Application Analysis of Artificial Intelligence in Library Network Security

Sheng Li1,*

1Shandong Polytechnic College, Jining 272100 Shandong, China

*Corresponding author e-mail: lisheng@sdp.edu.cn

Abstract: In the information age, the rapid development of computer network technology and artificial intelligence technology has created conditions for the development of traditional libraries to digital libraries. While computer network technology and artificial intelligence technology continue to promote the development of digital libraries, library network security issues also arise. The purpose of this paper is to explore the application analysis of artificial intelligence technology in library network security. This article puts forward the problem of library network system security, describes the concept and characteristics of library network system security, and aims at the basic factors affecting the security of digital library network system, constructs artificial intelligence-based books from the perspective of management and technology Preventive mechanism for library network system security. In terms of situation prediction, this article uses artificial intelligence technology and analytic hierarchy process to evaluate the network security situation. This method can quickly extract parameters and can quickly evaluate the current network, thereby improving the real-time performance of the situation assessment. The optimized support vector machine method overcomes the subjectivity of support vector machine parameter selection, thereby improving the prediction accuracy of the network security situation. The simulation experiment verifies that the optimized support vector machine prediction model has a higher value for the network security situation. Forecast accuracy. Experimental research shows that the PSO-SVR prediction model used in this article can predict the future situation value relatively closely, help library management network administrators analyze the future network situation, and prompt library management network administrators to adjust defense strategies in time based on threat assessment to maintain library management network Safe and healthy operation.

Key words: Artificial Intelligence, Network Security, Situational Awareness, Anomaly Detection
1. Introduction
In recent years, our country has listed information system security technology as a key development area of the century, and as a key support subject of the national plan and the National Natural Science Foundation of China, this is enough to explain the importance of network security. The research on network security situational awareness based on artificial intelligence can promote the defense of the network against viruses and man-made attacks, so as to ensure the safe and healthy operation of the network. Harm, discovering potentially malicious intrusions, improving the system's counterattack capability and ensuring the security of national strategic development, etc. are of great significance.

Network security situational awareness has developed globally in recent years. Although this type of research is in its infancy, many universities have made certain research results on related technologies. Wilkins has done research from many aspects such as the realization technology of network security situation awareness, the method of determining the idle value of the network situation awareness system, the quantitative assessment of the network survival situation, and the network situation prediction [1]. Talwar R conducted an in-depth study on how to select network security status indicators and the indicator system of network situation awareness [2]. Bai Hongjun proposed a neural network-based security situation prediction method. This method can better predict the future network situation to a certain extent, but when the number of learning samples is large, it will fall into the disaster of dimensionality, and the generalization performance is not good. High [3].

This paper proposes a quantitative assessment method of network security situation based on artificial intelligence algorithm and analytic hierarchy process, which provides a basis for further rapid and real-time situation assessment. This paper analyzes the advantages and disadvantages of the support vector machine method, and proposes a support vector machine method based on particle swarm optimization algorithm, which improves the accuracy of network security situation prediction.

2. Application of Artificial Intelligence in Library Network Security

2.1 Intelligentization of Library Network Security Situation Awareness
(1) Model library
Input the existing model into the model library. Due to the different device topology in the network and the different focus of the relevant network, the situation assessment criteria used for the situation change of the entire local network are slightly different, according to the network and the administrator. Demand, adopt appropriate models to evaluate the situation, reflect the current network conditions more accurately, and maintain the safe operation of the library network [4-5].

(2) Knowledge reasoning library
From the situation assessment and threat assessment and human operations, the knowledge reasoning library performs self-learning, updates knowledge with the learned experience, and adjusts the response granularity and early warning preset settings, so that the response granularity is more reasonable, and the granularity of the response will not be excessive. Large, when the system is close to crash, it will not respond or because the response granularity is too small, and frequent response will affect the administrator's network management level [6-7].

(3) Visual interface
The current situation assessment and threat assessment can be displayed in intuitive graphics. On the basis of the situation assessment and threat assessment, response and early warning and situation prediction can be carried out. The situation prediction is related to the response and early warning. The situation change in the future period can affect the granularity of the response and early warning are displayed to the administrator in the form of graphics and sound through a visual interface.

2.2 Data Fusion
(1) Data preprocessing
Due to the diverse types of security devices in the network, and the data formats generated by
various security devices are not necessarily uniform, the overall processing of the data is quite difficult, so the data needs to be preprocessed. Data preprocessing is to complete data filtering and data. The functions of simplicity, data format conversion and data storage provide convenience for higher-level processing in the future [8-9].

2) Event filtering

Network events mainly refer to firewall log events, security scan logs, intrusion alarm events, host log events, etc. New technologies such as firewalls and security scanning have been applied to the network field, although the pressure on the network has been relieved, but it has also brought some new problems. These security incidents include not only the facts about the behavior of the intruder and the attacker, but also some irrelevant events used by the attacker and the intruder to hide their own behavior, as well as the events caused by the library network security equipment itself, such as Due to the small granularity of industrial alarms, an atomic attack will cause a series of false alarms, which greatly increases the number of attack alarms. If you directly correlate events, it will affect the accuracy and execution efficiency of the correlation. In order to detect the intrusion behavior of the attacker in time, the false and misreported events must be filtered in the process of library network security event correlation.

2.3 Based on Hierarchical Artificial Immune Situation Assessment

The research of artificial immune algorithm has become an important content in the field of artificial intelligence research. It penetrates into other fields of research technology, merges and promotes each other from all levels, and has important theoretical significance and practical application value for the development of information science and technology.

During the operation of the network, when no attack is received, the average number of alarms on the network remains the same, and fluctuates at a constant value during a certain period of time. When the network service receives an attack, the number of network alarms varies in response to changes in attack intensity, after the attack stops, the network returns to calm and the alarm system returns to its normal state. Therefore, the security situation of the network's service layer can be calculated according to the concentration formula [10].

Suppose that when a network service is not under attack, the number of attacks detected on a certain service $i$ is $c_i$, $n_i$ is the number of attacks detected on a certain service $i$ under the normal operation of the network, and $\sigma_i$ is the risk coefficient of a certain service attack, so the service attack concentration $d$ in the network can be expressed as:

$$d = 1 - \frac{1}{1+\ln(n_i-c_i+1)}$$ (1)

So the service risk index $F_s$ is:

$$F_s = d$$ (2)

The importance of host $i$ in a certain network segment is different. If the importance of a certain service of the host is $\theta_i$, then the host's attack concentration $F_H$ can be expressed as:

$$F_H = \sum_{i=1}^{n} v_i d$$ (3)

Where $d$ is the situational concentration of the service layer, and $v_i$ is the host weight normalized by $\theta_i$.

$$v_i = \frac{\theta_i}{\sum_{i=1}^{n} \theta_i}$$ (4)

The network layer situation value is a function of the security risk index of all hosts in the network and the host importance weight. If the importance weight of each host is $\mu_i$, the network layer situation concentration can be expressed as

$$N_s = \sum_{i=1}^{n} w_i f_i$$ (5)

Among them, $w_i$ is the weight normalized by the important weight of each host, and $f_i$ represents the situation concentration of each host.
2.4 Improved Algorithm Based on Population Cognition

(1) Role division
The population is divided into two roles: leader and follower. During each search, the leader in the population determines a good search direction, and other followers line up in a row to search in this direction, so that the calculation loss of determining the search direction is only borne by the leader, and other followers omit the search direction.

(2) Separation rules
The population is arranged in a row and searched along the direction determined by the leader. The initial position of each individual is evenly distributed on a straight line, and the appropriate spacing is maintained between each other to ensure that a face search is performed in the search space to prevent excessive clustering of the population, and lose the complexity of the group.

(3) Parallel rules
The direction of the population as a whole is the same, and the leader determines the search direction of the entire group. This direction is considered to be the optimal search direction for the population, and other groups search in this direction.

(4) Renewal of the leader and re-formation of the population
When the population is running, there is an individual acting as the leader, and the leader is rotated. After each round of search, if the leader cannot lead the entire population to find a better fitness value, the population will look for an individual to replace the head goose.

2.5 Security of Digital Library Network Layer
The core of the security problem of the digital library network layer is whether the digital library network is controlled, that is, whether any user from the source of the IP address can access the network. If the entire network of the digital library is compared to an office building, the security considerations for the network layer are like setting up gatekeepers for the building. The gatekeeper will carefully look at every visitor, and once a dangerous visitor is found, he will be turned away.

When accessing the network system through the network channel, each user will have an independent IP address, which can roughly indicate the user's source location and source system. By analyzing the source IP, the target website can preliminarily determine whether the data from this IP is safe, whether it will cause harm to the network system, and whether users from this IP have the right to use the data of this network. Once it finds that some data comes from an untrusted IP address, the system will automatically block the data out of the system, and most systems can automatically record those IP addresses that have caused harm in the past, so that their data will not be available for the second time Cause harm.

The products used to solve the security problem of the network layer mainly include firewall and VPN-virtual private network. The main purpose of the firewall is to determine the source IP, keep dangerous or unauthorized IP data out of the system, and let only safe IP data pass. Generally speaking, if the internal network of a digital library is to be connected to the public Internet, a firewall product should be configured between the two to prevent the internal data of the digital library from leaking.

3. Experimental Research on the Application of Artificial Intelligence in Library Network Security

3.1 Principle of Support Vector Machines
SVP can better solve the linear and nonlinear regression fitting problems, but whether the selection of the insensitive loss coefficient $\varepsilon$, the penalty coefficient $C$, and the width coefficient $\gamma$ of the kernel function in SVP is reasonable directly determines the learning accuracy and generality of the SVP
model. In the pure SVP application, the trial and error method is generally used to select better parameters, but this method lacks rationality and cannot select the optimal parameters, so it is particularly important to use optimization algorithms to select reasonable parameters.

3.2 SVR Library Network Security Situation Prediction Based on PSO Optimization

Use PSO to optimize the parameters to be determined in the SVR, and then pass the parameters to be optimized to the SVR. The SVR trains and predicts the samples, and outputs the prediction results. The fitness function is used to calculate the fitness. If it exceeds the number of iterations or this time If the fitness in the iteration is less than the preset fitness, then jump out of the iteration, otherwise update the particle swarm to continue training.

Fitness function:

\[ f_{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \]

Where \( n \) is the training sample, \( y_i \) is the actual value, and \( \hat{y}_i \) is the predicted value.

3.3 Experimental Data Acquisition and Processing

From the perspective of the attack process, a complete network attack is usually divided into four stages: detection, attack, hiding, and use of backdoors. Each stage has a different degree of harm to the current host and network. Among them, the attack is the most harmful and directly affects the current network. Operation quality, hiding and using backdoors is a potential high-risk threat, and detection is the detection of attack targets. In the network alarm, it is actually caused by detection and attacks. Hidden and used backdoors cannot cause alarms, so in the network in the calculation of the situation value, only detection and attack are considered, and potential threats should be considered in the analysis of the library network security status.

4. Application Experiment Analysis of Artificial Intelligence in Library Network Security

4.1 Library Network Security Honeynet Attack Data Analysis

This article adopts the HoneyNet attack data of the library management network security organization. Select the data from September 1 to 9, 2018 as the training data, to the day as the test data, and the number of attacks in a day as the library management network security situation value, using the formula to normalize the situation value, and the experimental results are shown in Table 1.

| Day | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  |
|-----|----|----|----|----|----|----|----|----|----|
| Test| 518| 209| 128| 417| 635| 98 | 880| 949| 661|
| Test| 468| 159| 274| 361| 487| 103| 943| 824| 719|
Figure 1. The actual attack line chart

It can be seen from Figure 1 that the peak value of the data between to is very high, indicating that there are many alarms, and the degree is much higher than other values. In the network, an atomic attack will often cause a series of false alarms. The number of attacks from day to day has increased significantly, and almost all attacks belong to the same kind of attack, and there are also sub-alarms at the same time. It can be considered that during this period, the false alarm rate due to attacks is relatively high.

4.2 PSO-SVR Model Library Management Network Security Situation Value Analysis

The PSO-SVR model is used to predict the security situation of the library management network for a week. The actual value and the predicted value are shown in Figure 2 below.

Figure 2. Curves of actual and predicted situation values

From the figure, we can intuitively see that the PSO-SVR prediction model can predict the future situation value relatively closely. It shows that the prediction model based on one has a good prediction effect, can correctly reflect the trend and change of the future library management network
security situation, help library management network administrators analyze the future network situation, and prompt library management network administrators to adjust defenses in a timely manner based on threat assessment strategies to maintain the safe and healthy operation of the library management network.

5. Conclusions
This article starts with the research on the library network security situation awareness architecture. Based on the existing library network security situation awareness framework research and intelligent methods, this article proposes an intelligence-based library network security situation awareness architecture, which is based on the overall system design. The existing framework has been improved to some extent, and the system's defense capabilities and intelligence have been enhanced. In this paper, particle swarm optimization algorithm is used to optimize the parameters of support vector machine, which overcomes the subjectivity of parameter selection of support vector machine, and uses the network attack data to further verify its performance.

References
[1] Wilkins, Jonathan. Is artificial intelligence a help or hindrance?[J]. Network Security, 2018, 2018(5):18-19.
[2] Talwar R, Koury A. Artificial intelligence – the next frontier in IT security?[J]. Network Security, 2017, 2017(4):14-17.
[3] Bai Hongjun, Zhang Xi. Application and Analysis of Artificial Intelligence Technology in Network Space Security Defense[J]. Digital Technology and Applications, 2019, 037(003):194-195.
[4] Wang Z, Fang B. Correction to: Application of combined kernel function artificial intelligence algorithm in mobile communication network security authentication mechanism[J]. The Journal of Supercomputing, 2019, 75(9):5965-5965.
[5] Barker J. A Quarter of People in EMEA Prefer Their Cybersecurity Managed by Artificial Intelligence.[J]. Software World, 2019, 50(4):18-18.
[6] Amin R U, Aijun L I, Ali M M. Fuzzy, Neural Network and Expert Systems Methodologies and Applications: A Review.[J]. Journal of mobile multimedia, 2015, 11(1-2):157-176.
[7] Sedjelmaci H, Guenab F, Senouci S M, et al. Cyber Security Based on Artificial Intelligence for Cyber-Physical Systems[J]. IEEE Network, 2020, 34(3):6-7.
[8] Liu Xiaolin. Prospects for the application of artificial intelligence in the field of public security[J]. Telecommunications Network Technology, 2019, 000(006):22-25.
[9] Jian-Hua L I. Cyber security meets artificial intelligence: a survey.[J]. Frontiers of Information and Electronic Engineering (English), 2018, 019(012):P.1462-1474.
[10] Bartos K, Rehak M. IFS: Intelligent flow sampling for network security–an adaptive approach[J]. International Journal of Network Management, 2015, 25(5):263-282.