

Research of Automatic Recognition Double Layers Intrusion Intention Algorithm Based on Attack Graph

Wang Guangze, Wang Peng, Luo Zhiyong, Zhu Suxia

Harbin University of Science and Technology, Harbin 150080, Heilongjiang Province, China
wanguangze_hust@sina.com

Abstract

The determination of network equipment weaknesses and the discovery of intrusion intention is one of the difficulties that troubled network security management personnel. Based on previous studies, further proposed a double attack graph based on domain-equipment. By the underlying network topology data collected and analyzed, using Bayesian theory to complete the quantify for the double attack graph and generation strategy in minimal power key set, with the cost of calculation of key equipment in the automatic recognition network topology, we provide an important basis for network maintenance. Experimental results show that the measure of using quantitative domain-equipment double attack graph to recognize the intrusion intention is not only effective and feasible, but also has the feature of easy promotion.

Keywords: network security; intrusion intention; attack graph; recognition algorithm

1. Introduction

Intrusion intention identification refers to a process in the Internet environment, through the analysis of a large number of feedback alarm data from the underlying intrusion detection system, and effectively judging the means of the invaders and the ultimate goal, which is the essence of the scientific analysis of the intrusion process[1~3]. The fast and effective method of identifying intrusion intention provide an important basis for network management personnel to better maintain network security, it is also the premise for early detection and prevention by the network security threats, an important part of network security situation analysis, and the emerging field of network security research focus.

In the computer field, artificial intelligence achieved the recognition of intention earliest. Bratman M used the recognition of intention for decision-making and action planning [4]. Swiler, who first introduced the attack graph, using the nodes and edges of graph to formalize association between network devices, providing a new approach [5] to identify intrusion intention. Sheyner, who used the greedy algorithm to optimize the attack graph, identifying the key weaknesses in the minimum set throughout network, and as a theoretical basis for the recognition of intrusion intention [6]. Ou X further improved the thoughts of attack graph, putting forward a fast method to identify intrusion intention [7]. Lippmann and others increased constraints on the attack graph nodes and weakness, thereby reducing the attack graph scale, providing a quick method to quantify [8]. Roschek used scanning tools NESSUS to find the key nodes in the network, generating a key node attack graph by the algorithm that search through the front and rear, effectively prevent the intrusion occurring [9]. Ye further quantified the attack graph by the probability of intrusion occurred to provide a theoretical basis [10] to identify intrusion intention. Chen Xiaojun, who proposed the idea of the probability and statistics to apply to the attack graph, the attack graph quantified provided the basis to further determine the intention of the network, effectively preventing the intrusion [11].

On the basis of years of research [12], the group further proposed the double attack graph model based on domain-equipment, and using Bayesian probability theory to carry out the research about the automatic recognition of intrusion intention.

2. Building Automatic Recognition Model for Intrusion Intention

2.1. Intrusion Intention and Recognition Model

Intrusion intention is an intruder original purpose, and it is expressed as an intruder brings the damage for the network which is invaded. Intrusion intention recognition is a process that is the network security people taking measures for the invaders intent which is expresses as forecast, analysis, evaluation and block. Its recognition model structure is shown in Figure 1.

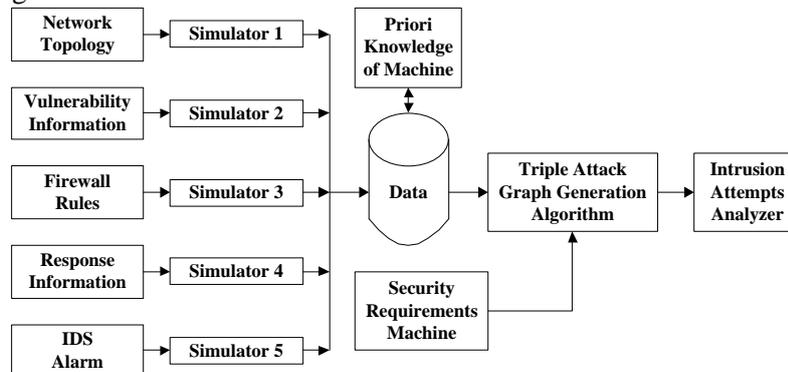


Figure 1. Architecture Diagram of Intrusion Intent Recognition Model

2.2. Domain-equipment Two-tier Attack Graph

To create a domain-equipment double attack graph model, we formalized the following constraints:

W represents a collection of weakness in the network device, then for any one weakness w in set W , w can be defined as a triple (w_{id}, w_p, w_c) . In the tuple, w_{id} is the number of the weakness in the network security CVE standard library; w_p is the set of precursor conditions that intruder can successfully exploit this vulnerability; w_c is the set of harm that an intruder using the vulnerability cause on the network.

IP represents the set of the equipment in network layer, then for any one network equipment ip in IP , ip defined as triples $(W, o, NetIP)$. In the triple, W is the set of network equipment weaknesses; o is the set of open ports in network equipment; $NetIP$ is the set of other devices connected to the network.

D represents the set of network topology domains, then for any one domain d in set D , d may be defined as triples $(IP, NetD)$. In the triple, IP is a set of network equipment in the domain; $NetD$ is a set of other domains connected to the domain.

E represents the set of node equipment, then for any one node device e in set E , e may be defined as a triples (e_{id}, D, IP) . In the triple, e_{id} is the node number of the device; D is the set of the node equipment in the domain; IP is a set of network equipment that belongs to the node equipment.

In the daily intrusion events, the intruder first scans all the weaknesses of a node equipment e in the network topology, forming weakness set w . By analyzing the set w , the intruders may invade the ip of equipment in some domain d , completing the controlling process that make the privilege of ip change from low to high. Secondly, by the set $NetIP$ of device ip , invading other equipment ip' in the domain d , the intruders completely control the domain d step by step eventually. Finally, by the set $NetD$ of domain d , the intruders continue invading other domains d' in the same way, realizing their invasion

intention eventually. Thus, the domain-equipment double attack graph model shown in Figure 2.

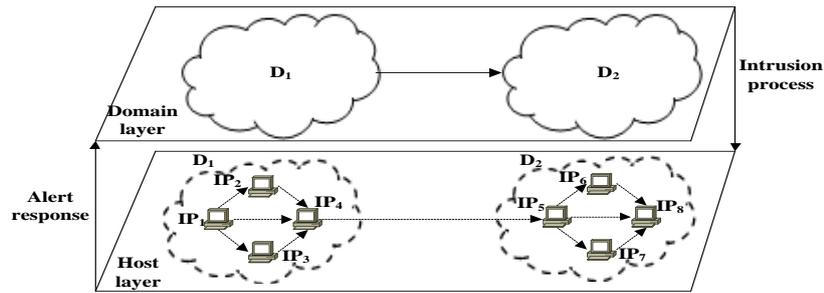


Figure 2. Domain-equipment Double Attack Graph Model

Definition 1: domain-equipment double attack graph V , it is a view and formalized as a triples $(V_D(E_d, L_d), V_{IP}(E_{ip}, L_{ip}))$, where: $V_D(E_d, L_d)$ is the domain layer attack graph; $V_{IP}(E_{ip}, L_{ip})$ is the equipment layer attack graph; E_q is a set of node equipment; L_q is the set of the routes between the node equipment; q is d or ip . In the definition 1, if there is a route that make node equipment e_i turn to the node equipment e_j , it means the intruders invade the node equipment e_j by the node equipment e_i successfully, where: $l_{ij} \in L_q, e_i \in E_q, e_j \in E_q, e_i \neq e_j, q = d$ or ip .

2.3. Attack Graph Generation Strategy

Domain-equipment two-tier attack graph V generation policy obeys the following restrictions:

- (1) If an intruder has gained a network equipment elevated privileges, and the intruder is no longer regain the network equipment lower privileges.
- (2) If the intruder has successfully invaded network equipment, the intruder is no longer re-invasion of the network equipment.
- (3) In the process of an intruder achieving his intention invasion, his invasion operations are necessary and non-redundant.
- (4) An intruder gets a network equipment's permission from low to high, and the orders are Null, Guest, Admin.

Generation strategy 1 of the domain-equipment double attack graph follows:

Step1: Initialize policy-related variables, and set Pow for permission variable;

Step2: Remove a equipment from the set IP of network equipment in device layer to store in variable ip ;

Step3: Set permission variable Pow corresponding the variable ip empty Null, computing the rank of the set of the weakness in network equipment and assigning the variable with Num ;

Step4: Make the loop variable $i = 1$;

Step5: Make loop variable $j = i + 1$;

Step6: If w_c that the weakness w_i was successfully invaded lead to meet the precondition w_p that weakness w_j be invaded, namely satisfy the relationship $w_j \times w_p \subseteq w_i \times w_c$, then assigning the permission variable Pow corresponding variable ip with Guest or Admin and perform Step9, otherwise perform Step7;

Step7: Let the variable $j = j + 1$, determine j whether the variable Num equal, if not perform Step6, otherwise perform Step8;

Step8: Let the variable $i = i + 1$, judge i whether the $Num-1$ equal to, if not perform Step5, otherwise perform Step9;

Step9: Remove the next device from the set IP of the network equipment in device layer to store in the variable ip ;

Step10: If the set IP is not empty, then perform Step3, and otherwise perform Step11;

Step11: Recover the set IP of the network equipment in device layer and set Num to the rank of set IP ;

Step12: Make the loop variable $i = 1$;

Step13: Make loop variable $j = i + 1$;

Step14: If the permission of the equipment ip_i is not null, namely the Pow corresponding equipment ip_i is Guest or Admin, then perform Step15, otherwise perform Step17;

Step15: If the equipment ip_i and ip_j can be connected through port o and it can invade equipment ip_j via the equipment ip_i and access the privilege Guest or Admin of the equipment ip_j , then perform Step16, otherwise perform Step17;

Step16: add the equipment ip_i and ip_j to equipment attack graph $VIP(E_{ip})$, and add the edge $ip_i \rightarrow ip_j$ to the attack graph $VIP(L_{ip})$ in the device layer;

Step17: Let the variable $j = j + 1$, determine j whether the variable Num equal to, if not perform Step14, otherwise perform Step18;

Step18: Let the variable $i = i + 1$, judge i whether the $Num-1$ equal to, if not perform Step13, otherwise perform Step19;

Step19: Add the attack graph $VIP(E_{ip}, L_{ip})$ in device layer to the double attack graph V ;

Step20: Make the loop variable $i = 1$;

Step21: Make loop variable $j = i + 1$;

Step22: Make d_1 as the protection domain for equipment ip_i , d_2 as the protection domain for equipment ip_j ;

Step23: If the two domains are not same, that $d_1 \neq d_2$, and the equipment ip_i has been invaded, then perform Step24, otherwise perform Step26;

Step24: Add the protection domain of the equipment ip_i to domain layer attack graph $VD(E_d)$;

Step25: If it can invade equipment ip_j through port o in equipment ip_i and improve the privilege of the control equipment ip_j to Guest or Admin, then add the protection domain of the equipment ip_j to the domain layer attack graph $VD(E_d)$, add the edge $d(ip_i) \rightarrow d(ip_j)$ to the domain layer attack graph $VD(L_d)$, otherwise perform Step26;

Step26: Let the variable $j = j + 1$, determine j whether the variable Num equal to, if not perform Step22, otherwise perform Step27;

Step27: Let the variable $i = i + 1$, judge i whether the $Num-1$ equal to, if not perform Step21, otherwise perform Step28;

Step28: Add the domain layer attack graph $VD(E_d, L_d)$ to the double attack graph V .

After analysis, the time complexity of the policy 1 is $O(|IP| \times |W_{ip}|^2)$.

3. Quantitative Analysis of Intrusion Intention and Response Strategies

3.1. Quantitative Analysis of Intrusion Intention

In the double attack graph of the domain-equipment, the property of weakness determines each node that intruder can invade successfully. Therefore, this article will divide the property of the device weakness into three parts: easiness α , privacy β and the rate of return γ . Depending on the complexity of the actual network operations, the author made assessment and assignment with each property of the weakness, as shown in Table 1, the probability of an intruder successfully exploited weaknesses w as follows:

$$(1) \quad \pi(w) = \delta_1\alpha + \delta_2\beta + \delta_3\gamma$$

Where: $\delta_1, \delta_2, \delta_3$ are weights assigned by the network management personnel in accordance with the actual situation.

Table 1. Properties and Assignment of the Weakness

Property	Degree	Assignment
α	easy	0.9
	secondary	0.6
	difficult	0.3
β	low	0.6
	mid	0.8
	high	0.9
γ	low	0.5
	mid	0.7
	high	0.9

Definition 2: Invasion route. In the domain-equipment double attack graph V , if there exists the set E' of node equipment, it may make the intruder start from node equipment e_0 along with the equipment in set E' to realize the intrusion intention, then the node equipment set E' and the link of them in domain-equipment double attack graph called a intrusion route, denoted by r . In the attack graph V , all the intrusion paths constitute a collection that is called invasion path set, denoted R .

In the attack graph of equipment level, if there exists j invasion path through k node equipment, and the intruder can achieve its intrusion intention i , then the probability of intrusion intention i that may happen as follows:

$$\pi(i) = 1 - \prod_j [1 - \prod_k \pi(\pi_k)] \quad (2)$$

By Bayesian formula, we can calculate the relative probability of an intruder may realize the intrusion intention by intrusion route t :

$$\pi(ip_t | i) = \frac{\pi(i | ip_t) \times \pi(ip_t)}{\pi(i)}, \quad t = 1, 2, \dots, j \quad (3)$$

If the relative probability of some invasion path is higher, it indicates the possibility of an intruder from the invasion path to achieve its intentions may be high. For this reason, network managers should focus on protecting the node equipment through the invasion route.

3.2 Minimal Right Key Set of Attack Graph

Definition 3: Minimal key right set. In the domain-equipment double attack graph V , $V(E)$ is the set of node equipment, it makes K_i the nonempty set that does not contain the start node and the destination node and meet $K_i \subset V(E)$, if any intrusion route r in set R through all node equipment in K_i , there call K_i is the minimal key right set K_{minW} .

For the generation strategy to design minimal right key set K_{minW} , this paper follows the formal constraint:

A_{rank} represents the rank of the node equipment set $V(E)$, there: $A_{rank} = |V(E)|$; D_{rank} represents the rank of the intrusion route set R , there: $D_{rank} = |R|$, and $D_{rank} = C_{A_{rank}-2}^1 + C_{A_{rank}-2}^2 + \dots + C_{A_{rank}-2}^{A_{rank}-2}$; r_i represents any intrusion route, and meet $r_i \in R$; E_i represents the set of all node equipment in the invasion route r_i ; In symbols r_i and E_i , the $i = 1, 2, \dots, D_{rank}$.

The generation strategy of the minimal right key set K_{minW} in the domain-equipment double attack graph V follows:

Step1: Initialize policy-related variables and set landmark variable $Flag$ is true;

Step2: Calculate the number of elements in the set R and assign to the variable i ;
 Step3: Set landmark variable $Flag$ is true;
 Step4: Calculate the number of elements in the set $V(E)$ and assign to the variable j ;
 Step5: Judge the intersection between the set K_i and K_j whether empty, if empty then there is a route that not through node in set K_i , set the landmark variable $Flag$ false and perform Step6; if not empty, then perform Step6;
 Step6: Variable $j = j-1$, determine j whether 0, if 0 perform Step7, otherwise perform Step5;
 Step7: Judge the landmark variable $Flag$ whether is true? If true add the set K_i to the set K ;
 Step8: Variable $i = i-1$, determine i whether 0, if 0 perform Step9, otherwise perform Step3;
 Step9: Find the minimum elements in the set K as the minimum right key set K_{minW} .
 After analysis, the time complexity of the strategy 2 is $O(D_{rank} \times A_{rank})$.

3.3 Intrusion Intention Response based on the Minimal Right Key Set

In the double attack graph, the main way is cutting off intrusion route to prevent intrusion intentions. Therefore, the intrusion intention response based on the minimal right key set is economically viable process.

According to the algorithm 2, you can get the minimal right key set K_{minW} in double attack graph. Suppose ip_i is any one node equipment in set K , the price of maintaining the node equipment ip_i recorded as $Cost(ip_i)$, including: labor costs, hardware costs and other costs.

Based on the above analysis, the optimal maintenance measures costs of the network managers cut off the intrusion intention follows (4), which the cost of maintain the equipment node in the minimal right key set.

$$Sum_{Cost} = \sum_{i=1}^{|K_{minW}|} Cost(ip_i) \quad (4)$$

4. Test and Analysis of the Experiment

To test the intrusion intention identified algorithm of the domain-equipment double attack graph, experimental environment that research group designed mainly consist of four domains and Internet. Among them, environment field is named: domain D_1 , domain D_2 , domain D_3 and domain DMZ . Access policy for each domain: network device in Domain D_1 and D_2 can access each other including the domain DMZ ; network equipment in domain D_3 can not access equipment in domain DMZ ; In domain D_1 and domain D_2 , network equipment IP_4 and network equipment IP_9 can access the database server equipment in domain D_3 ; the network equipment IP_4 in domain D_1 and the equipment in domain D_2 can access each other; in their respective domains, network devices can access each other; the inner internet equipment exchange data with Internet through domain DMZ ; other cross-domain access are disabled. Experimental environment, the weakness of network equipment in the network security CVE standard library and the network topology are shown in Figure 3.

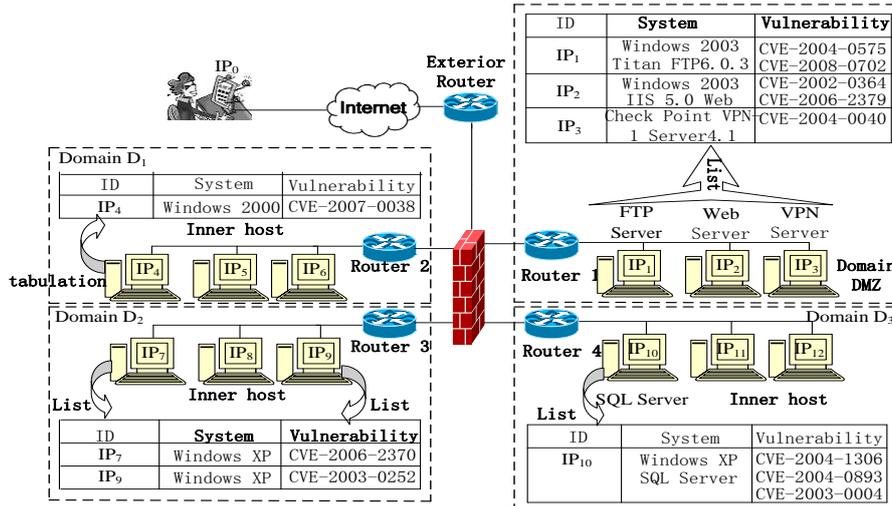


Figure 3. Network Topology and the Weakness of the Equipment

Since the server device IP_{10} stored a lot of sensitive data, so the device is a penetration goal by intruders and the majority of the intrusion refer to, it needs special protection. So the intrusion intention that intruders penetrate network equipment IP_{10} successfully is i , then take policy 1 can generate the intrusion intention, the corresponding domain-equipment double attack shown in Figure 4.

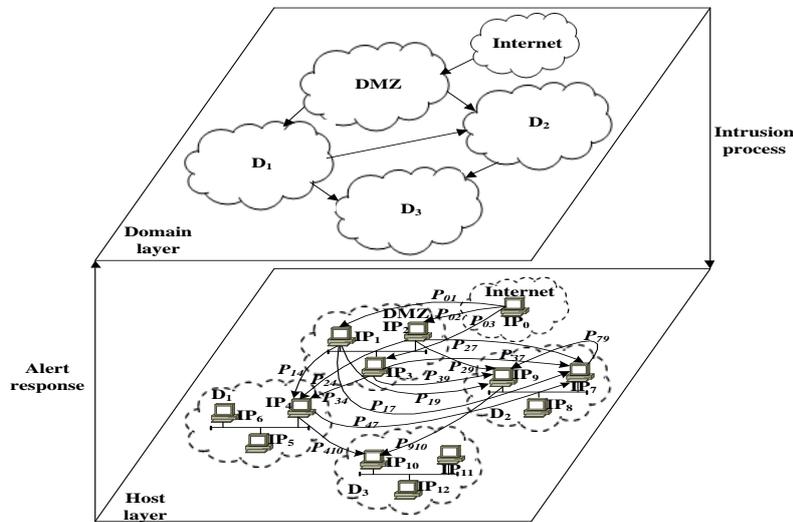


Figure 4. Domain-equipment Double Attack Graph of the Intrusion Intention i

In the double attack graph model shown in Figure 4, according to formula (1), the equipment $IP_1, IP_2, IP_3, IP_4, IP_7, IP_9$ and IP_{10} intrusion probability can be calculated by administrators, there: 0.3, 0.2, 0.5, 0.6, 0.7, 0.3 and 0.5, that in Figure 4, $\pi_{01}=0.3, \pi_{02}=0.2, \pi_{03}=0.5, \pi_{14}=\pi_{24}=\pi_{34}=0.6, \pi_{17}=\pi_{27}=\pi_{37}=\pi_{47}=0.7, \pi_{19}=\pi_{29}=\pi_{39}=\pi_{79}=0.3, \pi_{410}=\pi_{910}=0.5$. In the double attack graph model, there are total 12 kinds of intrusion route that achieve IP_{10} penetration for network equipment. Therefore, the probability of each route used by intruder is $1/12$, namely $\pi(r)=1/12 \approx 0.083$, each route and the probability distribution of the intrusion intention as shown in Table 2. Among them, the probability π represents the probability of each route is successfully invaded, according to the formula (2), the probability i of intrusion intention can be derived, where: $\pi(i) = 1 - (1-0.0900) \times (1-0.0189) \times \dots \times (1-0.0750) \approx 0.4749$.

According to the formula (3), it obtained the relative probability $\pi(r|i)$ for each intrusion route, as shown in Table 2, namely: $\pi(r_i|i) = (\pi_i \times \pi(r)) / \pi(i) = (0.0900 \times 0.083) / 0.4749 \approx 0.0157, \dots$. In relative probability, $\pi(r_9|i) = 0.0262$ is the maximum, so the intruder is most likely to adopt this intrusion route $IP_0 \rightarrow IP_3 \rightarrow IP_4 \rightarrow IP_{10}$ to achieve its intrusion intention. Strategy 2 can calculate the minimal right key set of each intrusion route in Table 2, where: (IP_4, IP_9) . To effectively prevent intruders achieve penetration intention i of the network equipment IP_{10} , administrators can strengthen the protection for equipment IP_4 and IP_9 , timely repairing related patches to restrict some users permission that access the database server IP_{10} . According to formula (4), we can calculate the cost of the maintenance of network security: $Cost(IP_6) + Cost(IP_9)$.

Table 2. Intrusion Route and the Probability of the Intention i

r	Intrusion route (r)	Probability π	Relative probability $\pi(r i)$
1	$IP_0 \rightarrow IP_1 \rightarrow IP_4 \rightarrow IP_{10}$	0.0900	0.0157
2	$IP_0 \rightarrow IP_1 \rightarrow IP_4 \rightarrow IP_7 \rightarrow IP_9 \rightarrow IP_{10}$	0.0189	0.0033
3	$IP_0 \rightarrow IP_1 \rightarrow IP_7 \rightarrow IP_9 \rightarrow IP_{10}$	0.0315	0.0055
4	$IP_0 \rightarrow IP_1 \rightarrow IP_9 \rightarrow IP_{10}$	0.0450	0.0079
5	$IP_0 \rightarrow IP_2 \rightarrow IP_4 \rightarrow IP_{10}$	0.0600	0.0105
6	$IP_0 \rightarrow IP_2 \rightarrow IP_4 \rightarrow IP_7 \rightarrow IP_9 \rightarrow IP_{10}$	0.0126	0.0022
7	$IP_0 \rightarrow IP_2 \rightarrow IP_7 \rightarrow IP_9 \rightarrow IP_{10}$	0.0210	0.0037
8	$IP_0 \rightarrow IP_2 \rightarrow IP_9 \rightarrow IP_{10}$	0.0300	0.0052
9	$IP_0 \rightarrow IP_3 \rightarrow IP_4 \rightarrow IP_{10}$	0.1500	0.0262
10	$IP_0 \rightarrow IP_3 \rightarrow IP_4 \rightarrow IP_7 \rightarrow IP_9 \rightarrow IP_{10}$	0.0315	0.0055
11	$IP_0 \rightarrow IP_3 \rightarrow IP_7 \rightarrow IP_9 \rightarrow IP_{10}$	0.0525	0.0092
12	$IP_0 \rightarrow IP_3 \rightarrow IP_9 \rightarrow IP_{10}$	0.0750	0.0131

5. Conclusion

Intrusion intention identification as a means of network intrusion situation analysis and identification has become an important research direction in security management; it provides an important basis for managers to effectively determine the network vulnerabilities, providing some protection to prevent the occurrence of network intrusion. Thus, intrusion intention identification technology is one of the hotspot problems of today's network security research. On the previous research results^[12], the group optimized and presented a identification method that analysis intrusion intention automatically, detecting and making response based on the results of previous research which is domain-equipment double attack graph. The method refers the probability analysis means of Bayesian in the double attack graph, quantifying for each intrusion route in the figure, and determining the minimal right key set in network. Administrators can set the focused maintenance on the network equipment in the set, thus effectively preventing the implementation of intrusion intention. Next, the group will focus on the weaknesses layer to study, dividing the double attack graph into the domain-equipment-weaknesses attack graph. Quantifying intrusion route from the weakness layer will further improve the accuracy of determining intrusion intention, provides important data basis for more simple and efficient managing intranet.

Acknowledgement

Supported by National Natural Science Foundation of China (NO: 61502123) and Heilongjiang Provincial Youth Science Foundation (NO: QC2015084).

References

- [1] W. Peng, C.-Z. Hu, S.-P. Yao and Z.-G. Wang, "A Dynamic Intrusive Intention Recognition Method Based on Timed Automata", *Journal of Computer Research and Development*, (2011), vol. 48, no. 7, pp. 1288-1297.
- [2] L.-X. Xiao, "Research on The Optimization of Enrollment Data Resources Based on Cloud Computing Platform", *International Information and Engineering Technology Association*, vol. 2, no. 2, (2015), pp. 9-12.
- [3] L. Wang, "An Ungreedy Chinese Deterministic Dependency Parser Considering Long Distance Dependency", *International Information and Engineering Technology Association*, vol. 2, no. 1, (2014), pp. 1-4.
- [4] Bratman M. *Intentions, Plans, and Practical Reason*, Massachusetts: Harvard University Press, (1987).
- [5] C. Phillips and L. Swiler, "A Graph-based System for Network Vulnerability Analysis", *Proceedings of the New Security Paradigms Workshop*, (1998); Charlottesville, VA, USA.
- [6] O. Sheyner, J. Haines and S. Jha, "Automated generation and analysis of attack graphs", *Proc. IEEE Symposium on Security and Privacy*, (2002); Oakland, California, USA.
- [7] X. Ou, S. Govindavajhala and A. W. Appel, "MulVAL: A logic-based network security analyzer", *Proceedings of the 14th Usenix Security Symposium*, ACM, (2005); New York.
- [8] R. Lippmann, K. Ingols and C. Scott, "Validating and Restoring Defense in Depth Using Attack Graphs", *Proc. the Military communications Conference*, IEEE Press, (2006); Washington, DC, USA.
- [9] S. Roschke, F. Cheng and C. Meinel, "A new alert correlation algorithm based on attack graph", *Proceedings of the 4th International Conference on Computational Intelligence in Security for Information Systems (CISIS)*, (2011).
- [10] Y. Yun, X.-S. Xu, J. Yan, Z.-C. Qi, "An Attack Graph-Based Probabilistic Computing Approach of Network Security", *Chinese Journal of Computers*, vol. 33, no. 10, (2010), pp. 1987-1996.
- [11] X. J. Chen, B.-X. Fang, Q. F. Tan and H.-L. Zhang, "Inferring Attack Intent of Malicious Insider Based on Probabilistic Attack Graph Model", *Chinese Journal of Computers*, vol. 37, no. 1, (2014), pp. 62-72.
- [12] Z.-Y. Luo, G.-L. Sun, J.-H. Liu and W.-B. Wang, "Application of Attack Graphs Algorithms in Invasion Prevention System", *Journal of Yunnan University (Natural Sciences Edition)*, vol. 34, no. 3, (2012), pp. 271-275.

Authors



Wang Guangze, He received a bachelor's degree in Mechanical Engineering (1987) from the University. Now he is an associate researcher at the university library. His current research interests include different aspects of computer network technology and network security.

