Network Intrusion Detection Model Based on Artificial Intelligence

Qingchuan Meng¹²*, Ziyou Zhang³, Fengzhi Wu⁴, Xiaoming Chen²⁵

¹ HeiLongJiang Province Cyberspace Research Center, Harbin, HeiLongJiang, 150090, China
² College of Computer Science And Technology, Harbin Engineering University, Harbin, HeiLongJiang, 150090, China
³ Zhejiang University of Technology, Hangzhou, Zhejiang, 310012, China
⁴ Ashikaga University, Ashikaga-shi, Tochigi, 326-8558, Japan
⁵ Harbin University of Science and Technology, Harbin, Heilongjiang, 150090, China
*Corresponding author’s e-mail: mqchappy365@163.com

Abstract. The Internet occupies a more and more important position in people's life. The society has become more convenient because of the progress of network technology, and the network security problem has attracted more and more attention. Therefore, the security technology based on network is more and more important. Intrusion detection technology is the main research direction of dynamic security tools. Aiming at the problem of low accuracy of network intrusion detection, this paper proposes a network intrusion detection model based on AI, introduces artificial intelligence neural network and algorithm, extracts data feature information through repeated training of the algorithm, and constructs network intrusion monitoring model based on neural network. According to the design goal of the model, the basic framework of the system is given, which provides accurate detection basis for network intrusion detection. The experimental results show that the detection accuracy of the model is high.

1. Introduction

The increasingly serious problem of network information security is seriously threatening the social and economic security of all countries and the world. Governments of various countries have taken actions to strengthen legislation and improve the legal guarantee system for cracking down on cyber crime, establish relevant institutions and take practical measures to crack down on cyber crime; increase the investment in network security technology and improve the level of network equipment to effectively solve the increasingly serious security technology problems. Intrusion detection (ID) is one of the effective methods to protect network information. This method judges the validity of the data by implementing certain security policies, that is, certain algorithms. Of course, the premise of implementing the strategy is to collect and process the data, and the detection process completes the data classification and preprocessing. Detection is to implement a certain algorithm to identify whether there is intrusion or not. Then it is the response control module, which gives the processing results to show whether there are intrusion attacks. At present, this technology is widely used and has a bright future.
Intrusion detection technology is a supplement to the traditional security mechanism. It has the characteristics of intelligent monitoring, real-time detection and dynamic response. It has become the main research and development direction of dynamic security tools. Neural network has the ability of self-adaptive, self-organizing and self-learning, and can deal with the problem of unclear background knowledge after training. Therefore, in intrusion detection technology, soft computing method is the development trend of detection technology. Network intrusion detection technology has been a popular application technology of current network security, but in the application process of network intrusion detection technology, there are problems of low detection efficiency, low detection accuracy and poor detection performance. In order to solve the above problems, a new method combining artificial intelligence with network intrusion detection is proposed. The network intrusion detection technology based on artificial intelligence refers to the application of artificial intelligence technology, such as neural network algorithm, to the network intrusion detection. The application of these artificial intelligence technologies makes the automatic detection of network intrusion detection model possible.

2. Artificial neural network

An artificial neural network (ANN) is a network formed by a large number of elements called neurons and the interaction of these elements. Artificially simulate the network model in which neurons are connected in a certain way, called artificial neural network model, which is also called connection model or distributed model. The method of simulating neural network processing information is called artificial neural network method.

The 3 elements of neural network are: neurons, the organizational structure of neurons (network structure) and the training method of the network. Neuron is the basic information processing unit of neural network operation, it is the design basis of artificial network. Neuron is a nonlinear element with multiple input and single output

(1) A set of connections. The connection strength is represented by the weight of each connection. If the weight is positive, it means activation; if it is negative, it means inhibition.

(2) A summation cell. It is used to calculate the weighted sum (linear combination) of each input signal.

(3) A nonlinear activation function. It plays a role of nonlinear mapping and limits the output amplitude of neurons to a certain range.

Artificial intelligence neural network is mainly composed of input layer, hidden layer and output layer. Each layer contains multiple neurons, namely data node signal. Figure 1 shows the structure of artificial intelligence neural network.

- **Input layer**: \( x_1, x_2, \ldots, x_n \)
- **Hidden layer**: \( z_1, z_2, \ldots, z_p \)
- **Output layer**: \( y_1, y_2, \ldots, y_m \)

Figure 1. Structure of ANN

Artificial intelligence neural network is a kind of forward network. Firstly, the signal of data node is input in the input layer, and the signal of data node is transmitted forward to the hidden layer. In the
hidden layer, it is activated by the activation function. Finally, the activated data node signal is transmitted to the output layer, and the output result is given by the output layer. The expression formula of activation function of artificial intelligence neural network is as follows:

\[ f(n) = -\frac{1}{1 + e^{-n}} \]  

In formula (1), \( n \) represents the signal of data node and \( e \) represents the hidden layer composed of neurons. The activation function of artificial intelligence neural network has the characteristics of linear amplification. It can transform the data node signal from negative infinity to positive infinity in the input layer into the data node signal in the range of \([-2, 2]\). For the larger input data node signal, the amplification coefficient of activation function is smaller; on the contrary, for the smaller input node signal, the amplification coefficient of activation function is larger, which reduces the complexity of data node signal, and makes the artificial intelligence neural network can obtain the output result approaching to nonlinear.

ANN uses artificial intelligence algorithm to train data packets. The training process of artificial intelligence algorithm in network intrusion detection model: firstly, part of the data signal is extracted from the data packet set, and the data signal is transmitted to the input layer \(\rightarrow\) hidden layer \(\rightarrow\) output layer of artificial intelligence neural network,

Then calculate the output of node signal in each layer, compare the calculated output value with the expected output value to get the error of the two. Finally, the output layer is calculated to the hidden layer in reverse. According to certain rules, the data error is reduced by adjusting the weights of each connection in the network. The artificial intelligence algorithm will repeat the above steps until the data error is zero.

The mathematical expression of artificial intelligence algorithm: suppose that there are \( k \) data node signals in the input layer and hidden layer, and \( m \) data node signals in the output layer, the initialization weight is \( W \), and the training data set is \( X \). Firstly, the artificial intelligence algorithm performs forward calculation. The output calculation formula of \( k \) data nodes in input layer and hidden layer is as follows:

\[ h_k = f(\sum WX) \] 

In formula (2), \( f \) is the activation function of ANN. The calculation formula of \( m \) data node signals in the output layer is as follows:

\[ y_m = f(\sum uh) \] 

In formula (3), \( u \) is the signal weight of data nodes in the output layer, and \( h \) is the output calculation results of \( h \) data node signals in input layer and hidden layer. The following is the error calculation function between the actual calculation results and the expected results.

\[ E(W) = \frac{1}{2} \sum (h - y) \] 

According to the above calculation results, the error of output result is corrected according to the change law of data node signal. The correction formula is as follows:

\[ \theta = f(\sum WX)(\sum uh) \otimes (h - y) \] 

Through the repeated correction of the correction function to eliminate the output error, get accurate data features, provide accurate detection basis for network intrusion detection model, improve the detection efficiency and accuracy of network intrusion detection model, and realize the construction of network intrusion detection model based on AI.

3. Network intrusion detection model based on AI

The network intrusion detection system designed in this paper adopts the modular design idea. The main design objectives are as follows:

(1) The system can adapt to the environment with large network traffic, and can intercept and process the network data in real time, so as to find the intrusion.
(2) The function of each module of the system is complete and has certain expansibility.

(3) Neural network detection module has a good accuracy, can effectively reduce the false positive rate and false negative rate, through continuous self-learning, can effectively detect unknown types of attacks.

According to the general model principle of intrusion detection system described in the general framework of intrusion detection, combined with the design objectives of the system, the system is divided into five modules: data acquisition, data preprocessing, feature extraction, neural network detection module and response module. The structure of each module is shown in Figure 2.

![Network intrusion detection model](image)

Figure 2. Network intrusion detection model

The workflow of the whole model is as follows: firstly, the packet capture module intercepts the network data stream according to certain rules, then sends the preprocessing module to analyze the data packet and carry out preliminary filtering. After transforming the data into standard format, the feature extraction module extracts the main feature vector of the packet, and then the neural network detection module combines with the feature database to detect, and then transmits the analysis results to the response module. The module will cooperate with the system management to make or write the log, or disconnect the network connection, or alarm and notify the system administrator.

### 3.1. Data acquisition

The selection of data source in network environment is an important problem. In order to detect intrusion, many types of input data can be used. Since all the attack information of network attacker is in the network packet, network packet is the reliable data source of intrusion detection. According to the different types of networks, network data can be captured in two ways: one is to set monitoring port or mirror port for routers or switches based on switching network environment mechanism, because network packets of switching mechanism will not be broadcast in the whole LAN; One is IEEE802.3 Ethernet protocol based on broadcast mechanism, which makes use of the characteristics of Ethernet sharing transmission medium, which can make the transmitted packets be received by all hosts in the same network segment.

### 3.2. Data preprocessing

In order to make full use of the effective data, we must sort out the required data from the captured data packets, while the data packets captured by WinPcap are only frames on the data link layer. In order to extract the effective content, we must analyze the captured packets according to the TCP/IP protocol. The parsing process is mainly based on the TCP/IP definition of various packet formats and packet types.
3.3. Feature extraction
The processing results of feature extraction module should be used as the input of neural network detection module, which requires effective feature selection and extraction of the parsed data to form feature vector and digitalize it. The feature vector should not only contain enough information to distinguish normal and abnormal data packets, but also cannot make the dimension of feature vector too large, otherwise it will lead to too many input neurons of neural network. The training is slow and can't even converge. This model adopts the data extraction method based on KDDCUP '99 standard, and uses the data set to test and evaluate the model.

3.4. Neural network detection module
Neural network detection module is the core component of network intrusion model, its quality directly relates to the operation efficiency of the whole system. The purpose of intrusion detection is to distinguish normal packets from attack packets in the mixed packets, which is an extension of pattern recognition and classification. From this point of view, Ann is mainly used in pattern recognition and classification, which has a wide range of adaptability and effectiveness.

3.5. Response module
The response module works according to the output of the neural network detection module. For network intrusion model, it is record and alarm. The purpose of the record is to record some information related to intrusion, which may be information about the time and level of attack, which may be kept in the database of log or audit record, which is mainly for convenience of post investigation or new learning of network. The alarm is to send some visible information to the user, so that the user can make a timely response.

The neural network must be trained before the actual work, and training is the process of learning. Through training, the connection weights of the network are determined, which store the pattern characteristics of the attack behavior. After training, the knowledge used for sample discrimination is memorized by neural network in the form of weight value, which improves the generalization ability of unknown samples.

4. Experiment
In order to prove the validity of the network intrusion detection model based on artificial intelligence, a comparative experiment is designed to compare it with the traditional network intrusion detection model to test the accuracy of the network intrusion detection model based on artificial intelligence.

The test data is from a standard data set, including 500 normal data and 500 intrusion data. The data set contains 8 kinds of data characteristics. Intrusion data is composed of five types of attacks, which are simulated by script or manual methods.
Figure 3. Experimental results

The model is tested through the above 500 data packets composed of 5 attack types. The 5 attack types are represented by attack A, attack B, attack C, attack D and attack E according to the order above. The experimental results are shown in Figure 3. It can be clearly seen from Figure 3 that all the data detection and recognition rates of the network intrusion detection model based on artificial intelligence are above 90%, while the traditional network intrusion detection model only has a high recognition rate for normal data, while the recognition rate for attack behavior is low, especially for long-term attack behavior, its recognition rate is the lowest about 45%. It is proved that the network intrusion detection model based on artificial intelligence has high accuracy for attack behavior detection in the application process.

5. Conclusion
The application of artificial intelligence technology to network intrusion detection is a new field of network intrusion detection technology, which can effectively improve the accuracy of network attack detection, realize the remote supervision of network security in real time, and provide security guarantee for complex network environment.

ACKNOWLEDGMENT
This work was supported by Heilongjiang Province of China (No. F2017016), Heilongjiang postdoctoral fund (LBH-Z17054), Heilongjiang provincial social science fund (No. 16GLC08).

References
[1] Wang, H. Q., Wang, Z. Q. Zhao, Q., et al. (2006) Mobile Agents for Network Intrusion Resistance. Lecture Notes in Computer ence., 3842:965-970.
[2] Luigi Pietro Cordella, Ro Limongiello, Carlo Sansone. (2004) Network Intrusion Detection by a Multi-stage Classification System. Lecture Notes in Computer ence., 3077:324--333.
[3] Casey E. (2005) Case study: Network intrusion investigation-Lessons in forensic preparation., 2(4):254-260.
[4] Dreiseitla S , Ohno-Machadob L. (2002) Logistic regression and artificial neural network classification models: a methodology review. Journal of Biomedical Informatics, 35( 5–6):352-359.
[5] Ghiassi M, Saidane H, Zimbra D K. (2005) A dynamic artificial neural network model for forecasting time series events. International Journal of Forecasting, 21(2):341-362.

[6] Guo Z, Zhao W, Lu H, et al. (2012) Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. Renewable Energy, 37(1):241–249.

[7] Garcia-Pedrajas N, Hervas-Martinez C, Ortiz-Boyer D. (2005) Cooperative coevolution of artificial neural network ensembles for pattern classification. IEEE Transactions on Evolutionary Computation, 9(3):271-302.