Intrusion Detection System Based on Improved Artificial Immune Algorithm

Jilin Wang(✉), Zhongdong Wu, and Guohua Wang

School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China
wangjilin515@163.com, {WUZHD,wangguohua}@mail.lzjtu.cn

Abstract. Artificial immunity is widely used in the field of intrusion detection by simulating the accurate identification function of biological immune system to foreign intrusions, among which negative selection algorithm is the most widely used. However, due to the large amount of network data and high dimensionality, it often leads to problems such as low detection accuracy. In this paper, the method of combining principal component analysis (PCA) with genetic algorithm (GA) and negative selection algorithm improves the accuracy of intrusion detection. Among them, principal component analysis performs dimensionality reduction and feature extraction on intrusion data, and genetic algorithm is used to optimize the generation part of detector. The performance test was performed on the NSL-KDD standard test data set. The results show that this method improves the accuracy of intrusion detection and reduces the false alarm rate, which proves the effectiveness of the method.

Keywords: Intrusion detection · Artificial immune algorithm · Negative selection algorithm

1 Introduction

As society enters the Internet age, the application of the Internet has changed the social form and people's lives. While enjoying the convenience of the Internet, people are also taking the risk of privacy leakage. Due to the complex interests involved in the Internet, the diversity of participant goals, and network security issues also affect social development and threaten the security and stability of the world's politics, military, economy, and culture, the security of the Internet has become an important guarantee. Internet infringement and cybercrime are spreading across various industries in various fields. The methods and frequency of infringement are often unexpected, the consequences and extent of the infringement are even more shocking. In the first half of 2020 alone, in February the US natural gas pipeline company was attacked and forced to close the compression facility. In April, the industrial control facility of the Israeli water supply department was attacked by a cyber attack. In May, Venezuela's national power grid was attacked, causing a large-scale blackout across the country [1]. Existing statistics show

J. Xiong et al. (Eds.): MobiMedia 2021, LNICST 394, pp. 563–576, 2021.
https://doi.org/10.1007/978-3-030-89814-4_41

that in terms of economy alone, the loss caused by cyber infringement in the United States has reached more than US$17 billion per year. Other Western countries such as France have exceeded 10 billion francs, and Britain and Germany also have billions of US dollars [2]. All sorts of incidents show that the problems of network security has reached a point that cannot be ignored. After nearly two decades of development, intrusion detection technology has become an important technology and research direction in the security field, and has been applied in many industries such as military, finance, government, commerce, transportation, and electric power, and plays a key role.

Immunity is a physiological function of the human body, and it is the third line of defense of the human body. When foreign antigens invade the human body and destroy healthy tissues and cause the internal balance of the human body to be out of balance, the immune defense line quickly functions to eliminate antigens and maintain human health. It can effectively deal with a large number of different types of virus invasion. The immune system's solutions are distributed, flexible and adaptive. These features are exactly what the field of intrusion prevention expects. Artificial Immune System (AIS) is a comprehensive intelligent scientific research direction that integrates control science, computer science and life science [3]. As an intelligent system developed by learning from and using various principles and mechanisms of the biological immune system, the artificial immune system is an improvement of the modern network behavior detection system from the perspective of biological immunity, and finally makes the system have a greater similarity with the biological immune system It can solve complex and changeable network viruses well. Modern researchers have made substantial improvements on the basis of the artificial immune system and applied it to the field of intrusion detection.

Zhang Ling [4] designed an intrusion detection algorithm based on random forest and artificial immunity (RFAIID), and proposed an antibody forest model; the clone selection algorithm was used to obtain an excellent large sample antibody set to improve the adaptability of intrusion detection. However, there is still the problem that the redundant attributes in the algorithm affect the detection speed of the algorithm. Xin Zhuang [5] combined rough set theory with artificial immunity, deleted redundant attributes, improved the operating efficiency of the algorithm, improved the stability of the model, accelerated the convergence speed, and ensured the superiority of the antibody. Feng Xiang [6] added segmentation technology and key bits to the negative selection algorithm to avoid the matching loopholes caused by the constant matching probability and reduce the system missed detection rate; and the clone selection algorithm in the genetic algorithm and the improved negative selection algorithm are compared In combination, the dynamics and diversity of detector generation are improved. Liu Hailong [7] proposed a high-dimensional real-valued detector distribution optimization algorithm, and used the principal component analysis method and affinity calculation to solve the problem of low detector distribution in the high-dimensional real-valued space in immune intrusion detection. The detector distribution was optimized through affinity comparison.

Amira Sayed A. Aziz [8] use detectors generated by immune algorithms to detect abnormal activities in the network. Minkowski distance function and Euclidean distance test detection process. Adeni jiuwashola David [9] improved the NSA detector generation stage, used neural network technology to build a model, and developed a new model called NNET-NSA, which has a higher detection rate. Soodeh Hosseini [10] proposed

a new combination of abnormal process detection technology. This technique unifies the negative selection algorithm and the classification algorithm. This method reduces the training time while improving the accuracy of the system. Based on the artificial immune system, a host-based abnormal process detection framework is established. Nguyen Thanh Vu [11] combines artificial immune system (AIS) and deep learning to classify benign and malignant documents. Use AIS to build a clone of the malware detector to improve the accuracy of the unknown virus detection rate, and then use the Deep Belief Network (DBN) to calculate and train the risk level of the file, and evaluate the performance of the system.

It can be seen from the above that inspired by the artificial immune system, many improved methods have been successfully applied to intrusion detection of computer network systems. However, there are some problems that hinder the wide application of this method [12]:

(1) Uneven distribution of attack types. The current types of attacks on the network include DOS, U2R, R2L, and Probe. Since the probability of each attack is different, there will be an uneven data distribution problem in the data set. In attack detection, the smaller the amount of data, the lower the probability of being detected.
(2) Intrusion detection speed problem. The training speed and detection time are not so ideal when dealing with large-scale data. So far, few intrusion detection systems have both accuracy and speed. How to improve the detection speed is also a problem that should be paid attention to in the field of intrusion detection.
(3) Large-scale data sets are easy to be missed. In high-speed switching networks, intrusion detection systems cannot detect all data packets well, and the accuracy of analysis is not high. Faced with the current increasing data dimensions and complex network behaviors, there are a large number of misjudgment warnings or a longer judgment time, which is likely to cause the problem of missing reports and missed detections.
(4) Poor flexibility. The fundamental reason why the system is frequently successfully attacked today is that the protection strategy of each system is passive and static, and it does not have autoimmune functions and flexibility, which causes the system to face risks [13]. Therefore, it is necessary for intrusion detection to have the ability of autonomous learning, which can be adjusted according to the changes of the environment, so as to optimize the performance effect.

Based on the above, the research of this paper starts from the perspective of the combination of artificial immune algorithm and intrusion detection technology, solves the current vulnerability attack problem of computer network, improves and perfects some problems in artificial immune algorithm, and makes it can play a greater role in the field of intrusion detection. In this article, the method of combining principal component analysis (PCA) with genetic algorithm (GA) and negative selection algorithm is adopted. We will build an architecture of an intrusion detection system based on an improved negative selection algorithm, and divide the architecture into three modules for detailed introduction.

The first section mainly introduces some major network security incidents currently encountered in the world, and then introduces the research results and existing problems in the research direction of network intrusion detection at home and abroad in recent years. On this basis, the article explains the significance. The second chapter introduces the definition of negative selection algorithm and the related formulas of genetic algorithm. The third chapter builds an intrusion detection system model of an improved artificial immune algorithm; makes a related flow chart, explains the overall structure of the system, analyzes the function of each module of the system, and the fourth chapter uses the NSL-KDD data set to test, verified the validity of the model. The fifth chapter summarizes the research work of this article, analyzes the problems existing in the experimental research, and prospects the follow-up work.

2 Definition of Related Algorithms

2.1 Negative Selection Algorithm

Inspired by the theory of immune model, researchers at home and abroad have proposed a variety of immune intelligent algorithms, represented by negative selection algorithms, clone selection algorithms, immune network algorithms, and so on [14]. Among them, in 1997, Forrest of the University of New Mexico in the United States proposed a negative selection algorithm [15]. She studied the process of human thymocytes to produce immune T cells, that is, by negating the lymphocytes that produce immune responses to the body, and divided the detection procedures into self-collection and non-self collection. By negating the detection program that matches the pattern with normal traffic, the collection of detection programs that do not match the normal traffic is used as a detector to filter network traffic. Negative selection algorithms are widely used in intrusion detection, and have high research value (Fig. 1).

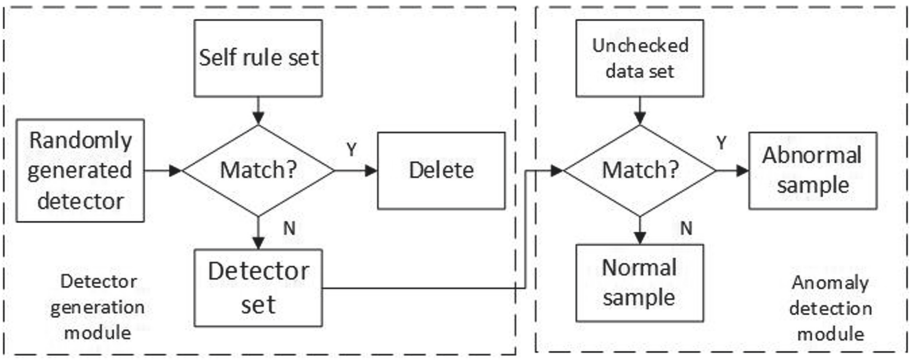

Fig. 1. Negative selection algorithm intrusion detection architecture.

Definition 1: Antigen. Antigen $g = \{g|g =< X1, X2, X3, \cdots Xn >\}$ means all samples in the space.

Definition 2: Self-set and non-self-set. Self (S) represents normal samples in space; Nonself (N) represents abnormal samples in space.

$$\begin{aligned} S \cup N &= \mathrm{g} \\ \mathrm{S} \cap \mathrm{N} &= \phi \end{aligned} \tag{1}$$

Definition 3: Detector set. Mature detector set D, Represents a mature detector obtained by judging whether it matches the self-set S. D exists in the non-self region. Among them, rd represents the detector radius.

$$D = \{\mathrm{d} | d =< y1, y2, y3, \cdots, yn, rd >\} \tag{2}$$

Definition 4: Affinity calculation. Affinity is the firmness of the binding of an antibody binding site to an epitope. The tighter the binding, the less likely the antibody is to separate from the antigen. Affinity function (fit): The value of the degree of match between an antigen and an antibody is generally represented by a real number between 0 and 1. Use function fit(i,j) represents the affinity between antigen i and antibody j. Obviously, affinity is related to the degree of matching between antibody and antigen. The affinity can be expressed by the following formula:

$$fit(i, j) = \frac{1}{1 + tij} \tag{3}$$

In the formula, tij represents the distance between antigen i and antibody j. The distance function can be Hamming distance or Euclidean distance. The matching function is as follows:

$$d = \sum_{i=1}^{L} \delta \tag{4}$$

Hamming distance:

$$d(i, j) = \sqrt{|xi1 - xj1|^2 + \cdots + |xip - xjp|^2} \tag{5}$$

2.2 Genetic Algorithm

GA is an algorithm with learning function, which is a random search optimization algorithm. Its algorithm principle draws on the natural selection and genetic mechanism of nature in Darwin's theory of evolution. It was proposed by Holland in 1975 and published in the book "Adaptation in Natural and Artificial Systems". In this book [16], the GA system machine was comprehensively discussed for the first time and the corresponding data theory proof was given. This kind of algorithm is obviously different from traditional and optimal solution algorithms. It is no longer highly dependent on gradient information like the algorithm. Instead, it uses a group search strategy and makes the individuals between them complete the corresponding exchange and mutation. According to the fitness function, the calculation of the corresponding value is completed, and

new individuals are formed by means of crossover, mutation and selection methods, that is, the problem solution is obtained, and the search space of the group is then realized to achieve the overall optimal solution. So this method can solve those non-linear problems, genetic algorithm puts forward a new idea and new method. Based on the theory of biological evolution, genetic algorithms are not directly related to specific problems. Instead, they receive the described objective function through a randomly generated population, and find the corresponding optimal solution with the help of genetic and natural selection models (Fig. 2).

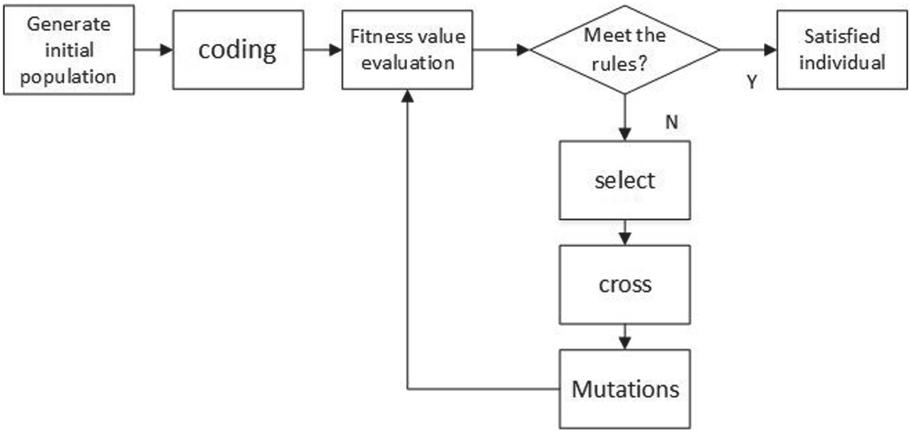

Fig. 2. Genetic algorithm flow chart

In the fitness function design link needs to strictly follow the rules of reasonableness, generality and specification. The quality of a fitness function should also meet the widest possible versatility, so that users do not need to change the fitness function when solving various problems. Generality requirements are higher requirements for fitness function design. In this research, the functions mentioned in reference [17] are used, as follows:

$$Fit(f(x)) = \begin{cases} 1 - 0.5x\left[\left|\frac{f(x)-b}{a}\right|\right]^a, & |f(x)-b| < a \\ \frac{1}{1+\left[\left|\frac{f(x)-b}{a}\right|^\beta\right]}, & |f(x)-b| \geq a \end{cases} \tag{6}$$

Generally, under an ideal background, the value of b is $f\ \min(x) = y*$, if the current fitness value is 0.5, then a represents the distance from f(x) to $f\ \min(x)$.Since the applicable environment of the fitness function needs to be considered, it is necessary to set β to 2.

In the above formula, both a and b will be optimized for the next generation under the action of genetic algorithm. b is f(x)xi, At this time, the formula of a is:

$$a = \max X\left[0.5, \frac{|f\ \max(i) - f\ \max(i)*|}{30}\right] \tag{7}$$

3 Intrusion Detection System Based on Improved Negative Selection Algorithm

3.1 Data Set

The NSL-KDD data set is used in the selection of experimental data. The NSL-KDD data set is a network environment established by Lincoln Lab to simulate the US Air Force LAN. It collects 9-week TCPdump network connection and system audit data to simulate various user types, various network traffic and attack methods. Make it like a real network environment. This data is based on a large number of improvements on the traditional authoritative intrusion detection data set. NSL-KDD removes the redundant data in the KDD99 data set. All types of attack samples retain only one record, so that the data will not be repeated records and deviations. NSL-KDD data is widely used in intrusion detection research on computer networks, so it is feasible to conduct intrusion detection experiments based on the NSL-KDD data set. Each network connection in the data set is marked as normal (normal) or abnormal (attack), and the abnormal types are subdivided into 4 categories, a total of 39 attack types, of which 22 attack types appear in the training set, and 17 unknown attacks The type appears in the test set. The purpose of this design is to test the generalization ability of the classifier model. The ability to detect unknown attack types is an important indicator for evaluating the quality of an intrusion detection system [18].

3.2 Related Functional Modules

(1) Data Preprocessing Module. The first step is to preprocess the data set, and then divide the input data set into self-set and non-self-set. Also, because the characteristic attributes of the sample records are messy, the values are referenced after comparison according to 41 attribute standards, which will lead to inconsistent values, and the final test results will be disordered. Therefore, in order to improve the versatility of this experiment, we preprocess the collected data before testing the data. The specific method is: first standardize the collected data, use one-hot encoding to transform the discrete data in the data set in the data preprocessing part, and then use the principal component analysis algorithm to reduce the dimension of the features, reducing the complexity of the calculation and save time.

Convert the string to discrete numbers, use one-hot encoding to transform the features, convert the discrete type to a numeric type to normalize its value, and map the values to [0,1] to reduce error. After finding the maximum and minimum values of the attributes of each record, use the following formula to normalize them:

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

After the values are normalized, the relationship between the original data is retained, and the errors caused by the large value difference of each attribute are also eliminated. And to ensure that the program converges faster at runtime (Table 1).

Table 1. Conversion of coding features.

Before mapping	After mapping
TCP	1, 0, 0
UDP	0, 1, 0
UDP	0, 0, 1

Principal component analysis (PCA) is a commonly used method of feature dimensionality reduction and feature extraction. Through orthogonal transformation, a group of potentially correlated variables is converted into a group of linearly uncorrelated variables. The converted group of variables is called principal component. Principal component analysis analyzes the data model, reduces the dimensionality of the data set while ensuring the minimum loss of information, and projects the feature space onto a smaller subspace to better describe the data.

The steps for preprocessing the data using principal component analysis are as follows:

Step1: Standardize the data set.

Step2: The eigenvectors and eigenvalues are extracted from the covariance matrix. The calculation formula of the covariance matrix is as follows:

$$Cvjk = \frac{1}{n-1}\sum_{i=1}^{n}(xij - xj')(xik - xk') \tag{9}$$

Where the mean vector:

$$x' = \frac{1}{n}\sum_{k=1}^{n} xi \tag{10}$$

The covariance between the two features is calculated as follows:

$$CM = \frac{1}{\mathrm{n}-1}((X - x')^T(X - x')) \tag{11}$$

Step3: Arrange the eigenvalues in descending order, and select the k eigenvectors corresponding to the k eigenvalues, and k is the dimension of the new eigenspace.

Step4: The projection matrix W is constructed by the selected k eigenvectors.

Step5: Pass the original data set through the projection matrix W to generate a k-dimensional feature subspace,$Y = X * W$.

(2) Detector Generation Module. The creation and selection of the detector is based on the NSA algorithm, which generates a random detector, matches the generated detector with the self-set, if it matches, deletes the detector, if it does not match, then generates a mature detector set D. Because the genetic operator has the characteristics of selection, mutation, crossover, fitness, etc., the genetic algorithm is used to improve the traditional negative selection algorithm, generate an optimized and balanced subset, and optimize the generation and distribution of detectors. When the algorithm runs to the set genetic total algebra, the algorithm will stop running. The specific process is as follows:

Step1: In the GA algorithm, scientific coding will affect the performance of the algorithm and its population diversity to varying degrees. Compared with real number coding, binary has a higher search ability. Binary coding is used here. Binary coding is the structure of the original problem is transformed into the bit string structure of the chromosome.

Step2: Use the fitness function mentioned in the above formula (6) to calculate the corresponding values of different populations.

Step3: Select. The selection function means to select certain individuals from the parent population for inheritance. In this article, random competitive selection method is used to perform selection operations. According to the roulette gambling selection method, a pair of individuals are selected each time, and then the fitness values of the two individuals are compared. The one with the higher fitness will be selected. Repeat this process until the total number reaches the specified number.

Step4: Cross. The group obtained after the selection is processed in a uniform crossover method according to the predetermined crossover rate. For randomly selected individuals X1, X2, two intersection points are determined by a random method, and then three integers of 0, 1, and 2 are randomly generated. When the random number is 0, the front part of X1, X2 crosses; when the random number is 1, the middle part of X1, X2 crosses, and when the random number is 2, the back part of X1, X2 crosses.

Step5: Mutations. The variogram is used to change the value of one or more bits in the chromosome. The probability of mutation is generally 0.02–0.03. The variogram aims to improve the fitness of the chromosome by introducing new characteristics.

Step6: Repeat steps 1 through 5 above until the set maximum genetic algebra is completed.

For the detection device, its main function is to detect foreign intrusions, which is similar to lymphocytes in the immune system. The optimized and balanced subset is obtained through the above steps, which is defined as a detector, and the detector matches with the self set. If it matches, the detector is deleted, and if it does not match, a mature detector set is formed. The implementation process of this mature detection is shown in Fig. 3.

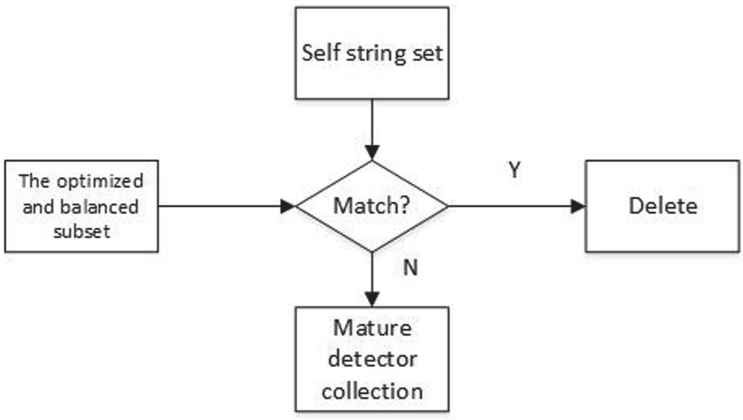

Fig. 3. The production process of the mature detector.

(3) Intrusion Detection Module. The mature detector is sent to the detection module, and the match between the sample to be tested and the generated detector is calculated

and processed. If it matches, the intrusion data is found, and the relevant processing is performed immediately. Euclidean distance is used here, the match between detector d and antigen g is defined as:

$$mdg = 1/\left(1 + \sqrt{\sum (di - gi)^2}\right) \tag{12}$$

Calculate the Euclidean distance between the data to be detected and the self-set in the space to determine whether a match occurs. By setting a threshold r, compare the size of r and l to determine whether they match. When $r \leq l$, the sample is far away from the detector, and there is no intersection or mismatch. If the result is the opposite, then match. When generating the detector, for its radius, a fixed radius can be artificially set.

3.3 The Whole Frame

See Fig. 4.

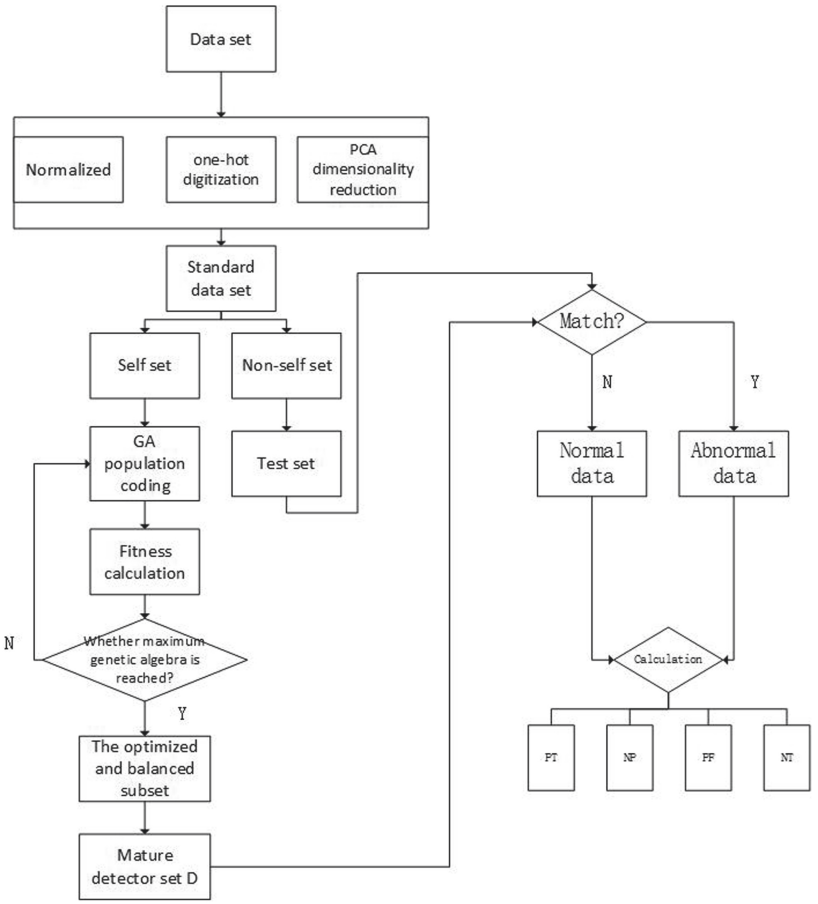

Fig. 4. Improved negative selection algorithm intrusion detection structure

4 Experiment

In order to verify this algorithm, the authoritative NSL-KDD data set is selected, which is a more classic data set in the current intrusion detection field. Each record data covers normal and abnormal features, and the latter can be subdivided into four types of attacks. The evaluation indicators used in the experiment are Accuracy (ACC), Detection rate (DR) and False positive rate (FAR). The meanings and calculation expressions of each indicator are shown in Table 2. Among them, TP represents the number of samples that actually classify abnormal samples as abnormal, TN represents the number of samples that actually classify normal samples as normal, FP represents the number of samples that actually classify samples that are actually normal as abnormal, and FN represents The samples that are actually abnormal are classified as the number of normal samples.

Table 2. Performance evaluation indicators.

	Meaning and computational expression
DR	Detection rate (DR) refers to the proportion of aggressive behaviors that are correctly classified in a concentrated antigen: $DR = \frac{TP}{TP+FN}$
FAR	False positive rate (FAR) refers to the ratio of the number of all normal antigens misidentified as attack in the antigen set to the total number of normal antigens in the test set: $FAR = \frac{FP}{TN+FP}$
Accuracy	Classification accuracy rate (ACC) refers to the ratio of the number of correctly classified antigens in the antigen set to the total number of samples in the antigen set: $ACC = \frac{TP+TN}{TP+TN+FP+FN}$

Now use the improved negative selection algorithm proposed in Chapter 3 to conduct experiments, and select a part of the NSL-KDD data set, which contains 20,000 normal and intrusive values. Before calculation, set the corresponding parameters as follows (Table 3):

Table 3. Experimental parameter settings.

Data	Number of mature detectors	Generations	The self set	Non-self test set	Non-self training set
20000	1000	200	93	6637	13270

Set the number of test sets as 1000. When the value of detector radius Rd is determined and the range of radius variation is maintained from 0.2 to 0.8, the accuracy and false

positive rate of the algorithm are tested to see if they are still within the acceptable range. In each case, the experiment was run 20 times, and the average value was taken. The results are shown in the following Table 4:

Table 4. Values of detection rate and false alarm rate under different self-radius

rd	DR (%)	FAR (%)
0.2	89.25	9.3
0.3	94.97	6.7
0.4	95.83	2.3
0.5	97.98	1.1
0.6	96.51	1.9
0.7	94.97	1.3
0.8	92.41	2.1

It can be seen from the above table that when the self-radius is small, the space occupied by the normal sample is relatively small, and the non-self space occupied by the corresponding detector is relatively large. Therefore, during intrusion detection, most abnormal behavior samples can be detected, but the self-region is too small, causing some new normal samples to fall in the non-self-region, making the false alarm rate higher. With the increase of the self-body radius, the self-body area becomes larger and larger, making the non-self-body area occupied by the detector smaller and smaller, which will correspondingly cause a decrease in detection rate and false alarm rate. On the whole, when the self-body radius rd = 0.5, the overall performance is the best.

Table 5. Comparison of experimental results

The experimental method	DR (%)	FAR (%)	ACC (%)
Improved negative selection algorithm	96.43	0.97	97.88
The original negative selection algorithm	88.92	1.01	90.12
Single-SVM	89.42	3.62	95.32
BP + GA	82.18	2.74	89.75

From the results obtained in Table 5, it can be seen that the accuracy and detection rate of the intrusion detection behavior of the improved negative selection algorithm is much higher than that of the original negative selection algorithm, and the false alarm rate is also lower. Compared with other methods, the performance of this algorithm is also significantly better than several other algorithms. Through multiple sets of comparative experiments, using the NSL-KDD data set to test, the results show the effectiveness of the method.

5 Conclusion

Nowadays, intrusion detection systems are widely used as a line of defense for security protection. Artificial immune algorithms simulate the characteristics of biological immune systems against foreign intrusions, which have huge application space in intrusion detection. However, due to the large amount of network traffic and unobvious features, the correct rate of intrusion detection is often not high, and the detection effect is poor. Therefore, this article improves the Negative Selection Algorithm (NSA) from two aspects of feature extraction and detector generation, and discusses the application of the method of combining PCA and GA with negative selection algorithm in intrusion detection. When facing a large number of high-dimensional features of network data, PCA is used as a feature extraction method to achieve the purpose of dimensionality reduction. The GA is used to optimize the detector generation. However, this algorithm also has some shortcomings. For example, it consumes a lot of time during detection and is easy to fall into local optimal problems. These aspects need to be taken seriously in the future.

Acknowledgments. This work was supported by State Grid Gansu Electric Power Research Institute Project No. 520012 Intelligent Recognition and Defense of Intrusion Behavior of Electric Power Information and Physical Fusion System in Cyber Attack Environment.

References

1. Bi, R., Chen, Q., Chen, L., Xiong, J., Wu, D.: A privacy-preserving personalized service framework through Bayesian game in social IoT. Wireless Commun. Mobile Comput. **2020**, 1–14 (2020)
2. Xiong, J., Ma, R., Chen, L., et al.: A personalized privacy protection framework for mobile c2rowdsensing in IIoT. IEEE Trans. Industr. Inf. **16**(6), 4231–4241 (2020)
3. Cooper, E.L.: Evolution of immune systems from self/not self to danger to artificial immune systems (AIS). Phys. Life Rev. **7**(1), 55–78 (2010)
4. Zhang, L., Zhang, J., Sang, Y., et al.: Intrusion detection algorithm based on random forest and artificial immunity. Comput. Eng. **46**(8), 146–152 (2020)
5. Xin, Z., Wan, L.: Iintegrated intrusion detection model based on artificial immunity. Comput. Eng. Design **40**(10), 2799–2804 (2019)
6. Feng, X., Ma, M., Zhao, T., Yu, H.: Intrusion detection system based on hybrid immune algorithm. Comput. Sci. **41**(12), 43–48 (2014)
7. Liu, H., Zhang, F., Xi, L.: High dimensional real-valued detector distribution optimization algorithm for intrusion detection based on immunity. J. Tsinghua Univ. (Sci. Tech.) **52**(10), 1415–1419 (2012)
8. Amira Sayed, A., et al.: Artificial immune system inspired intrusion detection system using genetic algorithm. Informatica: Int. J. Comput. Inf. 347–357 (2012)
9. David, A.O., Joseph, U.J.: A novel immune inspaired concept with neural network for intrusion detection in cybersecurity. Int. J. Appl. Inf. Syst. **12**(30), 13–17 (2020)
10. Hosseini, S., Seilani, H.: Anomaly process detection using negative selection algorithm and classification techniques. Evol. Syst. (2019). https://doi.org/10.1007/s12530-019-09317-1
11. Thanh Nguyen, V., Hoang Dung, L., Dinh Le, T.: A combination of artificial immune system and deep learning for virus detection. Int. J. Appl. Eng. Res. **13**(22), 15622–15628 (2018)

12. Zhang, Y., Wang, L., Sun, W., et al.: Distributed intrusion detection system in a multi-layer network architecture of smart grids. IEEE Trans. Smart Grid **2**(4), 796–808 (2012)
13. Zhao, J., Zhang, X., Di, F., et al.: Exploring the optimum proactive defense strategy for the power systems from an attack perspective. Secur. Commun. Netw. **6**, 1–14 (2021)
14. Yang, H., Li, T.: Intrusion detection based on T cell receptor principle. Int. J. Performabil. Eng. **15**(9), 2407–2413 (2019)
15. Forrest, S., Perelson, A.S., Allen, L., et al.: Self-nonself discrimination in a computer. Comput. Soc. Symp. Res. Secur. Privacy 202–212 (1994)
16. Holland, J.H.: Adaptation in natural and artificial systems. Ann. Arbor. (1975)
17. Liu, Y.: Research on fitness function in genetic algorithm. J. Lanzhou Polytech. College **3**, 1–4 (2006)
18. Weikai, W., Lihong, R., Lei, C., et al.: Intrusion detection and security calculation in industrial cloud storage based on an improved dynamic immune algorithm. Inf. Sci. 43–557 (2018)
19. Aldhaheri, S., Alghazzawi, D., Li, C., Barnawi, A., Bandar, A.: Alzahrani. Artifificial immune systems approaches to secure the internet of things: a systematic review of the literature and recommendations for future research. J. Netw. Comput. Appl. 157 (2020)
20. Al-Qatf, M., Lasheng, Y., Al-Habib, M.: Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. IEEE Access **6**, 52843–52856 (2018)