Feature Contour Construction Methods Based on Neural Network for Network Abnormal Behavior Detection

Ma Jiwei, Lin Zhipeng, Lv Xuejiao
Hebei Normal University Of Science & Technology, China, Hebei, 066604.

Abstract. With the development of network application, how to detect the abnormal behaviors from network data messages gets to be a hotspot of network security management. This paper puts forward network abnormal behavior detection methods based on neural network. As a typical network abnormal behaviors, network intrusion has attracted much more attention in the past years. In order to detect network intrusion, the feasibility and effectiveness of neural network are investigated in detail. On the basis of neural network, two feature contour construction methods are presented to conduct the network intrusion detection. Experiment results indicate that the proposed methods can be used to detect the network abnormal behavior effectively. In the future, comparisons with other abnormal behavior detection methods will be scheduled.

1. Introduction
Nowadays, as the computer network technology develops, the amount of network information is increasing significantly, accompanying with a large number of network abnormal behaviors. People surfing on the line have to attach great importance to network information security, while these security-related technologies have become mature gradually after a period of rapid development. They are primarily composed of data encryption technology, firewall technology, intrusion detection technology and virtual private network technology, as well as anti-virus technology. Thereinto, to deal with network intrusion, many commercial intrusion detection systems are established by using different intrusion detection technologies, with their own fundamental theories improved constantly.

For nework intrusion detection, it refers to detection of all intrusion behaviors from the Internet. The data information is collected and analyzed at all critical points of the network, so as to find all the behaviors violating the security regulations. It is James Anderson who put forward this concept in the early 1980s [1]. Since then, this technology has successfully changed from the host to the network, and the structure has also switched from the previous centralized one to the distributed agent [2]. Now there are mainly two intrusion detection methods: One is abuse detection, a detection technology based on specific rules, which is able to detect all known intrusion patterns; the other is anomaly detection, a detection technology based on behavior, which treats the feature contour corresponding to normal behavior as the starting point, and compares the detected activity with the normal feature contour. This detection method has a good adaptability. In the late 1990s, Forrset et al. first proposed the method of using system call sequence to construct feature contour for anomaly detection [3]. The feasibility of their method was verified by some convincing experiments. This analysis method based on program behavior can reduce greatly the false alarm caused by unpredictable user behavior. As for short sequences, two classical statistical methods are commonly used, i.e. time-delay embedding (TIDE) and time-delay embedding (STIDE) methods. The TIDE method stores each unique short sequence to form a feature database, used as the detection standard; the STIDE method uses TIDE as the basis to determine
the anomaly by mismatch rate, which judges whether the number of mismatches can reflect the abnormal behaviors well. Though the tree storage method is adopted, the algorithm still needs a lot of storage space for feature data to form a feature database. Meanwhile, it lacks the ability to handle with intrusion variation and accidental events, which is prone to cause false reports [4].

In order to develop a system with strong adaptability and low false alarm rate, a few researchers introduced intelligent algorithms in intrusion detection. Lee et al. adopted data mining technology to sample and analyze the system call data for the first time [5]. RIPPER described the pattern features of normal data with the aid of small rule sets [6]. During monitoring, if these characteristics are violated, the sequence will be abnormal. After several years of development and improvement, China has successfully established a fuzzy expert system based on detection model. This model improves the first mock exam method by absorbing the superiority of above-mentioned detection methods, and it can reduce false alarm and missing alarm effectively. As a typical intelligent technology, neural network can also be applied for intrusion detection. Ghosh et al. designed a strategy of program behavior analysis based on neural network [7]. They transformed a short sequence into the x-dimensional space, and then implemented anomaly detection via coordinates of the short sequence.

To detect network intrusion, this paper proposes two feature contour construction methods based on neural network. One method is reducing the false alarm by using neural network, and the other method is improving the actual detection rate based on regression network while reducing the false alarm.

2. Feasibility and Effectiveness of Neural Network

For anomaly detection, how to form feature contour is always the key point. Because of its self-learning habit and self adaptability, neural network has attracted attention of many scholars. They began to analyze how to leverage it in the process of intrusion detection, and then neural network became the preferred technology for feature data analysis. Through learning and training, as well as adjusting feature patterns of the subject, a feature contour implementing the adaptive target is established. The contour is able to respond correctly and timely to some accidental, unknown intrusion behaviors while varying in a normal range, reducing false reports and adapting to network environment changes. In addition, neural network can record the characteristics of intrusion behavior without requiring an extra memory space to store feature data [8].

Considering neural network has a habit of self-learning, the generated system can construct the feature contour while the data is not complete and easy to change. Therefore, It is directly related between algorithm selection and network function design. If the classification function needs to be constructed, it is more desirable to employ the back error propagation network (BP) network. The basic principle of BP algorithm is forwarding propagation data, and calculating the output of each layer; after calculation, the error between the expected and actual outputs can be determined, and this is error back-propagation. It can execute the weights and thresholds adjustment for each layer, and ensure the network output error maintain at a minimum. During the data forwarding propagation process, the forwarding output of each layer can be worked out via Sigmoid activation function, aslo called S-type function, which can be formulated as:

\[ F(x) = \frac{1}{1 + \exp(-x + b)} \]  

(1)

where \( x \) is a S-type neuron input, \( b \) an imposed bias, and \( F(x) \) a S-type neuron output.

During the error back propagation process, the weight and threshold of each layer can be adjusted with gradient descent. The following equation is responsible for weight adjustment:

\[ w_{ij}(t + 1) = w_{ij}(t) + \eta * \delta_{ij} * y_j + \alpha * (w_{ij}(t) - w_{ij}(t - 1)) \]  

(2)

where \( w_{ij}(t) \) is the weight at current time \( t \), \( w_{ij}(t+1) \) the weight at next time \( t+1 \), \( w_{ij}(t-1) \) the weight at previous time \( t-1 \), \( \eta \) a network learning factor, \( \delta_{ij} \) an error generated by a node in the mode, and \( y_j \) the node’s output, as well as \( \alpha \) is a momentum factor. During network training, the output error is computed
continuously, and the network weight is dynamically adjusted by calculational results until the error falls into a standard range [9].

In order to reflect dynamic features of the generated system better, regression neural network can be used to implement data analysis. In this network, part of its output is fed back to its input to participate in the subsequent stage of training. Through this adjustment, the global memory can be generated in the network, and the memory information can meet the requirements of network prediction. In this paper, the network achieving the connected target is adopted, making the error back-propagation between different levels more practical in weight adjustment, with the feedback value determined as 1. In addition, the memory information is closely relevant to the entire process; weights corresponding to different neurons can represent the long-term memory of the network, and store network behavior rules; the activation function corresponding to the neurons can reflect all kinds of short-term memories of the network, and store the current sequence information. By utilizing this network, the next potential sequence could be predicted, and then the anomaly analysis would be implemented.

3. Feature Contour Construction Methods

During sequence analysis, TIDE method can be used to generate a short sequence, with its length marked as \( K \). Based on neural network functions, the two schemes designed for network intrusion detection are described below.

3.1. First scheme for intrusion detection

In the first scheme, intrusion detection is considered as a classification problem, and all program behaviors can be divided into normal class and intrusion class. For neural network, it is able to determine the short sequence corresponding to a normal behavior. During network training, each point of the sequence corresponds to different input nodes. The expected output is marked as 1 for the normal sequence, while the expected output is set as 0 for the abnormal sequence. Only one short sequence is output at each time, and the weight is adjusted by this output. This process has to proceed iteratively until the error falls within the allowable range.

During actual network training, it is necessary to predefine a threshold to distinguish the behaviors of mismatch and anomaly. In other words, it is required to determine what degree of deviation exists between the short sequence and the normal pattern, and then compute the mismatch rate value. This method is a bit simple on the whole, and it makes less use of the learning ability of neural network. Furthermore, the training time is too long to facilitate the algorithm convergence.

An experiment shows that the network sensitivity resulted from this scheme is relatively low, and there is no significant difference between the actual mismatch rates of normal and intrusion behaviors. Therefore, this scheme cannot improve the false alarm of the system greatly and increase the intrusion detection rate [10].

3.2. Second scheme for intrusion detection

The second scheme is able to predict the next probable short sequence by using regression neural network, with the purpose to implement the intrusion detection. Its basic idea can be summarized as follows: At first, the number of nodes in the input layer is determined as \( n+1 \), where \( n \) represents the number of programs in the input node, with each node corresponding to the system’s feedback. The nodes are numbered on the basis of the system call code. If the input object is a short sequence sample, all system call input nodes of the system are 1 while any others are 0. Then, the number of neurons in the output layer is \( n \), with different neurons corresponding to their respective system calls. And the actual output value represents the call probability in other sequences. Next, depending on the feedback network which can achieve the goal of complete connection, the weights will be adjusted dynamically through error back-propagation in the network. Finally, during intrusion detection, the predicted and actual short sequences are compared with each other, and the distance between the two sequences is calculated, with the result recorded as \( d \). As a result, the maximum value can be determined precisely. Based on this value, whether there exists a mismatch can be made clear, and the reaction system will be
used to calculate the mismatch rate. At such time, the final detection result can be obtained by the mismatch rate value.

The second scheme gives full play to the learning ability of neural network, assuring the detection rate on the basis of reducing the false alarm rate. However, the network developed by this scheme has a large scale, and it is often difficult to determine the number of neurons in the input layer. In order to avoid the increase of computational burden, it is anticipated to reduce the number of neurons. Nevertheless, the number of neurons can only be determined by specific experimental analysis currently.

4. Conclusion
Compared with other network security technologies, intrusion detection not only provides static and passive protection functions, but also detects actively all kinds of network intrusion behaviors. This lays a good foundation for formulation and application of the corresponding measures. Under the current network environment, a variety of network intrusion means can be employed by network attackers, and information security is suffering from great threats. Therefore, the introduction of intrusion detection technology can meet the requirements of network security. In this paper, two feature contour construction methods are proposed on the basis of neural network, which can enhance the intrusion detection ability to some extent, and cope with the increasing network security challenges. In the future, these two methods will be compared with other intrusion detection methods to demonstrate their superiority.

Acknowledgments
This work was financially supported by the Applied Research Project of Public Welfare Technology of Zhejiang Province (2017C33013).

References
[1] J. P. Anderson. Computer Security Threat Monitoring and Surveillance[B]. James P. Anderson Co., Fort Washington, 1980.
[2] Li Tao, Duan Lixiang, et al. Application of adaptive convolution neural network in fault diagnosis of rotating machinery[J]. Vibration and shock, 2020,39(16):275-282,288.
[3] Forrest S., Hofmeyr S.A., et al. A sense of self for unix processes[C]. in Proceeding of IEEE Symposium on Security and Privacy, 1996.
[4] Yao Lei. Research on road settlement prediction based on neural network model optimized by mind evolutionary algorithm[J]. Science and technology innovation, 2020,10(27):122-123.
[5] Lee W., Stolfo S.J., Mok K.W. A data mining framework for building intrusion detection models[C]. in Proceeding of IEEE Security and Privacy, 1999: 120-132.
[6] Frank E., Witten I.H. Generating accurate rule sets without global optimization[C]. in Proceedings of the 15th International Conference on Machine Learning, Madison, WI. (1998)
[7] Ghosh A.K., Schwartzbard A., et al. Learning program behavior profiles for intrusion detection[C]. in Proceedings of the 1st USENIX Workshop on Intrusion Detection and Network Monitoring, Santa Clara, CA IEEE Computer Society, 1999:9-12.
[8] Wang Yingli, Tao Shuai, et al. Short term photovoltaic power generation prediction based on GA-BP neural network based on MIV analysis[J]. Acta solar Sinica, 2020,41(08):236-242
[9] Du Lin, Wang Guangxia, Li Ke. Research on multimodal landform recognition algorithm based on deep learning[J]. Surveying and spatial geographic information, 2020,43(08):21-25.
[10] Li Weiyi, Ye Wenhua, Xiong Tianzhong. Dual energy X-ray transmission metal recognition algorithm based on BP neural network[J]. Nonferrous metals engineering, 2020,10(08): 124-130.