Journal of Electronic Research and Application
OPEN

Simulation Research on Optimal Detection of Intrusion Node Information in Network

Ma Yuanyuan, Zhou Cheng, Li Qianmu

Abstract: The optimal detection of intrusion node information in the network can guarantee the safe and stable operation of the network. When the intrusion node information is detected, it is necessary to obtain the optimal parameters of SVM according to the optimal acquisition path of the node to complete the detection of intrusion node information. The traditional method uses the ant colony to find the network node path, get the support vector machine parameters, but ignores the optimization of the parameters, resulting in the information detection results are not accurate. An improved detection method of intrusion node information based on attribute attack graph is proposed. The intrusion signal with intrusion node information is decomposed into IMF single frequency intrusion signal, and the state transition equation of the network intrusion detection system is obtained. The ant colony theory and the support vector machine parameters are merged, and the network intrusion check rate is used as the objective function, and the ant is changed to the ant, and the nodes on the optimal path are connected to get the SVM optimal parameters. Based on this parameter, the intrusion node information detection is completed. The experimental results show that the proposed method has high accuracy and can improve the accuracy of embedded computer network intrusion detection.

Key words: network; intrusion node information; optimal detection; attribute attack graph;

1 Introduction

In the country, with the continuous improvement of computer technology, the network has been widely used in life and work, but the network security problem is also increasing, greatly affecting people's lives and quality of work [1 ~ 3]. At present, for the diversification of network attacks and intrusion methods, the traditional firewall technology has been unable to deal with, as a network information system security, the second line of defense can check the network traffic and log audit data to determine whether the network intrusion, Network security has important significance, therefore, in this case, how to effectively identify the network intrusion node information has become a major problem in the field of computer problems, and invasive node information optimization detection method can be network intrusion check rate as Objective function, the optimal path of the nodes connected to get the SVM optimal parameters, based on the parameters to complete the network intrusion node information optimization is to solve the problem of the fundamental way, causing a lot of experts and scholars attention [4 ~ 6].

In [8], an intrusion node information detection method based on improved ant colony method is proposed. This method uses the ant colony to find the network path node on behalf of the support vector machine parameters, the network intrusion detection rate as the objective function, based on the attack map to build the intrusion node information detection problem model, the use of ant colony method to solve the global optimal solution to the network detection of intrusion node information. This method has the problem that the calculation process is cumbersome and time consuming. In [9], an intrusion node information detection method based on distributed learning is discussed emphatically. This method divides all the data in the network into several small data blocks, sends each small data block to the neural network to learn, and uses the benchmark data to test the learning ability of each data block. Based on the test result Complete the detection of intrusion node information. The time complexity of this method is low, but when the intrusion node information is detected by the current method, the envelope feature of the intrusion node information cannot be extracted, and the detection error is large. In [10], an intrusion

(Global Energy Internet Research Institute, Nanjing 210003)

node information detection method based on kernel Fisher discriminant analysis is studied emphatically. The method divides the normal node and the intrusion node, obtains the mean vector of each sample, obtains the linear discriminant need to be maximized, gives the sample characteristics of the normal node and the intrusion node, and completes the detection of the intrusion node based on the binary tree KFDA. The method has high detection efficiency, but there are some problems with the limitation of the method.

Aiming at the above problems, an intrusion node information detection method based on attribute attack graph is proposed. The experimental results show that the proposed method has high accuracy and can improve the accuracy of embedded computer network intrusion detection.

2 Network intrusion node information optimization detection principle

In the process of detecting the intrusion node information in the network, the data information is collected, the interclass divergence matrices and intraclass divergence matrices of different network intrusion node information samples are set up, and the network intrusion feature data is mapped to a high Dimensional feature space, on the basis of which the sensor node data and the established intrusion behavior characteristic data are compared, and the result of the comparison is used to detect the intrusion node information. The specific steps are as follows:

Suppose that *d* represents the number of network attributes of X_i on the network node to form the dimension of a vector, define *x* as the number of monitored attributes, and *n* for the number of nodes to be monitored within a cluster. The interclass divergence matrices and intraclass divergence matrices of different network intrusion node information samples are established by using equations (1) and (2):

$$\left(\varpi^T S_w \varpi\right) = \frac{J\left(\varpi\right)\times \beta\left(e\right) \mp R\left(\nu\right)\times d}{\left[n_1 \otimes n_2\right] \otimes \left[X^{*}, X^{m}\right]}\left[\varpi_1, \varpi_2\right] \quad (1)$$

$$\left(\varpi^T S_b \varpi\right) = \frac{J\left(\varpi\right)\left(\varpi^T S_w \varpi\right)}{\left[n_1 \otimes n_2\right]}\left[\varpi_1, \varpi_2\right] \quad (2)$$

In the above equation, $\beta(e)$ represents the log-likelihood function that maximizes the complete data, $R(\nu)$ represents the log-likelihood function that maximizes the missing data, identifies n_1 belonging to the collected normal node w_1 as X^+ , n_2 identifies the samples belonging to the intrusion node w_2 class as X'''', ($w^T S_b w$) represents the interclass divergence matrix of the sample, ($w^T S_w w$) on behalf of the internal divergence matrix, *J (w)* represents the maximized generalized Rayleigh quotient.

Suppose that m_i represents the mean vector of the various samples, ϕ represents the non-linear mapping of the input space *F* to any feature space, the network intrusion feature data is mapped to a high-dimensional feature space by non-linear mapping using Eq. (3)

$$\left(\varpi * \phi\left(x\right)\right) = \frac{S_b^{\phi} \otimes S_w^{\phi} \otimes k\left(x, y\right)}{\varpi^{\phi} \otimes \varpi^{T}} \oplus \frac{\left(m_i \otimes \phi\right)}{F} \quad (3)$$

In the above equation, *k (x, y)* represents the RBF kernel function, S_b^{Φ} and S_w^{Φ} represent the divergence feature sample class divergence matrix and intraclass divergence matrix in *F*, w^{Φ} represents the kernel function instead of the dot product, w^T represents the linear combination of training samples in space.

Assuming that *E(V)* represents the discrete random variable of the intrusion sample, the sensor node data is compared with the established intrusion behavior characteristic data using equation (4) :

$$\phi\left(E\right) = \frac{E\left(V\right) \otimes \partial\left(s\right)}{\sigma\left(j\right) \otimes p_k^{*}} \times \frac{\hbar\left(q, p\right) \otimes x\left(y^{*}\right)}{\varpi\left(S\right)} \times \varepsilon\left(d\right) \quad (4)$$

In the above equation, p_k^{*} represents the *p* attribute on each data *i* on the network dataset *h*, $\partial(s)$ represents the standard deviation of the *s* attribute on the dataset, and б(j) represents the judgment of the heterogeneous distance function threshold, $\hbar(q,p)$ represents the basis of the intrusion data vector, $x(y^{*})$ represents the optimal classification function, $\varpi(S)$ represents the Euclidean distance of different attribute components, $\varepsilon(d)$ represents the correlation between different intrusion profiles.

Assuming that C_i represents the type of intrusion feature, $\varepsilon(E)$ represents the probability that the sample of type C_i appears in the set, then the detection of intrusion node information is done by using equation (5)

$$\lambda^{*}\left(e\right) = \frac{C_i}{\mathrm{N}\left(R\right)\times \varepsilon\left(E\right)} \otimes \wp\left(\mu\right) \quad (5)$$

In the above equation, N (*R*) represents the optimal attribute set, and $\wp(\mu)$ represents the discriminant error function for intrusion detection.

In summary, it can be explained that the intrusion node information optimization detection principle, the use of this principle can be completed on the intrusion node information detection optimization.

3. Optimizatio Method of Intrusion Node Information Based on Attribute Attack Graph

3.1 Extraction of information characteristics of intrusion node

In the process of network intrusion node optimization, the time-frequency state transition model of signal characteristics is established by using HTT method, the evolution map of network potential attribute intrusion attack is defined, and the intrusion node information is analyzed to get the envelope feature of intrusion node information. The specific steps are as follows:

Suppose that p represents the linearly stable Cauchy frequency characteristic, *x (t)* represents the original intrusion node information, τ represents the characteristic time scale of the intrusion node information, then the state space intrinsic modality of the potential intrusion node information is obtained by using equation (6) function:

$$y(t)=\frac{1}{\pi}p\int\frac{x(\tau)}{t-\tau}d\tau \quad (6)$$

It can be concluded from (6) that *x(t)* contains two parts of both empirical modal decomposition and Hilber spectrum, respectively, using equation (7)

$$\begin{cases} c_i(t)=\left\|X_S-\sum_{i=1}^{n}\omega_i X_i\right\| \\ h_i(t)=\dfrac{v_s}{v_j}\otimes\lambda\otimes a(t) \end{cases} \quad (7)$$

In the above equation, v_s represents the deviation of the state holding time X_s and w_i under the specific network potential intrusion. The more the number of state nodes blocked by the network in the potential intrusion state, the greater the value of v_s, the network intrusion model is set to *a (t)* for the initial time, λ of the invasion in the course of mathematical evolution.

Assuming that f represents the initial instantaneous frequency of the signal at different stages of network intrusion node information, x^{Φ} represents the convolution of the original intrusion signal, then the complex intrusion signal of network intrusion node information is decomposed into IMF single frequency intrusion signal by using equation (8) :

$$WD_x(t,f)=\int x^*\left(t+\frac{\tau}{2}\right)f\left(t-\frac{\tau}{2}\right)e^{-j2\pi ft} \quad (8)$$

Assuming that *x (t)* represents the latent intrusion signal square integrable function, $\psi(t)$ represents the fundamental function, the state transition equation of the intrusion detection system is established by using equation (9)

$$WT_f(\alpha,\tau)=\frac{1}{\sqrt{a}}\int x(t)\psi^*\left(\frac{t-\tau}{a}\right)dt\times\psi(t) \quad (9)$$

In the above equation, ψ^* represents the fundamental function, $\sqrt{a}$ representing the scale factor.

Assuming that *z (t)* represents the intrusion signal, then use the formula (10) to construct an information model with intrusion node information signal analysis:

$$z(t)=x(t)+iy(t)\otimes a(t)e^{i\theta(t)} \quad (10)$$

In the above equation, *x (t)* represents the real part of the signal resolution model, *y (t)* represents the intrinsic modal function of the intrusion signal, and *a (t)* represents the maximum and minimum of the data sequence with intrusion node information The upper and lower envelope obtained after three spline interpolation, *Θ(t)* represents the high frequency component.

According to the result of Eq. (10), it can be concluded that the original intrusion signal can be decomposed into multiple narrowband signal IMF components by EMD decomposition, then the envelope feature of the intrusion node information signal is extracted by Eq. (11)

$$a(t)=\sqrt{x^2(t)+y^2(t)}\otimes\theta(t)\frac{y(t)}{x(t)} \quad (11)$$

In the above equation, *a (t)* and Θ(t) represent the envelope and phase of the analytic form of the potential intrusion signal, respectively, and *a (t)* and Θ(t) are functions of time, therefore, the analysis of the intrusion node information signal *z (t)* is carried out with *1 / t* empirical mode decomposition convolution, thus preserving the local characteristics of the intrusion node information initial signal *x (t)*.

In this paper, we can explain that the intrusion signal is decomposed into IMF single frequency intrusion

signal and the state transition equation of the network intrusion detection system is obtained by defining the attribute attack graph of network potential intrusion in the process of network intrusion node optimization detection. Which has laid a foundation for the optimal detection of network intrusion node information.

3.2 Detection of intrusion node information based on optimal parameters

Based on the envelope characteristics of the intrusion node information signal acquired in section 3.1, the ant colony theory and the support vector machine parameters are merged to calculate the transition probability of the ants, and the ant is different from the ants in the process of network intrusion node optimization. And the nodes on the optimal path are connected to get the SVM optimal parameters, and the information of the network intrusion node is detected based on the parameters. The specific steps are as follows:

Assuming that the ant colony scale, representing each ant has a one-dimensional array, then the 3.1 node to obtain the intrusion node information signal envelope based on the characteristics of the use of (12) in turn save the first ant through a node

3.2 Detection of intrusion node information based on optimal parameters

Based on the envelope characteristics of the intrusion node information signal acquired in section 3.1, the ant colony theory and the support vector machine parameters are merged to calculate the transition probability of the ants, and the ant is different from the ants in the process of network intrusion node optimization. And the nodes on the optimal path are connected to get the SVM optimal parameters, and the information of the network intrusion node is detected based on the parameters. The specific steps are as follows:

Assuming that m is the ant colony scale, $Path_k$ represents each ant *k* has a one-dimensional array, then the intrusion node information obtained in Section 3.1 is based on the envelope feature ***a (t)***, using the formula (12) in turn save the first *k* ant through the *n* nodes of the vertical coordinates of the composition of the ant crawling path:

$$\mu^{*}(q)=\frac{n*m}{k\otimes a(t)}\otimes Path_k \tag{12}$$

Assuming that *O* is the starting point of all ants, *N* is the number of cycles, N_{max} represents the maximum number of iterations, then the probability of the ant is calculated using Eq. (13)

$$p_x\left(x_i,y_{i,j},t\right)=\frac{N_{\max}\times N}{O}\otimes path_k\otimes\Delta\tau\left(x_i,y_{i,j},0\right) \tag{13}$$

In the above equation, $\Delta\tau\left(x_i,y_{i,j},0\right)$ represents the amount of information on the initialization node.

Suppose that the parameters of SVM are represented by *C* and б, and the intrusion detection rate is used as the objective function by using Eq. (14) to induce the hormone material to be left on each node where the ant travels:

$$\Delta\tau_k\left(x_i,y_{i,j}\right)=\frac{\Delta\tau\left(x_i,y_{i,j},0\right)\otimes knot\left(x_i,y_{i,j}\right)}{(C,\sigma)} \tag{14}$$

In the above formula, $knot(x_i, y_{i,j})$ represents a node.

Assuming that $p_k(x_i, y_{i,j}, t)$ represents the probability of the *c*th of the *k*th-only ant, the ant is used to change the ant to select the optimal ant:

$$\varepsilon(d,j)=\frac{p_k\left(x_i,y_{i,j},t\right)*k}{(C,\sigma)}\otimes E_k \tag{15}$$

In the above equation, E_k represents the average of k times the intrusion detection rate.

According to the calculation result of Eq. (15), the *C* and б values of the corresponding ant are calculated by the formula (16)

$$\theta_i(C,\sigma)=\frac{\varepsilon(d,j)}{p_k\left(x_i,y_{i,j},t\right)*k}\otimes\mu(e,r) \tag{16}$$

In the above equation, $\mu(e,r)$ represents all the node information that the ant has crawled.

The optimal SVM optimal parameter is obtained by using equation (17)

$$\alpha(e,p)=\frac{\theta_i(C,\sigma)\oplus E_k}{\varepsilon(d,j)}\times\Delta\tau_k\left(x_i,y_{i,j}\right) \tag{17}$$

Based on the calculation results of Eq. (17), it can effectively perform the optimal detection of network intrusion node information.

4 Experimental and simulation proved

In order to prove the validity of the proposed intrusion node information detection method based on the attribute attack graph, an experiment is needed. In the Mtlab platform to build intrusion node information detection simulation platform. Experimental data from the KDD CuP 1999, which contains 5 million data, there are 5 categories of 40 kinds of attack types. Each data has two category labels and 40 attributes.

4.1 Comparison of the importance of different methods of feature extraction

The intrusion node information detection experiment is carried out by using the attribute attack graph method and the traditional distributed learning method respectively. The importance of extracting the information characteristics of intrusion nodes is compared with two different methods. The comparison results are shown in Fig. 1

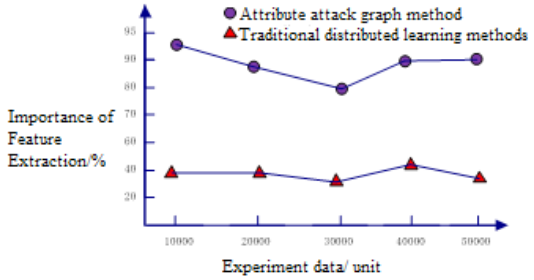

Figure 1 Comparison of different methods of feature extraction

It can be concluded from Fig. 1 that the importance of extracting the characteristics of intrusion node information based on the attribute attack map method is much higher than that of the distributed learning method, which is mainly due to the use of the proposed. When the intrusion node information is detected by the attribute attack map method, the complex intrusion signal with intrusion node information is decomposed into IMF single frequency intrusion signal, and the state transition equation of the network intrusion detection system is obtained. So as to enhance the effectiveness of the attack information based on the attribute attack graph.

4.2 Comparison of the overall superiority of different methods of intrusion node information detection

The intrusion node information detection experiment is carried out by using the attribute attack graph method and the traditional distributed learning method respectively. The missing report rate, false alarm rate and efficiency of the intrusion node information are compared by two different methods. The overall superiority of different methods for intrusion node information detection is measured by the results of the comparison. The results are shown in Fig. 2 and Fig. 3 and Fig. 4.

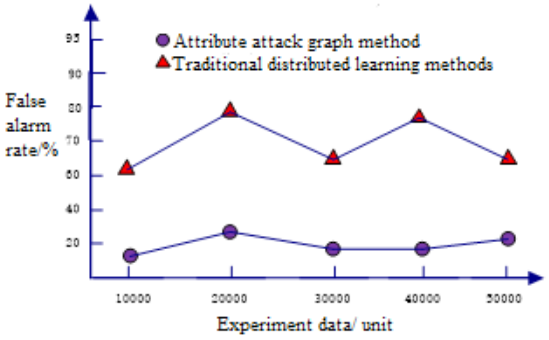

Figure 2 Comparison of false alarm rates for different methods of intrusion node information detection

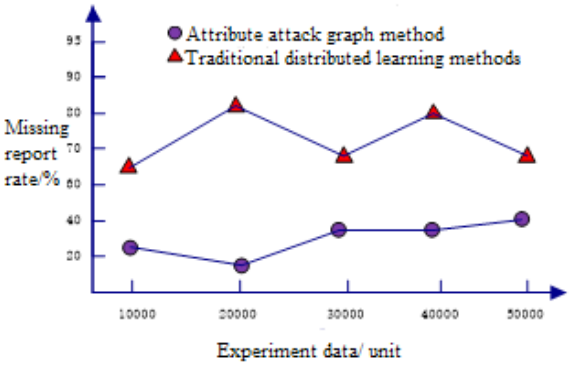

Figure 3 Comparison of different methods of intrusion node information detection missing report rate

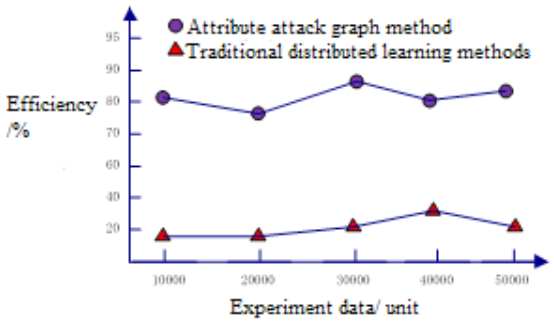

Figure 4 Comparison of the efficiency of different methods of intrusion node information detection

It can be concluded from Fig. 2 to Fig. 4 that the overall superiority of the intrusion node information detection using the proposed attribute attack graph method is higher than that of the distributed learning method for the detection of intrusion node information. This is mainly because when the intrusion node information is detected by the proposed method, the ant colony theory and the support vector machine parameters are merged, the network intrusion detection rate is taken as the objective function, and the ant is changed by Gaussian mutation, and the nodes on the optimal path are connected. The optimal parameters of SVM are obtained, and the information of network intrusion node is detected based on this parameter, which guarantees the overall superiority of the proposed method to detect the intrusion node.

The experimental results show that the proposed method has high accuracy and can improve the accuracy of embedded computer network intrusion detection.

5 Conclusion

When the information of the network intrusion node is detected by the existing method, the latest information attack type cannot be detected, and there is a problem that the detection deviation is large. This paper proposes an intrusion node information detection method based on attribute attack graph. The experimental results show that the proposed method has high accuracy and can improve the accuracy of embedded computer network intrusion detection.

References

[1] Luo Zhiyong, You Bo, Xu Jiazhong, et al. Intrusion - based Automatic Recognition Model Based on Three - layer Attack Graph [J]. Journal of Jilin University: Engineering Science, 2014, 44 (5): 1392-1397.

[2] Liu Lei, Liu Shuwei. Research on the Method of Identifying Driver ‘s Braking Intention [J]. Agricultural Equipment and Vehicles Engineering, 2015, 53 (11): 27-30.

[3] DING Qi-chuan, XIONG An-bin, ZHAO Xin-gang, et al.Study on the Research and Application of Motion Intention Recognition Method Based on Surface Electromyography [J]. Acta Automatica Sinica, 2016, 42 (1): 13-25.

[4] Ma Guocheng, Liu Zhaodu, Pei xiaofei, et al. Method of line-of-line vehicle alignment and line recognition based on fuzzy support vector machine [J]. Automotive Engineering, 2014, 36 (3): 316-320.

[5] Chen Xiaojun, Fang Binxing, Tan Qingfeng, et al. Study on Inference Algorithm of Internal Attack Intention Based on Probability Attack Graph [J]. Journal of Computers, 2014, 37 (1): 62-72.

[6] Liu Yuan, Li Qun, Wang Xiaofeng. Network defense strategy based on attack graph and improved particle swarm algorithm [J]. Computer Engineering and Applications, 2016, 52 (8): 120-124.

[7] Zhang Jian, Wang Jindong, Zhang Hengwei, et al. Network risk analysis method based on node game vulnerability attack graph [J]. Computer Science, 2014, 41 (9): 169-173.

[8] Chen Xiaojun, Shi Jinqiao, Xu Fei, et al. Study on Optimal Security Strategy Algorithm for Internal Threat [J]. Journal of Computer Research and Development, 2014, 51 (7): 1565-1577.

[9] HUANG Ting, LIU Zheng-Lian, WANG Si. Study on Watermark Attack Algorithm for Visible Digital Image Based on Curvature Diffusion Model [J]. Journal of Central University for Nationalities (Natural Science Edition), 2014, 23 (3): 49-55.

[10] Zhang Mu, Pu Cunlai. A new image encryption fusion algorithm simulation research [J]. Computer simulation, 2014, 31 (4): 402-406.