Simulation on the Effective Detection Model for Network Differentiation Intrusion

Liu Chun

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

52236025@qq.com

Keywords: network intrusion; differentiation; signature database

Abstract. The problem of network differentiation intrusion is researched. Network intrusion has the characteristics of complex changes such as concealed, randomness, difference and abruptness. Traditional method cannot describe the change rules, leading to low correct rate of detection. For this, a detection method for network differentiation intrusion based on artificial immune algorithm is proposed. The dynamic change equation of network differentiation intrusion characteristics is established, to obtain the cross point distribution condition of network differentiation intrusion characteristics. The network differentiation intrusion signature database is updated, and the network differentiation intrusion feature in the database is selected. The results show that the artificial immune algorithm solves the problem existing in traditional algorithm, improves the correct rate of network differentiation intrusion detection.

Introduction

With the increasing of network intrusion events, once the network is vulnerable to malicious attacks, network security will suffer great damage, leading to private data leakage of user, malicious tampering to the rights of user, network paralysis, however the intrusion detection system as the last protection of safe defense, is able to detect intrusion behavior of various forms, therefore network intrusion detection is always a hotspot in research of network security[1,2]. During the process of network intrusion detection, network intrusion detection methods commonly used include the method based on BP neural network algorithm, fuzzy clustering algorithm, and neighborhood model k-means algorithm [3-5]. Among them, the most commonly used is the BP neural network algorithm. But this algorithm used for network intrusion detection, when the difference of network is strong, the change rules of which cannot be accurately described [6]. Therefore, further research should be carried out for the network differentiation intrusion detection problem. Thus, the network differentiation intrusion detection method, has very broad prospects for development, has received great attention.

The principle of network differentiation intrusion detection method

A establish network differentiation intrusion feature dynamic equation. During network differentiation intrusion detection process, H_c is set to be data collection of network differentiation intrusion feature, H is the number of samples in feature collection in specified time period. The following formula is utilized to describe dynamic situation of the network differentiation intrusion characters:

$$H(u+\Delta u) = H(u) - \mu \cdot \Delta u - (\frac{\partial H_a}{\partial y_a} + \frac{\partial H_e}{\partial y_e}) \cdot \Delta u \tag{1}$$

In the formula, $\mu \cdot \Delta u$ is the new network differentiation intrusion feature appeared in the network, $\frac{\partial H_e}{\partial y_e} \cdot \Delta u$ is the update time for network differentiation intrusion feature, denoted as ν,

affinity coefficient is χ , features replaced in network differentiation intrusion feature update process described by $H_e = \{y | y_{\cdot} a > \nu, y_{\cdot} d < \nu\}$, the normal operation behavior in network described by y . The updating process of each network differentiation intrusion feature includes network update cycle parameter age and affinity parameter $affinity$, which are calculated with the following formula:

$$age(u+1) = age(u) + 1 \quad u < \mu$$
$$affinity(u+1) = affinity(u) + 1 \tag{2}$$

In the above formula, the affinity parameter between network differentiation intrusion behavior features and normal network operation behavior is obtained by calculation. After treatment of each iteration, accumulation processing is needed for cycle parameter values of the network update. Assuming the affinity registration success, affinity parameters are needed to be accumulated. Generally divided into the following three kinds of circumstances:

(1) assuming that $affinity > \vartheta \wedge u < \mu$, then the network differentiation intrusion feature is activated, the network intrusion feature is viewed as the sample characteristic.

(2) assuming that $affinity \le \vartheta \wedge u \le \mu$, then the accumulation result of network differentiation intrusion feature affinity is too low, need to continue accumulating processing.

(3) assuming that $affinity < \vartheta \wedge u > \mu$, then network differentiation intrusion feature cumulative results in accordance with the affinity measure standard, so it can tell that network differentiation intrusion feature have been completely replaced by the new features. Among them, ϑ is the affinity measure standard of network differentiation intrusion feature update.

In network differentiation intrusion detection, assuming that update speed V of network differentiation intrusion feature increases, then the operation data transmission speed v of network increases, $V = l_1 \cdot v$ is utilized to describe the relationship between these. In the formula, l_1 is the correlation coefficient of network differentiation intrusion feature updating speed and operation data transmission speed.

The distribution of crossing point when mutation processing is applied to network differentiation intrusion feature can be described using the following formula:

$$P_i(Y = \lambda) = \frac{\mu^2 e^{-1}}{\lambda!}, \lambda = 1, 2, 3, \cdots \tag{3}$$

In the formula, Y is the number of cross point of network differentiation intrusion feature variation.

B extraction of network intrusion data feature. in this paper, the idea of chaos synchronization is introduced to detect network differentiation intrusion characteristic behavior.

Assuming network data flow vector of given monitoring is:

$$U = \{U_1, U_2, \cdots, U_N\} \tag{4}$$

Where U_i is the random variable of d dimension, each random variable U_i is independent of each other, wherein, assuming U meets mixture distribution of K Gauss density function, the probability density function is expressed as:

$$p(U | \Theta) = \sum_{k=1}^{K} \alpha_k G(U | u_k, \sum_k) \tag{5}$$

$$\Theta = \left[\alpha, u, \sum \right] \tag{6}$$

Among them, $\alpha_k \ge 0$, $\sum_{k=1}^{K} \alpha_k$ =1,and

$$G(U|\mu_k,\sum_k)=(2\pi)^{-d/2}\left|\sum_k\right|^{-1/2}$$
$$\times\exp\left[-\frac{1}{2}(U-u_k)^T\sum_k^{-1}(U-u_k)\right] \quad (7)$$

$G(U|\mu_k,\sum_k)$ is the Gauss density function; $p(U|\Theta)$ the weighted sum of multiple Gauss density function, Θ is the collection of α, u, Σ three parameters. α is weight vector of the Gauss mixed model, u is the mean vector of Gauss density function, Σ is the covariance matrix. The goal of Gauss mixed model learning is to obtain the collection of α, u, Σ parameter vector:

$$\alpha=[\alpha_1,\alpha_2,\cdots,\alpha_k] \quad (8)$$

$$u=[u_1,u_2,\cdots,u_k] \quad (9)$$

$$\sum=[\sum_1\ \sum_1\ \cdots\sum_k\ ] \quad (10)$$

C Adding chaotic disturbance factor. In the design, the chaotic differential theory is introduced to classify the feature in feature set, so as to improve the quality of system identification, process is as follows:

(1) classify features in the intrusion feature set, and refresh the clustering center, with each update of cluster center, fuzzy classification matrix able to be updated in a refresh process is set;

(2) constrain the diversity of intrusion characteristic, processes the variance calculation on the characteristics, so as to confine the appearance of some false features, the variance is mf, then

$$mf=\frac{1}{NP}\sum_{i=1}^{NP}(f(x_i)-\overline{f(x)})^2$$

Among them, NP is the size of the intrusion collection, $f(x_i)$ is the fitness value of i-th intrusion feature, $\overline{f(x)}$ is the average fitness value.

(3) in addition, in the iterative process, in order to avoid the algorithm involved early into the local extremum, perturbation theory of chaotic time series is introduced to improve algorithm.

D update database of the network differentiation intrusion feature. The settings of V is the data set of network operating characteristics, including two data subsets T and ST, wherein, T is a data subset of normal network operation, ST is a data subset of network differentiation intrusion operation. The normal network operation feature information and affinity parameters included in the T. Network normal operation feature information can be described through the following form: $\{1000111010000010101110111100101010\}$, the relationship between T and ST can be expressed by the following formula:

$$\{T\}\cup\{ST\}=V$$
$$\{T\}\cap\{ST\}=\varnothing \quad (11)$$

According to the characteristics in the data set T, to select the sample $T\cdot B_j(j=1,2,\cdots)$, the following formula can be chosen to process reverse transform operation:

$$P_j=\overline{T\cdot B}(j=1,2,\cdots) \quad (12)$$

The above results is used to establish a data collection, described with $ST=\{P_1,P_2,\cdots\}$. The number of elements in the collection is 1.

The arbitrary element P_j is selected from the ST data collection as the network differentiation intrusion detection candidate detectors, calculating the sample in P_j and samples in T, $P_j=\{b_1,b_2,\cdots,b_{12}\}$, $T\cdot B_j=\{c_1,c_2\cdots,c_{12}\}$ can be obtained. Assuming the measure is ϕ, and ϕ is a constant, the correlation coefficients can be calculated with the following formula:

$$E=\sum_{j=1}^{12}\varepsilon \quad (13)$$

Assume that $\beta_j = c_j$, then $\varepsilon = 1$, assuming $\beta_j \neq c_j$, else $\varepsilon = 0$.

Assuming $E < \varepsilon$, then the sample is network differentiation intrusion operation, and vice versa.

Simulation experiments

In the process of the experiment, data is from the data set of network intrusion standard test set KDD CUP 99, which comprises 4 kinds of intrusion types: DoS, Probe, U2R and R2L, including normal samples at the same time, each sample has 41 characteristics, 7 symbol field and 34 numeric fields. The data distribution of randomly selected data are shown in table 1.

Table 1 distribution of sample set

Intrusion types	Meaning	training sample	test sample
DoS	denial of service attack	1000	500
Probe	Monitoring and other exploration activities	800	400
R2L	Illegal access to remote machines	600	300
U2R	Illegal access of ordinary users to local superuser privileges	400	200

The experimental results are shown as below. According to the results, it can be learned that the proposed algorithm, effectively reduce the false alarm rate of intrusion detection.

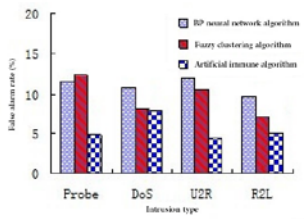

Fig. 1 Comparison of the false alarm rate obtained with different algorithms

In order to further verify the performance of the algorithm, network differentiation intrusion detection is processed for sample test data, the detection results shown in Table 2

Table 2 intrusion detection data information

No.	Start time	deadline	Types of intrusion	Detection continuity
1	0	0.15	smurf	continued
2	0.33	0.42	nmap	unsustainable
3	0.87	1.01	ipsweep	continued
4	1.00	1.50	back	continued
5	1.55	1.65	nmap	unsustainable

ROC diagrams of detection probability curve of the proposed detection system and the traditional ARMA detection system are shown in figure 2. It can be seen that the detection probability of proposed system have been improved comparing to the traditional method.

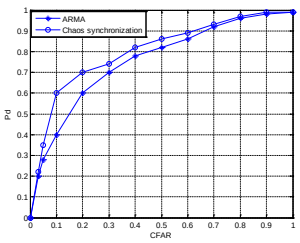

Figure 2 Comparative analysis of different methods

Conclusion

A network differentiation intrusion detection method based on artificial immune algorithm is proposed. The establishment of the differential equations of network intrusion dynamic change characteristics, obtain the distribution condition of cross point differentiation intrusion characteristic

variation. To update the network differential intrusion signature database, select the network differential intrusion feature in the database. In order to realize the network differential intrusion detection. The experimental results show that the false alarm rate of the artificial immune algorithm is significantly reduced, and has a broad application prospect.

References

[1] Zhang Haichun, Li Yuan, Zhang Zili. Realization of an intelligent intrusion detection system design and simulation [J]. Mathematics in practice and theory, 2009, 39 (6): 162-169.

[2] Wang Tao, Gong Huili. The application of support vector machine in the intrusion detection system [J]. Microcomputer information, 2006, 22 (12): 89-91.

[3] D. Koutsonikolas, SM. Das, YC. Hu. Path Planning of Mobile Landmarks for Localization in Wireless Sensor Networks [J]. In Proc.ofIEEE is Dtributed Computing Systems Workshops, 2006: 80-86

[4] B Waters. Efficient identity-based encryption without random ora-cles[C]. Advances in Cryptology-EUROCRYPT 2005 Berlin: Springer-Verlag, 2005.114-127.

[5] Niu Honghui, Liu Lingxia. The application research of neural network in information security risk assessment [J]. Computer simulation, 2011.6:117-120.

[6] Qing Sihan, Jiang Jianchun, Ma Hengtai, et al. Review on intrusion detection technology [J]. Journal on Communications, 2004, 25 (7): 19-21.