Research and Simulation of Network Intrusion Detection Algorithm Based on Fuzzy Classification

LIU Chun[1,a]

[1]Sichuan College of Architectural Technology,Network Management Center, Deyang 618000, China

[a]52236025@qq.com

Keywords: fuzzy classification; abnormal data; intrusion detection

Abstract. The network intrusion accurate detection problem is researched. In the network intrusion detection process, the classification features of network operation data has poor uniformity, and the error of features classification is large, a network intrusion detection method based on fuzzy classification algorithm is proposed, principal component analysis method is used, the dimensions of network operation data are reduced. The redundant data are reduced, and fuzzy classification method is used, network intrusion featuresare classified, the network intrusion detection is realized. The simulation results show that the algorithm can effectively improve the accuracy of detection, it has perfect results.

Introduction

With the development of computer network technology, the network security is more and more important. Assumed that the network has intrusion behavior, it will bring great threat to network security. Therefore, network intrusion detection method has become a hot problem in the field of computer network research[1]. It has received moreand more attention, and it becomes the main subject need to study. At present, network intrusion detection method mainly includes K mean clustering algorithm in network intrusion detection method, and the network intrusion detection method based on neural network algorithm, the network intrusion detection method based on genetic algorithm[2-4]. Among them, the K means clustering algorithm is most commonly used. Because the network intrusion detection method has wide application range[5,6], therefore, it has received a lot of attention of the experts. It have very large development space and practical value in application.

Relative principle of network intrusion detection method

A Features dimensions reduction process of network intrusion data

Typically, the computer network operation data has high dimension, so the network operation data contains the massive characteristics of the network intrusion data without any relational data, before the detection design, the principal components analysis method is used to reduce the dimensions of the network intrusion data, the characteristics of the network intrusion data detection are extracted, and the redundant data is reduced, the efficiency of calculation is improved.

In order to effectively remove the difference between network operation data and the original variable data, the network operation data interference is reduced the dimension reduction processing is taken. The standardization process is taken by the following formula:

$$z_{jk} = \frac{y_{jk} - \overline{y_k}}{t_k} \in Z \quad (1)$$

In the formula, $\overline{y_k}$ is the mean value of different dimension vector network operation data, t_k is the corresponding standard deviation, $y_j \in Y, j = 1,2,\cdots,N, k = 1,2,\cdots,P$.

The following formula is used to calculate the covariance matrix of network operation data of different dimension vectors:

$$T = \frac{1}{P}[Z - \overline{Zm})[Z - \overline{Zm}]^T \quad (2)$$

The following equation is used to describe the relationship between network intrusion feature value and their corresponding value U:

$$(\mu m - T)V = 0 \quad (3)$$

According to the characteristics of the network operation data value, the principal component features of network operation data are determined. Assumed the network intrusion feature value is greater than 1, the cumulative contribution rate of the operation data principal component is more than 90%.

B Network intrusion detection algorithm

Fuzzy classification method is a classification method based on statistics data, the network intrusion detection uses this method for classification, according to the original data, the classification hyperplane is calculated, and the hyper plane classification is obtained based on the data.

Use the following formula to detect network intrusion feature for two times planning:

$$z(y) = sign(\sum_{j=1}^{p} \beta_j z_j l(y_j, y) + c) \quad (4)$$

Use the following formula to calculate the duality programming, and the two times planning of network intrusion feature is obtained:

$$\begin{cases} \max \sum_{k=1}^{m} \beta_k - \frac{1}{2} \sum_{j=1}^{m} \sum_{k=1}^{m} z_j z_k \beta_j \beta_k (y_j \cdot y_k) \\ s.t. \sum_{k=1}^{m} z_k \beta_k = 0 \\ 0 \le \beta_k \le v(y_k) D \ k = 1, 2, \cdots, m \end{cases} \quad (5)$$

According to the method described above, the optimal hyper plane problem is converted to the duality programming problem of two quadratic programming. The solution of the optimal plan can be described by the following formula:

$$\beta^* = (\beta_1^*, \beta_2^*, \cdots, \beta_m^*)^T \quad (6)$$

Use the following formula to express fuzzy optimal classification function:

$$g(y) = \text{sgn}\{(x^* \cdot y) + c^*\} \quad (7)$$

The above formula must conform to the following constraints:

$$x^* = \sum_{k=1}^{m} \beta_k^* z_k y_k$$

$$b^* = y_i - \sum_{j=1}^{l} y_j \alpha_j (x_j \cdot x_i)$$

$$j \in \{ j \left| 0 < \beta_j^* < v(y_j) D \right. \} \quad (8)$$

Use the following formula to obtain the optimal classification, the fuzzy classification function is:

$$g(y) = \text{sgn}\{\sum_{k=1}^{m} \beta_k^* z_k L(y, y_k) + c^*\} \quad (9)$$

Where, $c^* = z_j - \sum_{k=1}^{m} z_k \beta_k L(y_k, y_j)$, $j \in \{ j \left| 0 < \beta_j^* < v(y_j) D \right.\}$.

According to the method described above, principal component analysis is used to reduce the dimension of the network operation data, so as to reduce the redundant data. The fuzzy classification method is used to classify the characteristics of network intrusion, so as to realize the network intrusion detection.

Experiment results and analysis

In order to verify the validity of this algorithm, we need for an experiment, the experiment environment is Visual C++6.0., the network operation data number is n, the network intrusion feature number is p, the data set of composed of network operation data features is $\{g_1, g_2, \cdots, g_m\}$. The distribution uniformity coefficient of network operation data is μ. Use the following formula to calculate the network intrusion detection accuracy:

$$\psi = \frac{\sqrt{n-p}}{g_i^2 - \mu} \times 100\% \quad (10)$$

According to the accuracy of the network intrusion detection, the performance of intrusion detection of different algorithms can be measured.

In the experiment process, it needs for 10 network intrusion detection experiments, and the related experimental data samples can be described by the following table 1.

Table 1 Experimental data samples

Experiment number	Number of samples	Network intrusion feature number
1	907	67
2	889	77
3	980	98
4	896	57
5	960	74
6	987	86
7	890	70
8	970	45
9	707	74
10	970	39

Using different algorithms for 10 time the network intrusion detection, the results of net work intrusion detection are analyzed, and the detection accuracy is calculated. The results of the experiment are shown in Table 2.

Table 2. Detection results of different detection algorithms

Experiment number	Network intrusion feature number	Detection accuracy of K means clustering algorithm (%)	Detection accuracy of neural network algorithm (%)	Detection accuracy of genetic algorithm (%)	Detection accuracy of fuzzy classification algorithm (%)
1	67	78	79	82	89
2	77	77	77	82	87
3	98	79	81	83	89
4	57	77	80	82	89
5	74	76	80	83	88
6	86	78	82	81	89
7	70	75	81	82	87
8	45	77	80	81	86
9	74	78	82	82	88
10	39	76	81	84	86

Through the above experiments, it can be learned that the new algorithm has better network intrusion detection performance, and the accuracy of detection can be effectively improved.

Conclusions

In this paper, the network intrusion accurate detection problem is researched. In the network intrusion detection process, the classification features of network operation data has poor uniformity, and the error of features classification is large, a network intrusion detection method based on fuzzy classification algorithm is proposed, principal component analysis method is used, the dimensions of network operation data are reduced. The redundant data are reduced, and fuzzy classification method is used, network intrusion features are classified, the network intrusion detection is realized. The simulation results show that the algorithm can effectively improve the accuracy of detection, it has good application value in practice.

References

[1] LI Feng, WU Chun- ming. Research on Prevention Fluctuation Control method of Network Intrusion Based on Energy Management[J]. Computer simulation, 2013,30(12): 45-48, 335.

[2] Alfaro V M, Vilanovab R. Robust tuning of 2DoF five-parameter PID controllers for inverse response controlled processes[J]. Journal of Process Control, 2013,23(4): 453-462.

[3] Ou Shi-feng, Gao Ying, Zhao Xiao-hui. Adaptive Combination Algorithm and Its Modified Scheme for Blind Source Separation[J]. Journal of Electronics & Information Technology, 2011, 33(5): 1243-1247.

[4] LUO Liang, WU Wen-Jun, ZHANG Fei. Energy Modeling Based on Cloud Data Center [J]. Journal of Software, 2014,25(7):1371-1387.

[5] Song Minghong，Yu Huafeng，Chen Haiyan. Application of Improved Quantum Evolutionary Algorithm in Computer Network’s Routing Choice[J]. Bulletin of Science and Technology, 2014,30(1):170-173.

[6] LvYong-fang, SUN Ling-fang. The Research and Application of Intrusion Detection Model Based on Petri Net[J]. Science Technology and Engineering,2011; 11(34): 8514-8518.