https://doi.org/10.31449/inf.v48i21.6257 Informatica **48** (2024) 51–66 **51**

SD-WSN Network Security Detection Methods for Online Network Education

Zhenpeng Zhang
Beijing Open University, Beijing 100081, China
E-mail: ttbb1234315@163.com

Keywords: online education, SD-WSN, ABC algorithm, cluster detection algorithm, information security

Received: May 24, 2024

Online network education faces serious information security problems, and traditional network security detection methods are insufficient to deal with new types of attacks and unknown threats. To improve the security level of online education, the paper proposes an information security detection method for online education based on improved clustering detection algorithm and software-defined wireless sensor network to improve the detection accuracy. The study combines the improved clustering detection algorithm and software-defined wireless sensor network technology to construct a novel network security detection model. The improved clustering detection algorithm achieved a training error of 0.031 and a test error of 0.040. The proposed model demonstrated a detection accuracy of 90.3% and an average detection time of 2.6 seconds, outperforming other models in comparative tests. The above data validate that the designed novel network security detection model can better safeguard the information security and user data privacy of online network education. The study not only improves the information security level of online education, but also provides strong support for promoting its healthy development.

Povzetek: Razvit je izboljšan model za zaznavanje varnosti v spletni izobraževalni mreži s pomočjo gručenja in SD-WSN tehnologije, kar izboljšuje točnost zaznavanja in varnost uporabnikovih podatkov.

1 Introduction

The growth of network technology makes online network education become an essential platform for more and more people to pursue knowledge and skill improvement [1]. However, this emerging education model has also triggered serious information security problems, such as the leakage of user data and the frequent occurrence of cyber-attacks, which not only harms the legitimate rights and interests of users, but also restricts the healthy development of online education [2, 3]. Traditional network security detection methods mainly rely on known attack patterns, which are often overstretched for the ever-changing new types of attacks and unknown threats [4]. At the same time, in the face of the massive data generated by online education, many detection methods have bottlenecks in computational complexity and detection efficiency. Therefore, there is an urgent need for an efficient and accurate information security detection method that can adapt to the characteristics of online education [5]. Software-defined wireless sensor networks (SD-WSN), as an emerging network technology, can provide better support and guarantee for online education [6]. In addition, clustering detection algorithm is a clustering-based anomaly detection algorithm, which aims to improve the detection accuracy and efficiency, and can also improve the security of online web education [7]. Therefore, this research proposes an information security detection method for online e-education based on improved clustering detection algorithm and SD-WSN, which aims to improve the detection accuracy and achieve real-time monitoring (RT-M) of the educational environment. The innovations of this research are the introduction of an improved clustering detection algorithm, which can adaptively identify abnormal data and greatly improve the accuracy of detection, and the combination of SD-WSN technology, which can monitor the educational environment in real time through wireless sensor networks, so as to detect and deal with potential security risks in a timely manner. This research is not only expected to lift the information security level of online network education and guarantee the security of user data, but also help to stimulate healthy progress of online education.

The article is discussed in four parts, the first part is the research related to SD-WSN technology, clustering detection algorithms and online education information security; the second part is the construction process of the network security detection model incorporating the improved clustering intrusion detection algorithms and SD-WSN; the third part is the performance validation of the improved algorithms and the proposed model; and the fourth part is the conclusion of the full article.

2 Literature review

With the gradual emergence of SD-WSN technology, its applications are becoming more and more widespread. Aiming at the energy efficiency problem of multicast

transmission in wireless sensor networks, Banerjee's team proposed an intelligent green multicast scheme based on SD-WSN framework. The scheme divides the network into multiple regions, each of which is managed by an SDN controller capable of calculating energy efficient paths. This method could effectively improve network throughput and save energy [8]. Younus et al. proposed a solution based on reinforcement learning for the routing optimisation problem in SD-WSN. By designing a reward function and deciding the next action based on the received reward, the SDWSN controller is able to improve the routing paths based on previous experience. The method outperformed other techniques in network lifetime and packet delivery rate [9]. With the development of intrusion detection techniques, the application areas of clustered intrusion detection algorithms are becoming more and more widespread, for example, Nagaraja's team proposed a novel clustered intrusion detection algorithm to improve classification accuracy by addressing the data dimensionality challenge and similarity computation problem in intrusion and anomaly detection. Experiments show that the proposed intrusion detection algorithm outperforms CANN, CLAPP, SVM, and KNN in terms of recognition attack accuracy [10]. Liang et al. proposed an intrusion detection method based on collaborative clustering and feature data fusion to address anomalous intrusions in blockchain systems. The method uses mathematical models for data fusion and combines artificial intelligence techniques to train and analyse data clusters in blockchain networks. This method is beneficial for high accuracy and real-time detection, and effectively copes with abnormal intrusion behaviours in blockchain systems [11].

With the rise of online network education, its information security problem has received widespread attention, and related scholars have conducted a great deal of research on online network education network security problems. To solve the problem of security risks posed by online network education, Dincelli et al. proposed a gamification design-based approach to security education, training and awareness, using both textual and visual forms. The textual format was more likely to improve users' attitudes and behaviours, while the visual format was more likely to improve users' experience and memory. The study provides a useful reference for the field of online education [12]. To address a series of challenges in digital copyright management in online education platforms, such as digital copyright infringement of multimedia learning resources, and insecurity of digital education certificates, Guo's team proposed a blockchain-based digital copyright management system. The system demonstrates as a promising solution for blockchain-based multimedia data protection in online education environments [13]. To address the issue of the impact of online security information on user behaviour, Kamar et al. proposed a methodology to assess the effectiveness of security information. The results of the study show that security information and deliberate decision-making processes do not directly affect SMiShing fraud victimisation, but security information is effective in preventing victimisation when the user's level of deliberate decision-making process is high. This highlights the importance of criminological theory and interdisciplinary research in online security practice [14]. In order to address the comprehensive evaluation of online education security culture, Arbanas' team proposed a comprehensive framework for evaluating information security culture, which was validated by empirical research and multivariate analysis. The results show that the framework contains technical, organisational and social factors and can effectively evaluate online education information security with high reliability and validity [15].

Based on the above relevant studies, Table 1 is summarized, in which the research theme, main index methods and shortcomings of relevant studies are summarized.

Table 1: Summary of relevant information of relevant studies

Author	Research theme	Main index	method	Insufficient
Banerjee et al. [8]	SD-WSN Smart Green multicast	Network throughput, energy savings	Partition areas, SDN controller management, calculation of energy efficient paths	Lack of optimization of dynamic networks and complex structures
Younus et al. [9]	SD-WSN route optimization	Network lifetime, packet delivery rate	Reinforcement learning, design reward function, improve routing path	Performance may be limited in large scale and highly dynamic environments
Nagaraja et al. [10]	Cluster intrusion detection	Attack identification accuracy	New clustering algorithm improves classification accuracy	Face challenges with high-dimensional data and complex attack patterns
Liang et al. [11]	Block chain abnormal intrusion detection	Detection accuracy, real-time	Collaborative clustering, feature data fusion, artificial intelligence technology	Lack of specialized intrusion detection methods for blockchain systems
Dincelli et al. [12]	Online education security	User attitude, behavior improvement, user experience	Gamified design, text and visual forms of safety education	Limited in terms of increasing user continuous learning and

		and memory		safety awareness
Guo et al. [13]	Digital rights management	Multimedia data protection	A blockchain-based digital rights management system	Face challenges when dealing with large-scale multimedia resources and cross-platform copyright management
Kamar et al. [14]	The influence of network security information on user behavior	Safety information effectiveness to prevent injury	Evaluate the impact of security information on user behavior	Lack of in-depth understanding of user decision-making processes and behavior change mechanisms
Arbanas [15]	Online education safety culture evaluation	The reliability and validity of the evaluation framework	Comprehensive information security culture evaluation framework, empirical research and multivariate analysis	There is a lack of evaluation systems that take into account technical, organizational and social factors
This text	SD-WSN network security detection method	Detection accuracy, detection time	Improved clustering detection algorithm combined with SD-WSN	/

Based on the above related studies, it can be found that although the current research has good performance in the face of most network intrusions, the effect is not ideal in the face of some new attacks and unknown threats. At the same time, faced with the massive data generated by online education, many detection methods also have bottlenecks in computational complexity and detection efficiency. Therefore, this research innovatively introduces an improved clustering algorithm to identify abnormal data adaptively and improve the accuracy of detection. A new detection model is proposed by combining SD-WSN technology and clustering detection algorithm, so that the proposed model can better adapt to the characteristics of online network education and prevent intrusion efficiently and accurately.

3 Construction of SD-WSN network security detection model for online network education

This chapter focuses on how SD-WSN and improved clustering detection algorithm can be applied in online educational network security detection. By introducing SD-WSN technology, RT-M of the educational environment is achieved to identify and deal with potential security risks in a timely manner, while a novel network security detection model is proposed in conjunction with the improved clustering detection algorithm to lift the detection efficiency.

3.1 Application of SD-WSN in educational security detection in online education networks

With the popularity of online education, network security problems are gradually highlighted, and how to guarantee the security of online education network has become an urgent problem.SD-WSN, as an emerging network technology, has the characteristics of flexibility and programmability, which has an important application value in the security detection of online education network. The framework diagram of SD-WSN is shown in Figure 1.

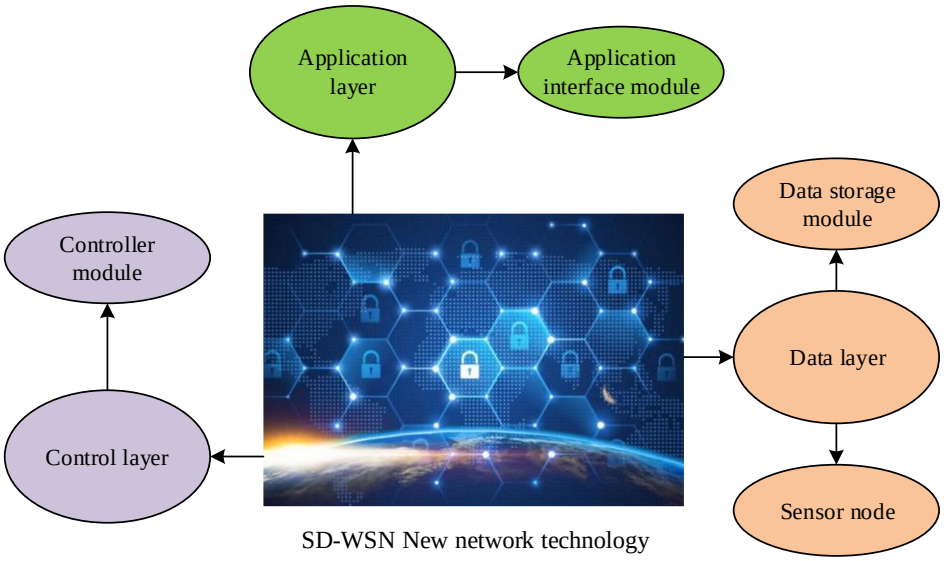


Figure 1: Frame diagram of SD-WSN

From Figure 1, SD-WSN can be divided into three main parts: application layer, control layer and data layer. The application layer is responsible for interacting with users; the control layer is responsible for parsing and executing commands; and the data layer is responsible for collecting and transmitting sensor data, and transmitting the data from the sensor nodes to the control and application layers. Specifically, the framework diagram of SD-WSN contains important components such as sensor nodes, controllers, application interfaces and data storage. Among them, the sensor nodes are responsible for collecting and transmitting data, while the controller is responsible for managing the network and processing data, and the sensor node data transmission formula is shown in Eq. (1).

$$D_{sensor} = \sum_{i=1}^{n} S_i \times R_i \tag{1}$$

In Eq. (1) D_{sensor} denotes the total amount of data transmitted by the sensor node, S_i is the amount of data collected by the i th sensor node, R_i is the data transmission rate of the node, and n is the total number of sensor nodes. The application interface is responsible for interacting with the user, while the data storage is used to store the data and historical information in the network. Through the collaboration and co-operation of these components, SD-WSN can achieve flexible control and management of the wireless sensor network and improve the reliability and security of the network. The specific workflow of SD-WSN in online education network security detection is shown in Figure 2.

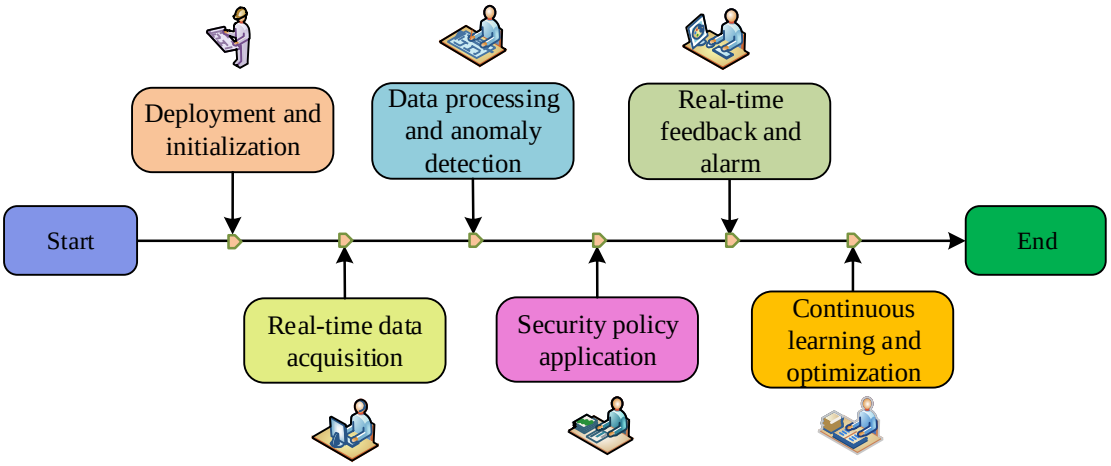


Figure 2: Specific work flow of SD-WSN

In Figure 2, the workflow of SD-WSN in online education network security detection is mainly divided into six steps, firstly, deploying sensor nodes in the network environment that needs to be monitored, and carrying out the necessary initialisation settings to ensure that the nodes work normally and establish a connection with the central control plane. Afterwards, the sensor nodes are used to monitor network traffic, device status and other key details in real time, and transmit these data to the central control plane in real time. These data are then received and analysed at the central control plane, and abnormal traffic and behaviour are identified through algorithms and models, and the abnormal traffic detection threshold formula is shown in Eq. (2).

$$T_{threshold} = \mu_{traffic} + k' \times \sigma_{traffic} \quad (2)$$

Eq. (2) where $T_{threshold}$ is the anomaly traffic detection threshold; $\mu_{traffic}$ and $\sigma_{traffic}$ are the mean value and standard deviation of network traffic. k' is a constant. The fourth step automatically or semi-automatically takes corresponding defence measures against the detected risks or anomalies according to the pre-defined security policy. The risk score calculation formula is shown in Eq. (3).

$$S_{risk} = \sum_{a=1}^{m} I_a \times P_a \quad (3)$$

In Eq. (3), S_{risk} is the total score of network security risks, I_a is the impact degree of the a th security risk, P_a is the probability of occurrence of the risk, and m is the total number of security risks. After that, for the confirmed security events or risks, the system generates alarm information and provides real-time feedback to the relevant management personnel for timely intervention. Finally, it needs to continuously learn and summarise from the actual operation to optimise its detection and processing mechanism and improve its ability to respond to future security events. This six-step process covers the main aspects of SD-WSN's online education network security detection, from data collection and processing to the application and continuous optimisation of security policies, forming a complete set of working mechanisms.

3.2 Clustered online education network intrusion detection algorithm construction based on ABC algorithm

The network technology has made online education gradually become mainstream; however, network security problems have also come to the fore. Some experts have proposed to apply ABC algorithm in clustering online education network intrusion detection [16-20]. ABC algorithm is a kind of optimization search algorithm simulating the foraging behaviour of bees (FBoBs) in nature, which combines the intelligent behaviour of bees and the mechanism of collegial collaboration, and seeks for the optimal solution of the problem through continuous iterative search. The detailed process of ABC algorithm mainly includes the following five stages. Initialization phase: sets the number of initial solutions, each solution is called a bee. lead bee stage: each lead bee searches the solution space according to the current solution, finds a new solution. Following phase: each following bee searches according to the leader's solution and selects a new solution. Scout phase: If the leader and the follower do not find a better solution during the search, is called a scout. Each scout bee randomly selects a new solution in the solution space, and updates the current best solution. If the new solution generated by the scout bee is better than the current optimal solution, replaces the new solution with the current optimal solution. end condition stage: set the end condition, for example, the maximum number of iterations or the value of the objective function reaches a certain threshold. The graphical representation of the FBoBs in nature is shown in Figure 3.

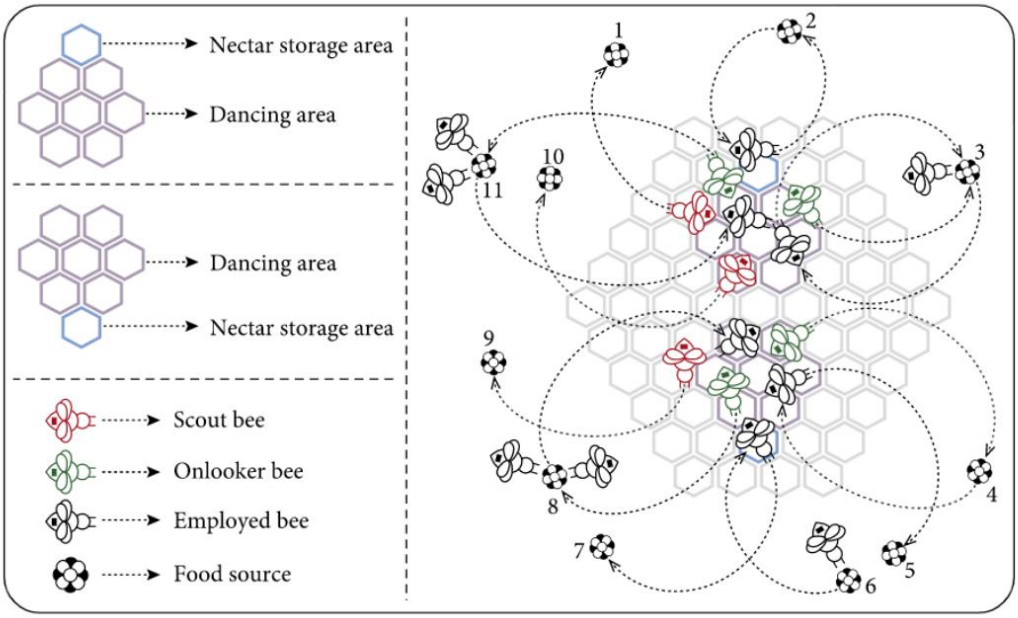


Figure 3: Illustrates foraging behaviour of bees in nature

In swarm algorithm, worker bees are responsible for searching within the known resources, observation bees select based on the information of worker bees and scout bees search for new resources randomly. The worker bee search formula is shown in Eq. (4).

$$x_b^{t+1} = x_b^t + \varphi(x_b^t - x_j^t) \tag{4}$$

In Eq. (4), x_b^t denotes the position of the b th worker bee in the t th iteration, x_j^t denotes a randomly selected position from the neighbourhood of the current worker bee, and φ is a random number. The observation bee search formula is shown in Eq. (5).

$$\begin{cases} p_b = \dfrac{fitness_b}{\sum_{j=1}^{N} fitness_j} \\ x_b^{t+1} = x_b^t + \varphi(x_b^t - x_k^t) \end{cases} \tag{5}$$

In Eq. (5), p_b is the probability that the b th bee is selected as an observer bee; x_k^t is a position selected from the current observer bees according to the roulette wheel selection mechanism. The scout bee search formula is shown in Eq. (6).

$$x_b^{t+1} = x_{\min} + rand(0,1)(x_{\max} - x_{\min}) \tag{6}$$

In Eq. (6), $x_{\min}$ and $x_{\max}$ are the upper and lower bounds of the problem search space, respectively, and the fitness calculation formula is shown in Eq. (7).

$$fitness = f(x) \tag{7}$$

Traditional intrusion detection is inefficient and has a high false alarm rate when facing large-scale network data, while ABC algorithm can solve the above problems by clustering analysis, which improves the detection efficiency and accuracy. The flow of clustering online education network intrusion detection model based on ABC algorithm is shown in Figure 4.

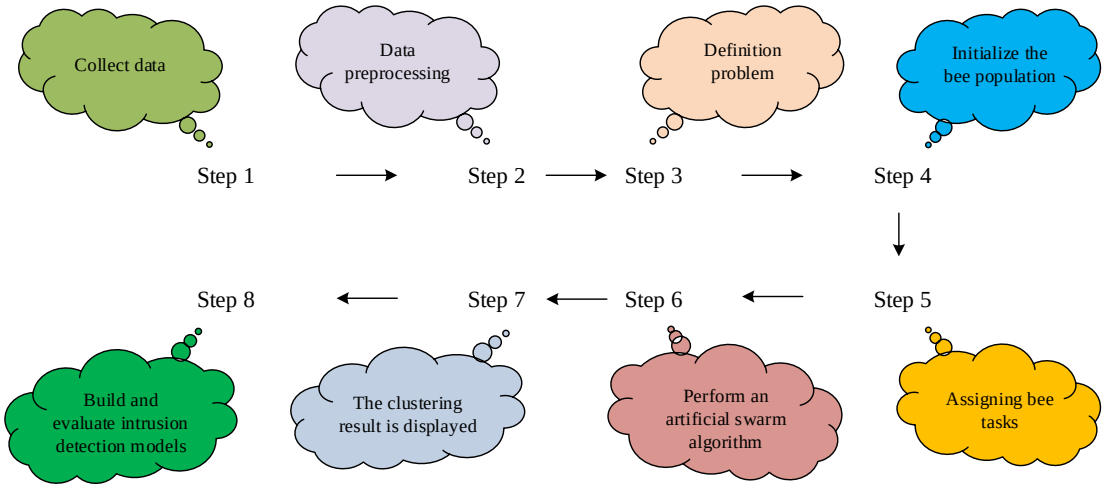


Figure 4: Intrusion detection algorithm flow of clustering online education network based on ABC

As shown in Figure 4, the running process of ABC-based clustered online education network intrusion detection algorithm is divided into eight steps. Firstly, online education network data is collected to train the intrusion detection algorithm. After that, the network data are preprocessed, and the data preprocessing includes cleaning, feature extraction and normalization to better present the network situation. The commonly used normalization formula is shown in Eq. (8).

$$x'_{normalized} = \frac{x' - x'_{\min}}{x'_{\max} - x'_{\min}} \tag{8}$$

In Eq. (8), x' is the original data, and $x'_{\min}$ and $x'_{\max}$ are the min and max values of the feature, respectively. After that, the intrusion detection target needs to be specified and the clustering effect is evaluated by defining the fitness function. In order to evaluate the effectiveness of clustering, the distance of each

data point to the centre of the cluster it belongs to needs to be calculated. Euclidean distance is the most commonly used distance calculation formula and its expression is shown in Eq. (9).

$$d(x'', c) = \sqrt{\sum_{h=1}^{n'} (x''_h - c_h)^2} \tag{9}$$

In Eq. (9), x'' is the data point, c is the clustering centre and n' is the dimension of the feature. After that the honey bee population is initialised and the bees are divided into honey picking bees and observation bees, the former searching near the current centre and the latter searching globally based on the former's information. The ABC algorithm is executed in the next step to find the better clustering centre by calculating the fitness function, updating the pheromone, etc. The pheromone update rule is shown in Eq. (10).

$$\tau_{i'j'}(t+1) = (1-\rho)\tau_{i'j'}(t) + \frac{\rho}{d_{i'j'}} \tag{10}$$

In Eq. (10), $\tau_{i'j'}(t)$ is the amount of pheromone from location i' to j'. ρ is the volatility of pheromone, and $d_{i'j'}$ is the distance between i' and j'. Then comes the clustering results of the output algorithm, representing normal network traffic and user behaviour patterns. Based on these results, intrusion detection algorithms are constructed and evaluated that can be used to identify and classify abnormal traffic using

classifiers or other machine learning methods. In order to determine whether network traffic or user behaviour is abnormal a threshold needs to be set. The threshold expression is shown in Eq. (11).

$$Threshold = \mu + 3\sigma \tag{11}$$

In Eq. (11), μ is the mean value of normal traffic and σ is the standard deviation. If the statistical value of a traffic or behaviour exceeds this threshold, then it is considered abnormal. Through the above steps, the ABC-based clustered online education network intrusion detection algorithm can be constructed to achieve RT-M and anomaly detection of network traffic and user behaviours, and to safeguard the security of the online education network.

3.3 Construction of online education security detection model by fusing improved clustering algorithm and SD-WSN

To better protect the information security in online education, this research uses ABC algorithm clustering intrusion detection method and SD-WSN technology to construct a set of efficient security detection model. Intrusion detection is first enhanced with the help of ABC algorithm. This algorithm can simulate the FBoBs and intelligently cluster the network traffic data so as to identify abnormal traffic quickly and accurately. By combining it with the clustering algorithm can improve its clustering effect and detection speed, making the model able to quickly respond to various network attacks. Second, SD-WSN, as a flexible wireless sensor network architecture, can deploy sensor nodes in educational networks to monitor network traffic and the use of educational resources in real time. By combining SD-WSN with improved clustering intrusion detection algorithms, comprehensive collection and intelligent analysis of network traffic can be achieved, so as to discover and isolate abnormal traffic in a timely manner and effectively prevent potential security threats. In this model, the parameters of ABC algorithm are set as follows: the population size is set to 50, the number of food sources is set to 25, the limiting times are set to 100, the maximum number of iterations is set to 500, the neighborhood radius is dynamically adjusted, and the initial setting is 10% of the data feature space, which gradually decreases with iteration. The selection probability is set to 0.8. The SD-WSN technical parameters are set as follows: One sensor node is deployed for every 100 users or devices. Each node stores at least one hour of traffic data to ensure that there is enough data for anomaly detection. The sampling rate is set to once per second to balance real-time data collection and system load. The data transfer rate is set to 1Mbps. The energy management strategy uses the adaptive sleep mechanism, which automatically enters the sleep state when the flow around the node is low to save energy. Use the 6LoWPAN lightweight communication protocol to reduce communication overhead and latency. By setting the above parameters, the effectiveness and efficiency of ABC algorithm and SD-WSN technology in the education security detection model of online education network can be ensured. The specific modules of the online education network education security detection model constructed by the research are shown in Figure 5.

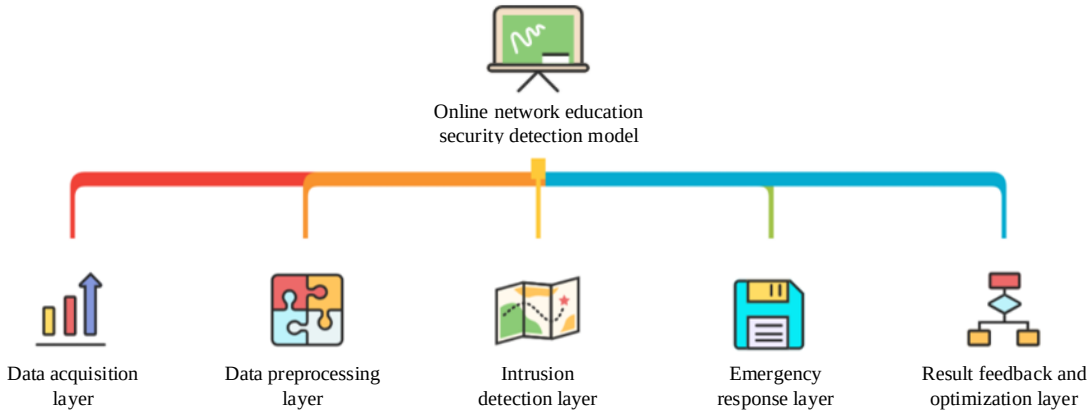


Figure 5: Specific composition of online education network education security detection model

In Figure 5, this model has five main modules. First is the data acquisition layer, which uses the sensor nodes in the SD-WSN to collect data such as network traffic and user behaviour in real time. The data acquisition formula for sensor node data acquisition in the data acquisition layer is shown in Eq. (12).

$$D_{raw} = \sum_{i=1}^{n} S_i \times T_i \tag{12}$$

In Eq. (12), D_{raw} denotes the raw data obtained from the sensor nodes in the SD-WSN, S_i is the reading of the i th sensor, and T_i is the timestamp of the i th sensor. Next is the data preprocessing layer, which has the task of eliminating noise and redundancy from the raw data, thus improving the quality of the data. The data cleaning is Eq. (13).

$$X_{clean} = X - X_{noise} - X_{redundant} \tag{13}$$

In Eq. (13), X denotes the original dataset, X_{noise} denotes the noisy part of the data, $X_{redundant}$ denotes the

redundant part of the data, and X_{clean} denotes the cleaned dataset. The core part of the model is ABC clustering intrusion detection layer. In this part ABC clustering algorithm is used for intelligent clustering and in-depth analysis of the pre-processed data. Optimisation of algorithms and policies ensures that the model accurately identifies anomalous network traffic. When a threat is detected, the emergency response layer is activated quickly to isolate the problem, stop the attack from spreading, and take urgent measures to maintain network security and stability. Finally, the feedback layer analyses the detection results in depth and adjusts the parameters and policies according to the actual situation to improve the detection effect. The above five modules work together to monitor and intelligently analyse network traffic in real time, effectively respond to security challenges, and ensure the safety of educational resources and user data. This will provide a safe and reliable environment for online education and promote its sustainable development.

4 Comparative performance analysis of improved clustering algorithms and empirical analysis of online educational testing models

The chapter first compares the performance of different intrusion detection algorithms on the KDD Cup 99 dataset, reflecting the effectiveness of the research algorithm. Afterwards, the superiority of the research model is verified by comparing it with similar security detection models in a comparative test.

4.1 Comparative performance analysis of improved clustering intrusion detection algorithms

To verify the performance of the improved clustering intrusion detection algorithm proposed in the study, the study conducts comparative experiments with the K-mean intrusion detection algorithm and DBSCAN intrusion detection algorithm, and Table 2 shows the specific environment.

Table 2: Specific experimental environment configuration

Experimental environment name	Environment configuration
Operating system	Windows 10 64-bit
CPU	Intel Core i7-8700K, 3.7GHz
GPUs	GTX 9500
Internal memory	16 GB DDR4 RAM
Video memory	4 GB
Programming language	Python 3.7
Data set	KDD Cup 99 data set

For this experiment, the study chose the KDD Cup 99 dataset as the test dataset. The KDD Cup 99 dataset is one of the most widely used benchmark datasets in the field of intrusion detection and contains many types of network traffic data and attack data. The data set consists of several files, including a training set and a test set. Each data point contains 41 features, of which 34 are continuous features and 7 are discrete features. The types of attacks in the data set are divided into four broad categories: DOS (denial of service attacks), R2L (unauthorized access from a remote host), U2R (unauthorized local superuser privileged access), and PROBING (monitoring and other probing). In this experiment, the research considers all four of the above types of attacks to fully evaluate the performance of the algorithm. The Python programming language and related tool libraries are used in the experiments to implement and improve the clustering intrusion detection algorithms. To ensure a fair comparison, they use the same parameter settings and data preprocessing steps. The pre-processing steps are as follows: First, the data is cleaned to remove duplicate records and missing values in the data set. Then, the most useful feature for intrusion detection task is selected from 41 features, and the continuous features are standardized to have the same scale. Then the discrete features are encoded to facilitate the algorithm to process. Finally, the data set is divided into a training set and a test set for training and evaluating the algorithm. The performance of the three algorithms is compared by comparing the error, accuracy, classification effect and other metrics of the three algorithms. The error results of the three algorithms in the training set and test set are displayed in Figure 6.

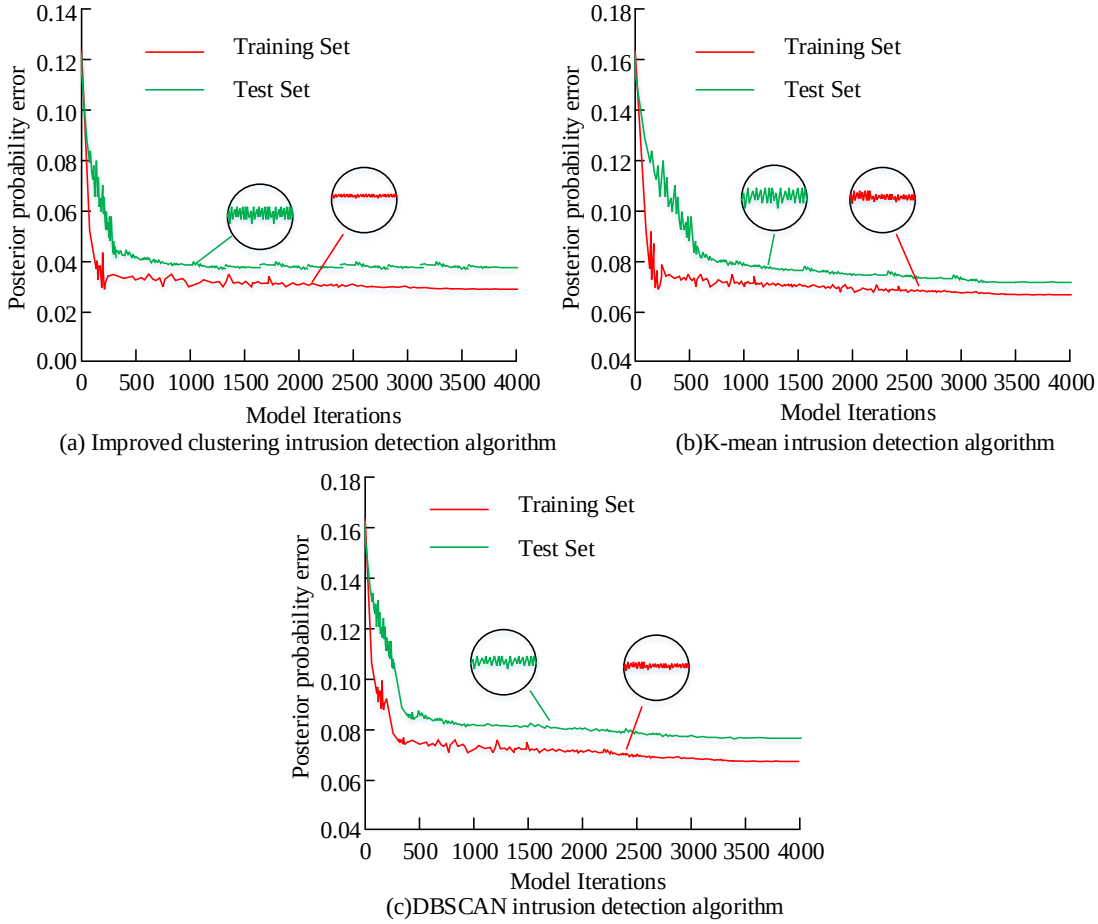


Figure 6: Comparison of error results of the three algorithms

From Figure 6(a), the training error of the improved clustering intrusion algorithm in the test set and the training set is as low as 0.040 and 0.031, respectively; from Figure 6(b), the training error of the K-mean intrusion detection algorithm in the test set and the training set is as low as 0.077 and 0.070, respectively; and from Figure 6(b), the DBSCAN intrusion detection algorithm in the test set and the training set is as low as 0.082 and 0.073, respectively. The above data prove that the overall error level of the improved clustering intrusion algorithm is lower than that of the K-mean intrusion detection algorithm and the DBSCAN intrusion detection algorithm. This means that the overall error level of the improved algorithm is lower than that of the K-mean intrusion detection algorithm and the DBSCAN intrusion detection algorithm. Afterwards, the accuracy and F1 value data of the three algorithms are collected and plotted to obtain Figure 7.

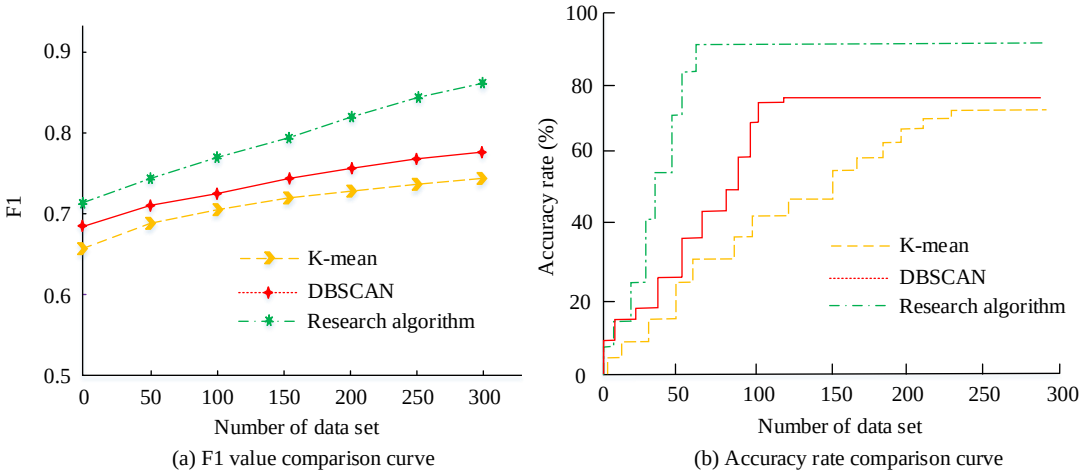
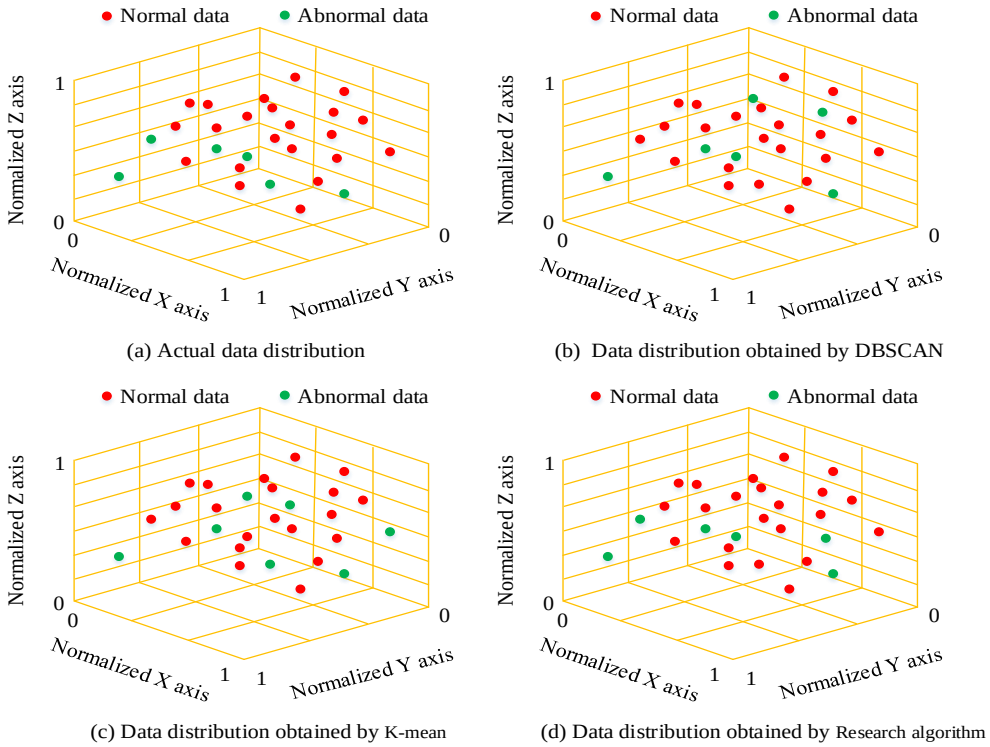


Figure 7: Comparison results of accuracy and F1 value of the three algorithms

From Figure 7(a), the F1 value curve of the improved algorithm is obviously higher than the other two algorithms, and the maximum F1 value of this algorithm is 0.86, which is significantly better than 0.74 of the DBSCAN intrusion detection algorithm and 0.72 of the K-mean intrusion detection algorithms. from Figure 7(b), the accuracy curve of the proposed improved clustering intrusion algorithm is also significantly better than the two compared algorithms. algorithms, and the maximum accuracy of this algorithm is 93.7%, which is significantly better than 77.8% for DBSCAN intrusion detection algorithm and 75.1% for K-mean intrusion detection algorithm. The above results indicate that the improved algorithm performs better in terms of the dimensions of F1 value and accuracy. Finally, Figure 8 is used to show the effectiveness of the three algorithms in detecting network intrusion data.



(a) Actual data distribution

(b) Data distribution obtained by DBSCAN

(c) Data distribution obtained by K-mean

(d) Data distribution obtained by Research algorithm

Figure 8: Comparison of detection effects of three algorithms on network intrusion data

Figures 8(a), (b), (c), and (d) show the actual distribution of network data, the distribution of network data obtained by the DBSCAN algorithm, the distribution of network data obtained by the K-mean algorithm, and the distribution of network data obtained by the proposed algorithm, respectively. Comparing Figures 8(a) to 8(d), the network data distribution obtained by the proposed algorithm has the smallest difference from the actual network data distribution, whereas the network data distributions obtained by the DBSCAN algorithm and the K-mean algorithm both have a large difference from the actual network data distribution. This shows that the improved clustering intrusion algorithm is more effective in detecting network intrusion data. In addition, in order to better illustrate the effectiveness of the proposed algorithm, the performance of the proposed algorithm was compared with other three novel clustering algorithms. The comparison results are shown in Table 3.

Table 3: Performance comparison results of different clustering algorithms

Algorithm name	Training set error	Test set error	Accuracy rate	F1 value	Detection effect score
Improved clustering intrusion algorithm	0.031	0.040	93.7%	0.86	9.0
FCM fuzzy clustering algorithm	0.045	0.060	88.5%	0.80	7.5
Multi-density clustering algorithm based on Spark	0.052	0.068	85.2%	0.76	6.8
K-prototypes clustering algorithm	0.048	0.065	87.1%	0.78	7.2

As shown in Table 3, the error of the improved clustering intrusion algorithm proposed in the study was 0.031 on the training set, which performed better than 0.045 for FCM fuzzy clustering algorithm, 0.052 for Spark based multi-density clustering algorithm and 0.048 for K-prototypes clustering algorithm. Similarly, on the test set, the error of the improved clustering intrusion algorithm is the lowest, which is 0.040. In terms of accuracy, the improved clustering intrusion algorithm reaches 93.7%, which is significantly higher than the other three algorithms. F1 value as an important index of classification performance, the improved clustering intrusion algorithm also obtained the highest 0.86. Finally, from the point of view of detection effect score, the improved clustering intrusion algorithm gets a high score of 9.0, which further proves its superior performance in network intrusion data detection. In summary, the improved clustering intrusion algorithm proposed in this study has excellent performance in many indexes, and has high practical value and application prospect.

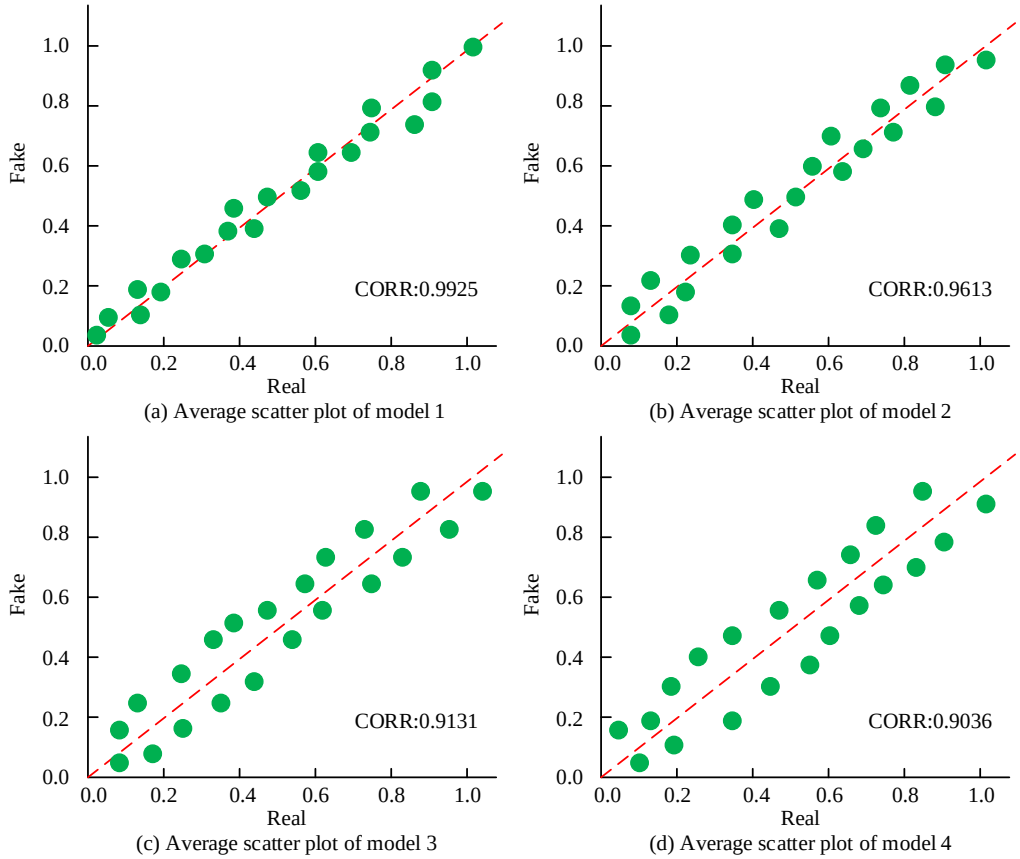
4.2 Performance analysis of online education security detection model incorporating improved clustering algorithm and SD-WSN

To analyze the practical effectiveness of the proposed online education security detection model (Model 1) incorporating the improved clustering algorithm and SD-WSN, the study compares it with the online education security detection model based on the improved clustering algorithm (Model 2), the K-mean algorithm (Model 3) and the DBSCAN algorithm (Model 4) for comparison experiments. Table 4 is the specific environment.

Table 4 Experimental environment of network education security detection model performance test

Serial number	Experimental environment	Description
Hardware equipment	Server	High-performance computer configured with Intel Xeon E5-2650 v4 processor, 128GB memory, and 2TB SSD
	Network equipment	Cisco Catalyst 3850 switch, Cisco ASA 556-X firewall
Software platform	Operating system	Ubuntu Server 20.04 LTS
	Programming language	Python 3.8.5
Network environment	Sensor node deployment	There are 30 sensor nodes deployed in the experiment site, including various types of sensors such as temperature, humidity, and network traffic
	Network topology	The topological structure of star and tree is adopted to ensure efficient data transmission and stability
Data set	Network education security dataset	Using a publicly available network education security dataset, including 500,000 normal traffic samples and 500,000 abnormal traffic samples, a total of 1 million sample data were collected. Using a publicly available network education security dataset

In this research, the dataset is first pre-processed, including data cleaning, feature extraction and labelling, in order to be suitable for the training and testing of individual security detection models. Afterwards, the performance of each security detection model is evaluated and compared using metrics such as correlation coefficient, detection accuracy, and detection time. The correlation coefficient results of the four detection models are shown in Figure 9.



(a) Average scatter plot of model 1 (b) Average scatter plot of model 2

(c) Average scatter plot of model 3 (d) Average scatter plot of model 4

Figure 9: Correlation coefficient results of the four detection models

From Figure 9(a) to 9(d), the correlation coefficients of model 1, 2, 3, and 4 are 0.9925, 0.9613, 0.9131, and 0.9036. Comparing Figure 9(a) with Figures 8(b), 8(c), and 8(d), it can be found that the Model 1 has the highest correlation coefficient result, which is much better than the comparison models. This result indicates that the research-proposed online education security detection model performs better in terms of the correlation coefficient dimension. Afterwards, the detection accuracy and detection time of the four models are compared, and the comparison results are shown in Figure 10.

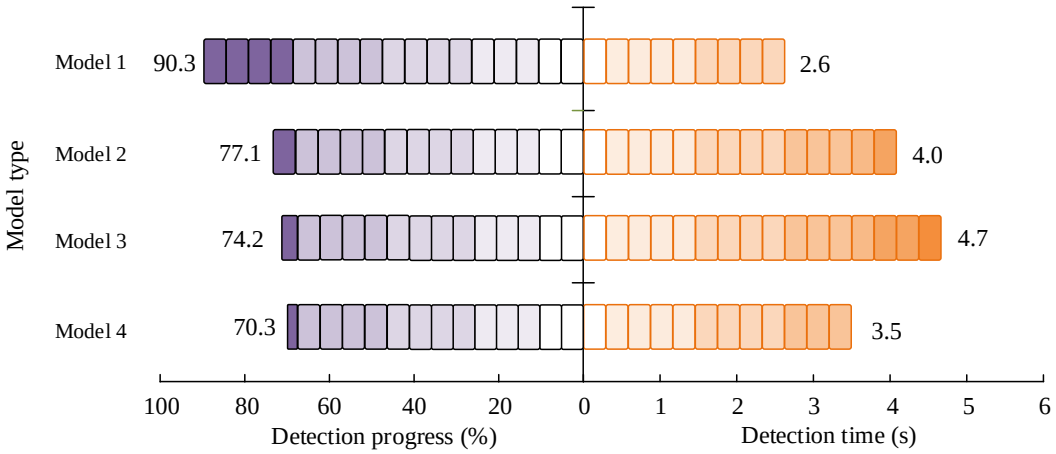


Figure 10: Comparison results of detection accuracy and detection time of the four models

From Figure 10, the detection accuracy of Model 1 is 90.3%, which is significantly better than 77.1% of Model 2, 74.2% of Model 3, and 70.3% of Model 4. In addition, from Figure 10, it can also be found that the detection time of Model 1 is 2.6 s, which is lower than that of Model 2 (4.0 s), Model 3 (4.7 s), and Model 4 (3.5 s). The above results indicate that the proposed online education security detection model performs better in terms of both detection accuracy and detection time dimensions. In a complex network education environment, the high detection accuracy indicates that the model can more accurately identify potential security threats,

including known and unknown attack types. This is essential to protect the privacy and integrity of user data and can effectively reduce security vulnerabilities caused by false or missed positives. The high-precision detection capability makes the model more reliable for educational institutions and users to use the online platform for teaching activities and learning. In addition, in the network environment with high real-time requirements, the length of detection time directly affects the speed of security response. The proposed model can complete the detection task in a very short time, which indicates that once a security threat is detected, the system can react quickly and take necessary defensive measures to prevent the attack from spreading further. The low detection time also improves the user experience, reduces delays or interruptions caused by security detection, and ensures continuity and stability of online education activities. In summary, the proposed high-precision and low-latency network education security detection model can greatly improve the security protection capability of education networks in actual deployment, and provide strong technical support for the healthy development of online education.

5 Discussion

This paper proposes an online education information security detection method based on improved cluster detection algorithm and software-defined wireless sensor network, and verifies its effectiveness through experiments. Compared to existing technologies, Banerjee's team's research focuses on improving network throughput and energy savings through the calculation of energy-efficient paths through zone partitioning and SDN controller management. In contrast, although SD-WSN technology is also adopted in this study, the focus is to realize real-time monitoring of educational environment through SD-WSN to improve information security detection capabilities. This difference is due to the different purposes of the study, Banerjee's team's research focused on network efficiency, while this study focused on information security. In addition, Younus et al. 's study optimizes routing paths through reinforcement learning, which improves network lifetime and packet delivery rate. Although this study also takes advantage of the flexibility of SD-WSN, the main goal is to combine improved clustering detection algorithms to detect security threats in the network in real time. This difference reflects the innovation in technology integration and application scenarios of this research. In addition, Nagaraja's team and Liang et al. proposed new clustering algorithms for intrusion detection and anomaly detection respectively, both of which improved classification accuracy to a certain extent. However, the improved clustering algorithm proposed in this study not only optimizes classification accuracy, but also realizes real-time monitoring of network environment and rapid response to abnormal traffic by combining with SD-WSN. This kind of real-time and adaptive improvement is not available in the existing clustering detection algorithms. Finally, a comparison with online education security-related research shows that Dincelli et al. 's research improves users' security awareness and behavior through gamification design, while Guo's team proposes a blockchain-based digital rights management system. Although the above research also focuses on the security issues in the field of online education, it does not directly focus on the real-time detection and analysis of network traffic. In contrast, this study provides an efficient and accurate information security detection scheme for online education through the combination of SD-WSN and improved clustering algorithm.

Compared with the above related research, the innovation of this research is mainly reflected in the following two aspects: the innovation of technology integration and the innovation of application scenarios. The innovation of technology integration refers to the combination of SD-WSN and improved clustering detection algorithm for the first time to realize real-time monitoring of network environment and rapid identification of abnormal traffic. This method not only improves the accuracy and real-time detection, but also overcomes the bottleneck of the traditional detection methods in computational complexity and detection efficiency. The innovation of application scenario refers to the application of the proposed method in the field of online education, which provides an effective information security guarantee for the emerging education model. This not only solves the information security problems faced by online education, but also promotes its healthy development. In summary, by integrating SD-WSN and improving cluster detection algorithm, a new information security detection method is proposed, which effectively makes up for the shortcomings of the existing technology in dynamic network monitoring and real-time detection.

6 Conclusion

As the importance of online network education information security increases, the problem of insufficient performance of the current network information security detection model becomes more and more serious. Aiming at the problem of online network education information security, the study introduces the ABC algorithm to get the improved clustering intrusion detection algorithm, and then combines it with the SD-WSN technology to get a new online education information security detection model. The performance test of the improved clustering intrusion detection algorithm shows that the accuracy of the proposed algorithm is 93.7%, which is significantly better than that of the DBSCAN algorithm (77.8%) and the K-mean algorithm (75.1%). In addition, a comparative test of the online educational information security detection model revealed that the model has a detection accuracy of 90.3%, which is significantly better than the 77.1% of Model 2, the 74.2% of Model 3, and the 70.3% of Model 4. The above data showed that the overall detection performance of this detection model was greater than the comparative models, and it could better protect the information security and data privacy of online network education. The innovative points and practicality of this study provide strong support for

improving the information security level of online education and help promote the healthy development of online education. However, there are still shortcomings in this study, such as the need to further explore how to optimise the algorithm parameters and strategies and improve the generalisation ability of the model to adapt to a wider range of educational environments.

Data availability statement

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

Conflict of interest

The authors declare that they have no competing interests.

References

[1] Tamil Selvi M, and Jaison B. Lemuria: a novel future crop prediction algorithm using data mining. The Computer Journal, 65(3), 655-666, 2022. https://doi.org/10.1093/comjnl/bxaa093

[2] Hebbi C, and Mamatha H R. Comprehensive dataset building and recognition of isolated handwritten kannada characters using machine learning models. Artificial Intelligence and Applications, 1(3): 179-190, 2023. https://doi.org/10.47852/bonviewAIA3202624

[3] Dhinakaran D, and Prathap P J. Protection of data privacy from vulnerability using two-fish technique with Apriori algorithm in data mining. The Journal of Supercomputing, 78(16), 17559-17593, 2022. https://doi.org/10.1007/s11227-022-04517-0

[4] Cho W, Kim J, and Lee C. Extension of simultaneous Diophantine approximation algorithm for partial approximate common divisor variants. IET Information Security, 15(6), 417-427, 2021. https://doi.org/10.1049/ise2.12032

[5] Wahyudi T, and Arroufu D S. Implementation of Data Mining Prediction Delivery Time Using Linear Regression Algorithm. Journal of Applied Engineering and Technological Science (JAETS), 4(1), 84-92, 2022. https://doi.org/10.37385/jaets.v4i1.918

[6] Shafique K, Khawaja B A., Sabir, F., Qazi, S., and Mustaqim, M. Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. Ieee Access, 8(3), 23022-23040, 2020. https://doi.org/10.1109/access.2020.2970118

[7] Chen J I Z, and Zong J I. Automatic vehicle license plate detection using K-means clustering algorithm and CNN. Journal of Electrical Engineering and Automation, 3(1), 15-23, 2021. https://doi.org/10.36548/jeea.2021.1.002

[8] Banerjee A, and Sufian A. Smart-Green-Mult (SGM): Overhear from topological kingpins in software defined wireless sensor networks. Journal of Ambient Intelligence and Humanized Computing, 12(1), 387-404, 2021. https://doi.org/10.1007/s12652-020-01984-2

[9] Younus M U, Khan M K, and Bhatti A R. Improving the software-defined wireless sensor networks routing performance using reinforcement learning. IEEE Internet of Things Journal, 9(5), 3495-3508, 2021. https://doi.org/10.1109/JIOT.2021.3102130

[10] Nagaraja A, Uma B, and Gunupudi R K. UTTAMA: an intrusion detection system based on feature clustering and feature transformation. Foundations of Science, 25(4), 1049-1075, 2020. https://doi.org/10.1007/s10699-019-09589-5

[11] Liang W, Xiao L, Zhang K, Tang M, He D, and Li K C. (2021). Data fusion approach for collaborative anomaly intrusion detection in blockchain-based systems. IEEE Internet of Things Journal, 9(16), 14741-14751, 2021. https://doi.org/10.1109/jiot.2021.3053842

[12] Dincelli E, and Chengalur-Smith I. Choose your own training adventure: designing a gamified SETA artefact for improving information security and privacy through interactive storytelling. European Journal of Information Systems, 29(6), 669-687, 2020. https://doi.org/10.1080/0960085x.2020.1797546

[13] Guo J, Li C, Zhang G, Sun Y, and Bie R. Blockchain-enabled digital rights management for multimedia resources of online education. Multimedia Tools and Applications, 79(8), 9735-9755, 2020. https://doi.org/10.1007/s11042-019-08059-1

[14] Kamar E, Howell C J, Maimon D, and Berenblum T. The moderating role of thoughtfully reflective decision-making on the relationship between information security messages and smishing victimization: An experiment. Justice Quarterly, 40(6), 837-858, 2023. https://doi.org/10.1080/07418825.2022.2127845

[15] Arbanas K, Spremic M, and Zajdela Hrustek N. Holistic framework for evaluating and improving information security culture. Aslib journal of information management, 73(5), 699-719, 2021. https://doi.org/10.1108/ajim-02-2021-0037

[16] Shafiq M, Tian Z, Sun Y, Du X, and Guizani M. Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city. Future Generation Computer Systems, 107(3), 433-442, 2020. https://doi.org/10.1016/j.future.2020.02.017

[17] Mei L. Model Construction of higher education quality assurance system based on fuzzy neural network. Informatica, 48(10), 2024. https://doi.org/10.31449/inf.v48i10.5676

[18] Rosenberg I, Shabtai A, Elovici Y, and Rokach L. Adversarial machine learning attacks and defense methods in the cyber security domain. ACM Computing Surveys (CSUR), 54(5), 1-36, 2021.

https://doi.org/10.1145/3453158

[19] Dhiman V, and Kumar M. In search of cluster-based routing protocol for WSN using consensus algorithm. International Journal of Communication Networks and Distributed Systems, 29(3), 290-314, 2023. https://doi.org/10.1504/ijcnds.2023.130571

[20] Ikotun A M, Ezugwu A E, Abualigah L, Abuhaija B, and Heming J. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information Sciences, 622(3), 178-210, 2023. https://doi.org/10.1016/j.ins.2022.11.139