Research on Network Security Quantitative Model Based on Probabilistic Attack Graph

Yimin Cui [1,*], *Junmei* Li[1], *Wei* Zhao[1], and *Cheng* Luan [1]

[1] Beijing Institute of System Engineering, Anxiangbeilu 10# Chaoyang District, Beijing, China

Abstract. In order to identify the threat of computer network security and evaluate its fragility comprehensively, the related factors of network security are studied, and the methods based on attack graph are improved. Based on the attribute attack graph, the probabilistic attack graph model is generated by adding various factors which affect network security. The model uses security equipment performance data, common vulnerability scoring system data and etc. to calculate priori probability, finally obtains the network security index, and carries on the exploratory analysis. The experimental results show that the model is feasible and effective. Compared with other vulnerability assessment methods, the model has the characteristics of comprehensive evaluation and concise calculation.

1 Introduction

Network attacks persist and rise, which pose a great threat to the Internet and various business networks. In order to protect the network effectively, it is necessary to analyse the network security comprehensively and realize the threat from attack and defence. In order to achieve this goal, the model based on probabilistic attack graph is improved.

There are many researches on security evaluation methods, which have been developed based on attack graph [1-9], but none of them consider some uncertainties in the network. Network attack is a complex behaviour, and the result is not convincing because the network security is quantitatively analysed based on the combination of attack. The main contributions of this paper are as follows: (1) The factors that affect the attack effect are added in the probabilistic attack graph, (2) The new conversion method of attribute attack graph to probabilistic attack graph is given.

The 2nd section of this paper introduces the relevant research work, and the 3rd section gives the definition of the probabilistic attack graph model, then introduces the construction method of probabilistic attack graph, the probability calculation formula and the network security quantification method, the 4th section shows the calculation process and the validity of the probabilistic attack graph model by experiment. Finally, it summarizes the work of this paper and forecasts the future research.

2 Related work

The related research work started from the research of attack graph, Sheyner et al. [1] gave the method of automatic generation of state attack graph. Wang et al. [3] proposed a network security measurement method based on attribute attack graph. Chen Sisi[10] proposed a quantitative evaluation method for the vulnerability assessment based on the calculation of Bayesian network. Chen Feng et al. [5], Ye Yun et al. [4] used the maximum probability to solve the problem of the attribute attack graph including the loop. Jia Wei et al. [6] quantitatively evaluated the cost of an attacker using a vulnerability attack and carried out the analysis on a minimal attack cost path.

This paper is similar to the research work in the literature [8, 9]. Fang Yan et al. [8] proposed to avoid the loop-containing algorithm removes some of the state nodes, which would cause the loss of some information for security assessment. In addition, there is a mixed relation between the nodes in the graph, which leads to an unclear description of the graphs. The Bayesian attribute attack graph model established by Wang Xiujuan et al. [9] didn't consider the factors that affected the attack effect in the network, so there were some defects.

3 Probabilistic Attack Graph Model

The attack graph is mainly divided into two categories: a state attack graph [1] and an attribute attack graph [3]. In recent years, scholars have tended to use attribute attack graphs. It is a directed graph that contains two types of nodes, namely, attribute nodes and atomic attack nodes. Atomic attacks occur when the conditional attributes of an atomic attack are fully met, and an atomic attack succeeds, and the resulting attribute is used to represent the new attack condition that the attacker has acquired. This paper uses attribute attack graph to generate probabilistic attack graph.

3.1 Definition of probabilistic attack graph

[*] Corresponding author: tsui-min@163.com

The probabilistic attack graph is a directed loop-free graph, which has causality and probability semantics; the state and occurrence probability of the node in the graph are only related to its parent node. A probabilistic attack graph is defined as a directed loop-free graph $AG = (N, E, P)$.

N represents a collection of nodes, $N = S \cup A \cup I \cup C \cup D$, S represents a set of attribute nodes, A represents an atomic attack node set, I represents a collection of scan nodes, C represents an operation control node set, D represents a set of security protection nodes, and all nodes are valued as 1 or 0.

P is a probability set. $\forall s_j \in S, P(s_j)$ represents the probability that a attribute condition s_i is satisfied. $\forall a_j \in A, P(a_j)$ represents the probability of success of an atomic attack. $\forall i_{a_j} \in I, P(i_{a_j})$ represents the probability that information about the target network obtained by an attacker through a tool, such as a probe scan can help an attacker to successfully implement an atomic attack a_j. $\forall c_{a_j} \in C, P(c_{a_j})$ represents the probability that an attacker would ensure the success of an atomic attack through control or other tools. $\forall d_{a_j} \in D, P(d_{a_j})$ represents the probability that an atomic attack a_j would succeed if the protection in the target network fails.

E is a set of directed edges that represent causal relationships between various nodes. E can be expressed as $E = E_s \cup E_i \cup E_c \cup E_a \cup E_d$, where $E_s \subseteq S \times A$, $E_I \subseteq I \times A$, $E_c \subseteq C \times A$, $E_a \subseteq A \times S$, $E_d \subseteq D \times A$, where in a probabilistic attack graph, there are ‘AND’ and ‘OR’ relations between edges pointing to the same node (see Fig. 1), which is defined as follows:

(1) The edges pointing to the attack node are in the relation of "AND". $\forall e_m, e_n \in \{E_s, E_i, E_c, E_d\} \wedge \mathrm{end}(e_m) = \mathrm{end}(e_n) = a_j$, and the relation between e_m and e_n is "AND";

(2) The edges pointing to the attribute nodes are in the relation of "OR", $\forall e_m, e_n \in E_a \wedge \mathrm{end}(e_m) = \mathrm{end}(e_n) = a_j$ the relation between e_m and e_n is "OR".

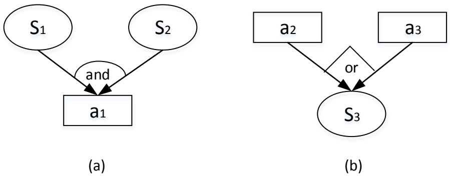

Fig. 1 Basic dependence relation among nodes in the probabilistic attack graph

3.2 Generation of a probabilistic attack graph

To generate a probabilistic attack graph, the loop path in the attribute attack graph should be eliminated at first. The loop in the attack graph can occur in real-world network attacks, such as the repeated infection of a virus to a machine. But in the study of the attack graph, most of the scholars used the monotonicity hypothesis [2], that is, the attacker's income is monotonically increasing, the attacker will not repeatedly attack a machine. There are many methods to eliminate the attack graph loop, the literature [4] deletes some nodes of the loop, and literature [9] deletes the most difficult atomic attack in the loop. This paper holds that: (1) The least secure path should be retained, and (2) all states that can be reached by attackers should be kept as far as possible. The algorithm of generating attribute attack graph and the algorithm of eliminating loop are not the focus of this paper, which is not explained in detail here.

After deleting the loop, additional nodes need to be added according to the security assessment requirement. First, find out the uncertain factors related to the attack effect, according to the relationship between the location that they deploy and the attack path, judge the atomic attack that it may affect, then add the protection node, the information node, the control node and so on in the graph. Second, there are a number of attacks which don’t exploit software vulnerabilities, such as guessing password attacks, traffic-based Denial-of-service attacks. Such attacks should be added to the graph as atomic attacks. After the above treatment, the basic structure of the probabilistic attack graph is determined.

According to the previous definition, there are only two relationships between edges pointing to a node in a probabilistic attack graph: "AND" and "OR". In accordance with the Bayesian theory, $P(a_i) = P(a_i | P_a(a_i))$ or $P(s_i) = P(s_i | P_a(s_i))$ can be obtained, where s_i is the middle node or leaf node, $P_a(a_i)$ and $P_a(s_i)$ are the parent node sets of a_i and s_i respectively; thus, $P(a_i)$ and $P(s_i)$ are defined as

$$\begin{aligned} P(a_i) &= \prod\nolimits_{x_i \in P_a(a_i)} P(x_i) \\ P(s_i) &= 1 - \prod\nolimits_{y_i \in P_a(s_i)} \big(1 - P(y_i)\big) \end{aligned} \qquad (1)$$

The following is a discussion of the assignment method of the root nodes in a probabilistic attack graph. It is assumed that S_o is a collection of initial attribute nodes, S_m is a collection of intermediate nodes, and S_f is a collection of endpoint states. The initial attribute state is an initial condition for an attacker, so $\forall s_j \in S_o, P(S_j) = 1$. In this paper, we use the intrusion path AV, identity authentication AU and attack complexity AC, which are provided by CVSS, as the basis for calculating $P(c_{a_j})$. In CVSS, the availability metric for a vulnerable point is defined as $E = 20VCU (0 \leq E \leq 10)$. The smaller the value of E, the greater the difficulty of an atomic attack, the more difficult it is to manipulate control. So in general, the following Formula (2) is used to assign the value to $P(c_{a_j})$

$$P(c_{a_j}) = 2 \times AV_{a_j} \times AC_{a_j} \times AU_{a_j} \qquad (2)$$

Assign a value to $P(i_{a_j})$ by using a popular scanning tool (such as NMAP) accuracy index. The computation of $P(d_{a_j})$ is a little complicated, and the probability $P_{detecting\&blocking-up}$ is used to express the influence of network protection equipment and security mechanism on atomic attack.

$$P\left(d_{a_j}\right) = 1 - P_{detecting\&blocking-up}(a_j) \qquad (3)$$

3.3 Network security exploration and analysis

Using probabilistic attack graph to evaluate security, the contents include: (1) Evaluate network security after obtaining some evidences, (2) Do exploring analysing on network security. The literatures [8,9] [11] carried on the researches to (1), and this paper mainly aims at the realization of (2).

The basis of exploratory analysis is the quantification of network security. The quantified value is called the Network Security Index (NSI), which is defined as the maximum gain that an attacker can gain in a network, that is, the weighted sum of all the available states of an attacker or of all atomic attacks. The weight values depend on the value of the network assets, and the NSI is calculated as follows:

$$NSI = \sum_{s_i \in S} P(s_i)\, \omega_i \qquad (4)$$

Or to calculate an atomic attack:

$$NSI' = \sum_{a_i \in A} P(a_i)\, \omega_i \qquad (5)$$

In the Formula (5), ω_i is the weight value of the asset.

The basic process of exploratory analysis is: first, generate exploration cases according to the threat of the network and the possible security plans, and then determine the impact of different security schemes on potential attacks; then calculate the NSI (or NSI') value of each case and do comparative analysis; finally, determine the optimal security improvement program to do actual deployment.

4 Experimental verification

Literature[11] is referred to establish the experimental environment and to detect the feasibility and effectiveness of the model. The network contains 3 subnets: An Internet zone, an isolated zone, and a trusted zone. The quarantine includes a Web server, a mail server, and a DNS server. The trust zone includes database server, FTP server, and gateway server.

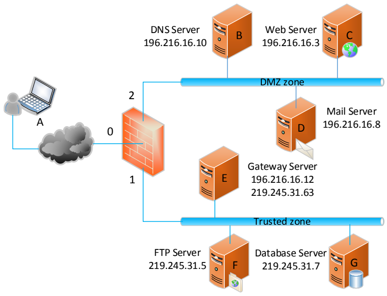

Fig. 2. Experimental Network topology

Use the tool (Mulval) to generate an attribute attack graph, and then create a probabilistic attack graph (see Fig. 3).

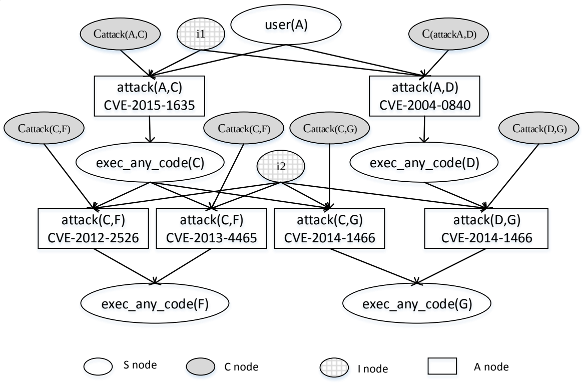

Fig. 3. Probabilistic attack graph

The C-node value is a priori probability calculated according to the Formula (2) as follows:

Table 1. C-node values associated with each atomic attack

Atomic attack	AV	AC	AU	C-node value
Attack(A,C) CVE-2015-1635	1.0	0.71	0.704	1.0
Attack(A,D) CVE-2004-0840	1.0	0.71	0.704	1.0
Attack(C,F) CVE-2012-2526	1.0	0.61	0.704	0.86
Attack(C,F) CVE-2013-4465	1.0	0.35	0.56	0.39
Attack(C,G) CVE-2014-1466	1.0	0.71	0.704	1.0
Attack(D,G) CVE-2014-1466	1.0	0.71	0.704	1.0

The attack network of an attacker typically requires two probes, so there are two I nodes, and this paper assumes that each I node has the same effect on the corresponding atomic attack. The value of the Node I depends on the accuracy and completeness to obtain the vulnerability information and related information. Generally, $i_1 < i_2$. It is more difficult to get information from outside the network than to get it inside the network.

After the probabilistic attack graph model is established, the security of the network can be evaluated under different circumstances, and the security of the network can be evaluated after obtaining some evidence. This paper quantifies the changes of network security before and after the detection and scanning measures of each host in the network. It is supposed $i_1 = 0.85$, $i_2 = 0.95$ before taking measures. After taking a series of protective measures, the value of the I node is greatly reduced, $i_1 = 0.45$, $i_2 = 0.90$. Because the attackers have dominion over their own machines, thus $\mathrm{P}(\mathrm{user(A)}) = 1$.

Table 2. The unconditional probability of each atom attack

No.	Unconditional probability	Before	After
1	Attack(A,C) CVE-2015-1635	0.85	0.45
2	Attack(A,D) CVE-2004-0840	0.85	0.45
3	Attack(C,F) CVE-2012-2526	0.69	0.35
4	Attack(C,F) CVE-2013-4465	0.31	0.16
5	Attack(C,G) CVE-2014-1466	0.81	0.41
6	Attack(D,G) CVE-2014-1466	0.81	0.41
NSI'		4.32	2.23

A probabilistic attack graph can be used to quantitatively evaluate the probability of the computer network being attacked, and can be applied to do security analysis in various situations. The implementation of security measures and the enhancement of attack skills will result in the change of the value of I, C and D, which results in the change of the whole network security value, so we can explore all kinds of schemes and find the best safety decision by analyzing the size of various effects quantitatively.

5 Conclusions

Security evaluation has always been a hot topic in the field of network research, and as an effective method of analysis, and the scholars have been developing new research on attack graphs. In order to analyse some uncertain factors in the process of network security improvement, this paper quantifies the network security index by probabilistic attack graph. The former scholars' models have been improved, mainly in the structure and the node type. First, the structure is more concise, directly conversed from the attribute attack graph, and only "AND" and "OR" relations are allowed between edges. Then new type nodes are added according to the need of exploratory analysis. Feasibility and effectiveness are proved through the experiment. The next step will further optimize the probabilistic attack graph generation, automatically joining the new type node, improving the description ability of the probabilistic attack graph, to improve its application efficiency in the large-scale network.

References

1. Sheyner O, Haines J, Jha S, et al . Automated generation and analysis of attack graphs [C]// Proceedings 2002 IEEE Symposium on Security and Privacy, 2002:273-284.
2. AMMANN P, WIJESEKERA D, KAUSHIKS. Scalable, graph based network vulnerability analysis [C]//Proceedings of the 9th ACM Conference on Computer and Communications Security. New York: ACM Press, 2002:217-224.
3. Wang Lingyu, Yao Chao, Singhal A, et al. Interactive analysis of attack graphs using relational queries [M]. Data and Applications Security XX: Springer Berlin Heidelberg, 2006:119-132.
4. Ye Yun, Xu Xishan. An attack graph based probabilistic computing approach of network security [J]. Chinese Journal of Computers, 2010(10):1987-1996.
5. Chen Feng, Zhang Yi. Research of quantitative vulnerability assessment based on attack graphs [J]. Computer Engineering and Science, 2010, 32(10):8-11.
6. Jia Wei. The research on computer network vulnerabilities assessment methods [D]. Hefei: University of Science and Technology of China, 2012.
7. Poolsappasit N, Dewri R, Ray I. Dynamic security risk management using Bayesian attack graphs [J]. Dependable and Secure Computing, IEEE Transactions on, 2012, 9(1):61-74.
8. Fang Yan, Yin Xiaochuan, Li Jingzhi. Research of quantitative network security assessment based on Bayesian-attack graphs [J]. Application Research of Computers, 2013, 30(9):2763-2766.
9. Wang Xiujuan, Sun Bo, Liao Yanwen, Xiang Congb in, Computer Network Vulnerability Assessment Based on Bayesian Attribute Network, Beijing University of Technology, Beijing 100124, China, 2015, 38(4):106-112
10. Chen Sisi, Lian Yifeng, Jia Wei. A network vulnerability evaluation method based on Bayesian networks [J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2008, 25(5): 639-648.
11. GAO Ni, GAO Ling, HE Yiyue, et al. Optimal security hardening measures selection model based on Bayesian attack graph. Computer Engineering and Applications, 2016, 52(11):125-130.