International Journal of Security and Its Applications
Vol. 10, No. 4 (2016), pp.423-430
http://dx.doi.org/10.14257/ijsia.2016.10.4.37

Application of Improved BP Neural Network with Correlation Rules in Network Intrusion Detection

[1]*Yongfeng Cui, [2]Xiangqian Li Ma and [3]Zhijie Liu

[1]School of Science and Technology, Zhoukou Normal University, Zhoukou Henan 466001, China;
[2]Network Centre, Zhoukou vocational and technical college, Zhoukou Henan 466000, China;
[3]Library, Zhoukou Normal University, Zhoukou Henan 466001, China
cuiyf@zknu.edu.cn

Abstract

To detect various network attacks in real time, this paper developed a network intrusion detection system based on artificial neural network. This paper first introduced the recent development of neural network, BP algorithm and structure of a simple perceptron. Then, this paper developed an improved BP neural network algorithm to detect anomaly network traffic with adjusted correlation rules. Finally, the network intrusion system in this paper was tested in a real network situation; the improved BP algorithm neural network with adjusted correlation rules shows a reduction in total error and increment in alarm rate compared to the traditional basic BP algorithm model.

Keywords*: network intrusion detection, BP neural network, correlation rules, anomaly network traffic*

1. Introduction

With the rapid development of the network, the problem of network information security is becoming increasingly prominent. Network information leakage, as well as a series of chain reactions caused by the invasion of the virus[1]. Network information security is an important issue that affects the state and the interests of the country. Network information security capability is an important part of the new century, which reflects the comprehensive national strength, competition ability and survival strength[2]. Today, simple usage of traditional firewalls, routers, and common host access control, effective authentication and data encryption technology have been unable to effectively resist the developing modes of invasion. The development of network security technology, especially the development of intrusion detection technology has become an indispensable part of network security architecture.

Artificial neural network (ANN) is a computer simulation network of the intelligent biological impulses. In the network, a node is a neuron, it can store memory data, process information, work together with other nodes. Using artificial neural network technology, input data to a single node, through the neural network collaborating with other nodes, the final output after processing results[3]. The advantage of artificial neural network is the complex nonlinear analytic ability and great fault tolerance and robustness, with parallel processing method for big data, ANN has an adaptive learning ability toward uncertain or unknown system. In this paper, BP algorithm is a widely used neural network algorithm[4].[1]

Yongfeng Cui is the corresponding author.

ISSN: 1738-9976 IJSIA

Neural network is a bionic system, it simulates the human brain complex analysis and reasoning abilities, and establishes a close mathematical model. Human brain has the powerful information processing ability, because human brain neuron can deal with nonlinear signals. Artificial neural network simulates the structure of the human brain, the data processing process, to construct the model structure quite similar to information acquire and data processing process in the real life[5,6]. Figure 1 is a simple model of multiple inputs and single output perceptron. Weight W is the intensity of neuron connection, f(x) is a nonlinear function.

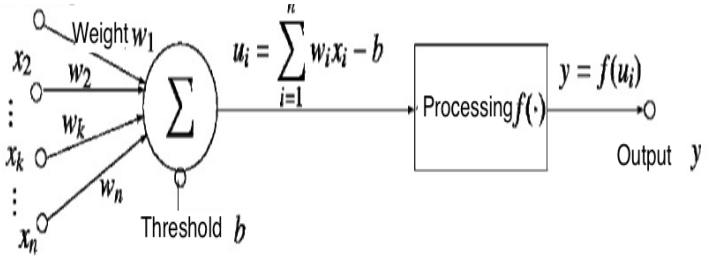


Figure 1. Structure of a Simple Perceptron Machine

2. Methods

2.1. Improvement of the Basic BP Algorithm

Add the effect of momentum factor involves the error and the influence of the surface change, if not increase the momentum effect, can lead to network into a local minimum problem[7-9]. It is the basic principle of back propagation of weights and threshold change before plus a weights and threshold of change, to generate new weights or threshold value.

Assume is number of training, is the momentum factor, then the weights of increment momentum factor and fixed threshold formula is:

$$\Delta w_{ij}(k+1) = (1-mc)\eta\delta_i p_j + mc\Delta w_{ij}(k) \tag{1}$$

$$\Delta b_i(k+1) = (1-mc)\eta\delta_i + mc\Delta b_i(k) \tag{2}$$

The essence of increment momentum factor method is conducting the changes of weights or threshold value. If dynamic quantity factor is 0, using the gradient descent method to obtain the changes of weights or threshold value; If the momentum factor is 1, the new weights or threshold value is the last change of weight and threshold value. Therefore, the increment momentum factor can adjust the direction of the average error of curved surface, value will decrease, and, it can effectively prevent the occurrence of, in the result of preventing the minimum of partial surface. Momentum factor is a reflect of the rich adjustment experience, it hinders the adjustment in time. When there is a sudden fluctuation of the error surface, the system can improve the training speed through reducing oscillation trend[10].

Assume is the error square sum of step, the judging condition is set to:

$$mc = \begin{cases} 0, E(k) > E(k-1)\times 1.04 \\ 0.95, E(k) < E(k-1) \\ mc, \text{others} \end{cases} \tag{3}$$

It is not easy to set the learning rate. Learning rate is actually the step length, in the view of error surface, a small will increase the number of training in a flat region, and in the area where the error changes violently, will across a narrow sunken place making the training oscillation, increasing the number of iterations. To choose a adaptive adjustment method for, in the condition of total error increase after the changes of weighs, must multiply; in the condition of total error decrease, must multiply.

Set to the square sum of errors, the adaptive adjusting learning rate is:

$$\eta(k+1)=\begin{cases}1.05\eta(k), E(k+1)<E(k)\\0.7\eta(k), E(k+1)>1.04E(k)\\\eta(k), \text{others}\end{cases} \quad (4)$$

All in all, the increment momentum factor can get the optimal solution of BP algorithm, and use learning rate adaptive adjustment section, shorten training time can make the BP algorithm.

2.2. Correlation Analysis

Correlation analysis can find out the relationship within the data items in the database records. Correlation analysis is given a set of Item and a record collection, through the analysis of the recorded set, to find the correlation between the items. As an important part of the correlation analysis, the support degree and the reliability are the important contents of the correlation analysis[11-13]. If the association rules satisfy the minimum support threshold and the minimum confidence threshold, then it is considered that the association rule is meaningful. The association analysis can find the effective rules that satisfy the minimum support degree.

Association rules can be found in the same database, the relationship between different groups or objects, and data mining a deeper level of information on the relationship between the rules, accessing to the relationship between different attributes of the record.

Assume is a set of data records, set is a collection of data set, and, X is a subset of I.

Assume correlation rule has support degree and confidence degree in the set. X, Y is the data item in the records, that is and; is the proportion of in the data set D, that is; is the case that probability of containing X and Y, that is. Both degree of support and confidence take value in the range of 0 and 1. Strong rules can generate both minimum support and confidence rules. If the minimum support degree is met, it is called a frequent item set.

A collection of data items containing items is called k-set, the item set count refers to the frequency of the item set. If the item set number is equal to the minimum support degree, the collection is called a frequent item set. Data mining of the correlation rules generally have two processes:

a. looking for complex set process: through the user's minimum support degree to find all the frequent item sets, which meet the minimum support and some additional criteria, find out all the frequent item sets. When frequent item sets that may have a relationship, they are not included in the other frequent item sets.

b. generate strong correlation rules: according to user defined minimum set of reliability, generate the strong correlation rules same time to meet the minimum degree of reliability and confidence.

The process of generating strong correlation rules is relatively simple, so the overall performance of correlation rules is determined by finding frequent item sets.

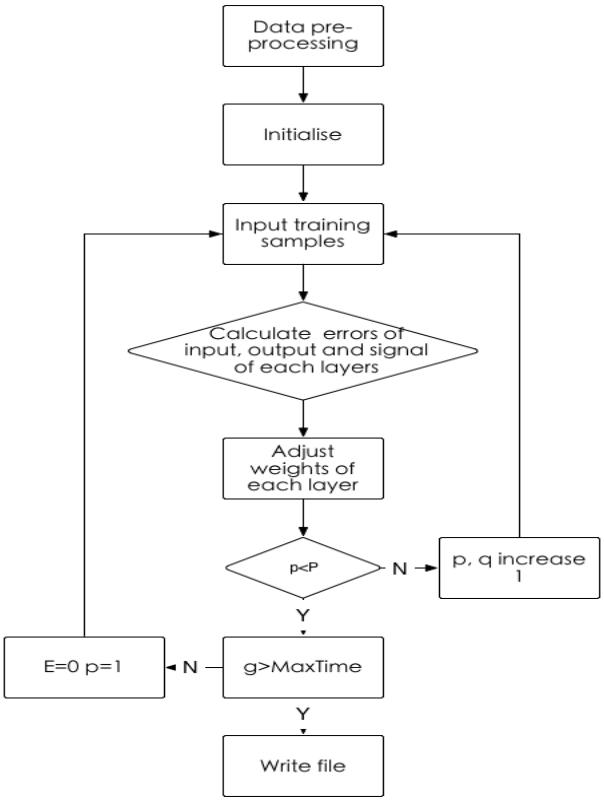
2.3. The Process of Network Intrusion Detection

Network intrusion detection system works in hybrid mode, using sensors to detect network data packets. Network sniffer is common information collection tools used by attackers, therefore it is an effective way to gather the information of network events. System works in the network layer and transport layer on the network to achieve interception, their main task is to pass all the packets to the upper layer.

IEEE802.3 Ethernet protocol is widely used in local network; the Ethernet card is usually works in normal mode or hybrid mode. In normal circumstances, if the network card receives the packet from the host address, then the packet is put into the buffer, if the network card receives the packet from other address, then the packet is discarded. Under normal circumstances, the hosts only process the data packets of its own, while in the mixed mode, the Ethernet card can receive all the data packets in the network transmission, regardless of the packet's destination address. At present, vast majority of Ethernet network cards can provide this kind of setting, therefore, as long as the Ethernet network card is properly set, the packets go through the Ethernet card can be monitored in order to monitor the network. At the same time, different operating systems have different data link access interface, the data link access interface must choose carefully based on the operation system[14-18].

Anomaly network traffic detection module can detect abnormal network traffic of a certain sub network port in a specified time period. Through one month of observation to a certain sub network, this paper trained the traffic logs with BP algorithm, and then got the results through weight matrix, if the error value is greater than the minimum setting error, the network traffic was considered abnormal. Specific steps are as follows:

a. Data pre-processing. The characteristics of traffic record were converted to digital features, and the data was normalized.
b. Initialization. Set the weight , to random number, the sample mode counter and the number of training times were set to 1, the total error is set to 0, define the number of hidden layers, set the learning rate between (0, 1), define maximum training times.
c. Input the sample data, calculate the output component of each layer. Vector X and D were calculated with the sample and, component vector of Y and O were calculated with the transit function.
d. Calculate the output error. For the training sample P, assume the error of sample P is, the total error can be represented with.
e. Calculate the error signal of each layer. Using the (3) and (4) to calculate the error of the output layer and the hidden layer error signal.
f. Adjust the weights of each layer. Applying (3) and (4) to calculate the weights of each layer.
g. Check whether all samples completed a rotation, if, calculator, increased 1. Return to step c, otherwise turn to step h.
h. Check the network total error is stable or whether achieve the maximum number of operation times, if q>MaxTime, turn to step i, otherwise set E to 0, p set to 1, return to step 3. And if the maximum number of operation times are reached, but the error is still not standard, then consider increasing the number of operations.
i. After training, store the related parameters in the model file.
j. Enter the unknown records, predict the results according to the parameters stored in the model file.



k.

Figure 2. The Flowchart of Anomaly Network Traffic Detection

3. Results

The generation of correlation rules to find all the frequent item sets were based on the support degree of the prior setting. In general, the degree of support and the production of frequent items are inversely related, that is, the lower the degree of support, the higher the frequent of item sets, and the higher the degree of support, the lower the frequent of item sets. The same degree of confidence will affect the number of the correlation rules, that is, the lower the confidence, the more the number of rules generated but less accuracy, and the higher the confidence, the higher the number of rules generated, but the accuracy is high. The degree of support and frequent items set, degree of confidence and accuracy are shown in Figure 3 and 4.

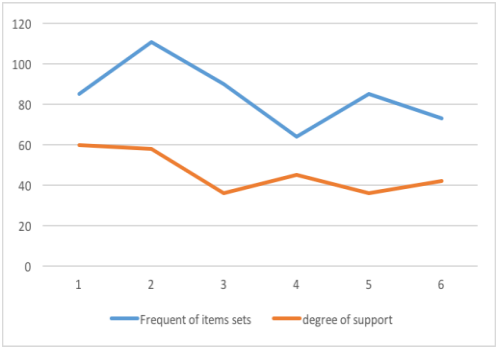


Figure 3. Relationship of the Degree of Support and Frequent Items Set

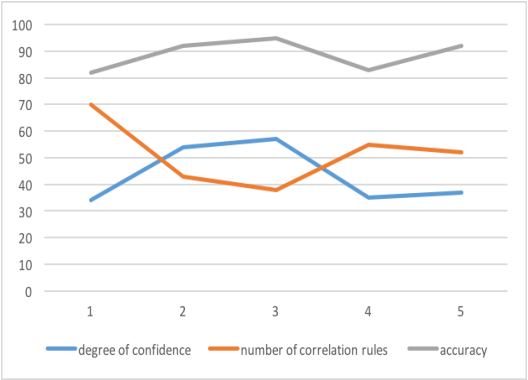


Figure 4. Relationship of Degree of Confidence and Accuracy

In the actual operation, the intrusion rules can be predefined, which can decrease the false negative rate, though it is still available if the intrusion rules not predefined. At present, most of the behavior of the computer system is realized by the network system, which can be used to check the network data packets in the system. Almost all denial of service attacks are based on the network attack. Compared with the traditional network intrusion detection system, the detection accuracy of the BP neural network with correlation rules is much higher, shown in table 1.

Table 1. Test Results Compared to the Traditional Network Intrusion Detection System

	Maximum Run-time	Total Error	Alarm Rate	Accuracy
Traditional Detection	10000	0.000026	81.26%	89.34%
Neural Network Detection	10000	0.000013	92.57%	96.42%

The intrusion detection system uses the network communication packet as the data source. If the intrusion detection system is used as a filtering router, the original content of TCP or IP packets is analyzed using the pattern matching and signature analysis. However, an application grid to analyze and broadcast protocol related data and need to deeply analyze the data, the coast of manpower and material resources and financial resources will be higher. In summary, this model can solve the following problems of the traditional network intrusion detection system:

a. Real-time detection and alarm. When a suspicious access or attack occurs, the system can detect intrusion in time and quickly take measures. This kind of real-time performance reduces the damage caused by malicious access or attacks to the system, and ensures the security of the system.
b. Effectively detect attack attempts. The goal of some of attacks is the protected resources of the firewall, and the model can be used to detect attacks in time.
c. The system is independent to the operating systems. The system is independent to the operating system, so the type of operating system will not affect the results of the test.
d. Low cost and easy to implement. The model can set up one or more access points to detect network communications, it can protect many key devices, and do not need to install on the host, the system can greatly reduce the complexity of security and management with a relatively low cost.

The system model has many characteristics, for example, effectively use of data mining technology to detect the existing intrusion, the operation is close to the real situation with the advanced neural network algorithm, with adaptive adjustment the database can be updated according to the current environment, the system can predict the new attacks and detect the unknown attack.

4.Conclusion

Based on the analysis of the traditional BP neural network intrusion detection method, this paper combines the technology of neural network and data mining to analyze the network traffic. Through the improvement of the original BP algorithm, the setting of the initial weights, the determination of the number of hidden layer, the setting of the maximum operating frequency, the calculation method of the total error of the network, the minimum support threshold and the minimum confidence threshold, this new network intrusion detection system proved to be valid in the real simulation with great accuracy.

The main limitation of the system in this paper is poor performance in encryption environment. In order to carry out the security communication, the network must encrypt the data. So the intrusion detection system in this paper is unable to decrypt the captured network data packets. Therefore, in the encryption environment, the realization of intrusion detection is a direction of future research.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant no. U1504602), Postdoctoral Science Foundation of China (2015M572141), the Key Research Projects of Universities in Henan Province for contract 15A520035 and 15A520124, under which the present work was possible.

References

[1] Aburomman, A.A., Ibne Reaz, M.B., n.d. A novel SVM-kNN-PSO ensemble method for intrusion detection system. Applied Soft Computing. doi:10.1016/j.asoc.2015.10.011
[2] Agrawal, V., Panigrahi, B.K., Subbarao, P.M.V., 2015. Review of control and fault diagnosis methods applied to coal mills. Journal of Process Control 32, 138–153. doi:10.1016/j.jprocont.2015.04.006
[3] Dai, Q., 2013. A competitive ensemble pruning approach based on cross-validation technique. Knowledge-Based Systems 37, 394–414. doi:10.1016/j.knosys.2012.08.024
[4] Dai, Q., Zhang, T., Liu, N., 2015. A new reverse reduce-error ensemble pruning algorithm. Applied Soft Computing 28, 237–249. doi:10.1016/j.asoc.2014.10.045
[5] Ganapathy, K., Vaidehi, V., Chandrasekar, J.B., 2015. Optimum steepest descent higher level learning radial basis function network. Expert Systems with Applications 42, 8064–8077. doi:10.1016/j.eswa.2015.06.036
[6] Kashyap, Y., Bansal, A., Sao, A.K., 2015. Solar radiation forecasting with multiple parameters neural networks. Renewable and Sustainable Energy Reviews 49, 825–835. doi:10.1016/j.rser.2015.04.077
[7] Kim, M., Choi, C.Y., Gerba, C.P., 2013. Development and evaluation of a decision-supporting model for identifying the source location of microbial intrusions in real gravity sewer systems. Water Research 47, 4630–4638. doi:10.1016/j.watres.2013.04.018
[8] Li, P., Xiao, H., Shang, F., Tong, X., Li, X., Cao, M., 2013. A hybrid quantum-inspired neural networks with sequence inputs. Neurocomputing 117, 81–90. doi:10.1016/j.neucom.2013.01.029
[9] Liu, S., Yamada, M., Collier, N., Sugiyama, M., 2013. Change-point detection in time-series data by relative density-ratio estimation. Neural Networks 43, 72–83. doi:10.1016/j.neunet.2013.01.012
[10] Mitchell, R., Chen, I.-R., 2014. A survey of intrusion detection in wireless network applications. Computer Communications 42, 1–23. doi:10.1016/j.comcom.2014.01.012
[11] Modi, C., Patel, D., Borisaniya, B., Patel, H., Patel, A., Rajarajan, M., 2013. A survey of intrusion detection techniques in Cloud. Journal of Network and Computer Applications 36, 42–57. doi:10.1016/j.jnca.2012.05.003
[12] Olatunji, S.O., Selamat, A., Abdulraheem, A., 2014. A hybrid model through the fusion of type-2 fuzzy logic systems and extreme learning machines for modelling permeability prediction. Information Fusion, Special Issue on Information Fusion in Hybrid Intelligent Fusion Systems 16, pp. 29–45. doi:10.1016/j.inffus.2012.06.001

[13] Sundarkumar, G.G., Ravi, V., 2015. A novel hybrid undersampling method for mining unbalanced datasets in banking and insurance. Engineering Applications of Artificial Intelligence 37, pp. 368–377. doi:10.1016/j.engappai.2014.09.019
[14] Sweetlin Hemalatha, C., Vaidehi, V., Lakshmi, R., 2015. Minimal infrequent pattern based approach for mining outliers in data streams. Expert Systems with Applications 42, pp. 1998–2012. doi:10.1016/j.eswa.2014.09.053
[15] Yu, J., Kang, H., Park, D., Bang, H.-C., Kang, D.W., 2013. An in-depth analysis on traffic flooding attacks detection and system using data mining techniques. Journal of Systems Architecture, Advanced Smart Vehicular Communication System and Applications 59, pp. 1005–1012. doi:10.1016/j.sysarc.2013.08.008
[16] Zhao, J., Chen, S., Zuo, R., n.d. Identifying geochemical anomalies associated with Au–Cu mineralization using multifractal and artificial neural network models in the Ningqiang district, Shaanxi, China. Journal of Geochemical Exploration. doi:10.1016/j.gexplo.2015.06.018
[17] W Ya, S Qin. Aircraft fault diagnosis by solving map exactly. Review of Computer Engineering Studies, 2015, vol. 2, no. 1, pp. 1-8.
[18] Y Zhao, Cn Feng, J Yang, and L Wang. LITERATURE REVIEW OF NETWORK PUBLIC OPINION ABOUT THE ECOMMERCE. Review of Computer Engineering Studies, 2015, vol. 2, no. 2, pp. 25-30.

Authors

Yongfeng Cui. Yongfeng Cui received the BS degree in Computer Science and Technology from Henan Normal University and the MS degree in Computer Application Technology from Huazhong University of Science and Technology, China in 2000 and 2007 respectively. He is currently researching on Computer Application Technology (CAT).

Xiangqian Li. Xiangqian Li received the BS degree in Physics from Henan Normal University and the MS degree in Software Engineering from University of Electronic Science and Technology, China in 1996 and 2011 respectively. He is currently researching on Computer Application Technology (CAT) and Mobile Internet Technology (MIT).

Zhijie Liu. Zhijie Liu received the BS degree in Henan University for Library and Information in 2002. He is currently researching on Network consulting and Network Management.

