Advanced Engineering Forum *Online: 2012-09-26*
ISSN: 2234-991X, Vols. 6-7, pp 882-886
doi:10.4028/www.scientific.net/AEF.6-7.882

An Application-Layer Distributed Intrusion Detection Model Based on the C/S Mode*

Ding Zhi-guo[1,a] Zhu Xue-yong[1,b] Yuan Yuan[2,c]

[1]Center of Network and Information, Electronic Engineering Institute of Hefei, China

[2]Research Department , Electronic Engineering Institute of Hefei, China

[a]dhgzg@mail.ustc.edu.cn, [b]zhuxy@ustc.edu.cn, [c]yuanyuan@163.net

Keywords: Distributed Intrusion detection; C/S mode; The belief; Application-layer

Abstract: In order to overcome the disadvantages of the traditional distributed intrusion detection system, an application-layer distributed intrusion detection model based on the C/S mode is proposed here. The new model, is composed of a main system of server and several sub-systems of clients, fully utilizes detection abilities of the client by means of computing the belief dynamically, while the cost is not increased. Theoretical analysis and experimental results show that the model is a simple structure, reasonable design and higher accuracy than the traditional models.

Introduction

With the development of network attack techniques, the traditional protecting methods such as digital encryption, access control and security certification can not meet the security. Intrusion detection[1] based on the dynamic protection technology is being more and more paid attention to. Especially in the intrusion detection by means of fuzzy theory[2,3], neural networks[4,5] and immune algorithm[6,7], many scholars have made fruitful research results.

IDS (Intrusion Detection System) can monitor and detect the intrusions and attacks to the network by means of data interception, characteristic extraction, traffic statistics and so on. It can adjust security policies according to the security threats to the network dynamically. According to IDS's structure, it can be divided into two categories, host-based centralized system and network-based distributed system. Centralized system is simple and the low cost of deployment, but it will consume the host resources seriously, and it is also vulnerability to attack. Distributed system can tackle a large amount of data in short time and is easy to hide, but it needs additional hardware devices and is not easy to deploy. However, the distributed IDS has stronger monitoring capabilities than the centralized in spite of the more expense of hardware costs, and it is more adapted to the complex networks. Obviously that both are integrated with is a good idea.

For that reason, an application-layer distributed intrusion detection model based on the C/S mode is proposed. The new model, which has the properties of the C/S mode and host-based in the application-layer, fully utilizes the detected the abilities of client by means of computing the belief dynamically and overcomes the shortcomings of the traditional distributed IDS.

Application-layer distributed intrusion detection model based on the C/S mode

The distributed intrusion detection needs to set some monitors and a monitoring center on the network. The server can manage and configure the monitoring points dynamically, process data from the monitors and assess the security threats. The C/S mode is just composed by a server and some clients. When a client is installed to the specific software, it can act as a monitor, and similarly the server can act as the monitoring center.

In the application-layer, IDS can gather information from the application-layer, the detected content is definite and feature extraction is easy, so the detection accuracy is higher than of in other layers. In addition, for the application based on the C/S mode, the network environment of the client is uncertain and complex, if IDS only relies on the data from the network layer, the missed and false detections are inevitable.

Using of the advantages of IDS in the application-layer and distribution characteristic of the C/S mode, we propose an application-layer distributed intrusion detection model(ADID for short), which is in Fig.1.

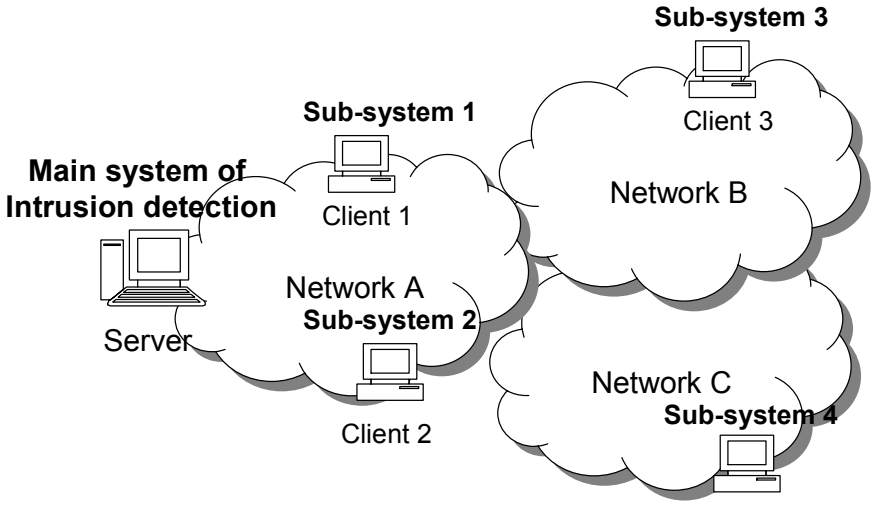

Fig.1 The topology of the application-layer distributed intrusion detection model

From Fig. 1, the ADID is generally composed of a server-side main system and several client-side sub-systems. Sub-systems are responsible for collecting and pre-processing data, they sent suspicious information to the main system after simple analysis. The main system is responsible for integration of different sources of information and feature extraction. It assesses the security threats and takes appropriate security policy according to the feature and rule database.

According to the location of the clients in network and detection accuracy of sub-systems, the different sub-systems would be assigned to the different degrees of trust, i.e. the belief.

At the time of t, the belief of i th sub-system is defined as:

$$c_i(t) = w_i(t) / d_i(t) \tag{1}$$

Where $w_i(t)$ is the weight of i sub-system. The weight of sub-system is adjusted according to its normalized detection accuracy $a_i(t)$ $(-1 \le a_i(t) \le 1)$:

$$w_i(t+1) = \begin{cases} 1 & w_i(t+1) > 1 \\ w_i(t) + \lambda a_i(t) & 0 \le w_i(t+1) \le 1 \\ 0 & w_i(t+1) < 0 \end{cases} \tag{2}$$

Where $0 \le \lambda \le 1$ is the rate of weight update. $d_i(t)$ is hops between the server and clients.

The module of the application layer distributed intrusion detection system is shown in Fig. 2, which includes the data monitor, data preprocessor, collaborative analyzer, risk evaluator and security manager. Their main functions are as follows:

Data monitor: It can intercept packets in the network and get some important information, such as MAC/IP address, protocol type, port number and data length. On the other hand, it can search the log files to find traces of the attacks and operation records of the system.

Data preprocessor: It extracts suspicious information from packets and operation records, and then assesses the threat degree of each suspicious information. For the sub-systems, they send the suspicious information to the main system finally.

Collaborative analyzer: It can gather data from the main system and sub-systems., and then it calculates the threat degree of each suspicious information according to the belief of main system and sub-systems. The threat degree of a suspicious information is defined as follows:

$$Z(j,t) = \frac{1}{N+1} \sum_{i=0}^{N} c_i(t) z_i(j) \tag{3}$$

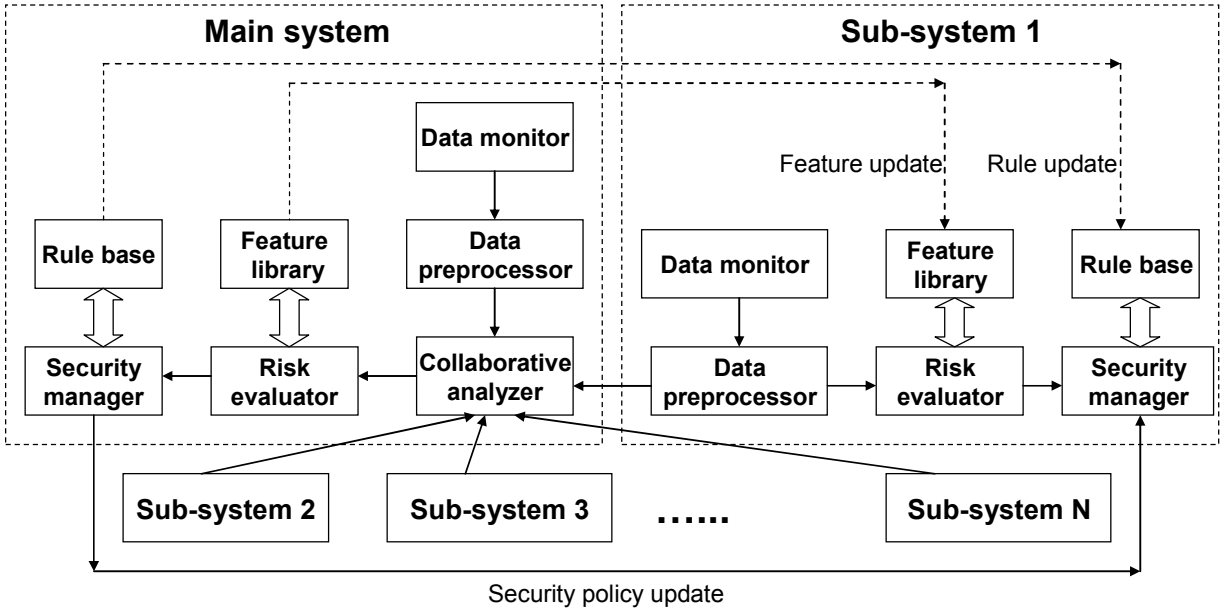

Fig.2 The module of application layer distributed intrusion detection system

Where $c_0(t)$ is the belief of the main system, $c_i(t)$ is the belief of i th sub-system ($1 \le i \le N$), $z_i(j)$ is the threat degree of j th information from the i th sub-system, N is the number of the sub-systems.

The main system summaries all suspicious information and calculates the average threat degree:

$$Z(t) = \sum_{j=1}^{M} Z(j,t)/M = \frac{1}{M(N+1)} \sum_{j=1}^{M} \sum_{i=0}^{N} c_i(t) z_i(j) \tag{4}$$

Where M is the number of suspicious information in unit time.

Risk evaluator: It is used to assess the accuracy of each suspicious information and the security threat level to the system according to the feature database. When the new attacks are found, the risk evaluator can get new feature and add it to the feature database. Meanwhile, the risk evaluator of main system calculates the accuracy of the i th sub-system:

$$s_i(t) = \frac{1}{M} \sum_{j=1}^{M} \frac{|\hat{z}(j) - z_i(j)|}{\hat{z}(j)} \tag{5}$$

Where $\hat{z}(j)$ is the thread degree of j th information after the main system assessment. The average accuracy of the sub-systems is as follows:

$$s(t) = \frac{1}{N} \sum_{i=1}^{N} s_i(t) = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{|\hat{z}(j) - z_i(j)|}{\hat{z}(j)} \tag{6}$$

Finally, the normalized detection accuracy of the i th sub-system is given as:

$$a_i(t) = s_i(t) - s(t) \tag{7}$$

Then, the main system updates the weight of each sub-system according to the formula (2).

Security Manager: It can update the security policy according to the security threat level to the network, the type of the attacks and rule database, such as update the key, enhance the protection level of firewall and adjust encryption algorithms.

The new model still includes the feature database and rule database. There are two kinds of features in the feature database, system features and invasion features. The former includes features of users, packets and operation records, such as the user's key and permissions, authentication methods, address and port of the packets. The later includes the address of the attacker, the mode and period of the attacks. The feature database are self-learning and self-updating. The feature

database of the main system and sub-systems must update synchronously in definite period. The rule database can be used to formulate risk assessment criterion and protect method against the attack. The ADID system will take suitable security strategies according to the rule database. Similar, the rule database of main system and sub-systems must also update synchronously in definite period.

Performance simulation and analysis

Because the new model does not provide the methods of data processing, in order to simplify problem, the difference between the sub-systems is not considered in the performance analysis. It is assumed that the sub-system only report suspicious information without distinguishing between the threat degree, the threat degree of suspicious information are fixed value, for example $z_i(t)=1$, and moreover, the detection capability of each sub-system is equal, the detection accuracy is p, every sub-systems are independent of each other and non-related, the number of the sub-systems is N.

Then, the threat degree of the main system *Z(t)* complies with the binomial distribution:

$$P(Z(t)=k)=\binom{N}{k}p^k(1-p)^{N-k} \tag{8}$$

If the main system's detection threshold is h, the detection probability of the main system:

$$P(Z(t)>h)=\sum_{h}^{N}\binom{N}{k}p^k(1-p)^{N-k} \tag{9}$$

According to Dermot Buddhism-Laplace theorem, we can get an approximation:

$$P(Z(t)>h)=P\{\frac{h-p}{\sqrt{p(1-p)/N}}<\frac{z(t)-p}{\sqrt{p(1-p)/N}}\}\approx 1-\Phi(\frac{h-p}{\sqrt{p(1-p)/N}}) \tag{10}$$

Similarly, while the false detection rate of each sub-system is $1-p$, the false detection rate of the main system is given as:

$$P_w(Z(t)>h)=P\{\frac{h-p}{\sqrt{p(1-p)/N}}<\frac{z(t)-p}{\sqrt{p(1-p)/N}}\}\approx 1-\Phi(\frac{h-(1-p)}{\sqrt{p(1-p)/N}}) \tag{11}$$

Then, the detection accuracy of the main system is given as:

$$P_s=P-P_w=\Phi(\frac{h-(1-p)}{\sqrt{p(1-p)/N}})-\Phi(\frac{h-p}{\sqrt{p(1-p)/N}}) \tag{12}$$

In order to verify the practicality and effectiveness of the new model, Matlab7.0 is used to simulate it. Let the parameters are as follows, $N=20$, $p=0.8$, $h=0.7,0.5,0.4$ respectively. The simulating performance of the new model is shown in Fig. 3

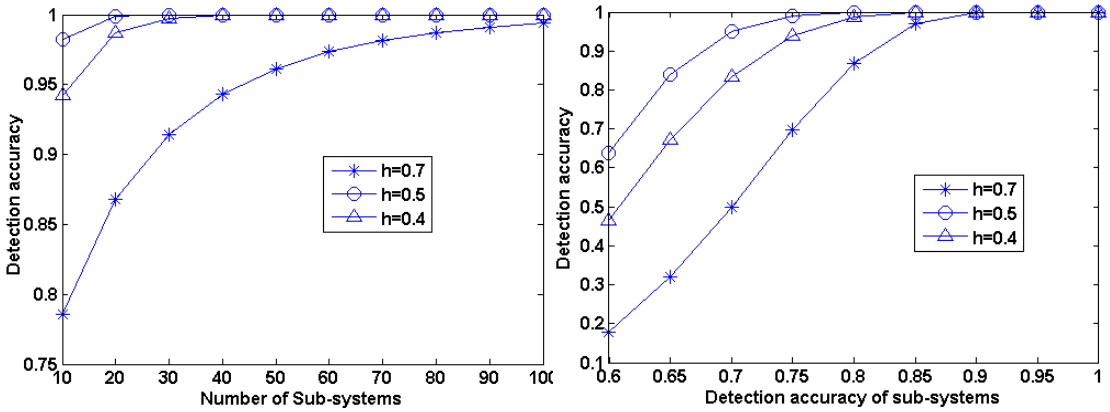

Fig.3 Performance of the new model

From Fig. 3, with increasing of the number and the detection accuracy of the sub- systems, the detection accuracy of the new model is increasing.

However, it is worthily noted that the detection threshold affects the performance of the new model significantly. If the detection threshold is higher, the detection rate of the model is higher, but the missed detection rate is also higher simultaneously. If the detection threshold is lower, the missed detection rate will be lower, but the false detection rate is also higher simultaneously.

So, In order to enhance the detection accuracy of the new model, the correct method is to increase the detection accuracy of the sub-systems rather than the number of the sub-systems. Because the sub-systems will increase the network load, the performance of detection system must be affected.

Conclusion

In order to overcome the disadvantages of the traditional distributed intrusion detection system, the ADID model is proposed in this paper. Theoretical analysis and experimental results show that it can enhance the detection accuracy with adjusting detection threshold, and the new model is a simple structure, reasonable design and higher accuracy than the traditional IDS model.

* This paper is supported by 090412055 program of the Natural Sciences Foundation of Anhui China

References

[1] Hu Changzhen, Network intrusion detection theory and technology, first ed, Beijing institute oftechnology press,2006.

[2] Chimphlee W, Abdullah A H. To Detect Misuse and Anomaly Attacks Through Rule Induction Analysis and Fuzzy Methods[J]. WSEAS Trans on Computers, 2006, 5(1):49-54.

[3] Huang Guoyan, Chang Xuliang, Gao Jianpei. Application research of fuzzy logic theory in intrusion detection systems[J]. Engineering and applications. 2010,46(98):110-113.

[4] Tong X J, Wang Z, Yn H N. A research using hybrid RBF/Elman neural networks for intrusion detection system secure model[J]. Computer physics Communications. 2009, 180(10):1795-1801.

[5] Xu qinzhen, Yangluxi. An optimized neural network tree based on anomaly intrusion detection method[J]. Signal processing. 2010,26(11):1663-1669.

[6] D.Dasgupta, F.Gonzalez． An Immunity-based Technique to Characterize Intrusions In Computer Networks[J].IEEE Transactions on Evolutionary Computation, 2005,6(3):281～291.

[7] Yu Yan, Huang Hao. An ensemble approach to intrusion detection based on improved multi-objective genetic algorithm. Journal of software. 2007，18（6）：1369-1378.

[8] Fu Desheng, Zhou Shu, Guo Ping. Design and Implementation of distributed network intrusion detection system based on date mining. Computer science..2009,26(3)：103-105.