining the instantaneous feature of the signal.

4. Experiment Results

In this section, experiments are conducted to verify the performance of the proposed deep learning-based SMI technique in the noncooperative FBMC/OQAM system. We also give compare the identification accuracy with conventional methods. In the experiments, 256 subcarriers are considered in the FBMC/OQAM system, and only 16 subcarriers are active. In addition, each subcarrier contains 8 symbols. As for the Rician channel, the Rician factor is 20 and the sampling frequency is 10 kHz, suffering from a Doppler frequency offset 400 Hz. In the experiment, 3 existing classification schemes are used for the comparison, which are all machine learning-based methods, i.e., CNN with IQ samples, CNN with AP samples, and deep neural network (DNN) with manmade features extracted by logistic regressive.

Figure 4 depicts the PCC comparison between the proposed scheme and the existing schemes, in which dataset θ_1 is used for training. From the results, the PCC of the proposed scheme is improving gradually, while the existing methods remain unchanged almost. When SNR >20 dB, the accuracy of modulation identification nearly achieves 100% by the proposed deep learning-based SMI technique. In addition, the CNN with AP samples can achieve better PCC accuracy than the DNN method with manmade features which is because that CNN can extract the data features automatically via the convolution kernel.

As shown in Figure 5, the PCCs for different modulation modes are depicted by the CNN with IQ samples. From the results, the BPSK modulation can be always identified correctly at the whole range of the SNR, while the other three modulations suffer the accuracy degradation at the SNR range below 15 dB. When the SNR comes to the high SNR range, the PCCs of all modulations stabilize gradually. Above all, the proposed technique can achieve high PCC accuracy for the signal modulation identification.

In Figures 4 and 5, dataset θ_1 is used for the signal identification to evaluate the proposed SMI technique. However, the performance may be affected by the introduction of a new modulation in practice. To demonstrate the robustness of the proposed scheme, dataset θ_2 is used for the causing the dataset mismatch. Figure 6 shows the PCCs of

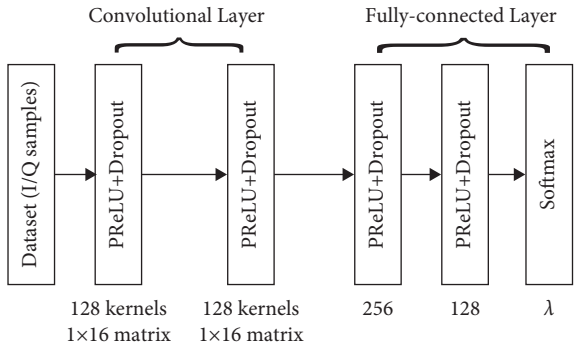


Figure 3: The structure of deep learning-based SMI.

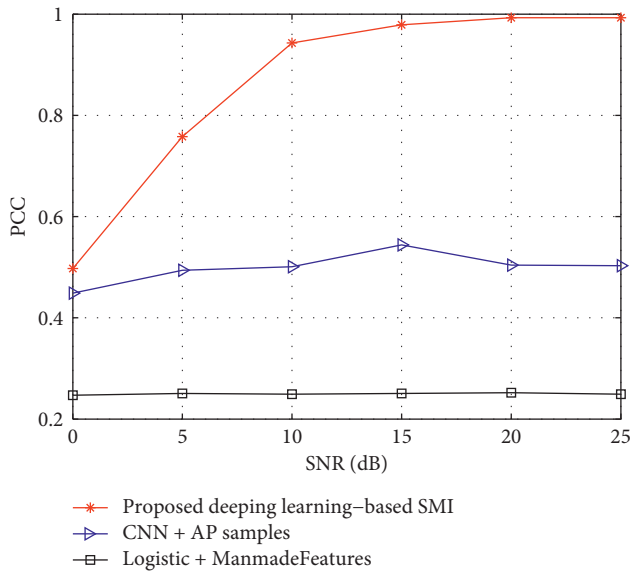


Figure 4: PCC comparison between the proposed scheme and the existing schemes, dataset θ_1.

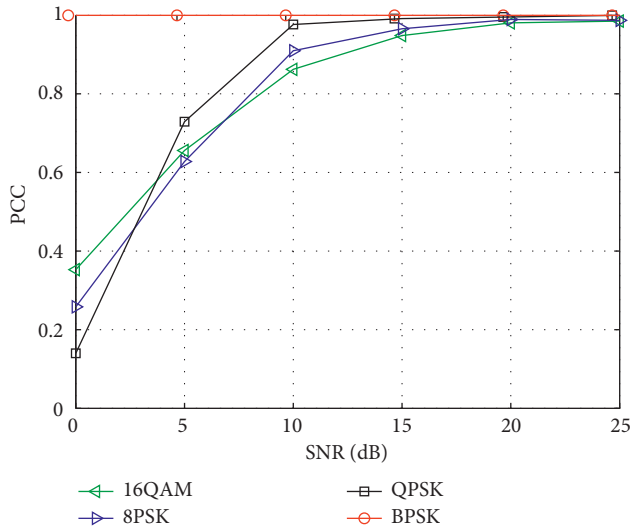


Figure 5: PCC comparison between different modulation types, dataset θ_1.

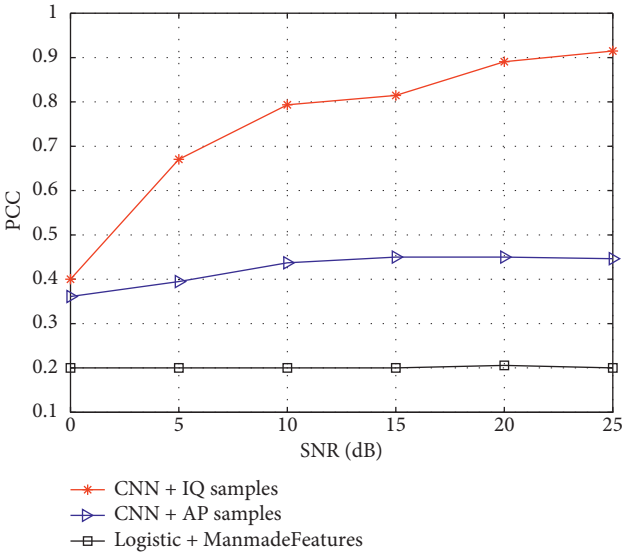


Figure 6: PCC comparison of different methods, dataset θ_2.

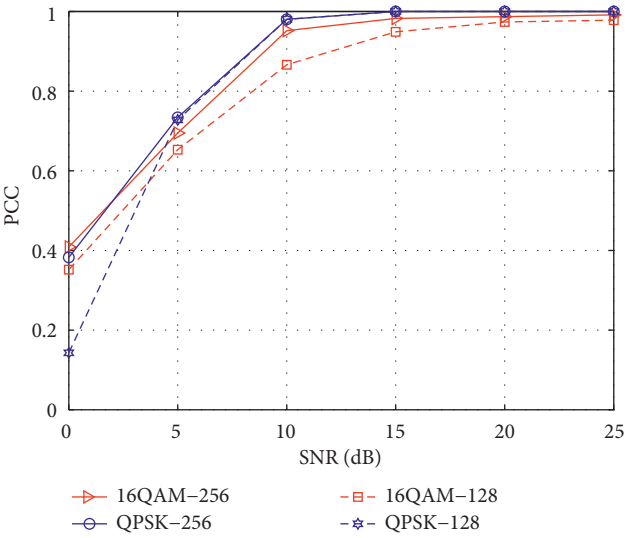


Figure 7: PCC comparison of the proposed scheme under different sample points, dataset θ_1.

the proposed scheme under the dataset mismatch. From the result, the curve trend of Figure 6 is similar to that of Figure 4, and the PCC accuracy remains steady in the proposed scheme, achieving the accuracy of above 90%. The mismatch of the datasets has no damage on the signal modulation identification.

In the following experiments, the sample points of the signal are increased to verify the proposed deep learning-based SMI technique. Figure 7 depicts the PCC comparison of the proposed scheme under different sample points, in which dataset θ_1 is employed. From the results, the proposed scheme with 256 points can achieve higher PCC accuracy than the method with 128 sample points. Especially, the performance difference is more obvious at low SNR. The reason is that, as the sample points increase, more features of

the signal can be extracted by the CNN algorithm, improving the identification accuracy effectively.

5. Conclusions

In this paper, a deep learning-based SMI technique was presented to identify FBMC/OQAM signals in a noncooperative system. To achieve the imaginary interference cancelation, the concept of block repetition was employed in the FBMC/OQAM system to generate signals. In the proposed scheme, 3 fully connected layers and 2 convolutional layers were designed. In addition, the dropout layer was designed to lower the interaction neurons of the same layer. To evaluate the proposed scheme, extensive experiments were conducted by employing datasets with different modulations. The results indicated that the proposed deep learning-based SMI scheme exhibited high accuracy of modulation identification and strong robustness.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was financially supported in part by the Scientific Research Fund of Wuhan Textile University under Grant 20200828.

References

[1] R. Nissel, S. Schwarz, and M. Rupp, "Filter bank multicarrier modulation schemes for future mobile communications," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 8, pp. 1768–1782, 2017.

[2] P. Siohan, C. Siclet, and N. Lacaille, "Analysis and design of OFDM/OQAM systems based on filterbank theory," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1170–1183, 2002.

[3] D. Kong, D. Qu, and T. Jiang, "Time domain channel estimation for OQAM-OFDM systems: algorithms and performance bounds," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 322–330, 2014.

[4] R. Nissel and M. Rupp, "Pruned DFT-spread FBMC: low PAPR, low latency, high spectral efficiency," *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4811–4825, 2018.

[5] F. Wen, Z. Zhang, K. Wang, G. Sheng, and G. Zhang, "Angle estimation and mutual coupling self-calibration for ULA-based bistatic MIMO radar," *Signal Processing*, vol. 144, no. 3, pp. 61–67, 2018.

[6] S. Hong, Y. Zhang, Y. Wang, H. Gu, G. Gui, and H. Sari, "Deep learning-based signal modulation identification in OFDM systems," *IEEE Access*, vol. 7, pp. 114631–114638, 2019.

[7] M. Liu, J. Chen, B. Li, and J. Li, "Fractional frequency offset estimation for OFDM systems in non-cooperative communication," *China Communications*, vol. 13, no. 9, pp. 65–71, 2016.

[8] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8440–8450, 2018.

[9] W. Li, H. Liu, Y. Wang, Z. Li, Y. Jia, and G. Gui, "Deep learning-based classification methods for remote sensing images in urban built-up areas," *IEEE Access*, vol. 7, pp. 36274–36284, 2019.

[10] H. Huang, W. Xia, J. Xiong, J. Yang, G. Zheng, and X. Zhu, "Unsupervised learning based fast beamforming design for downlink MIMO," *IEEE Access*, vol. 7, no. 1, pp. 7599–7605, 2018.

[11] F. Meng, P. Chen, L. Wu, and X. Wang, "Automatic modulation classification: a deep learning enabled approach," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10760–10772, 2018.

[12] W. Xie, S. Hu, C. Yu, P. Zhu, X. Peng, and J. Ouyang, "Deep learning in digital modulation recognition using high order cumulants," *IEEE Access*, vol. 7, pp. 63760–63766, 2019.

[13] M. W. Aslam, Z. Zhu, and A. K. Nandi, "Automatic modulation classification using combination of genetic programming and KNN," *IEEE Transactions on Wireless Communications*, vol. 11, no. 8, pp. 2742–2750, 2012.

[14] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2021.

[15] D. Kong, X.-G. Xia, P. Liu, and Q. Zhu, "MMSE channel estimation for two-port demodulation reference signals in new radio," *Science China Information Sciences*, vol. 64, pp. 169303:1–169303:2, 2021.

[16] C. Lélé, P. Siohan, and R. Legouable, "2 dB better than CP-OFDM with OFDM/OQAM for preamble-based channel estimation," in *Proceedings of the IEEE International Conference on Communication*, pp. 1302–1306, Xiamen, China, May 2008.

[17] C. Lélé, J.-P. Javaudin, R. Legouable, A. Skrzypczak, and P. Siohan, "Channel estimation methods for preamble-based OFDM/OQAM modulations," in *Proceedings of the European Wireless Conference*, pp. 59–64, Delft, Netherlands, March 2007.

[18] D. Kong, X. Zheng, Y. Zhang, and T. Jiang, "Frame repetition: a solution to imaginary interference cancellation in FBMC/OQAM systems," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1259–1273, 2020.

[19] D. Kong, J. Li, K. Luo, and T. Jiang, "Reducing pilot overhead: channel estimation with symbol repetition in MIMO-FBMC systems," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7634–7646, 2020.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 7003252, 14 pages
https://doi.org/10.1155/2021/7003252

Research Article

Edge-Cloud-Assisted Multiuser Forward Secure Searchable Encryption (EMFSSE) Scheme in the P2P Networking Environment

Yongli Tang, Jingran Li, Xixi Yan, and Qiang Zhao

School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, Henan, China

Correspondence should be addressed to Xixi Yan; yanxx@hpu.edu.cn

Received 13 July 2021; Accepted 26 August 2021; Published 20 September 2021

Academic Editor: Han Wang

P2P network enables users to share resources effectively. However, with the advent of the big data era, the sensitive data of users in P2P network are also increasing dramatically. In order to solve the contradiction between the huge amount of sensitive data and the limited local storage space, an increasing number of users choose to encrypt their sensitive data and store them in the cloud server. For the problem of the secure storage and flexible access of large amounts of user data in P2P networks, an edge-cloud-assisted multiuser forward secure searchable encryption scheme is proposed. The scheme uses the proxy reencryption technique to optimize the multiuser searchable encryption and prevent the decryption key from being directly transmitted between users. By introducing an edge-cloud architecture, the system achieves efficient communication and timely response capabilities. The security analysis proves that our scheme achieves the CPA (chosen-plaintext attack) security based on DBDH assumption and the forward privacy. Finally, the theoretical and experimental comparisons between this scheme and other schemes show that our scheme has high efficiency in the process of data update, search, and trapdoor generation. In addition, due to the use of edge-cloud architecture, our scheme reduces about 90% and 75% of the user's consumption in the encryption and token generation process.

1. Introduction

P2P network file sharing system has been widely used and has rapidly developed because of its various advantages, such as decentralization, good extendibility, strong robustness, high cost-effectiveness, and load balancing characteristics [1–3]. With the development of communication and computer technology [4–6], the amount of data owned by users is increasing, but the local storage space of users is limited. This contradiction has become one of the main reasons for restricting the development of P2P networks. The development of cloud storage [7] has brought about a turning point for this situation. Clients can choose to transfer a large amount of local data [8] to a cloud server and use the cloud server to store, manage, and utilize the data. In addition, searchable encryption technology can ensure data security and flexible access during remote storage, which not only reduces local storage and computational cost but also achieves more flexible and convenient data access. However, in reality, there is a huge physical and logical distance between the client and the cloud server, which reduces the efficiency of data collection, transmission, and analysis, causes the delay of the communication process, and hinders the timely response ability of the whole system. This brings new challenges to the combination of file transmission and searchable encryption technology in P2P networking environment. A lot of work has shown [9–11] that edge-cloud collaboration can successfully replace the only-cloud system and strike a balance between powerful computing power and rapid response. The edge-cloud architecture is an optimized cloud computing system that transforms cloud service from the "cloud" as the center to the Internet edge and cloud cooperative operation, which overcomes the problems of only-cloud computing such as distributed data collection, transmission and response delay, and poor data analysis efficiency. With the help of edge computing, clients store some data information on the nearest edge platform, and the edge platform retrieves the data according to authorization, thereby limiting unnecessary data upload to the cloud, saving communication costs, and providing more efficient

and timely data search. In addition, the edge platform can also assist clients in some expensive operations and use the efficiency of edge computing to reduce the amount of calculation on the client side. Therefore, the edge-cloud architecture and the P2P network can be combined. Clients upload part of their data information to the edge-cloud architecture and enjoy the ability to share and use data in time.

The combined use of searchable encryption and edge-cloud collaboration can largely solve the contradiction between data storage, data confidentiality, and data transmission efficiency. It ensures that neither the edge side nor the cloud side can learn any useful information without completely sacrificing searchability of the data or that search function can be improved with minimal disclosure. However, the search server is generally assumed to be an honest but curious entity in a searchable encryption scheme. In the real world, the malicious server will carry out illegal operation on its stored ciphertext. Attackers can even use a small amount of leakage to leak the user's query. Particularly in dynamic databases, the attacker can inject new documents into the cloud server to gain more advantages, and the impact of file injection attacks can even be devastating. The searchable encryption scheme that realizes forward privacy can effectively resist such attacks [12]. Therefore, from the perspective of usage, it is very necessary to construct a forward secure searchable encryption scheme to solve the information security and privacy issues in the process of file updating. How to design a searchable encryption scheme with forward security under the edge-cloud architecture and apply it to P2P networks is the first goal of this paper.

In the P2P network, a node can not only be the data user who sends the request to obtain the required information but also provide the data as the data owner. Specifically, the P2P network aims to share various services among users. But so far, only a few schemes extend forward secure searchable encryption scheme to multiuser [13, 14]. Wang et al. [13] introduce a semihonest proxy server that does not collude with the cloud server to solve the user query problem. Lu et al. [14] improve Wang's scheme, remove the state value transfer between the data owner and the proxy server, and add the user authentication function. However, both schemes require the data owner to send the decryption key to the data user directly. In order to prevent the disclosure of the decryption key, this paper introduces the proxy reencryption technique. The authorized proxy (obtains the reencryption key generated by the data owner) transforms the ciphertext provided by the data owner into another ciphertext that can be decrypted by the data user. Then, the data user can download the ciphertext alone and decrypt it with his own private key to get the original plaintext. How to combine forward secure searchable encryption with proxy reencryption to provide secure file sharing between users is the second goal of this paper.

1.1. Contribution. In order to achieve the two goals mentioned above, we study a series of existing schemes. First of all, users in P2P network upload a large amount of data to the cloud for storage and search. However, the distance between clients and server causes serious operation delay in this process, which affects the performance of end-to-end communication in the P2P network. Inspired by the work [15], edge-cloud architecture can effectively realize timely response between client and edge platform and greatly improve communication efficiency. Secondly, malicious attackers may carry out a variety of attacks to damage the confidentiality of user data by using the information leaked in the communication process, for example, file injection attacks. Therefore, constructing a searchable encryption scheme with forward security is particularly important, which can effectively ensure that the file update process does not disclose data information. Finally, in order to combine P2P network with forward secure searchable encryption technology, it is necessary to implement forward secure searchable encryption technology among multiple users. However, the existing multiuser schemes need to transmit the decryption key directly [13, 14]. In order to prevent the decryption key from leaking, the proxy reencryption technique is introduced to optimize the security of file transmission of each node in the P2P network. In summary, we have constructed a multiuser forward secure searchable encryption scheme EMFSSE based on the edge-cloud architecture in the P2P network. In the scheme, we store part of the index information in the edge server nearest to the user and store encrypted files and encrypted indexes in the cloud server. The communication efficiency of the whole system is also enhanced because it can realize fast update, search, and transmission between the user and the server by making full use of the storage and computing capacity of cloud edge. The main contributions are as follows:

(1) This scheme combines the P2P network with the multiuser forward secure searchable encryption. It realizes the encryption storage, ciphertext search, and update function of user data in the P2P network environment. And it supports data transmission between users.

(2) Proxy reencryption technique is introduced to solve the decryption key leakage problem in the scheme [13, 14]. In our scheme, the edge server reencrypts the ciphertext, and then the data requester uses its own private key to decrypt the reencrypted ciphertext.

(3) We take advantage of edge-cloud collaborative computing to spread out the workload of only-cloud server and reduce the burden of users. Firstly, the edge server maintains part of the keyword information and assists the authorized data user to generate trapdoor. The data owner can authorize the data user to query the encrypted database dynamically and effectively. Secondly, because of the excellent computing power of edge server, it can be used to realize the reencryption process and convert part of the heavy encryption operations to the edge, which can ensure data confidentiality and reduce the computational cost of the client and cloud.

(4) The scheme satisfies the correctness and indistinguishable forward security under the adaptive attack model. And the scheme is proved to be against the chosen-plaintext attacks (CPA) based on the DBDH assumption during the reencryption process.

(5) Comparative analysis and experimental results show that the proposed scheme has a good efficiency in the process of encryption and searching. Compared with the scheme without edge computing, the proposed scheme can save about 92% of the computational costs of the client in the process of ciphertext generation. It can save about 75% of the computation cost in the token generation process.

1.2. Organization. The rest of this paper is organized as follows. In the next section, we introduce the related work. The description of symbols and related knowledge are introduced in Section 3. Section 4 analyzes the design goals and basic model. Then the specific structure and safety analysis of the EMFSSE scheme are introduced in Section 5 and Section 6, respectively. Section 7 compares the EMFSSE scheme with other related schemes and evaluates its performance. In the end, Section 8 summarizes the full text.

2. Related Work

This paper studies the multiuser forward secure searchable encryption technology optimized by proxy reencryption technique in the edge-cloud-assisted P2P network. Therefore, this section focuses on the research progress of forward secure searchable encryption, edge-cloud architecture, and proxy reencryption technique.

2.1. Forward Secure Searchable Encryption (FSSE). The desired goal of forward secure searchable encryption (FSSE) is that the malicious server does not disclose any information of existing data in the database during the file update phase. The scheme of Chang Mitzenmacher [16] already supports forward security. This paper points out that when a file is updated, the newly added file may be related to the previously queried keywords, thereby leaking some information. Stefanov et al. [17] proposed the concept of FSSE and the concept of backward security. Subsequently, the first scheme with good communication costs and computational complexity was proposed by Bost [12], which hides the relationship between document identifiers and keywords through trapdoors replacement in the indexing process and does not disclose the information of previously queried files when new files are injected. Subsequently, scholars have successively proposed various forward secure searchable encryption schemes [18–20] on the basis of the $\sum_{o\phi o\varsigma}$ scheme. The application of the FSSE scheme has also become a research hotspot recently. Chen et al. [21] proposed a blockchain-based forward secure searchable encryption scheme and applied it to vehicle trading websites. The scheme uses smart contract to replace the cloud server for search. Wang et al. [22] proposed a verifiable forward security ranking search scheme in P2P network environment. The scheme realizes ranking search and achieves forward security by changing the file vector structure and using modular residual computation. However, most of these schemes cannot realize file transfer between clients. Wang et al. [13] proposed an FSSE scheme to support multiuser search, which realizes data access by using proxy server to maintain keyword state information. However, this scheme directly transmits decryption key through secure channel, which not only causes the problem of decryption key leakage but also requires the data owner to keep online during the whole search phase.

2.2. Edge-Cloud Architecture. It is an effective way to optimize cloud computing by using cloud-edge collaboration instead of only-cloud system, which has become a research trend in the near future. The edge platform is closer to the underlying data source and has the geographical advantage of being closer to the real clients, so it can effectively reduce the system delay. And edge server can replace the client side or cloud side to perform some complex operations, which can improve the efficiency of the client and cloud server, and provide on-demand services. Ren et al. [23] proposed that the cloud-edge cooperation method can effectively improve the delay performance. Zhang et al. [24] used edge-cloud cooperation to construct a multidevice management service system and applied it to the industrial Internet of things. The system improved the response speed of data collected by basic equipment to a certain extent and reduced the network bandwidth load pressure brought by data transmission. Mollah et al. [25] combined cloud-edge collaboration with public key encryption to construct a secure data sharing scheme, which supports searching on encrypted data (searchable encryption), uses cloud-edge collaboration to reduce computing and communication costs, and provides privacy and confidentiality functions for outsourced data. Wang et al. [15] constructed a lightweight scheme of edge-cloud-assisted public key searchable encryption, using lightweight cryptography and edge-cloud computing to reduce the encryption cost of infrastructure equipment by 70%. But the previous schemes are all applied to the industrial Internet of things, and few schemes combine edge computing with P2P network.

2.3. Proxy Reencryption. Proxy reencryption technique means a trusted third party that converts data encrypted by the data owner with his public key into ciphertext data that can be decrypted by the data user with his private key, so as to complete data sharing. It is important to note that, during this process, no plaintext information should be disclosed to the proxy. In 1998, proxy reencryption technique was proposed by Blaze et al. [26]. With the development of cloud storage, a large number of researchers focus on the combination of proxy reencryption technique and cloud storage to achieve the purpose of cloud data sharing [27–29]. Early proxy reencryption schemes have the problem of reencryption key abuse. In 2005, [30] introduced the concept of nontransferability, but it was not until 2015 that Guo et al. [31] completed the instantiation of this concept using

indistinguishable confusion and unforgeable authentication. On this basis, a public accountability proxy reencryption scheme was proposed by Guo et al. [32] in 2021. Different from the traditional reencryption key generation, the generation process of this scheme is noninteractive, which saves the communication cost, and the scheme can determine the operator of the decryption operation, which improves the security of the scheme.

3. Preliminaries

This section mainly introduces the symbols and encryption-related knowledge used in this paper.

3.1. Notations. The notation information of this paper is given in the Table 1

3.2. Cryptography Materials

3.2.1. Trapdoor One-Way Permutation Technique. We cite the definition of trapdoor one-way permutation in literature [33].

The group D based on trapdoor permutation family Π consists of three algorithms: Random generation algorithm (Gen): it inputs security parameter λ and outputs permutation description s and related trapdoor t. Evaluation algorithm (Eva): it inputs s and the original image $x \in D$ and outputs the image of x, $y \in D$. Inverse algorithm (Inv): it inputs s, trapdoor t, and y and output the preimage of y in the permutation. And the one-way permutation function has the following properties:

(i) For all $\text{Gen} \longrightarrow_R (s, t)$, the related $\text{Eva}(s, \cdot)$ should be a permutation of group D

(ii) For all $x \in D$, $\text{Inv}(s, t, \text{Eva}(s, x)) = x$

3.2.2. Bilinear Pairing. According to literature [34], we briefly introduce the definition of bilinear pairing. Let G_1 and G_2 be two cyclic groups of the same prime order $p \geq 2^\lambda$, g be the generator of G_1, and λ be the security parameters. Let e: $G_1 \times G_1 \longrightarrow G_2$ denote a map to satisfy the following properties:

(i) Bilinear: For any $x, y \in G_1$ and $a, b \in Z_p$, the equation $e(x^a, y^b) = e(x, y)^{ab}$ holds

(ii) Computable: for $x, y \in G_1$ that meets the condition, an algorithm can be found to calculate the value of $e(x, y)$ in polynomial time

(iii) Nondegenerate: $e(g, g) \neq 1$

3.2.3. Decisional Bilinear Diffie–Hellman (DBDH) Assumption. According to literature [35], the description of DBDH assumption is given. The definitions of groups G_1 and G_2 and map e: $G_1 \times G_1 \longrightarrow G_2$ are the same as above. For any adversary B, its advantages are defined as $\text{Adv}_B^{\text{DBDH}}(\lambda) = \left| \begin{matrix} \Pr[B(g, g^a, g^b, g^c, e(g, g)^{abc} = 1] \\ -\Pr[B(g, g^a, g^b, g^c, e(g, g)^z = 1] \end{matrix} \right|$, where $a, b, c, z \leftarrow_R Z_p$. We say that the DBDH assumption holds if, for any probability polynomial time (PPT) adversary B, its advantage $\text{Adv}_B^{\text{DBDH}}(\lambda)$ is negligible with security parameter λ.

3.2.4. Forward Privacy Security. This paper cites the definition of forward privacy technology in literature [12].

Common Leakage. We define the leak function $L = (L^{\text{Setup}}, L^{\text{Search}}, L^{\text{Update}})$, which represents the amount of information leaked to the adversary in the searchable encryption scheme. The leak function L treats the query list Q (including all queries) as status. List Q stores each query for keyword w, and its element is (i, w) or $(i, \mathbf{op}, \mathbf{in})$ for an update query $\mathbf{op}$ with input $\mathbf{in}$. The integer i represents the timestamp with the initial value of 0, which increases with the number of queries.

(i) Search pattern: $\text{sp}(x) = \{j: (j, x) \in Q\}$ (only matches search queries)

(ii) Query pattern: $\text{qp}(x) = \{j: (j, x) \in Q \text{ or } (j, \text{op}, \text{in}) \in Q \text{ and } x \text{ appears } in \ \mathbf{in}\}$

Forward Privacy. A L-adaptively secure searchable encryption scheme $\sum$ is forward private if the update leakage function L^{Update} can be expressed as

$$L^{\text{Update}}(\mathbf{op}, \mathbf{in}) = L'(\mathbf{op}, \{(\text{ind}_i, u_i)\}), \tag{1}$$

where $\{(\text{ind}_i, u_i)\}$ represents the modified document set and u_i represents the number of modified words.

3.2.5. Binary Index Tree. We use the binary index tree in literature [18] to store the encrypted file information, as shown in Figure 1.

Block. A block b is generated for each index (w, ind). The block b is divided into two parts, the block address b.id and the block data b.data, where b.data stores the data value b.value (the information of one index (w, ind) in this paper), the previous block key denoted by b.pre.key, and the previous block address b.pre.id. To sum up, $b = (b.\text{id}, b.\text{value}, b.\text{pre.key}, b.\text{pre.id})$.

Blocks Chain. Connect the blocks with the same keyword w into a chain C. Chain C represents the index list related to w, and its length is equal to the number of file indexes n_w containing w. The total number of chains is equal to the number of keywords $|W|$. There are three types of blocks on each chain C: the first block b.head, the last block b.tail, and the other blocks b.internal, where $b.\text{tail} = (b.\text{id}, b.\text{value}, \perp, \perp)$.

Binary Index Tree. Each node of the binary index tree T stores a block b. The node indexed by id is denoted by b.id, and the stored data is denoted by b.data. The value of the data block stores all indexes (w, ind) containing the keyword w. Finally, the tree T has a total of $|W|$ index lists (chains), including $|W| \cdot n_w$ nodes, and the depth is $\log(|W| \cdot n_w)$.

Table 1: Notations description.

Notation	Description
λ, μ	The security parameter
(q, p, G_1, G_2, e)	The bilinear parameters
G, F	The one-way permutation functions
H_1, H_2, H_3, H_4	The hash functions
$\sum$	The map stored on the data owner side
$\mathbf{T}$	The list stored on the edge server side
T	The tree stored on the cloud server side
(pk_i, sk_i)	i's public/private key
(fk_o, fk_o', wk)	Data owner's secret key
m	The plaintext
ind	The identifier of m
w	The keywords
W	The set of keywords associated with m
op	The addition/deletion operation
C	The ciphertext of m
$\mathrm{Tra}(w)$	The query trapdoor
$\mathrm{Tr}(w)$	The search trapdoor
b	A data block stored in the tree T is used to store index information
$(b.\mathrm{id}, b.\mathrm{key})$	An identifier/key pair matches the data block b
(qk_c, qk_u)	The query key generated by the data owner for a data user
$rk_{o \longrightarrow u}$	The reencryption key generated by the data owner for a data user
EDB	The encrypted database
$x \leftarrow_R X$	Uniformly and randomly select x from the finite set
$\|$	The concatenation operator
$\oplus$	The XOR operator
C	A chain
W	The keyword dictionary
$\|W\|$	The number of keywords in the keyword dictionary
n_w	The number of files contained in a keyword w

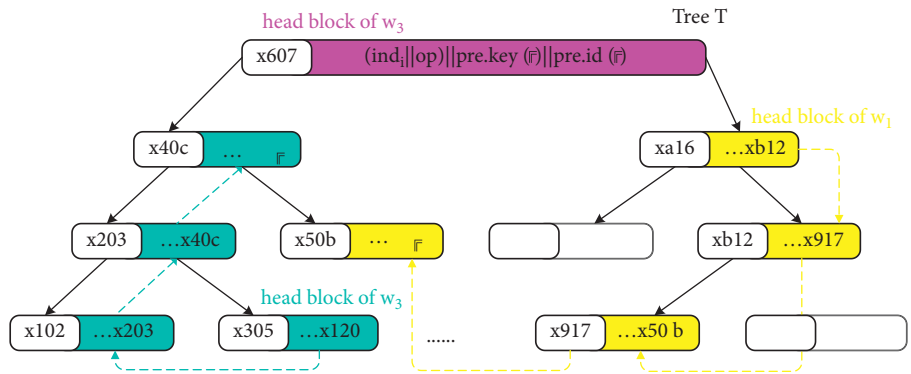


Figure 1: Binary index tree T (stores index information of three keywords (w_1, w_2, w_3)).

4. System Design

4.1. Design Goals. According to the information transmission requirements of P2P network and users, the scheme designed in this paper must meet the following goals:

Efficiency: the index structure in the searchable encryption scheme has an important impact on search efficiency. We store the index in a tree structure to improve search efficiency. In the phase of ciphertext generation and search query, we utilize the edge-cloud architecture to improve the fast response capability of the whole system.

Confidentiality: this is another key attribute. This scheme needs to ensure the confidentiality of data, index, and query. That is, the server cannot get the relevant information of the plaintext by analyzing the stored ciphertext or by reencrypting the ciphertext. And it also cannot get the relevant index information and data query when attacked by an adversary. In addition, this scheme provides forward privacy performance specifically against file injection attack [12].

Dynamic update: this scheme supports the dynamic update of files including add, delete, and update functions. When the user wants to modify the data, he

only needs to perform corresponding operations on the data index and the database.

Correctness: correctness is the basic property of a searchable decryption scheme, which requires the system to provide services correctly. Specifically, as long as the correct trapdoor is provided and the system is executed sequentially in a legal manner, the search results must be correct.

Multiusers: this attribute refers to the ability of the scheme to support data transmission between different users. And the data owner is able to authorize users and update their access rights at any time.

4.2. System Model. As shown in Figure 2, our scheme has four entities: data owner (DO), edge server (ES), cloud server (CS), and data user (DU). DO needs to outsource data while keeping the confidentiality and operation ability of data. At the same time, DO wants to have the right to control other users' access to the data. ES is responsible for storing part of the keyword information from the DO nearby, assisting the data user to generate trapdoor, and reencrypting the ciphertext found by CS. CS stores encrypted data, encrypts index, and responds to data queries. The scheme mainly includes the following processes:

(1) Step 1 (data update): DO first extracts index from the outsourced files and generates index information. He first uses a symmetric encryption algorithm to encrypt index and asymmetric encryption algorithm to encrypt documents. Then DO uploads part of the keyword information to ES and uploads the encrypted index and ciphertext to CS.

(2) Step 2 (user authorization). Whenever a new authorized user DU joins the system, the DO generates the query key and the reencryption key. DO sends the query key to DU and the reencryption key to the ES near the DU.

(3) Step 3 (search keyword). DU generates a query trapdoor containing the keyword. The query trapdoor is sent to the ES, and then the ES generates the trapdoor for search and sends it to the CS. CS decrypts the encrypted index from the search trapdoor to get the relevant document identifier and then obtains the required ciphertext from the searched file identifier. Finally, the CS sends the searched ciphertext to the ES near the DU.

(4) Step 4 (reencryption). After obtaining the matching ciphertext, the ES finds the stored reencryption key of the target user and reencrypts the ciphertext. Finally, the reencrypted ciphertext that the DU has the ability to decrypt is generated and sent to the DU.

(5) Step 5 (decryption). DU uses its private key to decrypt.

The basic framework of the FSSE scheme based on the edge-cloud architecture in P2P network EMFSSE is as follows:

(1) Setup (λ): it takes security parameters λ/μ as input and outputs public parameters *Para*.

(2) KeyGen (Para): enter the public parameter Para, and the ES and the DO generate the required key, respectively. Specifically, the ES generates its public/private key (pk_e/sk_e). The DO generates its public/private key (pk_o/sk_o), secret key of encryption index (fk_o/fk_o'), and keyword encryption key wk.

(3) Update (Para, m, ind, W, $\sum$): this operation is performed by the DO. It inputs public parameters Para, ciphertext m, document identifier ind, key set W related to the document, and keyword status map $\sum$ and outputs index information table T', encrypted index set **B**, and ciphertext C and updates the map $\sum$.

(4) Register (Para): when a new data user joins the system, first, his/her own public/private key (pk_u/sk_u) is generated. Then, DO authorizes the added user, inputs the public parameters*Para*, and generates the query key qk_u, query-related information qk_c, and reencryption key $rk_{o\longrightarrow u}$ for the new user DU. Finally, $(qk_c, rk_{o\longrightarrow u})$ is sent to the ES and qk_u is sent to the DU.

(5) Trapdoor (Para, qk_u, w): the DU takes public parameters Para, its query key qk_u, and query keyword w as input and outputs the trapdoor Tra (w) of w.

(6) Search (Para, Tra (w), **T**; T, EDB): this operation is performed by ES and CS in cooperation. ES inputs public parameters Para, trapdoor Tra (w), and index information table **T**, outputs query token token, and sends it to the CS. The CS inputs encrypted index tree T, encrypts database EDB, and finally outputs ciphertext set **C**.

(7) Re – Encryption $(rk_{o\longrightarrow u}, \mathbf{C})$: the ES inputs reencryption key $rk_{o\longrightarrow u}$ and ciphertext set **C** and outputs reencrypted ciphertext set C'.

(8) Decryption (C', sk_u): the operation is performed by the DU, which inputs the key sk_u and the reencrypted ciphertext set C' and outputs the plaintext m corresponding to each ciphertext C' in C'.

5. Scheme Construction

This section describes the structure of the scheme EMFSSE in detail.

5.1. System Initialization Phase. Setup $(\lambda, \mu) \longrightarrow$ (Para): taking security parameters λ and μ as input, the algorithm generates bilinear mapping e: $G_1 \times G_1 \longrightarrow G_2$, where G_1 and G_2 are two cyclic groups of order $p \geq 2^\lambda$. Then, it randomly chooses $g, g_1, g_2, h_1, h_2 \leftarrow_R G_1$, calculates $L = e(h_1, h_2)$, and lets $\mathbf{P} = (g, g_1, g_2, h_1, h_2, L)$; meanwhile, it selects some secure hash functions H_0: $\{0,1\}^* \longrightarrow \{0,1\}^{\lambda+\mu+1+\log N}$, $H_{1,1'}$: $\{0,1\}^* \longrightarrow \{0,1\}^\mu$, H_2: $\{0,1\}^* \longrightarrow \{0,1\}^\lambda$, and $H_{3,4}$: $\{0,1\}^* \longrightarrow \{0,1\}^*$ and selects one-way permutation functions G and F. Finally, the DO initializes an empty map $\sum$ that stores its state, the ES initializes the empty index information table **T**, and the CS

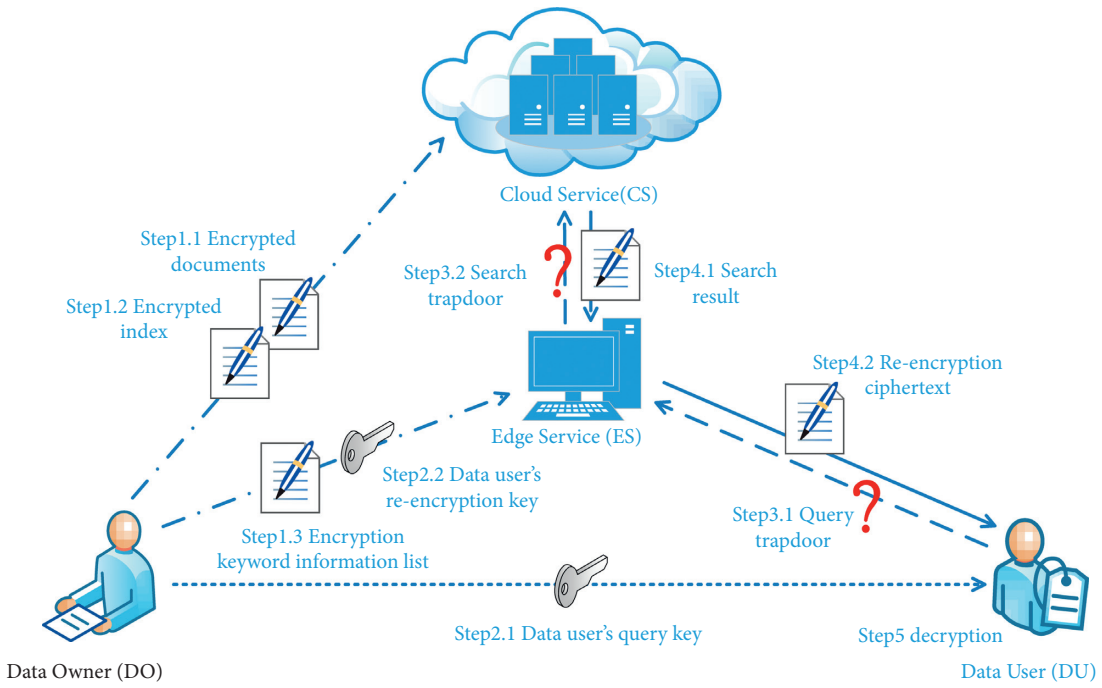


Figure 2: System model.

initializes an empty index tree T. Publish public parameters Para $= (\mathbf{P}, G, F, H_{0,1,1',2,3,4})$.

5.2. Key Generation Phase. KeyGen(Para) $\longrightarrow$ $(sk_e, pk_e; sk_o, pk_o, fk_o, fk_o', wk)$: it inputs the public parameters Para.

(1) Step 1: the ES randomly selects $z \leftarrow_R \mathbb{Z}_p$ and generates private key $sk_e = z$ and public key $pk_e = g_2^z$.

(2) Step 2: the DO randomly selects $x_o, y_o \leftarrow_R \mathbb{Z}_p$, respectively, generates private key $sk_o = (x_o, y_o)$ and public key $pk_o = (h_1^{x_o}, g_1^{y_o})$, respectively, selects key $fk_o, fk_o' \leftarrow_R \{0,1\}^\lambda$ of encryption index, and randomly selects key of encryption key $wk \leftarrow_R Z_p^*$.

5.3. Update Phase. Update(Para, m, op, ind, $W, \sum$) $\longrightarrow$ $(C, T', \mathbf{B})$: when a new file is added, the DO input public parameters. Para, file m, operator op = add (op = add/del, resp., for add or delete operation), file identifier ind, the file related keyword set W, and DO's status $\sum$, and then do operations as follows:

(1) Step 1 (documents encryption): the DO randomly selects $r \leftarrow_R Z_p$ and generates ciphertext $C = (L^r \cdot m, g_1^r, e(h_1, g_2)^{1/r}, x_o^r) = (c_0, c_1, c_2, c_3)$.

(2) Step 2 (index encryption): the DO initializes an empty set $\mathbf{B}$ and a table T'. For each keyword $w \in W$, DO first encrypts it to get $\mathrm{Tr}(w) = H_4(e(H_3(w), g)^{wk})$ and then uses function ReGen (as shown in Table 2) to calculate the identifier and key of the block currently indexed by keyword w, (id, key)$\leftarrow$ReGen$(w, \sum[w] = c)$. Then, DO makes pre.id = id, pre.key = key, and $c++$ and generates the identifier and key of the new block (id, key)$\leftarrow$ReGen(w, c). Finally, DO calculates the block key mask$\leftarrow H_0$(id, key), encrypts the index block with it $b \leftarrow$ (id, (//ind‖op//‖pre.key‖pre.id) $\oplus$mask), stores the index information (id, key) in the table T' indexed by $\mathrm{Tr}(w)$, where $T'[Tr(w)] =$ (id, key), and stores the block b in the set $\mathbf{B}$.

(a) The algorithm finally outputs index information table T', ciphertext C, and encrypted index set $\mathbf{B}$. The table T' is sent to the ES which stores its contents in an index information table. The ciphertext C and set $\mathbf{B}$ are sent to the CS which stores C in the encrypted database EDB and inserts the blocks into the encrypted index tree T.

Table 2: ReGen algorithm (note: i_w represents w or a unique identifier of w).

Algorithm ReGen$(w, \sum[w] = c)$
1: if $c == -1$ then id$\leftarrow\perp$, key$\leftarrow\perp$
2: else
3: $K_w \leftarrow F_{fk_o}(w), K_w' \leftarrow F_{fk_o'}(w)$
4: $ST_0(w) \leftarrow G(i_w)$
5: if $c == 0$ then
6: id$\leftarrow H_1(K_w, ST_0(w))$; key$\leftarrow H_2(K_w', ST_0(w))$
7: else
8: $ST_c(w) \leftarrow H_{1'}^c(ST_0(w))$
9: id$\leftarrow H_1(K_w, ST_0(w))$; key$\leftarrow H_2(K_w', ST_0(w))$
10: end if
11: end if
12: Return (id, key)

5.4. Authorization Phase. Register(Para) $\longrightarrow$ $(sk_u, pk_u; qk_u, qk_c, rk_{o\longrightarrow u})$: input the public parameter Para; the algorithm is completed by the new user DU and DO in order.

(1) Step 1: when a new user DU joins the system, it randomly selects $x_u, y_u \leftarrow_R \mathbb{Z}_p$ to generate a private key $sk_u = (x_u, y_u)$ and a public key $pk_u = (h_1^{x_u}, g_1^{y_u})$. Then, publish public key, indicating that a new user has joined the system.

(2) Step 2: the new user DU sends his public key to DO and then is authorized by DO to access the data. The DO randomly selects $qk_u \leftarrow_R Z_p^*$ as the query key, calculates $qk_c = g^{wk/qk_u}$, and generates the proxy reencryption key $rk_{o\longrightarrow u} = (h_2 g_1^{y^u} g_2^z)^{1/x_o}$ of DU.

5.5. Trapdoor Generation Phase. Trapdoor(Para, qk_u, w) $\longrightarrow$ Tra(w): the DU inputs the public parameter Para, the query key qk_u, and the query keyword w, then generates the query trapdoor $\mathrm{Tra}(w) = H_3(w)^{qk_u}$, and sends it to ES.

5.6. Search Phase. Search(Para, Tra(w), **T**; T, EDB) $\longrightarrow$ (**C**): the algorithm is completed by the cooperation of ES and CS.

(1) Step 1: input the query trapdoor Tra(w) and the encrypted keyword status table **T**, and the ES generates $\mathrm{Tr}(w) = H_4(e(\mathrm{Tra}(w), qk_c))$, then queries the table **T** to get the search trapdoor $(\mathrm{id}, \mathrm{key}) \leftarrow \mathbf{T}[\mathrm{Tr}(w)]$, and sends the search trapdoor token = (id, key) to the CS.

(2) Step 2: taking the search trapdoor token, index tree T, and encryption database EDB as input, the CS first initializes the empty sets **S** and **C** and sets $i \leftarrow 0$. Then, CS finds the latest block $b \leftarrow T[\mathrm{id}]$ by trapdoor token, calculates the block key mask$\leftarrow H_0$(key, id), decrypts the block with mask, and obtains $(\prime\prime \mathrm{ind} \| \mathrm{op} \prime\prime \| \mathrm{pre.key} \| \mathrm{pre.id}) \leftarrow b.\mathrm{data} \oplus \mathrm{mask}$, where b.value = (ind$\|$op). And the CS stores b.value in $\mathbf{S}[i \pm \pm]$ (merges the same ind in **S** according to op), sets id$\leftarrow$pre.id and key$\leftarrow$pre.key, and repeats the above process until id $\neq \perp$ and key $\neq \perp$. Finally, CS finds the corresponding encrypted data C in EDB by ind $\in$ **S**, stores C in **C**, and sends the search result **C** to the ES.

5.7. Reencryption Phase. Reencryption(**C**, $rk_{o\longrightarrow u}$) $\longrightarrow$ (C'): the ES finds the reencryption key $rk_{o\longrightarrow u}$ corresponding to the DU, inputs it and set **C**, and initializes an empty set C'. For each $C \in \mathbf{C}$, the ES reencrypts it to get reencrypted ciphertext $C\prime = (c_0', c_1', c_2') = (c_0, c_1, e(c_3, rk_{o\longrightarrow u})/c_2^z)$, stores the new ciphertext in C', and then sends C' to the DU.

5.8. Decryption Phase. Decryption(C', sk_u) $\longrightarrow$ (m): inputting reencrypted ciphertext C' and private key sk_u, DU decrypts each $C' \in C'$ to get all the required ciphertext $m = c_0' \cdot e(h_1, c_1')^{y_u} / c_2'$. The algorithm can get the final plaintext files m.

6. Security Proof

In this section, some important design objectives mentioned in Section 3 will be proved. Specifically, we will prove that the scheme can provide correctness, confidentiality, and forward security.

6.1. Correctness. An example is given to prove the correctness of the scheme. Suppose that there are files m_1, m_2 with keyword w_1, and their corresponding file identifiers are $\mathrm{ind}_1, \mathrm{ind}_2$.

Firstly, plaintext and index information are encrypted, respectively. We randomly select $r_1, r_2 \leftarrow_R Z_p$, generate ciphertexts $C_1 = (L^{r_1} \cdot m_1, g_1^{r_1}, e(h_1, g_2)^{1/r_1}, X_o^{r_1}) = (c_0^1, c_1^1, c_2^1, c_3^1)$ and $C_2 = (L^{r_2} \cdot m_2, g_1^{r_2}, e(h_1, g_2)^{1/r_2}, X_o^{r_2}) = (c_0^2, c_1^2, c_2^2, c_3^2)$, and generate encrypted keyword information $\mathrm{Tr}(w_1) \leftarrow H_4(e(H_3(w_1), g)^{wk})$. We also generate index blocks $b_1 = (\mathrm{id}_1, (\mathrm{ind} \| \mathrm{op} \| \perp \| \perp) \oplus \mathrm{mask}_1)$ and $b_2 = (\mathrm{id}_2, (\mathrm{ind}_2 \| op \| \mathrm{id}_1 \| \mathrm{key}_1)) \oplus \mathrm{mask}_2$ and store the client status $\sum(w) = c = 1$ (initialization value is -1) at this moment and encrypted keyword information list $\mathbf{T}[\mathrm{Tr}(w_1)] = (\mathrm{id}_2, \mathrm{key}_2)$.

Secondly, we search for the file containing the keyword w_1. The query token of w_1 is generated $\mathrm{Tra}(w_1) = H_3(w_1)^{qk_u}$, the encrypted keyword information $\mathrm{Tr}(w_1)$ is generated from the query token $\mathrm{Tr}(w_1) \leftarrow H_4(e(\mathrm{Tra}(w_1), qk_c))$, and then the search token token = $(\mathrm{id}_2, \mathrm{key}_2)$ is obtained from the $\mathrm{Tr}(w_1)$ query table **T**. Through token token, we find block b_2 in the index. Then, we generate the decryption secret key $\mathrm{mask}_2 \leftarrow H(\mathrm{key}_2, \mathrm{id}_2)$ of b_2 and do XOR operation to get the value of b_2, where b_2.value$\leftarrow T[\mathrm{id}]$.data$\oplus$mask, which contains the information of ind_2 and the information of the previous index block $(\mathrm{id}_1, \mathrm{key}_1)$. Similarly, we get b_1 and decrypt it to get ind_1 and previous index block information $(\perp, \perp)$ and then terminate the search. Through $\mathrm{ind}_{1,2}$, we can find the corresponding ciphertext $C_{1,2}$. In this scheme, the ES is required to reencrypt the ciphertext, and the DU can decrypt it. Taking C_1 as an example, we first generate the reencryption key $rk_{o\longrightarrow u} = (h_2 Y_u Z)^{1/x_o}$, then use it to encrypt the ciphertext, and get the reencrypted ciphertext $C_1' = (c_0^1\prime', c_1^1\prime', c_2^1\prime') = (c_0^1, c_1^1, e(c_3^1, rk_{o\longrightarrow u})/c_2^{1z})$, where $c_2^1\prime' = e(c_3^1, rk_{o\longrightarrow u})/c_2^{1z} = e((h_1^{x_o \cdot r_1}, (h_2 \quad g_1^{y_u} g_2^z)^{1/x_o})/e(h_1, g_2)^{1/r_1 \cdot z}) = e(h_1, h_2 g_1)^{r_1 y_u}$, and the final ciphertext form is $C_1' = (e(h_1, h_2)^{r_1} \cdot m_1, g_1^{r_1}, e(h_1, h_2 g_1)^{r_1 y_u})$.

Finally, we decrypt the final reencrypted ciphertext: $c_0^1\prime' \cdot e(h_1, c_1^1\prime')^{y_u} / c_2^1\prime' = m_1$.

To sum up, the scheme is correct.

6.2. Confidentiality. The EMFSSE scheme is proved to be secure according to the following theorem.

Theorem 1. *In the process of reencryption, the EMFSSE scheme is CPA secure based on the DBDH assumption. Specifically, let A be the adversary of attacking EMFSSE*

scheme with ε advantage in CPA game, and then there is an adversary B who at least solves DBDH assumption with ε/q_{okg} advantage.

Proof. B randomly selects five tuples $T = (g, A = g^a, B = g^b, C = g^c, Z)$ as input. The goal of B is to determine whether Z is random value $e(g,g)^z$ or $e(g,g)^{abc}$ by playing the following game with adversary A, where $a, b, c, z \leftarrow_R Z_p$.

(1) Setup(λ): randomly select $\delta \leftarrow_R Z_p$, set $g_1 = g, g_2 = g^\delta, h_1 = g^a, h_2 = g^b, L = e(g^a, g^b)$, and return $\mathbf{P} = (g_1, g_2, h_1, h_2, L)$. Run $ES - KeyGen(\mathbf{P})$: randomly select $e \leftarrow_R Z_p$, set $sk_e = e$ and $pk_e = g_2^e$, and return pk_e.

(2) Phase 1: A first performs polynomial private times key extraction query. When receiving the query from A, B replies as follows:

(a) $O_{\text{ekg}}(\lambda)$: return the private key of ES $sk_e = e$.
(b) $O_{\text{okg}}(o)$: select $x_o, y_o \leftarrow_R Z_p$ randomly.
(i) If B guesses that it is the target user for this guess, that is, $o' = o$, then calculate $pk_{o'} = (h_1^{x_{o'}}, h_2^{-1} g^{y_{o'}}) = (h_1^{x_{o'}}, g_1^{y_{o'}-b})$ and $sk_{o'} = (x_{o'}, y_{o'} - b)$, hide information $sk_{o'}$, and return $pk_{o'}$.
(ii) Otherwise, $o' \neq o$, calculate $pk_o = (h_1^{x_o}, g_1^{y_o})$ and $sk_o = (x_o, y_o)$, hide information sk_o, and return pk_o.
$O_{\text{ukg}}(u)$: randomly select $x_u, y_u \leftarrow_R Z_p$, let $sk_u = (x_u, y_u)$ and $pk_u = (h_1^{x_u}, g_1^{y_u})$, and return (sk_u, pk_u).
$O_{\text{rkg}}(pk_o, pk_u, pk_e)$, O_{okg} and O_{ukg} generate pk_o and pk_u, respectively; B runs $\text{Reencryption} - \text{KeyGen}(sk_o, pk_u, pk_e)$ and returns reencryption key $rk_{o \longrightarrow u} = (h_2 g_1^{y_u} g_2^e)^{1/x_o}$.

(3) Challenge: when A decides to complete Phase 1, it constructs two equal-length messages $m_{0,1}$ and public key $pk_{o'}$ and sends them to B. B performs the following operations:

(i) If $pk_o \neq pk_{o'}$, B outputs random string and terminates.
(ii) Otherwise, B randomly selects a value $\beta \in \{0,1\}$ and constructs a ciphertext $C = (c_0', c_1', c_2')$ for the challenge, where $c_0' = Z \cdot m_\beta$, $c_1' = g_1^r = g^c$, and $c_2' = e(c_3, W)/c_2^e = e(g,g)^{a(b+1)c^2 \cdot y^u}$, and hides information $r = c$.
Supposing that A does O_{okg} queries at most q_{okg} times, then in Challenge, the probability that B guesses the right o' is at least $1/q_{\text{okg}}$.

(4) Phase 2. B answers A's question, just like in Phase 1. A outputs its guess $\beta' \in \{0,1\}$. If $\beta = \beta'$, B outputs 1; otherwise, it outputs 0; then the advantage of B in solving the DBDH problem is at least $\varepsilon/q_{\text{okg}}$. The proof is complete. □

6.3. Forward Privacy. We refer to the form in the scheme [17] to prove the forward privacy of our scheme. The following is a simplified analysis:

Define $\text{Hist}(w)$ as the history of the keyword w which lists all the modifications to the $\text{DB}(w)$ over a period of time. F is the pseudorandom function controlled by the key, inputs λ bit string and key fk, and outputs λ bit string. $H_{0,1,1',2}$ are secure hash functions defined on random oracle model.

Theorem 2. *L – adaptively – secure of EMFSSE scheme. The leakage function $L_\sum = (L\sum^{Srch}, L\sum^{Updt})$ is defined. If $L\sum^{Srch}(w) = (sp(w), Hist(w))$ and $L\sum^{Updt}(w) = (w, ind) = \perp$, EMFSSE is an $L_\sum$-adaptive security scheme with forward security.*

Proof. Set security parameters λ and μ. Here, we use indistinguishable games derived from the real world to prove the theorem.

(1) Game G_0: G_0 is an SSE security game $\text{SSEReal}_A^{\text{ERMF}}(\lambda, \mu)$ in the real world. The relationship between the two can be expressed as $P[\text{SSEReal}_A^{\text{ERMF}}(\lambda, \mu) = 1] = P[G_0 = 1]$.

(2) Game G_1: G_1 uses random selection instead of calling F to generate keys K_w and K_w^* which are used for generating id and key in ReGen. When a new index (w, ind) is added, the keys K_w and K_w^* are randomly selected and added to the key table KEY. The next time someone asks for the keyword w, it checks the table KEY first. If there is a corresponding key in the table, the stored key is called directly; if not, randomly select the keys K_w and K_w^* and store them in the table. Thus, a reduction is established to distinguish F from real random functions. That is, there are effective adversaries A, $P[G_0 = 1] - P[G_1 = 1] \leq \mathbf{A}\ \mathbf{dv}_{F,A}^{\text{prf}}(\lambda, \mu)$.

(3) Game G_2: G_2 and G_1 have consistent behavior. Except for some inconsistent results of random oracle requests in G_2, the inconsistent time is recorded as Bad. Because the random string u of H_0 is generated in the update operation and updated in the search protocol, there is a time difference Bad. If the adversary asks H_0 before searching, H_0 randomly selects u' to return to the adversary according to the properties of random oracle. After this request, u is programmed to H_0.

(a) Through the consistent transition time node Bad, we can limit the distinguishing advantage between G_2 and G_1: $P[G_2 = 1] - P[G_1 = 1] \leq P[\text{Bad}]$. Probability of Bad: if and only if the adversary queries H_0 in the random oracle model with ind, because the output mask is randomly selected from $\{0,1\}^{\lambda+\mu+1+\log N}$, the probability of

the adversary selecting the corresponding mask is $(1/2^{\lambda+\mu+1+\log N})$. Suppose that a probabilistic polynomial time adversary can query with the random oracle for ploy(λ) times at most; then we can know that $P[\text{Bad}] \leq (\text{ploy}(\lambda)/2^{\lambda+\mu+1+\log N})$ is negligible; that is, $P[G_2 = 1] = P[G_1 = 1]$.

(4) Game G_3: G_3 is similar to G_2 in that it performs the same operation on H_0 in G_2 for $H_{1,1',2}$. The difference is the probability of Bad. If and only if the adversary can query $H_{1,1',2}$ in a random oracle model with correct input, the generated string length is μ, μ, λ, respectively. So the probability of the adversary choosing the right output value is $1/2^{\mu}$, $1/2^{\mu}$, and $1/2^{\lambda}$, respectively. Assuming that a probabilistic polynomial time adversary can perform ploy(γ) queries at most with a random oracle, then $P[\text{Bad}_1] \leq (\text{ploy}(\gamma)/2^{\mu})$, $P[\text{Bad}_2] \leq (\text{ploy}(\gamma)/2^{\mu})$ and $P[\text{Bad}_3] \leq (\text{ploy}(\gamma)/2^{\lambda})$ are all negligible; that is, $P[G_3 = 1] = P[G_2 = 1]$.

(5) Game G_4: G_4 is similar to G_3, but it removes information unrelated to hash functions $H_{0,1,1',2}$; that is, it is a single round trip protocol. Therefore, in the random oracle model, G_4 and G_3 cannot be distinguished; that is, $P[G_4 = 1] = P[G_3 = 1]$.

(6) Simulator S: The code in G_4 is divided into two parts: leak function and simulator. In G_4, the status$u$ of the block is stored, which means that when the time is u, the values in tables $B\text{mask}^*$, $B\text{key}^*$, and $B\text{id}^*$ are mask, key, and id, respectively, and their essence is random string. The simulator uses the counter κ instead of the keyword w. Its initial value is 0, and the value is increased by 1 every time the update operation is performed. The new value is stored in tables $B\text{mask}^*[\kappa]$, $B\text{key}^*[\kappa]$, and $B\text{id}^*[\kappa]$. In the view of adversary A, the update protocol only outputs three random strings, while the search protocol outputs the same number of random strings (key, id, and mask, resp.). So we get the equation $P[G_4 = 1] - P[\text{SSEIdeal}^{\text{ERMF}}_{A,S,L_{\sum}} = 1] = 0$.

Conclusion: According to the above games and simulator S and hash functions $H_{0,1,1',2}$, the following formula can be obtained: $P[\text{SSEReal}^{\text{ERMF}}_{A}(\lambda, \nu) = 1] - P[\ \text{SSEIdeal}^{\text{ERMF}}_{A,S,L_{\sum}} = 1] \leq \mathbf{A}\ \mathbf{dv}^{\text{prf}}_{F,A}(\lambda, \mu)$. □

7. Comparisons and Evaluation

In this section, the proposed scheme is compared with other schemes. And then, we implement the EMFSSE scheme and evaluate its performance.

7.1. Comparisons. Table 3 shows the differences in theoretical computation cost and other functions between EMRF and other schemes (Bost [12] scheme, Chen et al. [21] scheme, and Wang et al. [15] scheme). N is the number of indexes containing the keyword w; E_1 and E_2 represent the multiplication in groups G_1 and G_2, respectively; e is a bilinear pair; M represents secure map-to-point operation; H means the secure hash function; T/T^{-1} represents trapdoor function and its inverse operation (public key primitives, such as RSA), respectively; F/F^{-1} represents a pair of pseudorandom permutations (such as AES or DES).

As shown in Table 3, compared with other relevant schemes, the proposed scheme has better performance. For all schemes, in the update phase, the computational complexity is linearly related to the number of indexes N. Among them, the Bost scheme of [12] has a major flaw that it does public key operation T for N times. Because of its slow encryption, the scheme is limited in large data sets. The computational complexity of Wang's scheme is $2E_1 + e + H + N \times E_2$. Among them, operation e takes a lot of time, but it is independent of index. This scheme calls E_2 once for each index, which leads to relatively high consumption. The computational complexity of Chen's scheme and our scheme is $2E_1 + e + M + F + 6H + N \times (2H)$ and $e + E_2 + T + 7H + 2F + N \times H$, respectively. The operation related to the index number is only the hash function, so it is more efficient than the previous two schemes. However, it is worth noting that Chen's scheme uses map-to-point operation with high cost. To sum up, among the four schemes, our scheme has the best update efficiency.

In the trapdoor generation phase, the complexity of each trapdoor generated by Bost's scheme, Chen's scheme, Wang's scheme, and EMFSSE scheme is H, $e + M + E_1 + H$, $E_1 + H$, and $E_1 + H$, respectively. Obviously, Chen's scheme has a relatively high complexity of generating a trapdoor, while Bost's scheme has relatively optimal complexity. Wang's scheme and our scheme have the same complexity, which is slightly higher than Bost's scheme, but it is a must to extend functionality and improve security. In most cases, the diminished performance is acceptable.

Similar to the update phase, the computational complexity of all schemes in the search phase is also linearly related to N. In order to get each search result, Wang's scheme requires a time-consuming bilinear mapping, and Bost's scheme requires RSA decryption T^{-1}, both of them are expensive. Chen's scheme and EMFSSE scheme only need to perform the corresponding hash function operation, which is less time-consuming. Obviously, the consumption of the EMFSSE scheme is lower.

In terms of function, compared with other schemes, the EMFSSE scheme supports forward privacy and multiuser searchable encryption, which can resist file injection attacks and extend the performance of the scheme to multiuser. In conclusion, the EMFSSE scheme has relatively good performance and low computational costs.

7.2. Evaluation. We do the experiment on the personal computer with Intel Core i5-3337 CPU and 8 GB DDR3 RAM. We use C++ to realize the core function of EMFSSE.

Specifically, we set the security parameters of $\lambda = 128$ bits and $\mu = 64$ bits, use HMAC to implement the hash function, and choose pseudorandom permutation (CBC model, 128-bit key). And we use the 128-bit security level MIRACL CryptoSDK in [36] to calculate the efficiency of pairings and

Table 3: Comparison with other schemes.

Scheme	Update	Trapdoor	Search	Forward privacy	Multiuser
Bost [12]	$H + N \times (2h + T)$	H	$N \times (T^{-1} + 2H)$	√	×
Chen et al. [21]	$2E_1 + e + M + F + 6H + N \times (2H)$	$e + M + E_1 + H$	$4H + N \times (2H)$	√	×
Wang et al. [15]	$2E_1 + e + H + N \times E_2$	$E_1 + H$	$N \times e$	×	×
EMFSSE	$e + E_2 + T + 7H + 2F + N \times H$	$E_1 + H$	$N \times H$	√	√

Table 4: Time consumption of main operations.

Operation	e	E_1	E_2	M	H	F/F^{-1}	T	T^{-1}
Time cost (ms)	105.77	34.45	8.87	322.4	0.001	0.005	0.189	5.896

exponentials operation of bilinear pairing. The time cost of the main operation is shown in Table 4. It takes 105.77 ms to compute bilinear pairs, 34.45 ms to compute exponents in G_1, and 8.87 ms to compute exponents in G_2. It takes 322.4 ms to perform map-to-point operations, and the hash operation takes 0.001 ms. AES algorithm takes 0.005 ms for both encryption and decryption. The execution time of the RSA encryption/decryption algorithm took 0.189 ms and 5.896 ms, respectively.

Figures 3–5 show the approximate time costs of update algorithm, search algorithm, and token generation algorithm with different index numbers ($1.5 \times 10^4 \sim 5.0 \times 10^4$).

Update evaluation: Figure 3 shows the time-consuming comparison of updating algorithms of Bost's scheme, Chen's scheme, and EMRF scheme. (Wang's scheme is not shown in the figure due to its high computational complexity.) As shown in the figure, the abscissa is datasets containing $1 \times 10^4 \sim 5 \times 10^4$ indexes, respectively, with an interval of 0.5×10^4. When the number of indexes is 1×10^4, the cost of Bost's scheme, Chen's scheme, and EMFSSE scheme is 1.91 s, 0.52 s, and 0.12 s, respectively. Bost's scheme has a higher time consumption, Chen's scheme and EPMF have relatively lower time consumption, and EMFSSE has the lowest. The time-consuming of all three schemes increases as the number of indexes increases. Among them, the growth rate of Bost's scheme is the fastest, while that of Chen and EPMF schemes is low and basically flat. In general, the EMFSSE scheme has a lower time-consuming and growth trend during the update phase.

Search evaluation: Figure 4 shows the time-consuming comparison of the search algorithms of Chen's scheme and EMFSSE scheme. (Due to the large time consumption of the other two schemes, no comparison is made.) The abscissa is datasets containing $1 \times 10^4 \sim 5 \times 10^4$ indexes, respectively, with an interval of 0.5×10^4. When the number of indexes is 1×10^4, the cost of Chen's scheme and EMFSSE scheme is 20 ms and 10 ms, respectively. Similar to the update algorithm, the time consumption of the search algorithm also increases with the number of indexes, and the EMFSSE scheme is better than Chen's scheme in both time-consuming and growth trends.

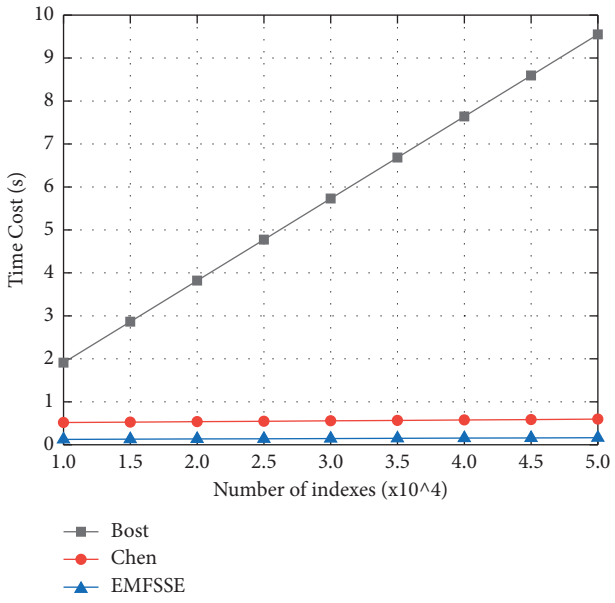


Figure 3: Time consumption of update algorithm.

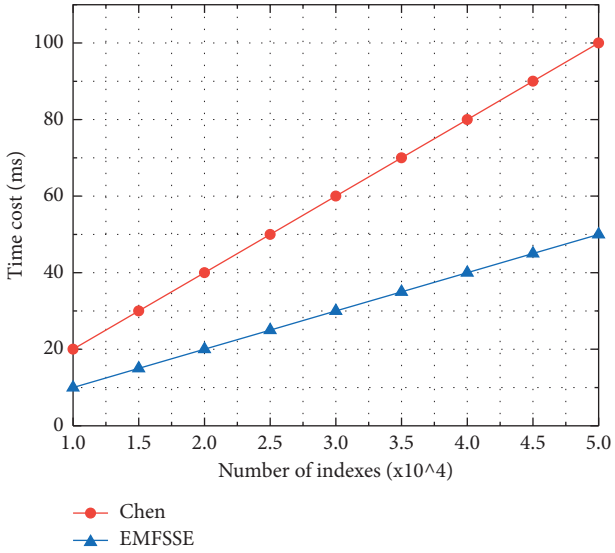


Figure 4: Time consumption of search algorithm.

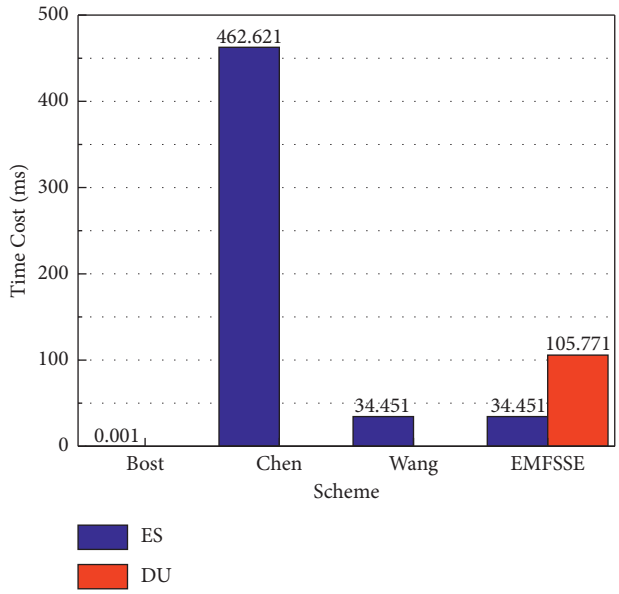


Figure 5: Time consumption of token generation algorithm.

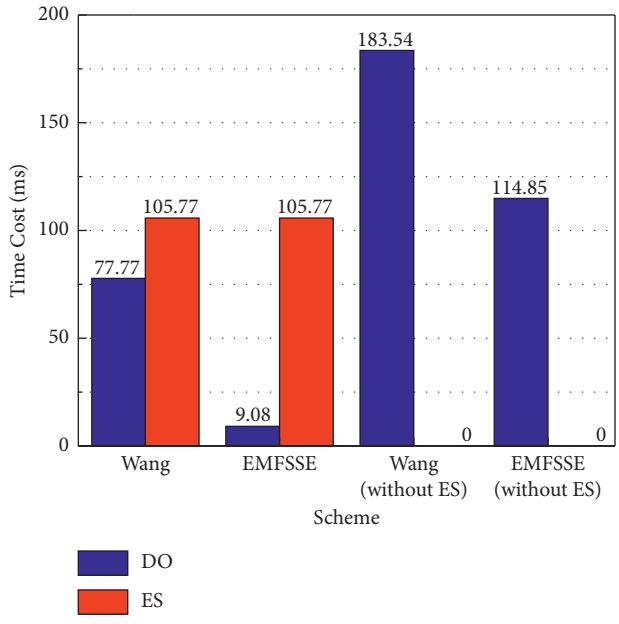


Figure 6: Time cost to generate one ciphertext.

Token generation evaluation: Figure 5 shows the time-consuming comparison of Bost's scheme, Chen's scheme, Wang's scheme, and EMFSSE scheme in a single trapdoor generation algorithm. Chen's scheme needs to do a map-to-point operation, so it is high time-consuming and inefficient. EMFSSE scheme consumes 140.222 ms to generate one trapdoor, but it is worth noting that EMFSSE transfers part of the operation of generating trapdoor to the edge platform, which saves about 75% of the computational costs of the user client. That is to say, the time consumed to generate trapdoor at the DU side is 34.451 ms, which is the same as Wang's scheme. Moreover, compared with the EMFSSE scheme, Wang's scheme can neither support data transmission between multiple users nor meet forward security. In addition, the time EMFSSE scheme consumes is significantly longer than that of Bost's scheme, which only uses the hash function once to generate a trapdoor. But we think that it is necessary to construct a multiuser searchable scheme.

In order to show the advantages of the EMFSSE scheme obtained by using edge computing, EMFSSE is compared with Wang's scheme that also uses edge computing, and the proposed scheme that does not use edge computing. Figure 6 shows the comparison of various schemes for generating a ciphertext. Firstly, the overall time of generating one ciphertext of the EMFSSE scheme is quicker than that of Wang's scheme. Secondly, after using edge computing, the client efficiency of Wang's scheme and the EMFSSE scheme are significantly improved compared to their corresponding schemes without using edge computing. Wang's scheme saves about 57.6% of the client's computing cost, and EMFSSE saves about 92% of the computing cost. As far as we know, in the real environment, edge computing will have better computing efficiency than personal computers, so in the real situation, the computing time of the ES side will be less than our experimental results.

The results show that the EMFSSE scheme has good computational efficiency and security.

8. Conclusion

In this paper, we construct a forward secure searchable encryption scheme supporting multiuser in the P2P network environment. We extend the traditional single-user scheme to multiuser to meet the application requirements of the P2P network environment. And the proxy reencryption technique is used to improve the multiuser scheme. Through reencrypting the searched ciphertext by the proxy party (edge platform), the data user can decrypt the ciphertext with his own private key to avoid direct transmission of the decryption key. In addition, this scheme also uses edge-cloud architecture to replace only-cloud computing and transfers part of the computing to the edge platform, realizing the lightweight computing of this scheme. In terms of security, this scheme satisfies CPA security based on the DBDH assumption in the ciphertext generation process and forward privacy in the index generation process. Theoretical and practical performance analyses show that, compared with the related schemes, this scheme has relatively better efficiency in update operation, search operation, and trapdoor generation operation, and our scheme rarely implements forward security that supports multiuser. In addition, this scheme saves the user's consumption in different degrees in the encryption process or in the token generation process. In the future, we will carry out researches on how to optimize the performance of multiuser searchable encryption scheme by using an attribute-based encryption method.

Data Availability

The data used to support the findings of this study are included within this article.

Conflicts of Interest

The authors declare there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Research Foundation of Young Core Instructor in Henan Province under Grant 2018GGJS058.

References

[1] N. N. Alleema and D. S. Kumar, "Volunteer nodes of ant colony optimization routing for minimizing delay in peer to peer MANETs," *Peer-to-Peer Networking and Applications*, vol. 13, no. 2, pp. 590–600, 2020.

[2] X. Meng and Y. Deng, "A time-aware resource search strategy with the ant colony optimization in MANETs," *Peer-to-Peer Networking and Applications*, vol. 12, no. 5, pp. 1013–1027, 2019.

[3] S. Zhou, T. Zhang, and X. Meng, "iForest: an informed resource search strategy in mobile P2P networks," *Peer-to-Peer Networking and Applications*, vol. 14, no. 5, pp. 1889–1904, 2021.

[4] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[5] X. Li, Q. Wang, M. Liu et al., "Cooperative wireless-powered NOMA relaying for B5G IoT networks with hardware impairments and channel estimation errors," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5453–5467, 2021.

[6] X. Wang, L. T. Yang, D. Meng, M. Dong, K. Ota, and H. Wang, "Multi-UAV cooperative localization for marine targets based on weighted subspace fitting in SAGIN environment," *IEEE Internet of Things Journal*, p. 1, 2021.

[7] G. D. Gonalves, I. Drago, A. B. Vieira, A. P. Couto da Silva, and J. Marques de Almeida, "A study of costs and benefits of content sharing in personal cloud storage," *Journal of Network and Systems Management*, vol. 29, no. 3, 2021.

[8] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2021.

[9] H. Teruo, Y. Hirozumi, H. Akihito, U. Akira, and Y. Keiichi, "Edge computing and IoT based research for building safe smart cities resistant to disasters," in *Proceedings of the IEEE 37th International Conference on Distributed Computing Systems*, Atlanta, GA, USA, June 2017.

[10] J. Zhong and X. Xiong, "Data security storage method for power distribution Internet of things in cyber-physical energy systems," *Wireless Communications and Mobile Computing*, vol. 2021, no. 1, 15 pages, Article ID 6694729, 2021.

[11] H. Zhong, Y. Zhou, Q. Zhang, Y. Xu, and J. Cui, "An efficient and outsourcing-supported attribute-based access control scheme for edge-enabled smart healthcare," *Future Generation Computer Systems*, vol. 115, pp. 486–496, 2021.

[12] R. Bost, "∑οφος: forward secure searchable encryption," in *Proceedings of the ACM Sigsac Conference on Computer & Communications Security*, ACM, Vienna, Austria, October 2016.

[13] Q. Wang, G. Yu, H. Huang, and X. Jia, *Multi-User Forward Secure Dynamic Searchable Symmetric Encryption*, Springer, Cham, Switzerland, 2018.

[14] B. Lu, J. Zhou, and Z. Cao, "A multi-user forward secure dynamic symmetric searchable encryption with enhanced security," *Journal of Computer Research and Development*, vol. 57, no. 10, pp. 2104–2116, 2020.

[15] W. Wang, P. Xu, D. Liu et al., "Lightweighted secure searching over public-key ciphertexts for edge-cloud assisted industrial IoT devices," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4221–4230, 2020.

[16] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in *Proceedings of the International Conference on Applied Cryptography and Network Security*, June 2005.

[17] E. Stefanov, C. Papamanthou, and E. Shi, "Practical dynamic searchable encryption with small leakage," in *Proceedings of the network and distributed system security symposium, NDSS*, San Diego, CA, USA, February 2014.

[18] Y. Wei, S. Lv, X. Guo, Z. Liu, Y. Huang, and B. Li, "FSSE: forward secure searchable encryption with keyed-block chains," *Information Sciences*, vol. 500, pp. 113–126, 2019.

[19] J. Li, Y. Huang, Y. Wei et al., "Searchable symmetric encryption with forward search privacy," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 1, pp. 460–474, 2021.

[20] X. Song, C. Dong, D. Yuan, Q. Xu, and M. Zhao, "Forward private searchable symmetric encryption with optimized I/O efficiency," *IEEE Transactions on Dependable and Secure Computing*, vol. 17, no. 5, pp. 912–927, 2020.

[21] B. Chen, L. Wu, H. Wang, L. Zhou, and D. He, "A blockchain-based searchable public-key encryption with forward and backward privacy for cloud-assisted vehicular social networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 5813–5825, 2020.

[22] H. Wang, K. Fan, H. Li, and Y. Yang, "A dynamic and verifiable multi-keyword ranked search scheme in the P2P networking environment," *Peer-to-Peer Networking and Applications*, vol. 13, no. 16, pp. 2342–2355, 2020.

[23] J. Ren, Y. He, G. Yu, and G. Y. Li, "Joint communication and computation resource allocation for cloud-edge collaborative system," in *Proceedings of the 2019 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2019.

[24] H. Zhang, S. Chen, P. Zou, G. Xiong, H. Zhao, and Y. Zhang, "Research and application of industrial equipment management service system based on cloud-edge collaboration," in *Proceedings of the 2019 Chinese automation congress (CAC)*, November 2019.

[25] M. B. Mollah, M. A. K. Azad, and A. Vasilakos, "Secure data sharing and searching at the edge of cloud-assisted Internet of things," *IEEE Cloud Computing*, vol. 4, no. 1, pp. 34–42, 2017.

[26] M. Blaze, G. Bleumer, and M. Strauss, "Divertible protocols and atomic proxy cryptography," *Advances in Cryptology — EUROCRYPT'98*, vol. 1403Springer, Lecture Notes in Computer Science, , pp. 127–144, 1998.

[27] L. Xu, X. Wu, and X. Zhang, "A certificateless proxy re-encryption scheme for secure data sharing with public cloud," in *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, ACM, Seoul, South Korea, May 2012.

[28] O. Blazy, X. Bultel, and P. Lafourcade, "Two secure anonymous proxy based data storages," in *Proceedings of the SECRYPT*, Lieusaint, France, July 2016.

[29] C. Zuo, J. Shao, J. K. Liu, G. Wei, and Y. Ling, "Fine-grained two-factor protection mechanism for data sharing in cloud storage," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 1, pp. 186–196, 2018.

[30] G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distributed storage," in *Proceedings of the Network and Distributed System Security Symposium,NDSS*, San Diego, CA, USA, February 2005.

[31] H. Guo, Z. Zhang, and J. Xu, "Non-transferable proxy re-encryption," *The Computer Journal*, vol. 62, no. 4, pp. 490–506, 2019.

[32] H. Guo, Z. Zhang, J. Xu, N. An, and X. Lan, "Accountable proxy re-encryption for secure data sharing," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 1, pp. 145–159, 2021.

[33] A. Lysyanskaya, S. Micali, L. Reyzin, and H. Shacham, "Sequential aggregate signatures from trapdoor permutations," in *Proceedings of the Advances in Cryptology-EUROCRYPT 2004, International Conference on the Theory and Applications of Cryptographic Techniques*, Interlaken, Switzerland, May 2-6, 2004.

[34] D. Boneh and X. Boyen, "Efficient selective-ID secure identity-based encryption without random oracles," in *Advances in Cryptology EUROCRYPT 2004*Springer, NY, USA, 2004.

[35] B. Waters, "Efficient identity-based encryption without random oracles," *Advances in Cryptology – EUROCRYPT 2005*, vol. 3494Springer, Lecture Notes in Computer Science, , pp. 114–127.

[36] J. Zhang, Z. Zhang, and Y. Chen, "PRE: stronger security notions and efficient construction with non-interactive opening," *Theoretical Computer Science*, vol. 542, pp. 1–16, 2014.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 6067678, 10 pages
https://doi.org/10.1155/2021/6067678

Research Article

Application of Adaptive Virtual Reality with AI-Enabled Techniques in Modern Sports Training

Yajun Zhang[1] **and Sang-Bing Tsai**[2]

[1]*Department of Physical Education, Henan University of Animal Husbandry and Economy, Henan 450046, China*
[2]*Regional Green Economy Development Research Center, School of Business, Wuyi University, Nanping, China*

Correspondence should be addressed to Yajun Zhang; pandajunjun03@126.com and Sang-Bing Tsai; sangbing@hotmail.com

Received 9 August 2021; Revised 31 August 2021; Accepted 7 September 2021; Published 18 September 2021

Academic Editor: Han Wang

Virtual reality technology has many characteristics, such as high immersion, dynamic interactive response, and multidimensional information digitization. These characteristics are needed by the combination of modern sports development and scientific and technological progress. The introduction of virtual reality technology into the field of sports will be conducive to the scientific training of sports competition and the digital research of technical theory and to the rapid development of modern sports. This paper constructs the framework of sports training based on virtual reality technology and designs a motion capture data algorithm based on behavior string, which successfully improves the advantages of virtual reality technology. In addition, the training experiment method is used to verify the effectiveness and superiority of the system, focusing on the high-school students learning tennis for the first time as the experimental object to study the differences in training effects between the traditional training method and the training method using virtual reality technology. The training effect of tennis is better than the traditional training method, and the difference is significant; the internal motivation of training psychology is better than the traditional training method, and the difference is significant. Compared with the traditional training method, this training method can stimulate students' interest in training, improve students' training effect, and promote students' psychological internal motivation to continue training. This study provides useful enlightenment for the further application of virtual reality technology and in the modern sports training for various sports events.

1. Introduction

Since the twentieth century, human society has entered the information age. Modern science and technology has a trend of high differentiation and high concentration and has also entered a new stage of multidisciplinary comprehensive utilization [1, 2]. A large number of modern research methods and technologies have been transplanted to sports training and scientific research. Many achievements of modern high technology have also begun to penetrate into the field of competitive sports and have had a series of profound effects on modern competitive sports training [3]. In particular, the rapid development of virtual reality technology provides a broader space for the application of high technology in the field of competitive sports [4]. Summarizing the successful experience of Beijing Olympic Games, it is found that using virtual reality technology for simulation training can realize the human motion measurement method from traditional human eye observation to high-precision motion capture and analysis [5]; The two changes from the experience-based method to the human motion analysis method of human motion simulation and simulation will be faster and more effectively improve China's training level and competitive level in these projects and realize the magic weapon of the Olympic glory plan. In order to ensure China's dominant position in the total number of gold medals and medals in the Olympic Summer Olympic Games, it is particularly important to timely track the development of virtual reality technology and deeply study the practical significance of virtual reality technology to the development of competitive sports [6].

Some researchers are committed to the application of virtual reality in physical training [7]. Akbas et al. [8] hold that "the application of virtual reality technology in physical education includes five aspects: knowledge learning, virtual sports experiment, skill training, physical education network distance teaching, and academic exchange. Wei et al. [9] and others proposed that the use of virtual reality technology can make physical education courseware and carry out physical training. Lohre et al. [10] hold that the application of virtual reality technology in physical education includes five aspects: physical education teaching and training, making physical education network distance teaching, accident prevention, and academic exchange. Arrighi et al. [11] hold that "the application of virtual reality technology in college teaching includes virtual campus, virtual classroom, virtual experiment, and FGH courseware design. Wu and Zhou [12] put forward that "there are four points in the application of virtual reality technology in Physical Education Teaching: one is to simulate sports scenes and break through the limitation of time and space; second, innovate the educational model to improve students' interest in learning and teaching effect; third, digital learning, tracking learning data, and effectively monitoring learning behavior; fourth, it has strong functional development and broad application prospects in college physical education [13]." Synovec et al. [14] hold that the application of computer-based "virtual reality" technology in college sports training has the following four points: first, the comparison between sports technical action and reality is completed through "virtual reality" technology; the second is to use "virtual reality" technology to realize the comparison between virtual action and virtual reality; third, build different training environments through "virtual reality" technology; fourth, the use of computer "virtual reality" technology can realize remote interactive training [15].

The application of virtual reality technology in college physical education teaching and training is the main trend of college physical education development. It can effectively stimulate students' sensory cognition and thinking mode and plays an important role in strengthening the effect of students' physical education learning and training. Therefore, the contributions of this paper are as follows: (1) design the motion training framework based on virtual reality technology; (2) design the motion data capture algorithm based on behavior string to capture the motion trajectory of the target; (3) the control group was designed and compared with the traditional training methods to verify the advantages of the modern training system based on VR. Through the research and understanding of virtual reality technology, find out the advantages of virtual reality technology and make it flexibly applied to high-school tennis training, which cannot only enrich the training content and training means but also improve the training efficiency and training effect of tennis class and create a good learning atmosphere and learning environment, so as to improve the training quality. Therefore, this paper constructs the framework of sports training based on virtual reality technology and designs a motion capture data algorithm based on behavior string, which successfully improves the advantages of virtual reality technology. With the improvement of students' interest in learning, the mastery of tennis skills is also improving, which is conducive to the further research of tennis training theory.

2. Structure and Design of Sport Training Based on Virtual Reality

2.1. Architecture Design of Cloud Platform. According to the problems existing in the current sports training system based on virtual reality, the motion capture technology is introduced into the human motion posture simulation and analysis system [16]. The data of the files obtained by motion capture are analysed by the least square method, and the results are applied to the simulation system, which will help to solve many problems and deficiencies in the human motion posture simulation system [17, 18]. For the research work of using motion capture technology to obtain data and extract human motion posture parameters, the development of this system is completed with the support of Bessel curve theory. The physical training system based on virtual reality is mainly composed of three functional modules: data acquisition module, motion posture capture module, interface management module, and system kernel module. The overall framework of the sports training system based on virtual reality is shown in Figure 1.

The physical training system based on virtual reality adopts a hierarchical architecture, which is divided into user visualization layer, signal processing layer, and signal acquisition layer from top to bottom. The user visualization layer is responsible for providing a "container" for all functional modules, receiving various messages provided by users, and transmitting the received information to the lower layer, that is, the signal processing layer, so as to complete the interaction between users. The user interface module is a portable module of the presentation layer, which is mainly used to present the user interface with certain personality and style. Although the user interface module is relatively single and the personality is not very distinctive, it has very good interaction and powerful function and can complete the basic functions required by the system. The three modules of model, Coleman filter, and human motion posture simulation in the signal processing layer belong to the system core. The data collected by 3D suit motion capture equipment are applied to human motion posture simulation, and combined with Bezier curve and numerical data to complete realistic human motion posture simulation, the realistic simulation of the whole human motion posture is completed in these three modules; the Bezier curve module is an extension module of the logic layer. It is responsible for freely editing curves to achieve the purpose of intuitive and realistic simulation of human motion. The motion data acquisition module is combined with the motion data acquisition module of the human body model to achieve the purpose of reading the motion data of the human body model.

Practice has proved that, with the hierarchical architecture, the system has clearer hierarchy, lower coupling, and stronger scalability. Therefore, the system is easy to maintain, improve, and expand. The independence of modules

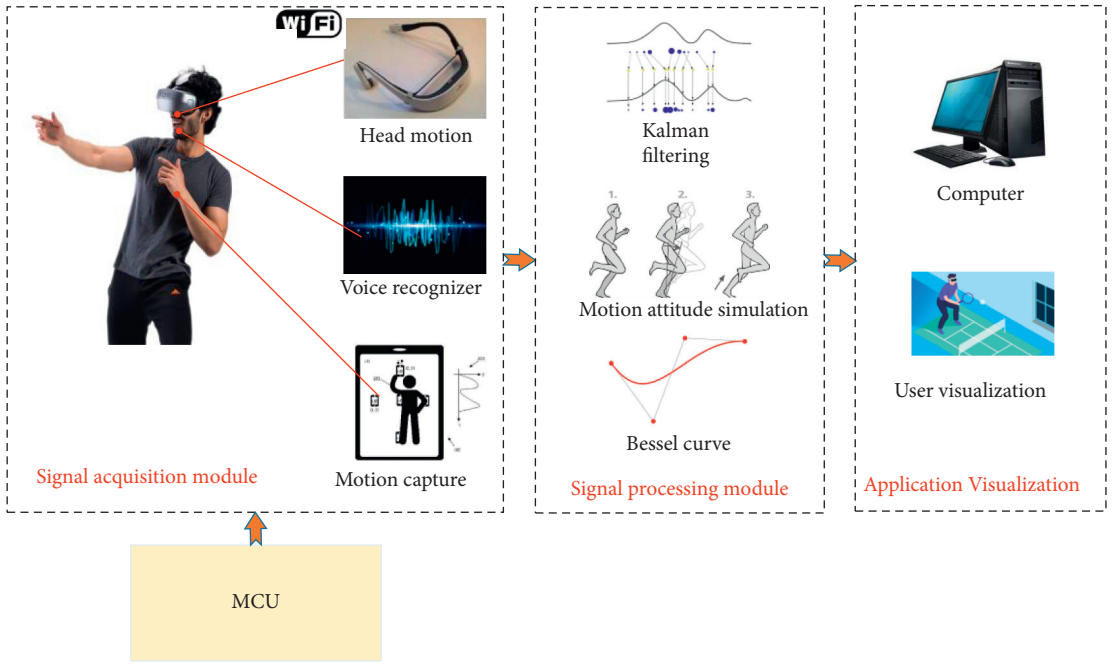


Figure 1: Structure design for sport training based on virtual reality.

makes it easy to modify, replace, or reuse modules in each layer. As long as the interface between modules remains unchanged, a layer can also provide a variety of interchangeable specific implementations. The design supported by the hierarchical system reflects the increasing level of abstraction, simplifies complex problems into a series of simple incremental steps, and makes the system highly scalable and flexible.

2.2. Overall Design of Whole Body Motion Capture System. In order to improve the accuracy of human motion capture data, the system integrates the data of two Kinect devices [18, 19] and adopts the client-server mode, and each device is connected with the corresponding client PC. Two client computers use SDK to obtain skeleton data and send it to server-side PC through local Ethernet. All bone data are processed in the server, and then, the fused bone data is transmitted to the Unity3d scene for visualization [20]. Because Kinect has the disadvantage of self-occlusion, it cannot obtain the full picture of human posture. In this paper, two Kinect devices are taken as an example to form a dual Kinect human motion capture system, which can obtain more accurate human motion data through data fusion and expand the range of human activities. When placing Kinect, try to obtain the whole picture of the human body. The total viewing angle of the two Kinect to obtain data is as large as possible, so the orthogonal layout scheme is adopted. At this time, the mutual interference of infrared rays is also the smallest. According to the height of the human body and the best capture range of Kinect, the two Kinect are placed at a vertical height of 70 cm from the ground and 2 m away from the fusion centre. The system scheme design is shown in Figure 2.

In addition, in order to realize the data transmission between the client and server, it needs to be completed through data transmission protocol [21]. The commonly used communication transmission protocols of the two hardware devices include TCP protocol and UDP protocol. The client converts the bone and joint data obtained from Kinect sensor into OSC information format for transmission and sends it through UDP protocol. The server takes Unity3d engine as the main body to realize data fusion and visualization and adds 0sc component to Unity to receive the data sent from the client. Set the IP address and the ports for sending and receiving information, establish a connection between the client and the server, connect the client and the server PC to the same LAN at the same time, and complete the data transmission through the UDP protocol.

3. Motion Data Capture Research Based on Behavior String

3.1. The Process Design for Behavior String Capture. The purpose of motion data behavior segmentation is to automatically separate human motion capture data containing different motion types to form motion fragments with independent semantic information, which is convenient for storage and reuse in the database and lay a foundation for motion analysis and synthesis. A motion training framework based on virtual reality technology is constructed, and a motion capture data algorithm based on behavior string is designed. The motion capture data algorithm based on behavior string can encode and reconstruct the motion trajectory according to the previous motion frames to form a behavior string and then capture the target motion data indirectly according to the historical behavior string, which can significantly improve the accuracy of motion capture.

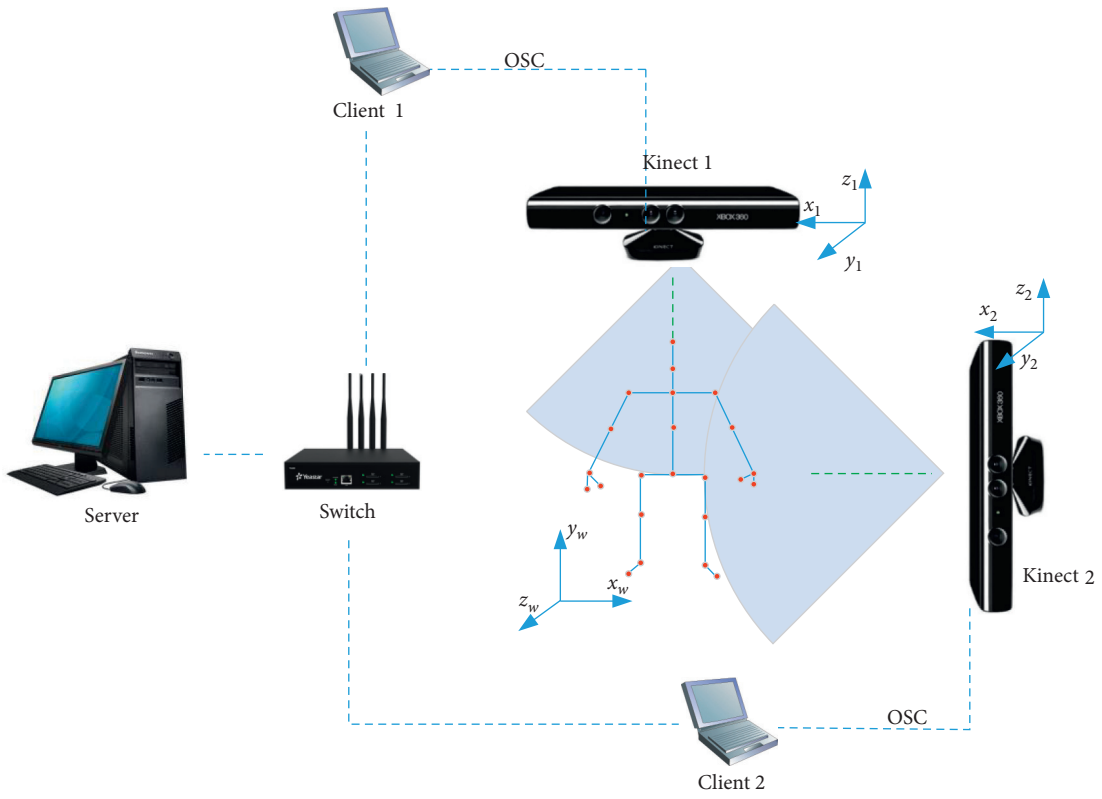


Figure 2: System scheme design based on two Kinects.

This paper proposes a new string representation method based on human motion capture data, which represents the high-dimensional human motion capture data sequence in the form of string. In this method, the human motion capture data is regarded as a high-dimensional data point set, the density-based clustering method w is used for clustering, and the obtained categories are represented by characters. Through time sequence recovery processing, the off-dimensional data point set represented by characters is reordered according to the time sequence of the original motion sequence to obtain the string corresponding to the original motion capture sequence, and this string is called behavior string (BS). Finally, by analysing the behavior string, the behavior segmentation of the human motion capture data sequence is realized and the motion cycle corresponding to various behaviors is extracted.

In this paper, the distance between the human body model and the local motion points of the human body model is regarded as the similarity between the two motion points of the human body model, and the distance between the two motion points and the virtual human body model is calculated, respectively, and according to the product of local density and distance, the cluster centers are determined, and the data points in the noncluster are clustered based on density. The nest dimension data point set represented by the clustered characters is reordered according to the timing of the original motion sequence to obtain the behavior string corresponding to the original motion capture sequence, so as to realize the string representation of human motion capture data sequence, which is shown in Figure 3.

3.2. Motion Behavior Cluster Based on Improved Mean-Shift Algorithm. In this paper, the quotient dimension data points representing the original motion sequence are clustered by clustering. There are many clustering methods. The real K-means and cure clustering methods need to manually specify the number of cluster centers, and the method in this paper can automatically find the number of clusters. The BIRCH clustering algorithm has good clustering performance for spherical clusters, but it cannot work well for irregular clusters. The method in this paper can calculate the local density by using the Gaussian kernel function, and it can also cluster nonspherical clusters.

Density estimation is to estimate the probability density distribution from a set of observations of unknown probability density distribution. There are usually two methods: parametric method and nonparametric method. Given n data points in d-dimensional Euclidean space, the kernel density is estimated with K, and the width of the window is h; therefore,

$$\widehat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right). \tag{1}$$

The mean-shift algorithm actually uses the gradient method to iteratively calculate the extreme points of the probability density function. According to the probability density function of the data, the gradient is

$$\nabla \widehat{f}(x) = \frac{n_x}{n\left(h^d c_d\right)} \frac{d+2}{h^2} \left\{ \frac{1}{n_x} \sum_{x_i \in S_h} (x_i - x) \right\}, \tag{2}$$

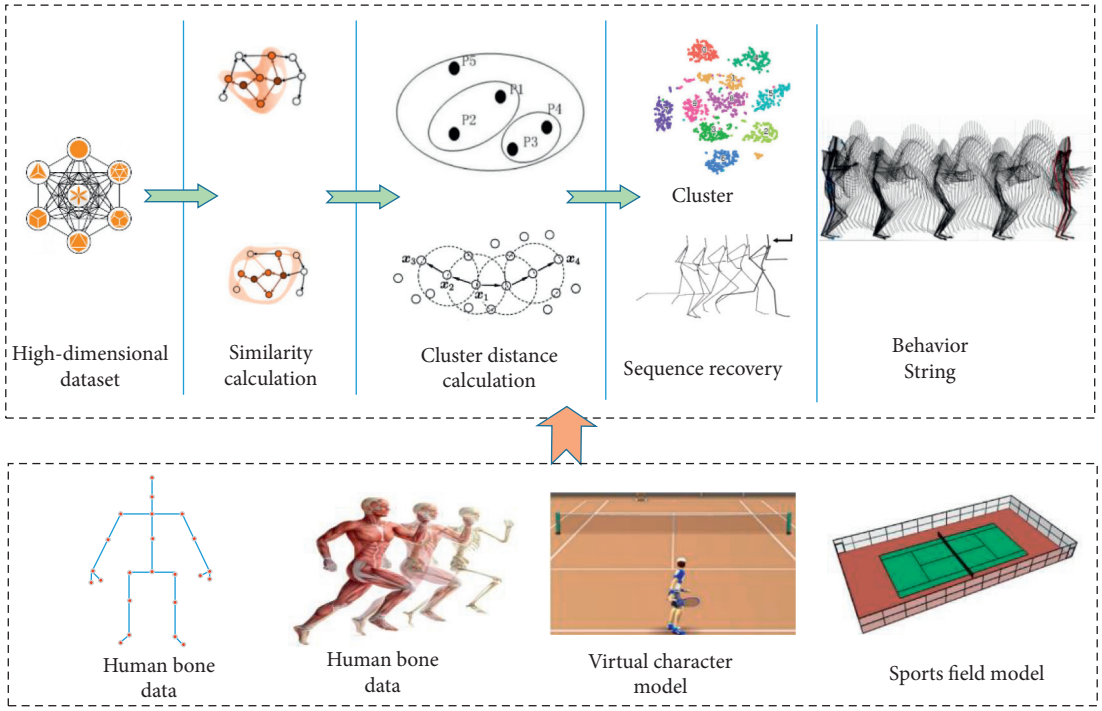


Figure 3: The process of teaching mode construction based on DBN-DELM.

where the range S_h and (z) is a hypersphere with radius h and volume $h^d c_d$, and its center is x and contains n_x data points.

The mean-shift vector $M(x)$ is defined as

$$M(x) = \frac{1}{n_x} \sum_{x_i \in S_h} (x_i - x) = \frac{1}{n_x} \sum_{x_i \in S_h} x_i - x. \quad (3)$$

Therefore, formula (3) can be rewritten as

$$M(x) = \frac{h^2}{d+2} \frac{\nabla \hat{f}(x)}{\hat{f}(x)}. \quad (4)$$

Therefore, the mean-shift vector is the difference between the local mean and the center of the window and the direction points to the peak or valley of the probability density function. The improved mean-shift algorithm is shown in Figure 4. The implementation of the mean-shift algorithm is shown: (a) calculate the probability distribution and similarity coefficient of the candidate target at $y0$ according to the initial position $y0$; (b) calculate the weight according to the formula; (c) calculate the new position $y1$ according to the formula; (d) calculate the probability distribution and similarity coefficient of the candidate target at $y1$; (e) if $|y1\text{-}y0| < C$, the iteration ends; otherwise, $y0 = y1$, skip to Step 1, and continue. The mean-shift algorithm rises (falls) at the fastest speed with variable step size and finally converges to the peak (Valley) point of the probability density function, see Figure 4.

4. Simulation and Results Analysis

4.1. Experimental Environment, Test Method, and Equipment. In order to verify the effectiveness of the physical training system based on virtual reality technology in this paper, the verification environment is designed as follows: run the algorithm in the software environment of MATLAB 2011, the running platform is Intel Core 2 Duo, the CPU is 3.00 GHz, and the memory is 2 GB. This paper verifies the superiority and reliability of the algorithm and system from three angles: (1) capture motion data in VR system and verify the accuracy of motion data capture evaluation based on behavior string; (2) compared with the traditional motion capture algorithm, the superiority of motion data capture evaluation based on behavior string is verified; (3) the control group was set up and compared with the traditional control group to verify the superiority of the system.

In addition, the two technicians need to make a normative judgment on the forehand and backhand strokes of the two groups of students under the same and relatively simple circumstances. The final grade is the average of the scores given by the two teachers. The scoring standard is 100, which is the full score, in which the action integrity and action coordination account for 50 points, respectively. The evaluation of action integrity is divided into preparation posture, backswing lead, swing, and follow-up, corresponding to the full score of 10, 15, 15, and 10, respectively; the scores of movement coordination are normative coherence, force coordination, natural pace movement, and stable centre of gravity, with full scores of 15, 15, 10, and 10, respectively.

Test method: the instructor stands two meters in front of the right of the tested students and throws the tennis ball to the tested students. The student stands at the centre line of the field with a racket and tests 10 balls in total. When the tested student tests the forehand, the forehand swings and hits the ball diagonally opposite the field, and the backhand needs to swing and hit the ball diagonally opposite the field.

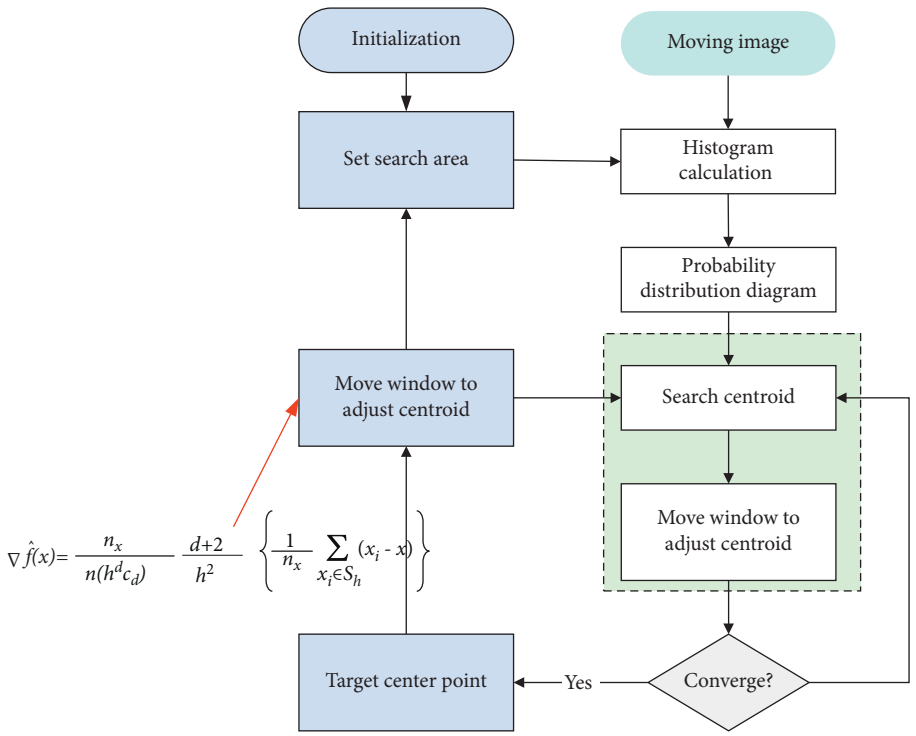


Figure 4: Implementation flow of the mean-shift algorithm.

The two tennis technicians scored strictly according to the standard.

The experimental equipment mainly includes a tape measure, a whistle, a stopwatch, a height tester, several markers, a weight tester, a vital capacity tester, 80 blowers, 40 tennis rackets, 200 tennis balls, 2 tennis courts, 1 laptop, 5 VR equipment, zero basic tennis introduction video, and highlights of tennis stars.

4.2. The Motion Data Capture Evaluation Based on Behavior String. The efficiency of the big data platform is not verified. This paper selects the common Hadoop framework platform for comparison; the frameworks of Hadoop are shown in Figure 5.

According to the components of elbow joints in x, y, and z directions, their motion is periodic. The law of human movement is not obvious, but it can be seen that the elbow joint is doing curve movement, and there are still some unsmooth and folding phenomena in the process of movement.

The arbitrary parameters of human motion mainly include joint displacement, linear velocity/acceleration, joint angle, and angular velocity/acceleration. From the human joint point sampling data, the parameters of each joint can be obtained through calculation and analysis. The position information of each joint can be obtained directly through the action capture equipment. The displacement of the joint can be obtained through calculation. The change of displacement can be obtained through the subtraction operation of the data of the first and second frames. The linear velocity = displacement ÷ time can be obtained by using the

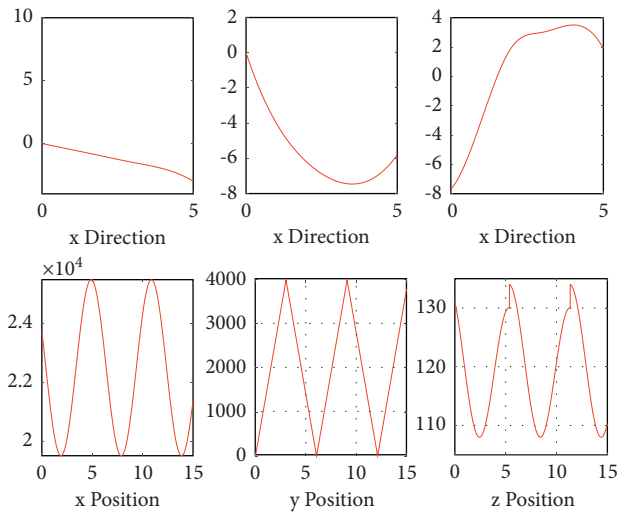


Figure 5: Motion curve based on behavior string.

motion trajectory of the joint. Acceleration reflects a change in the motion speed of human joints, which can also be obtained by using the change track of human motion speed. The kinematics of the human upper limb elbow joint is analysed, and its motion parameters are obtained. The motion trajectory of the elbow joint is fitted with its regular curve by the least square method. According to the principle of forward kinematics, because the wrist joint is the child node of the elbow joint, the motion of the elbow joint drives the motion of the elbow joint. Accordingly, when the wrist joint does not move, the motion parameters of wrist and elbow are consistent. However, in the process of data

acquisition, it cannot ensure that the wrist joint completely follows the movement of the elbow joint. In the actual processing process, it needs to be processed by matrix operation.

4.3. The Motion Data Capture Evaluation Based on Behavior String. In order to verify the superiority of the algorithm, we mainly verify it from the perspective of average error and average time. The results are shown in Figures 6(a) and 6(b).

The algorithm in this paper is compared with the traditional particle filter algorithm (PF), literature algorithm (VLMM), and literature algorithm (single GPDM) for 15 times. The comparison diagrams of average tracking accuracy and average time consumption are shown in Figures 6(a) and 6(b). As shown in Figure 6(a), under the same conditions, the algorithm proposed in this paper has higher tracking accuracy than other comparison algorithms and can well meet the accuracy requirements in 3D hand tracking. The precision of the traditional PF algorithm is the same as that of the VLMM algorithm. The single GPDM algorithm reduces the effect of real-time human hand parameters because it depends too much on the model. It has low accuracy and can only be used in the three-dimensional tracking process with low accuracy requirements. As can be seen from Figure 6(b), the traditional PF algorithm and VLMM algorithm consume a lot of time, which is a great bottleneck for the real-time tracking. The method in this paper has been greatly improved in time consumption. In the particle sampling stage, it effectively avoids blind sampling, realizes the purpose of reducing dimension, and can fully meet the needs of interaction in terms of real time. However, compared with the single GPDM algorithm and data glove algorithm, there is still a certain gap, which is also where the method in this paper needs to be further improved.

4.4. The Training System Evaluation Based on Virtual Reality Technology. Before tennis training, the physical function of the students participating in the experiment was evaluated according to the standards of height, weight, 50 m sprint, vital capacity, and standing long jump in the national student physical health standard. In the tennis skill test stage, the test content includes two items: forehand and backhand oblique stroke and forehand and backhand accuracy stroke.

The main research content of this paper is to test whether there is a significant difference in teaching effect between the new tennis training method and the traditional training mode using virtual reality technology. In order to eliminate the influence of other factors on the experiment, we should first test the five basic functions of the students in the experimental group and the control group before the experiment because the experimental samples are from students aged 15–17, and the students are in the period of growth and development, and it is very necessary to carry out statistical test and analysis. This experiment is divided into two groups, 20 people in each group, 10 male students and 10 female students. The comparison of the data in Table 1 shows that there is no significant difference in the physical state between the two groups before tennis class ($P > 0.05$), which can be used as the sample of this experiment.

40 students without tennis training foundation were randomly selected as the research objects. They were divided into two groups: experimental group and control group, with 10 male and 10 female students in each group. At the end of the two courses, the differences between the two groups were analysed.

Experiment 1. Ball feeling test.

The ball feeling test, as the name suggests, is the feeling of the ball. To be exact, it refers to the ability to predict the ball at the moment of hitting the ball as well as the physical coordination ability and feeling of the ball when hitting the ball. Before the experiment, all students participating in the experiment should be tested for the ball feeling. It should be emphasized that the ball feeling test in the static state is done in this experiment. The ball feeling test mainly includes two aspects: the first is to test the students' perception of the ball in motion, and the second is to test the students' sensitivity to the ball. In the process of learning tennis, the ball feeling and ball nature are very important. The better the students' sense of the ball, the stronger their control over the ball. At the same time, the better the sense of the ball, the stronger their belief in tennis and interest will increase, Therefore, cultivating students' ball sense is very important for tennis teaching. Therefore, in the teaching process, teachers should pay attention to cultivating students' ball sense and further enhance students' interest in learning tennis. The quality of ball sense will also affect the progress of teaching. Therefore, before doing the experiment, students should have a professional ball sense test. The results are shown in Table 2 and Figure 7:

As shown in Figure 7, 35% of the people in the virtual training experimental group have achieved excellence, while only 10% of the people in the control group (90–100) have achieved excellence, and only one third of the people in the virtual training group have achieved excellence. In the good range (80–89), the number of students from both sides is basically the same, 40% in the virtual training group, 35% in the control group, and 40% in the general grade range (70–79), while only 20% in the virtual training experimental group. Therefore, in general, most of the students in the control group will get medium grades and those in the pass range, and the virtual training group accounted for 5%, while the control group accounted for 10%.

Experiment 2. Accuracy test of forehand and backhand stroke.

After the forehand and backhand oblique stroke test, further judge the accuracy of students' stroke, so as to better explore students' mastery of tennis skills. The experimental results are shown in Table 3. According to Table 3, it can be judged that the scores of the accuracy of forehand oblique stroke of the experimental group and the students of traditional mode are 11.28 and 10.03, respectively, and the difference between the two is 1.25 points only; the average score of the backhand oblique shot was 9.32 in the experimental group, 8.18 in the traditional model control group,

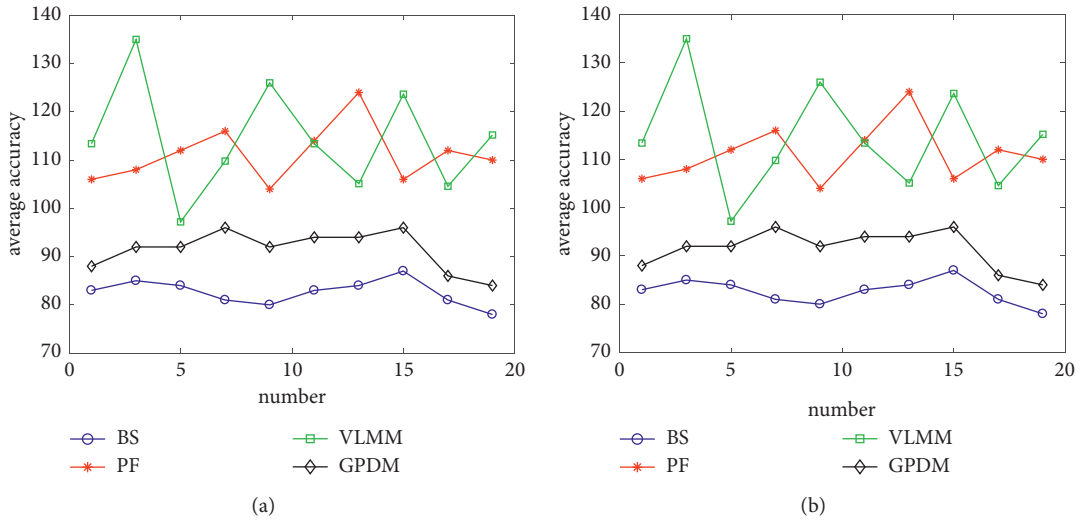


Figure 6: The comparison of motion data capture evaluation. (a) Comparison of average accuracy. (b) Comparison of average operation time.

Table 1: The table for the testing group.

	Experimental group		Control group		t	P
	$\overline{x}$	S	$\overline{x}$	S		
Height (man)	174.25	5.28	173.25	4.61	0.274	>0.05
Weight (man)	66.36	4.02	65.85	3.96	0.404	>0.05
Height (woman)	165.34	4.88	164.27	4.63	0.711	>0.05
Weight (woman)	50.05	3.74	49.76	3.38	1.002	>0.05

Table 2: The comparison of the ball feeling test.

	Experimental group		Control group		t	P
	$\overline{x}$	S	$\overline{x}$	S		
20 s racket in place	22.16	3.14	22.96	3.83	−0.722	>0.05
20 s smash	23.56	3.78	21.92	3.05	1.510	>0.05
20 s landing smash	17.26	4.69	16.87	4.15	0.279	>0.05

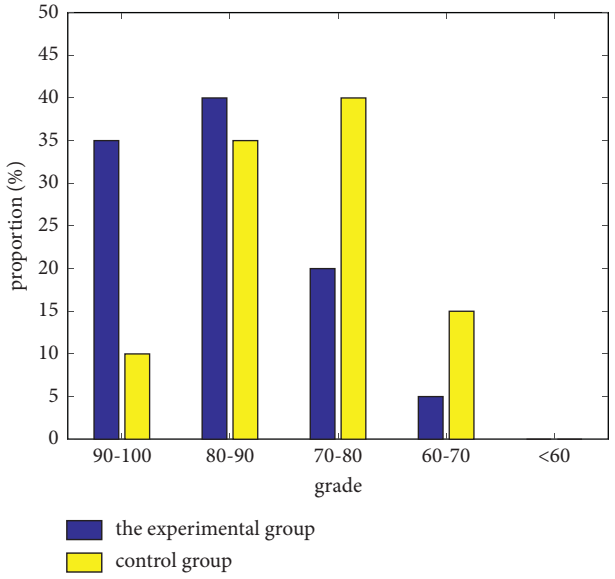


Figure 7: The score of the ball feeling test.

Table 3: The comparison of the ball feeling test.

	Experimental group		Control group		t	P
	$\overline{x}$	S	$\overline{x}$	S		
Forehand stroke	11.38	1.74	10.02	1.34	2.546	$<0.05^{*}$
Backhand stroke	9.32	1.12	8.30	1.05	3.34	$<0.05^{**}$

$P < 0.05^{*}$ means difference; $P < 0.05^{**}$ means obvious difference.

and 1.14 in the difference. Only the experimental results show that there are significant differences between the two teaching methods ($P < 0.05$). The accuracy of the students' forward and back hand oblique stroke is higher than that of the traditional teaching mode. The reason may be that when students carry out virtual teaching learning, the classroom content teaching is lively and interesting and has professional skills. Students not only pay attention to the training of skills but also cultivate the interest in learning tennis. The lively and interesting teaching mode makes students more focused and conscientious, which greatly mobilizes the enthusiasm and initiative of students' learning. Students in virtual tennis technology learning have greatly improved the ball processing and can adjust the body and ball distance and the degree of centre of gravity deviation according to their own situation. With the gradual improvement of proficiency in the technology, it is better and better to control themselves and the ball.

5. Conclusion

This paper constructs the framework of sports training based on virtual reality technology and designs a motion capture data algorithm based on behavior string, which successfully improves the advantages of virtual reality technology. In addition, the training experiment method is used to verify the effectiveness and superiority of the system, focusing on the high-school students learning tennis for the first time as the experimental object to study the differences in training effects between the traditional training method and the training method using virtual reality technology. The test results showed that, in terms of technical skill evaluation, the virtual reality experimental group achieved excellent results, which was significantly different from the traditional training group ($P < 0.05$). The average scores of forehand and backhand oblique stroke in the experimental group were 16.28 and 13.32, respectively, while the average scores of the traditional mode training group were 14.32 and 12.04, respectively ($P < 0.05$); in terms of the accuracy of forehand and backhand stroke, the average scores of the virtual reality experimental group were 11.28 and 9.32 and those of the traditional mode training group were 10.03 and 8.18, respectively ($P < 0.05$). In terms of the score of forehand and backhand hitting skills, the average scores of the subjects in the experimental group are 82.67 and 78.16, and the average scores of the students in the control group are 75.11 and 71.35 ($P < 0.05$). All these show that the training method of the experimental group is more suitable for students' tennis training. To process virtual reality, the video based on the existing framework can be compressed and optimized. In order to better match the processing characteristics of the virtual reality video, improve more coding efficiency and reduce more coding time, and reducing code stream output is the focus of future research.

Data Availability

The data used to support the findings of the study are included within the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] D. Thomas-Vazquez and D. Singh, "Virtual reality in surgical training," *International Journal of Surgery*, vol. 37, no. 4, pp. 34–38, 2021.

[2] C. Li and J. Cui, "Intelligent sports training system based on artificial intelligence and big data," *Mobile Information Systems*, vol. 2021, no. 1, Article ID 9929650, 11 pages, 2021.

[3] M. Zahabi and A. Razak, "Adaptive virtual reality-based training: a systematic literature review and framework," *Virtual Reality*, vol. 24, no. 1, pp. 725–752, 2020.

[4] F. Guo, F. R. Yu, H. Zhang, H. Ji, V. C. M. Leung, and X. Li, "An adaptive wireless virtual reality framework in future wireless networks: a distributed learning approach," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8514–8528, 2020.

[5] L. Rigamonti, U.-V. Albrecht, C. Lutter, M. Tempel, B. Wolfarth, and D. A. Back, "Potentials of digitalization in sports medicine," *Current Sports Medicine Reports*, vol. 19, no. 4, pp. 157–163, 2020.

[6] C. Pagé, P. M. Bernier, and M. Trempe, "Using video simulations and virtual reality to improve decision-making skills in basketball," *Journal of Sports Sciences*, vol. 37, no. 3, pp. 1–8, 2019.

[7] T. R. Kautzmann and P. A. Jaques, "Effects of adaptive training on metacognitive knowledge monitoring ability in computer-based learning," *Computers & Education*, vol. 129, pp. 92–105, 2019.

[8] A. Akbaş, W. Marszałek, A. Kamieniarz, J. Polechoński, K. J. Słomka, and G. Juras, "Application of virtual reality in competitive athletes - a review," *Journal of Human Kinetics*, vol. 69, no. 1, pp. 5–16, 2019.

[9] S. Wei, P. Huang, R. Li, Z. Liu, and Y. Zou, "Exploring the application of artificial intelligence in sports training: a case study approach," *Complexity*, vol. 2021, no. 8, Article ID 4658937, 8 pages, 2021.

[10] R. Lohre, A. J. Bois, J. W. Pollock et al., "Effectiveness of immersive virtual reality on orthopedic surgical skills and knowledge acquisition among senior surgical residents," *JAMA Network Open*, vol. 3, no. 12, Article ID e2031217, 2020.

[11] C. Arrighi, L. Rossi, E. Trasforini et al., "Quantification of flood risk mitigation benefits: a building-scale damage

assessment through the RASOR platform," *Journal of Environmental Management*, vol. 207, pp. 92–104, 2018.

[12] Z. Wu and C. Zhou, "Equestrian sports posture information detection and information service resource aggregation system based on mobile edge computing," *Mobile Information Systems*, vol. 2021, no. 3, Article ID 4741912, 10 pages, 2021.

[13] J. P. Hynes, J. Walsh, and T. P. Farrell, "Role of musculoskeletal radiology in modern sports medicine," *Seminars in Musculoskeletal Radiology*, vol. 22, no. 5, pp. 582–591, 2018.

[14] J. Synovec, L. Plumblee, W. Barfield, and H. Slone, "Orthopedic in-training examination: an analysis of the sports medicine section-an update," *Journal of Surgical Education*, vol. 76, no. 1, pp. 286–293, 2019.

[15] F. W. Otte, S. K. Millar, and S. Hüttermann, "How does the modern football goalkeeper train? – an exploration of expert goalkeeper coaches' skill training approaches," *Journal of Sports Sciences*, vol. 38, no. 4, pp. 1–9, 2019.

[16] J. Pang, X. Li, and X. Zhang, "Coastline land use planning and big data health sports management based on virtual reality technology," *Arabian Journal of Geosciences*, vol. 14, no. 12, pp. 1–15, 2021.

[17] A. Ślesicka and A. Kawalec, "An application of the orthogonal matching pursuit algorithm in space-time adaptive processing," *Sensors*, vol. 20, no. 12, p. 3468, 2020.

[18] X. Pan, "Brain-machine interface training system of motor imagery based on virtual reality," *NeuroQuantology*, vol. 16, no. 6, pp. 715–719, 2018.

[19] Y. Song, "Application of blockchain-based sports health data collection system in the development of sports industry," *Mobile Information Systems*, vol. 2021, no. 1, Article ID 4663147, 6 pages, 2021.

[20] P. M. Cogswell, J. D. Trzasko, and E. M. Gray, "Application of adaptive image receive coil technology for whole-brain imaging," *American Journal of Roentgenology*, vol. 216, no. 2, pp. 1–8, 2020.

[21] O. Besson, F. Vincent, and S. Matteoli, "Adaptive target detection in hyperspectral imaging from two sets of training samples with different means," *Signal Processing*, vol. 181, Article ID 107909, 2021.

Hindawi
Mobile Information Systems
Volume 2021, Article ID 4132622, 8 pages
https://doi.org/10.1155/2021/4132622

Research Article

Elman-Based Secure Data Transmission Quality Prediction for Complex IoT Networks

Fagen Yin

College of Physical Science and Engineering, Yichun University, Yichun 336000, Jiangxi, China

Correspondence should be addressed to Fagen Yin; ynx522@163.com

Received 18 June 2021; Revised 28 July 2021; Accepted 17 August 2021; Published 26 August 2021

Academic Editor: Qinglin Zhao

The information age has brought earth-shaking changes. For interconnection of all things, the data transmission has widely employed the Internet of Things (IoT). The IoT transmission faces complex environments. The secure data transmission is very important for mobile IoT networks. The secure data transmission quality prediction is investigated for mobile IoT networks. The probability of strictly positive secrecy capacity (SPSC) is used to evaluate the secure data transmission quality, and the expressions are first derived. Then, employing Elman network, a secure data transmission quality intelligent prediction approach is proposed. The extensive simulations are run to evaluate the proposed approach. The simulation results show that the Elman-based approach can achieve a higher quality precision than other methods. The Elman-based approach also can achieve a lower time complexity.

1. Introduction

With the explosive growth of mobile applications, Internet of things (IoT) networks are widely used to transmit data [1]. The fifth generation (5G) mobile communication also has been widely used in mobile IoT networks [2, 3]. Different 5G applications widely use sea-land-air mobile communication networks [4, 5]. The global and diversified application will provide quick and convenient services for IoT users. However, due to IoT mobility and the diversity of IoT networks, the physical layer security (PLS) of 5G mobile IoT networks is facing many challenges [6].

PLS of 5G IoT networks is a research hot spot [7]. Low-complexity schemes for IoT PLS were presented in [8]. In [9], power control mechanism and antenna transmission scheme were used to realize the secure data transmission in cognitive wiretap networks. Considering the mobile healthcare networks, Xu et al. [10] investigated the PLS performance using the deep learning method. In [11], the authors used compressed sensing and cooperative schemes to achieve the secure transmission. Considering the user and relay selection, Fan et al. [12] analyzed two criteria and investigated the achievable PLS performance. The authors of [13] analyzed the upper and lower bounds on PLS performance over dependent fading channels.

The IoT data transmission faces a wide variety of scenarios and complex environments. The PLS issue is more and more serious. However, predicting and evaluating the secure data transmission quality are very difficult. Recently, machine learning techniques are applied in 5G wireless communications [14, 15]. In medical IoT, support vector machine (SVM) model was used to train data privacy [16]. High-performance visual tracking was achieved by an extreme learning machine (ELM) model in [17]. In [18], general regression (GR) model was used to evaluate the video transmission quality. The radial basis function (RBF) network was optimized to reconstruct the image in [19].

The studies of secure data transmission quality prediction are rare. So, our paper investigates the secure data transmission quality prediction of mobile IoT networks. The main contributions are given as follows.

(1) With amplify-and-forward (AF) relaying scheme, we use SPSC to evaluate secure data transmission quality and derive the exact expressions.

(2) To realize real-time analysis of secure data transmission quality, we propose a secure data transmission quality prediction approach based on the Elman neural network. The proposed approach is compared with ELM, GR, and RBF methods.

(3) Through the extensive simulations, we verify the derived results. Compared with different methods, the quality assessment effect of Elman-based approach is better, and time complexity is lower.

2. The IoT System Model

The system has a mobile source (S), mobile destination (D), mobile eavesdropper (E), and mobile relay (R). Figure 1 shows the system model.

First, MR receives the signal r_{SR} as

$$r_{\mathrm{SR}} = \sqrt{M_{\mathrm{SR}}KE}h_{\mathrm{SR}}x + w_{\mathrm{SR}}, \tag{1}$$

where w_{SR} is Gaussian noise.

In the second time slot, D and E receive the signals r_{Rk}, $k \in \{\mathrm{D}, E\}$, as

$$r_{Rk} = \sqrt{M_{Rk}E}h_{\mathrm{SR}}h_{Rk}x + w_{Rk}. \tag{2}$$

The received SNR $W_{\mathrm{SR}k}$ is given as

$$W_{\mathrm{SR}k} = \frac{W_{\mathrm{SR}}W_{Rk}}{1 + \overline{W_{\mathrm{SR}}} + W_{Rk}}, \tag{3}$$

where

$$\begin{aligned} W_{\mathrm{SR}} &= M_{\mathrm{SR}}K\left|h_{\mathrm{SR}}\right|^2\overline{\gamma}, \\ W_{Rk} &= (1-K)M_{Rk}\left|h_{Rk}\right|^2\overline{\gamma}, \\ \overline{W_{\mathrm{SR}}} &= M_{\mathrm{SR}}K\overline{\gamma}. \end{aligned} \tag{4}$$

$W_{\mathrm{SR}k}$ is very complex. We approximate $W_{\mathrm{SR}k}$ as [22]

$$\begin{aligned} W_{\mathrm{SRA}k} &= \frac{W_{\mathrm{SR}}W_{Rk}}{1 + \overline{W_{\mathrm{SR}}} + \overline{W_{Rk}}}, \\ \overline{W_{Rk}} &= (1-K)M_{Rk}\overline{\gamma}. \end{aligned} \tag{5}$$

Bloch et al. [23] give the instantaneous secrecy capacity as

$$\gamma = \max\{\ln(1 + W_{\mathrm{SRAD}}) - \ln(1 + W_{\mathrm{SRAE}}), 0\}. \tag{6}$$

3. Secure Data Transmission Quality Analysis

The probability of SPSC F_{SPSC} is used to evaluate the secure data transmission quality. We will give the analysis.

According to the (6), F_{SPSC} is given as

$$\begin{aligned} F_{\mathrm{SPSC}} &= \Pr(\gamma > 0) = 1 - \Pr(\gamma < 0) \\ &= 1 - \int_0^{\infty} F_{\mathrm{SRAD}}(W_{\mathrm{SRAE}}) f_{\mathrm{SRAE}}(W_{\mathrm{SRAE}}) dW_{\mathrm{SRAE}}. \end{aligned} \tag{7}$$

With the help of [24], we obtain the PDF and CDF of $W_{\mathrm{SRA}k}$ as follows:

$$f_{W_{\mathrm{SRA}k}}(r) = \frac{1}{16r}G_{0,4}^{4,0}\left[\frac{81r}{\chi_k}\Big|_{3,3,3,3}^{-}\right], \tag{8}$$

$$F_{W_{\mathrm{SRA}k}}(r) = \frac{1}{16}G_{1,5}^{4,1}\left[\frac{81r}{\chi_k}\Big|_{3,3,3,3,0}^{1}\right], \tag{9}$$

$$\chi_k = \frac{\overline{W_{\mathrm{SR}}W_{Rk}}}{1 + \overline{W_{\mathrm{SR}}} + \overline{W_{Rk}}} \tag{10}$$

Substituting (8) and (9) into (7), F_{SPSC} is expressed as

$$\begin{aligned} F_{\mathrm{SPSC}} &= 1 - \frac{1}{256} \times \int_0^{\infty} \frac{1}{W_{\mathrm{SRAE}}} G_{1,5}^{4,1}\left[\frac{81W_{\mathrm{SRAE}}}{\chi_D}\Big|_{3,3,3,3,0}^{1}\right] \\ &\quad \times G_{0,4}^{4,0}\left[\frac{81W_{\mathrm{SRAE}}}{\chi_E}\Big|_{3,3,3,3}^{-}\right] dW_{\mathrm{SRAE}} \\ &= 1 - \frac{1}{256}G_{5,5}^{4,5}\left[\frac{\chi_E}{\chi_D}\Big|_{3,3,3,3,0}^{1,-2,-2,-2,-2}\right]. \end{aligned} \tag{11}$$

4. Secure Data Transmission Quality Prediction Approach

4.1. Data Sets. $T_i = (X_i, y_i)$. The input X_i includes 5 indicators. X_i is given as

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{i5}). \tag{12}$$

The output y_i is the SPSC. By using (11), the corresponding y_i can be obtained.

4.2. Network Design. Figure 2 shows the Elman neural network [25].

4.3. Predictive Evaluation. For PP testing data, MSE and AE are used to evaluate the prediction effect:

$$\begin{aligned} \mathrm{MSE} &= \frac{\sum_{z=1}^{PP}(d^z - y^z)^2}{PP}, \\ \mathrm{AE} &= \left|d^z - y^z\right|. \end{aligned} \tag{13}$$

5. Numerical Results

In this section, E = 1 and $\mu = W_{\mathrm{RD}}/W_{\mathrm{RE}}$ (in decibels).

With parameters in Table 1, we evaluate the SPSC performance with $\overline{\gamma} = 10$ dB in Figure 3. Simulation results show the following: (1) increasing u improves the SPSC

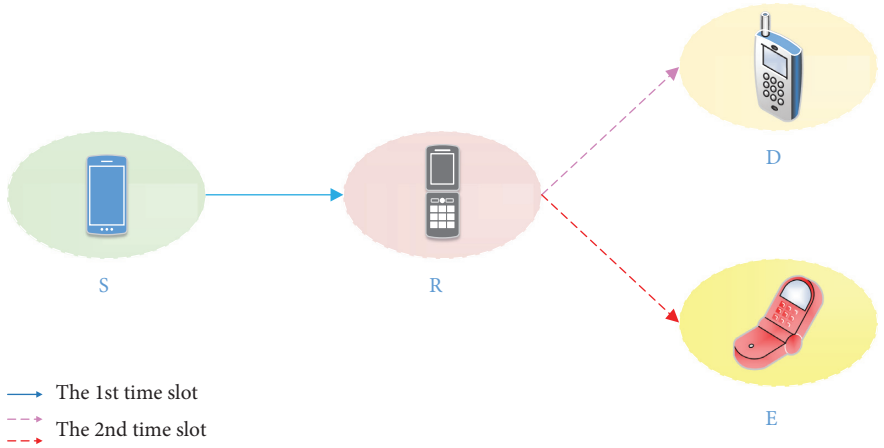


Figure 1: System model. E is the transmission power. For R and S, E is allocated with K. The channel coefficient h is 2-Nakagami distribution [20, 21]. M_{SR}, M_{RD}, and M_{RE} are the relative geometrical gains of S ⟶ R, R ⟶ D, and R ⟶ E links, respectively.

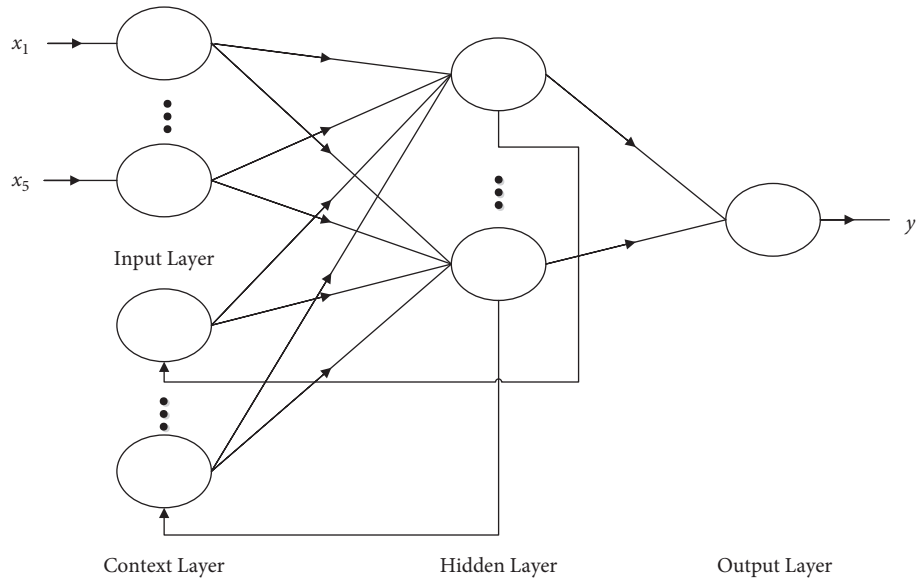


Figure 2: The Elman structure.

Table 1: Simulation parameters.

m_{SR}	1, 3
m_{RE}	1, 3
m_{RD}	1, 3
W_{SR}	5 dB
W_{RE}	5 dB
K	0.6

performance; (2) for Nakagami channels, the secure data transmission quality is the best. This is because a higher u improves the S ⟶ R⟶ D channel while degrading the S⟶R⟶E channel.

In Figures 4–11, ELM, GR, and RBF methods are compared with the Elman approach. Table 2 gives the simulation parameters. The MSE and AE of Elman approach are 0.00014 and 0.011, which are the lowest MSE and AE in the five methods. This is because Elman is a typical dynamic recurrent neural network and can adapt to the time-varying characteristics by adding a context layer.

The MSE is compared in Figure 12. Compared with GR, Elman has a better MSE performance, but the running time is longer than GR. Furthermore, compared with other methods, Elman has a higher quality precision and a lower time complexity.

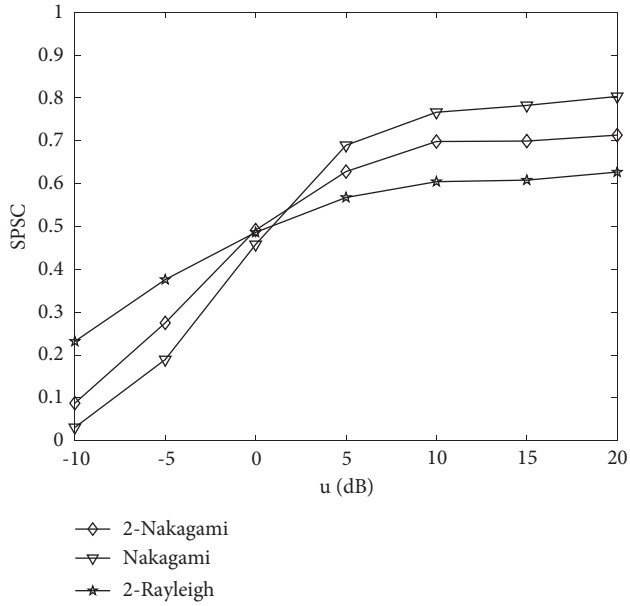


Figure 3: The SPSC performance versus u for different channels.

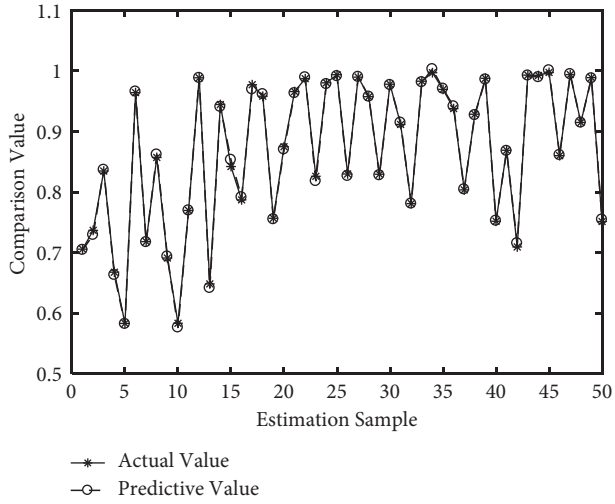


Figure 4: Prediction of elman.

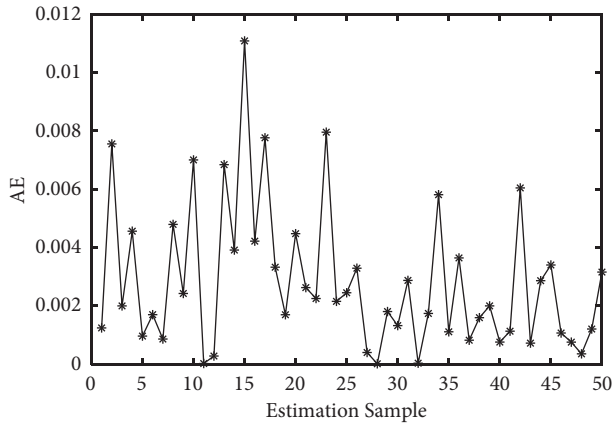


Figure 5: Ae of elman.

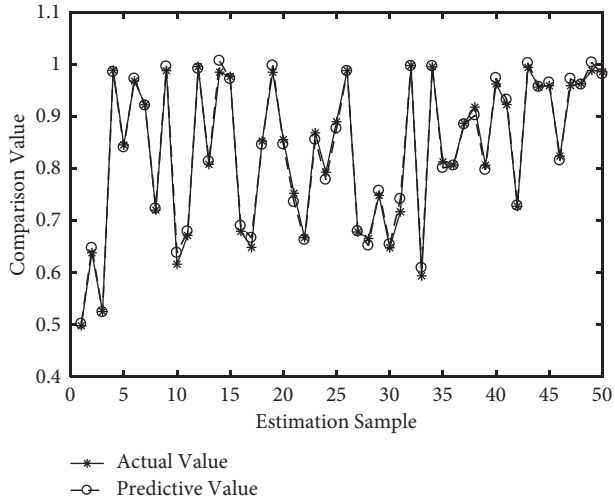


Figure 6: Prediction of RBF.

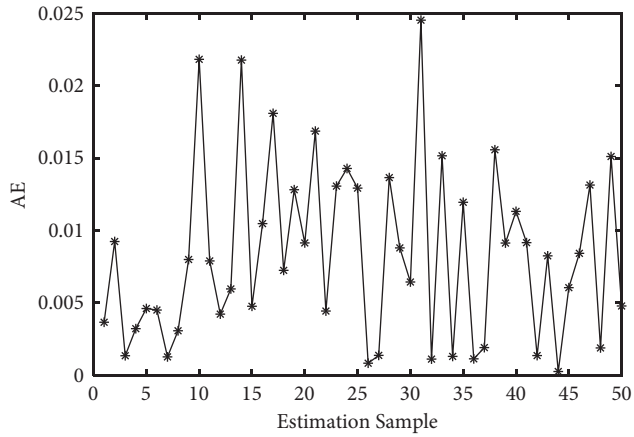


Figure 7: Ae of RBF.

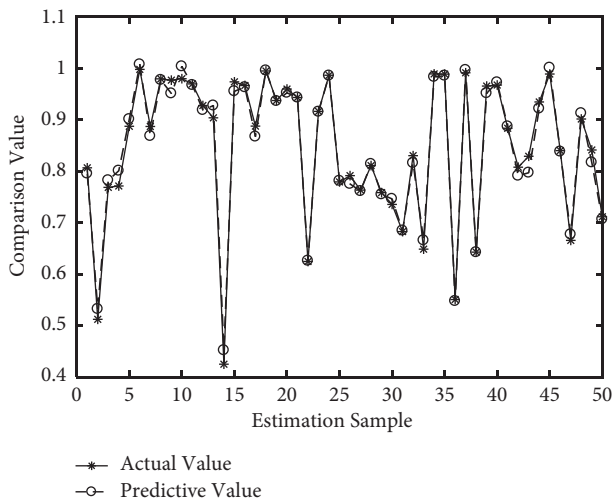


Figure 8: Prediction of ELM.

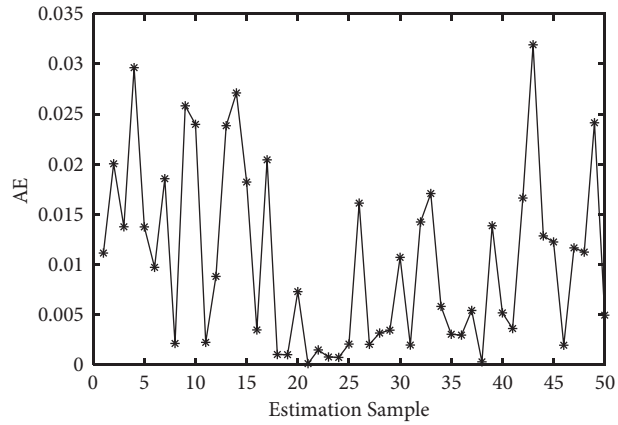


Figure 9: Ae of ELM.

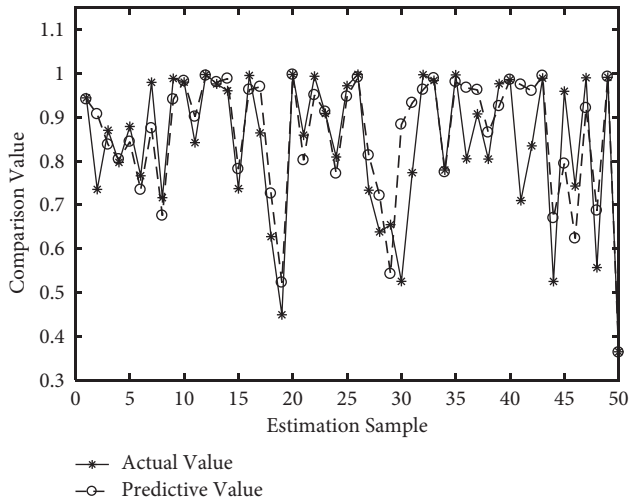


Figure 10: Prediction of GR.

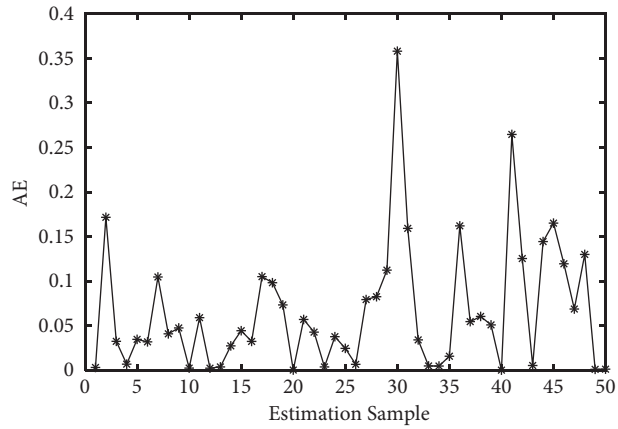


Figure 11: Ae of GR.

Table 2: The parameters for different methods.

Algorithm	Elman	ELM	RBF	GR
	Training data: 3000		Testing data: 50	
	q:10 η_1:0.01 η_2:0.01	q:15000	τ:20	τ:0.01

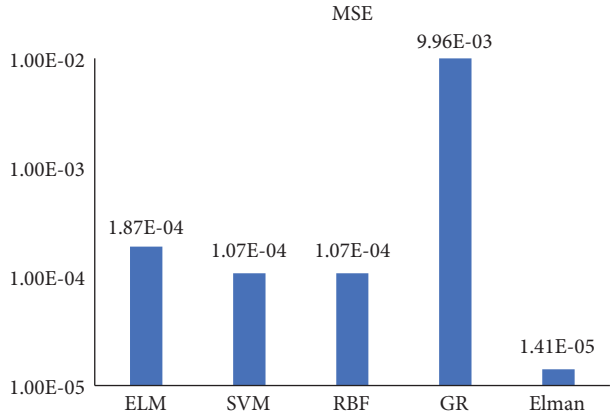


Figure 12: The MSE comparison.

6. Conclusion

This paper investigated the SPSC prediction of mobile IoT Networks. The exact expressions for SPSC were derived. Furthermore, based on the Elman network, we proposed an intelligent secure data transmission quality prediction algorithm. The theoretical analysis showed the following: (1) the SPSC performance over Nakagami channels was the best; (2) compared with different methods, the Elman algorithm can achieve a higher quality precision.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon reasonable request and with permission of funders.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (no. 11664043).

References

[1] M. B. Mollah, J. Zhao, D. Niyato et al., "Blockchain for the internet of vehicles towards intelligent transportation systems: a survey," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4157–4185, 2021.

[2] L. Chettri and R. Bera, "A comprehensive survey on internet of things (IoT) toward 5G wireless systems," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 16–32, 2020.

[3] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2021.

[4] L. Xu, H. Wang, and T. A. Gulliver, "Outage probability performance analysis and prediction for mobile IoV networks based on ICS-BP neural network," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3524–3533, 2021.

[5] G. Liu, Y. Liu, K. Zheng et al., "MCS-GPM: multi-constrained simulation based graph pattern matching in contextual social graphs," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 6, pp. 1050–1064, 2018.

[6] H.-M. Wang, Q. Yang, Z. Ding, and H. V. Poor, "Secure short-packet communications for mission-critical IoT applications," *IEEE Transactions on Wireless Communications*, vol. 18, no. 5, pp. 2565–2578, 2019.

[7] L. Sun and Q. Du, "Physical layer security with its applications in 5G networks: a review," *China Communications*, vol. 14, no. 12, pp. 1–14, 2017.

[8] A. Mukherjee, "Physical-layer security in the internet of things: sensing and communication confidentiality under resource constraints," *Proceedings of the IEEE*, vol. 103, no. 10, pp. 1747–1761, 2015.

[9] Y. Chen, T. Zhang, X. Qiao, H. Wu, and J. Zhang, "Secure cognitive MIMO wiretap networks with different antenna transmission schemes," *IEEE Access*, vol. 9, pp. 5779–5790, 2021.

[10] L. Xu, X. Zhou, Y. Tao, L. Liu, X. Yu, and N. Kumar, "Intelligent security performance prediction for IoT-enabled healthcare networks using improved CNN," *IEEE Transactions on Industrial Informatics*, p. 1, 2021.

[11] L. Qing, H. Xiaomei, and X. M. Fu, "Physical layer security in multi-hop AF relay network based on compressed sensing," *IEEE Communications Letters*, vol. 22, no. 9, pp. 1882–1885, 2018.

[12] L. Fan, N. Yang, T. Q. Duong, M. Elkashlan, and G. K. Karagiannidis, "Exploiting direct links for physical layer

security in multiuser multirelay networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3856–3867, 2016.

[13] K.-L. Besser and E. A. Jorswieck, "Bounds on the secrecy outage probability for dependent fading channels," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 443–456, 2021.

[14] L. Xu, X. Yu, and T. A. Gulliver, "Intelligent outage probability prediction for mobile IoT networks based on an IGWO-Elman neural network," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1365–1375, 2021.

[15] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, "Fast beamforming design via deep learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 1065–1069, 2020.

[16] J. Wang, L. Wu, H. Wang, K.-K. R. Choo, and D. He, "An efficient and privacy-preserving outsourced support vector machine training for internet of medical things," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 458–473, 2021.

[17] C. Deng, Y. Han, and B. Zhao, "High-performance visual tracking with extreme learning machine framework," *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2781–2792, 2020.

[18] L. Xu, H. Wang, H. Li, W. Lin, and T. A. Gulliver, "QoS intelligent prediction for mobile video networks: a GR approach," *Neural Computing & Applications*, vol. 33, no. 9, pp. 3891–3900, 2021.

[19] H. Wang, K. Liu, Y. Wu et al., "Image reconstruction for electrical impedance tomography using radial basis function neural network based on hybrid particle swarm optimization algorithm," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1926–1934, 2021.

[20] G. K. Karagiannidis, N. C. Sagias, and P. T. Mathiopoulos, "N∗Nakagami: $N{\ast}$Nakagami: a novel stochastic model for cascaded fading channels," *IEEE Transactions on Communications*, vol. 55, no. 8, pp. 1453–1458, Aug 2007.

[21] Z. X. Li, L. Z. Jia, F. Li, and H. Y. Hu, "Outage performance analysis in relay-assisted inter-vehicular communications over double-Rayleigh fading channels," in *Proceedings of the CMC*, pp. 266–270, Shenzhen, China, September 2010.

[22] F. K. Gong, Y. Wang, N. Zhang, and P. Ye, "Cooperative mobile-to-mobile communications over double Nakagami-m fading channels," *IET Communications*, vol. 6, no. 18, pp. 3165–3175, 2012.

[23] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, 2008.

[24] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic, San Diego, CA, USA, 7th edition, 2007.

[25] H. Y. Jin and X. M. Zhao, "Complementary sliding mode control via Elman neural network for permanent magnet linear servo system," *IEEE Access*, vol. 7, pp. 2169–3536, 2019.