Network Security Situation Prediction Based on BP and RBF Neural Network

Yaxing Zhang, Shuyuan Jin, Xiang Cui, Xi Yin, and Yi Pang

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
zhangyaxing@software.ict.ac.cn,
{jinshuyuan,cuixiang,yinxi}@ict.ac.cn,
pangyi@software.ict.ac.cn

Abstract. With tremendous complex attacks on the network, network analysts not only need to understand but also predict the situation of network security. In the field of network security, the research on predicting network security situation has become a hot spot. The prediction of network security situation can dynamically reflect the security situation of the entire network and provide a reliable reference to ensure the network safety. This paper predicts the network security situation using the BP and the RBF neural networks, and then makes a comparison between the two methods. The results show that the effect of the model based on the BP neural network is better than that of the model based on the RBF neural network on predicting the network security situation.

Keywords: network security situation awareness, prediction of situational value, BP Neural network, RBF Neural network.

1 Introduction

As rapid development of network technology, researchers have paid more attentions to the issues of network security situation prediction. The capability of understanding the situations of network security can help network managers know whether the current network is under attack or not and if so how much the strength of attack is.

The method of neural network is widely used in the prediction of network security situation. [1] proposed a network security situational awareness model based on audit log and network security posture correction algorithm. In order to take comprehensive network security factors into consideration, [2] provided a method for network security situation assessment using Honeynet. [3] proposed a method of network security prediction based on dynamic BP neural network with covariance to resolve the limitations of depending on experts giving weights. [4] predicted the network security situation using the method of RBF neural network with hybrid hierarchy genetic algorithm. Most of work has not focused on the difference between the BP and the RBF neural networks on predicting security situation. This paper provides two methods based on the BP neural network and the RBF neural network respectively to predict the network security situation. More importantly, it compares the two methods from different aspects, for example, time, correct rate, and mean-squared error.

Y. Yuan, X. Wu, and Y. Lu (Eds.): ISCTCS 2012, CCIS 320, pp. 659–665, 2013.

2 Related Concepts

Network security situation is the macroscopic reflection of network state, reflecting the past and the present situation and then predicting the status and the trend of next phase [5]. The meaning of “network security situation” is the same as the concept of “situation” used to explain all the states of an object with complex structure. The prediction of network security situation can help decision-makers do situation awareness faster and better [6]. The value of network security situation is an important quantitative indicator to illustrate the level of the network security. The value of network security situation could manifest the operating situation of the network, varying with the frequency and quantity of security incident and the degree of threat [5]. We can use intrusion detection system or network management system to collect the original information of network security situation.

3 Artificial Neural Network Profile

Artificial Neural Network (ANN) is a kind of artificial intelligence technology, which has a rapid development in 1980s. In recent years, the neural network has made a huge breakthrough both in theory and practice [7].

The neural network is an ideal tool used to predict the network security situation. The BP and the RBF neural networks are most commonly used in information processing, pattern recognition and optimization problems [8].

4 Methods

This section shows the methods of the BP neural network and the RBF neural network used in the prediction of the network situation.

4.1 BP Neural Network Method

The BP network is a feedforward network with three or more layers, spreading error from back to front while adjusting the parameters. Generally speaking, the BP neural network contains the input layer, the hidden layer and the output layer. After our providing the network with a learning sample, the neurons’ values of activated will be transferred from the input layer to the output layer. The neurons in the output layer should modify their connection weights and thresholds according to the error to minimize the mean squared error between the actual output and the desired output. In this paper we use the BP neural network with three layers which is shown in Fig. 1.

4.2 RBF Neural Network Method

A new and effective feedforward neural network with three layers called radial basis function (RBF) neural network, which has fine characteristics of approximation performance and the global optimum [3]. Generally speaking, the RBF network consists

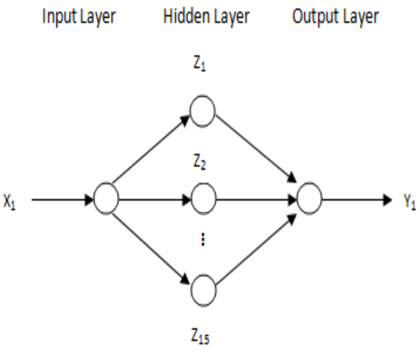

Fig. 1. BP Neural Network Structure

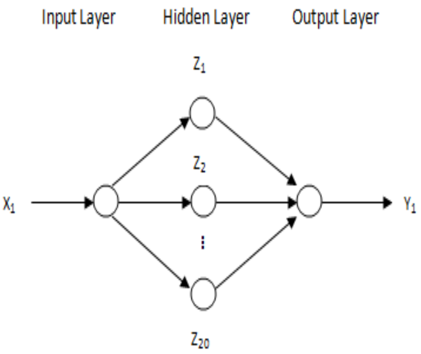

Fig. 2. RBF Neural Network Structure

of the input layer, the hidden layer and the output layer. The neurons in the input layer are only responsible for transferring the input signal to the hidden layer. In the hidden layer, we often use the radial basis function as the transfer function, while we usually adopt a simple linear function in the output layer. The RBF neural network with three layers used in the paper is shown in Fig. 2.

5 Experiment

In this paper, we use a neural network toolbox provided by MATLAB to establish the BP and the RBF neural networks.

5.1 Sample Data

The data used in the experiments is collected by “the network security situation analysis system”, developed by the laboratory where the authors work. The system will give a parameter called nsas_score, which is the value of network security situation used in this paper.

The system collected data every five minutes during the dates from 2011-05-13 to 2011-05-18. During this period, the system was taken a complex artificially attack. There are 1728 sets of sample data, which is shown in table 1 (The data is too much to be shown). 864 sets are used to train the neural network models , while the others are used to test.

Table 1. Data used in the experiments

number	nsas_score	number	nsas_score	number	nsas_score
1	1.36	8	1.38	15	1.37
2	1.28	9	1.35	16	1.37
3	1.4	10	1.39	17	1.34
4	1.34	11	1.29	18	1.32
5	1.38	12	1.39	19	1.28
6	1.36	13	1.35	20	1.36
7	1.35	14	1.37	21	1.32

5.2 Pretreatment

In the neural network models, the input vector P is 1-dimensional vector, while the output vector T is 1-dimensional vector. For the collected data were not in the same order of magnitude, the data must be normalized in the pretreatment period. The normalization function used in the experiments is mapminmax.

5.3 Experimental Results Analysis and Discussion

5.3.1 BP Neural Network Model

The prediction model based on the BP neural network contains three layers: the input layer, the hidden layer and the output layer. The number of neurons in the hidden layer is set to 15, the Mean Squared Error (mse) is 0.1 according to the actual training process, the maximum number of training is 1000, and the learning function is trainlm. The error curve is shown in Fig. 3 and the fitting curves between the predicted output and the expected output is shown in Fig. 4.

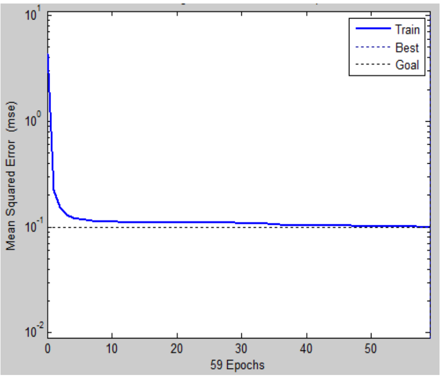

Fig. 3. Error curve in the network training

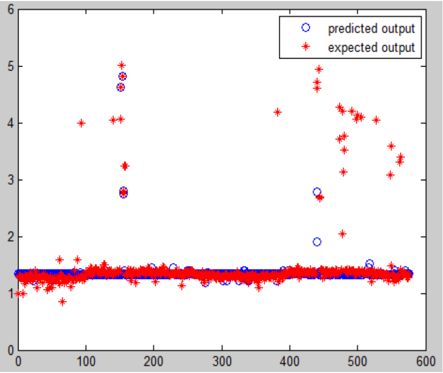

Fig. 4. Fitting curve between the predicted output and the expected output

Fig. 3 shows the error curve of the BP neural network model during the training process. In Fig. 3, the horizontal axis shows the number of training data, and the vertical axis represents the mean-squared error. The mean-squared error has the trend of becoming smaller as the growth of the number of training data. The curve shows that the learning algorithm has a fast convergence speed so that the model reaches the requirement after 59 epochs.

Fig. 4 illustrates the differences between the predicted and the expected outputs. The horizontal axis shows the number of the test data, and the vertical axis represents the predicted output and the expected output. The curve shows the fitting degree of the two kinds of output is high, and the correct rate is 85.42%.

5.3.2 RBF Neural Network Model

This paper uses the newrbe function to establish the RBF neural network model. The extension constant is set to 0.8 and the maximum number of neurons is 20. The other parameters such as the mean-squared error, the maximum number of training data, the learning function are set as the model based on the BP neural network. The error curve is shown in Fig. 5 and the fitting curves between the predicted output and the expected output is shown in Fig. 6.

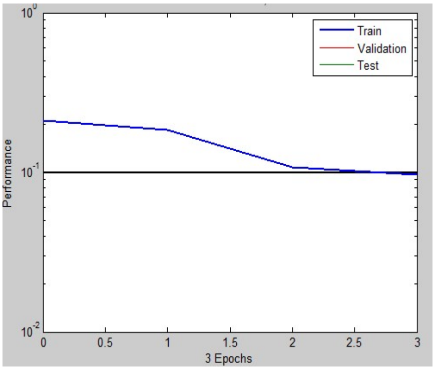

Fig. 5. Error curve in the network training

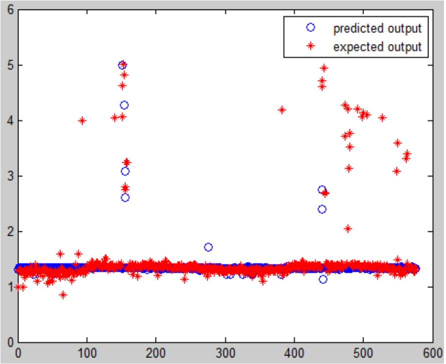

Fig. 6. Fitting curve between the predicted output and the expected output

Fig. 5 shows the error curve of the RBF neural network model during the training process. The horizontal axis in Fig. 5 shows the number of training data, the vertical axis represents the mean-squared error. The mean-squared error also becomes smaller with the growth of the number of training data. The curve shows that the learning algorithm has a faster convergence speed so that the model reaches the requirement after 3 epochs.

The fitting degree of the predicted output and the expected output is shown in Fig. 6. The horizontal axis shows the number of test data, the vertical axis represents the predicted and the expected outputs. The curve shows that the predicted output is similar to the expected output. The correct rate is 84.2%.

5.4 BP Network and RBF Network

- The time consumed by the BP neural network model is shorter than the RBF neural network model: the time used by the BP network model is less than 1 second, while the RBF network model uses 13 seconds. But the training time of the BP network is longer than that of the RBF network. This is because that the BP network reaches the goal after 59 epochs, while the RBF network makes it after 3 epochs. This shows that the test time of the BP network is less than that of the RBF network. This is partly because in the test of the RBF network, almost each sample has to be compared with the center vector of each neuron in hidden layer. This increases the time of test.

- The correct rate of the BP neural network model is higher than that of the RBF neural network model. The BP network reaches 85.42%, while the RBF network reaches 84.2%. Since the correct rate is calculated with a function, the predicted output of the RBF network may not be suitable for the processing function.
- The mean-squared error of these two methods is relatively large. The reason may be that the sample data used in the experiments is too small.

In a word, the model based on the BP neural network is more effective than the model based on the RBF neural network in the prediction of network security situation.

6 Conclusions and limitations

The prediction technology of network security situation could reflect the situation of the whole network and predict its trend. It gives network managers a better understanding of the network security situation, and effectively protects the network in the complex network environment. The experimental results show that in the prediction of network security situation, the model based on the BP neural network is more effective than the RBF neural network model. Due to the lack of the sample data and the mismatch between the test data and the training data, the mean-squared error that two networks could reach is relatively large. The performance of the experiments could be improved at this point.

References

1. Wei, Y., Lian, Y.: A Network Security Situational Awareness Model Based on Log Audit and Performance Correction. Chinese Journal of Computers 32, 763–772 (2009)
2. Xia, W., Wang, H.: Prediction model of network security situational based on regression analysis. In: 2010 IEEE International Conference on Wireless Communications, Networking and Information Security, WCNIS (2010)
3. Tang, C., Xie, Y., Qiang, B., Wang, X., Zhang, R.: Security Situation Prediction Based on Dynamic BP Neural with Covariance. In: Advanced in Control Engineering and Information Science, CEIS 2011 (2011)
4. Meng, J., Ma, C., He, J., Zhang, H.: Network Security Situation Prediction Model Based on HHGA-RBF Neural Network. Computer Science 38, 70 (2011)
5. Ren, W., Jiang, X., Sun, T.: The Prediction Method of Network Security Situation Based on RBF Neural Network. Computer Engineering and Applications 31, 136–144 (2006)
6. Xu, B.: Network Security Situation Prediction. Dalian University of Technology 1, 6–8 (2008)
7. Haykin, S.: Neural Networks, a Comprehensive Foundation, 2nd edn., pp. 161–175, 183–221,400–438. Prentice Hall (1998)
8. Xi, R., Jin, S., Yun, X., Zhang, Y.: CNSSA: A Comprehensive Network Security Situation Awareness System. In: 2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom (2011)

9. Onwubiko, C., Owens, T.: Situational Awareness in Computer Network Defense: Principles, Methods and Applications. Information Science Reference Press (January 2012) Ebook
10. Hu, W., Li, J., Chen, X., Jiang, X.: Network security situation prediction based on improved adaptive grey Verhulst model. Journal of Shanghai Jiaotong University (Science) 15(4), 408–413 (2010)
11. Jajodia, S.: Cyber situation awareness: issue and research (advanced in information security) (2009) Ebook