Research on the Detection Method of Information System Access Abnormal Behaviour

Linghao Zhang, Sheng Wang, Siyu Xiang, Jie Zhang and Yunhui Liang*

Electric Power Research Institute of State Grid Sichuan Electric Power Company
ChenDu, Sichuan, 610000, China
*Corresponding author’s e-mail: Lianghh2115@sc.sgcc.com.cn

Abstract. The security of the information system is very important to the production of electric power companies. However, at this stage, there are still problems such as the inability to detect abnormal/malicious behaviors and traffic of the terminals in the network, and the inability to detect and deal with abnormal behaviors in a timely manner, leading to increased risks. This paper chooses to start with the detection of abnormal traffic as the research goal, and realizes the detection and classification of abnormal network traffic through the convolutional neural network algorithm to achieve the detection of abnormal behavior of information system access. It plays a fundamental role in many network security-related tasks, such as the design of a network traffic anomaly detection system, and has very important practical significance.

1. Introduction

Information system access abnormal behaviour detection methods have always been an important direction of information network security. And abnormal network traffic refers to a network traffic pattern that has a negative impact on the normal use of the network. It is quite different from the normal traffic and can cause network performance to decline or even become unavailable. Network trafficked anomaly detection refers to the application of various anomaly detection methods to analyse network traffic and discover traffic with abnormal behaviour in time, which plays an important role in enhancing network situation awareness and maintaining cyberspace security. Starting with abnormal traffic can effectively distinguish whether the information system access behaviour is safe.

Aiming at the actual operating environment and real problems of a certain electric power company's network security system, this paper uses a convolutional neural network algorithm to build a model. The model is combined with the current mainstream network traffic abnormal detection methods to realize the detection of abnormal behaviours of information system access.

2. Research status of abnormal network traffic detection methods

So far, domestic and foreign scholars have proposed many different types of network traffic anomaly detection methods. According to the research results of Ahmed et al. [1], network traffic anomaly detection methods can be divided into four categories: statistics-based, classification-based, clustering-based, and information theory-based. Figure 1 shows the general framework of network traffic anomaly detection.
Figure 1. The general framework of network traffic anomaly detection

Statistics-based is the earliest algorithm to solve anomaly detection or outlier problems, and has been extensively studied. As early as the 19th century, Edgeworth et al. [2] proposed an outlier detection algorithm based on statistics. From a statistical point of view, an anomaly is a distribution that does not obey the assumed random model, and is completely or partially irrelevant to the observation result. Ye et al. [3] proposed an anomaly detection technology based on the chi-square distribution theory. This technology creates an overview of normal events in an information system. The basic assumption of anomaly detection is that abnormal events and normal events have a comparatively large statistical distribution. Big deviation. The distance measure based on the chi-square test statistic can be defined as:

$$\chi^2 = \sum_{i=1}^{n} \frac{(X_i - E_i)^2}{E_i}$$

Classification-based on the network anomaly detection technology usually relies on experts' extensive knowledge of network anomaly characteristics. Poojitha et al. [4] used a backpropagation algorithm on a specific data set to train a forward neural network to determine network traffic abnormalities, and obtained relatively high accuracy. In recent years, with the rapid development of deep learning, related applications of deep neural networks in the field of network traffic anomaly detection have emerged. For example, Gao et al. [5] proposed an intrusion detection model using deep belief network DBN, and verified it with the NSL-KDD dataset.

Aiming at the problem of low accuracy of traditional network traffic anomaly detection methods based on clustering analysis, Li et al. [6] proposed a traffic anomaly detection method based on improved k-means clustering. Through the pre-processing of various traffic characteristic data, the k-means algorithm can be applied to enumerated data detection, and then a high-dimensional data feature screening method based on numerical distribution analysis method is provided, which effectively solves the problem of excessive dimension.

Noble et al [7] conducted experiments on benchmark DARPA and UNM data sets, which proved the practicality of information theory and concluded that information theory can be used to create effective anomaly detection models and explain its performance.

3. Method for detecting abnormal network traffic

This paper uses a method based on convolutional neural network to detect abnormal network traffic.

The first one is the Input layer. Its role is to input data, and it usually does some data processing, such as: removing the mean, centering all dimensions of the input data to 0; normalizing, normalizing the amplitude to the same range; PCA/whitening: use PCA to reduce dimensionality. Whitening is to normalize the data on each feature axis of the data. Here we use $H_i$ to represents the feature input of the
i-th layer in the convolutional neural network.

Then there is the convolutional layer (CONV layer) and the pooling layer (Pooling layer). The convolutional layer is composed of multiple feature maps. Each feature surface is composed of multiple neurons, and each of its neurons passes The convolution kernel is connected to the local area of the feature surface of the previous layer. The convolution kernel is a weight matrix. The convolution layer of CNN is convolved The operation extracts different features of the input.

The neurons in the convolutional layer are organized into various feature surfaces, and each neuron is connected to the local area of the previous feature surface through a set of weights, that is, the neurons in the convolutional layer and the feature surface in the input layer Perform a local connection. Then pass the local weighted sum to a non-linear function such as the ReLu function to obtain the output value of each neuron in the convolutional layer. If $H_i$ is the convolutional layer, $H_i$ can be expressed as:

$$H_i = f(H_{i-1} * W_i + b_i)$$  \hspace{1cm} (2)

Where $W_i$ represents the weight vector of the convolution kernel of the i-th layer, * represents the convolution operation, and $b_i$ represents the offset vector between the output layer of the convolution and the i-th layer.

The pooling layer immediately follows the convolutional layer and is also composed of multiple feature surfaces. Each feature surface of it uniquely corresponds to a feature surface of the layer above it, and the number of feature surfaces will not change.

The following is the fully connected layer (FC layer). The fully connected layer, that is, all neurons between the two layers have the right to reconnect. Usually, the tail of the convolutional neural network will only transmit signals to other fully connected layers.

4. Implementation of abnormal detection method for network traffic

Data is transmitted in a byte stream on the network, and the value range of each byte is huge \([0, 255]\). The grayscale digital image is an image with only one sampled colour per pixel. The sampled value is represented by 0 to 255, with 0 representing black and 255 representing white. The range of the two values is consistent.

Let's take the Moore dataset as an example. Moore traffic data set statistics are shown in Table 1:

| Serial number | Category       | Quantity |
|---------------|----------------|----------|
| 1             | WWW            | 328092   |
| 2             | MALL           | 28567    |
| 3             | FTP-DATA       | 5797     |
| 4             | FTP-CONTROL    | 3054     |
| 5             | FTP-PASV       | 2688     |
| 6             | DATABASE       | 2648     |
| 7             | SERVICES       | 2099     |
| 8             | P2P            | 2094     |
| 9             | ATTACK         | 1793     |
| 10            | MULTIMEDIA     | 576      |
| 11            | INTERACTIVE    | 110      |
| 12            | GAME           | 8        |
|               | Total          | 377526   |
After labelling the 12 different types of traffic data in turn, randomly shuffle the order, take 80% of each traffic as the training set, and the last 20% as the test set. According to the needs of the experiment, construct a 16*16 matrix based on the 249-bit features of the Moore dataset. Since the feature dimension is less than the number of matrix elements, the corresponding zero-filling operation is performed at the end of the matrix. After labelling the 12 different types of traffic data in turn, randomly shuffle the order, take 80% of each traffic as the training set, and the last 20% as the test set. According to the needs of the experiment, construct a 16*16 matrix based on the 249-bit features of the Moore dataset. Since the feature dimension is less than the number of matrix elements, the corresponding zero-filling operation is performed at the end of the matrix.

Since our input is a 16*16*1 matrix, 1 is the number of channels. Since the amount of calculation is not very large, here we set the convolution kernel of the C1 layer to a 3*3*1 convolution kernel; the step size is set to 1, and the padding is 1. Then the output of our C1 layer is 16*16*8. 8 convolution kernels generate 8 16*16 feature surfaces. Considering an additional bias parameter of the convolution kernel, the C1 layer has a total of 80 parameters. We use the ReLu function mentioned above as the activation function.

Subsequently, the pooling layer S2 adopts the maximum pooling process, that is, the maximum value sub-sampling technology; the pooling layer is used to reduce the size of the 8 feature planes generated by the convolutional layer. Sampling technology makes the system only pay attention to the relative position between features and no longer pay attention to the specific location of features. Here we set the sampling window to 2*2; the step size is 2; and the filling is 0. The input of S2 is an 8*16*16 feature map, and the output is an 8*8*8 feature map.

The C3 layer is the second convolutional layer. Similar to the first convolutional layer, the size of the convolution kernel is set to 3*3*1; the step size is 1; the padding is 1; and 16 convolution kernels are set. The input of this layer is the output of the previous layer: 8*8*8 feature map, and the output is 8*8*16 feature surface. The trainable parameters of the C3 layer are 16*(3*3*8+1) = 1168.

As the second pooling layer, the S4 layer still uses maximum pooling processing. The sampling window is set to 2*2; the step size is 2; the filling is 0. The input here is 8*8*16 feature surfaces, and the output is 4*4*16 feature surface.

The F5 layer is a fully connected layer that uses 128 4*4*1 convolution kernels to map 16 4*4 feature surfaces to a 128*1 vector, and finally uses a soft-max classifier to output 12 types of results.

The structure of the convolutional neural network in this article is shown in Figure 2:

![Convolutional neural network structure](image)

In the model which based on convolutional neural network, when the device has traffic behaviour range alarm, as shown in Figure 3.
Figure 3. Generate traffic behaviour range alarm

Under the model based on the convolutional neural network, the system can automatically learn the network traffic behaviour characteristics of the devices in the network, and generate an alarm when the device has abnormal traffic; as shown in Figure 4.

Figure 4. Warning diagram of abnormal flow of equipment

5. Conclusion
The research of information system access abnormal behaviour detection methods has been an important direction of people's attention to information security in recent years. Convolutional neural networks, as an excellent algorithm today, can effectively detect abnormal network traffic. This paper implements a full-flow network security intelligent analysis platform based on convolutional neural network. A series of steps such as system network traffic data collection, pre-processing, cleaning, feature selection, convolutional neural network training and prediction to identify the abnormal behaviour of power network information system access, have important theoretical and practical significance for realizing the detection of abnormal behaviour of information system access value.

Acknowledgments
This work was supported by the State Grid SiChuan Electric Power Company of China (Project Name: Research and platform development of full flow Advanced Sustainable threat intelligent analysis technology based on deep learning).
References
[1] Ahmed M, Mahmood A N, Hu J. A survey of network anomaly detection techniques[J]. Journal of Network and Computer Applications, 2016, 60: 19-31
[2] F.Y. Edgeworth. On discordant observations[J]. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1887, 23(143): 364-375
[3] N. Ye, Q. Chen. An anomaly detection technique based on a chi-square statistic for detecting intrusions into information systems[J]. Quality and Reliability Engineering International, 2001, 17(2): 105-112
[4] Poojitha G, Kumar K N, Reddy P J. Intrusion detection using artificial neural network [C]//Computing Communication and Networking Technologies (ICCCN T), 2010 International Conference on. IEEE, 2010: 1-7.
[5] Gao N, Gao L, Gao Q, et al. An intrusion detection model based on deep belief networks [C]//Advanced Cloud and Big Data (CBD), 2014 Second International Conference on. IEEE, 2014: 247-252.
[6] LI Hong-cheng, WU Xiao-ping, JIANG Hong-hai. Traffic anomaly detection method in networks based on improved clustering algorithm [J]. Chinese Journal of Network and Information Security
[7] C. C. Noble, D. J. Cook. Graph-based anomaly detection[C]. Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, 2003, 631-636