Real-Time Risk Assessment of Network Security Based on Attack Graphs

Xiaochuan Yin[1,a], Yan Fang[1,b] and Yibo Liu[1,c]

[1] Information and Navigation College of Air Force Engineering University, Xi-an 710077, China

[a]yinxiaochuan21@sina.com, [b]fangyan1989217@163.com, [c]liuyibo0123456789@163.com

Keywords: real-time risk; security situation; attack graphs; Bayesian theory; attack evidence

Abstract: Facing hackers' intelligent attacks and multi-source information from various security equipments, evaluating real-time risk of the network becomes more and more complicated to handle. This paper proposes a new attack graphs model(NAG)-based analysis method in order to assess the impact on the network system made by multiple vulnerabilities. Aiming at simplify the attack graphs, we combined attack graphs with Bayesian theory and put forward an optimized algorithm to remove the cycles in attack graphs. By importing Common Vulnerability Scoring System (CVSS) and attack evidence, the assessment method in this paper can dynamically evaluate the partial or entire network security. Experimental results show that the method can reflect the changing situation of the network security.

Introduction

In face of sophisticated and multi-step network attacks, vulnerabilities bring a lot risk to the network systems. Though the scanners such as Nessus [1] can find individual vulnerabilities, the attackers usually combine multiple vulnerabilities to penetrate networks with devastating impact [2] which we should take into account when evaluating the security of the network. As a model of risk assessment, attack graphs can find out all possible actions of an attacker who may use the relationship of the vulnerabilities and then analysis the threats the network faced.

There exists many approaches for evaluating network security based on attack graphs. Marcel et al. [3] propose a Dynamic Bayesian Networks-based model which provides a theoretical foundation and a practical framework for continuously measuring network security. But as time goes on, there will be an enormous space to maintain which goes aginst to realtime assessment. Nayot et al. [4] combined Bayesian theory and attack graphs in risk assessment, however they ignored the cycles that exist in attack graphs. Yun Ye et al. [5] evaluated network security based on probability of nodes in attack graphs but they didn't give a measure of evaluating the impact on network system brought by attacks. Xi Zhang et al. [6] computed the risk score using the CVSS [7] in attack graphs, however they didn't concern the dynamic changing of the network.

In this paper, by combining Bayesian theory, we proposed a new attack graphs model (NAG) based on which we designed an effective assessment method by importing CVSS and attack evidence. The rest of this paper is organized as follows. We firstly define the new model of attack graphs, secondly give the method of risk assessment based on attack graphs, then present experimental evaluation and in the end conclude the paper.

Attack Graphs model

The new attack graphs model is a structure of using five elements group to describe information. Its structure: NAG=$< S, A, E, P >$.

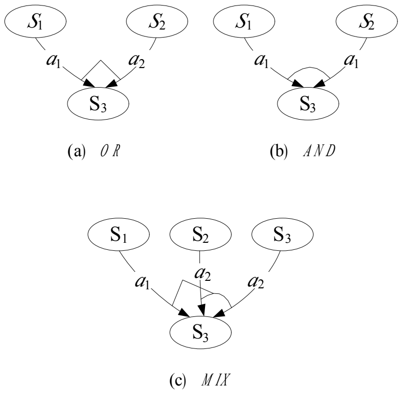
Among the model: S is the set of attributes; A is the set of atomic attacks and also the edges in NAG; E is the set of relationships; P is the set of probabilities. And the model should abide the followings:

(1). $A \in S \times S$. $\forall a_m \in A$, $a_m = pre(a_m) \rightarrow post(a_m)$, $pre(a_m)$ is the source attribute of a_m and $post(a_m)$ is the destination attribute of a_m.

(2). $S=So\cup Sin\cup Sf$. $\forall S_i\in S$, S_i has two property value: $S_i=0$ or $S_i=1$. $\forall S_i\in So$, $\exists\, a_m$ where $a_m\in A$ and $S_i=post(a_m)$; $\forall S_i\in Sin$, $\exists a_j$, $a_k\in A$ where $S_i=post(a_j)=pre(a_k)$; $\forall S_i\in Sf$, $\exists\, a_m$ where $a_m\in A$ and $S_i=pre(a_m)$.

(3). $\forall S_i\in S$, $P(S_i)$ denotes the probability of $S_i=1$; $\forall a_m\in A$, $P(a_m)$ denotes the probability of $pre(a_m)\rightarrow post(a_m)$.

(4). $\forall S_i\in Sin\cup Sf$, $\exists e_i\in E$ correspond to S_i and $e_i\in\{AND, OR, MIX\}$. In traditional attack graphs, the parents of atomic attacks has the relationship of *AND*. Similarly, the parents of attributes has the relationship of *OR*. Since this paper combined the edges with the atomic attacks, there are three relationships in NAG: *AND*, *OR*, *MIX*, as shown in Fig. 1 and the relationship is denoted by the set of *E*.

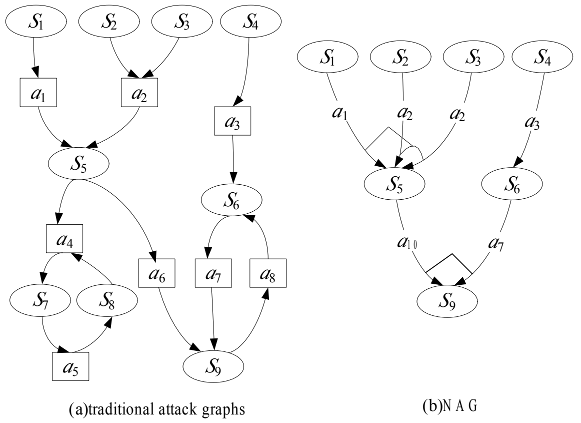


(a) *OR*

(b) *AND*

(c) *MIX*

Fig. 1 three types of relationship



(a)traditional attack graphs

(b)NAG

Fig. 2 comparison of attack graph and NAG

One difficulty in combining Bayesian theory and attack graphs lies in the cycles in attack graphs which deeply affect the probability computing of the nodes. There are mainly two types of cycles. As shown in Fig. 2(a), the ovals stand for attribute nodes and rectangles stand for atomic attack nodes. In the attack path $a_4\rightarrow S_7\rightarrow a_5\rightarrow S_8\rightarrow a_4$, the execution of a_4 requires both S_5 and S_8 to be satisfied. However, S_8 can be only satisfied by the execution of a_5, which requires a_4 to be executed first. So this type of cycles cannot exist in the real network. The other type of cycles is like the attack path $S_6\rightarrow a_7\rightarrow S_9\rightarrow a_8\rightarrow S_6$ which may be found in network, however, based on the study of paper [8], attackers won't lose its ability, so they won't try to obtain the attributes they already got.

In order to remove the two types of cycles, this paper proposed an algorithm to optimize attack graphs. The algorithm was shown as follows:

```
Input:    $A$—set of atomic attack nodes in original attack graph
          $S$—set of attribute nodes in original attack graph
Output：$A'$—set of atomic attack nodes in optimized attack graph
          S'—set of attribute nodes in optimized attack graph
Algorithm:
$InitQueue(Q)$；
Foreach $S_i\in So$ do   //push all initial attribute nodes in $Q$ and $S'$
    $EnQueue(Q, S_i)$；
    $Add\_to(S', S_i)$；
Foreach $a_i\in A$ do
    Foreach $S_j\in pre(a_i)$ do
        $a_i.pre[S_j]=0$;   //reset the atomic attacks
While $EmptyQueue(Q)=0$ do
    $q=DeQueue(Q)$;
    Foreach $a_i\in post(q)$ do
        $a_i.pre[q]=1$;   //set the atomic attack's flag when satisfied
        if $\forall q_j\in pre(a_i)$, $a_i.pre[q_j]=1$ then
            Foreach $S_k\in post(a_i)$ do
                If $S_k\notin Ancestor(pre(a_i))\cup pre(a_i)$ then
                    If $S_k\notin S'$ then    //push the attribute nodes in $Q$ and $S'$ when satisfied
            $EnQueue(Q, S_k)$;
            $Add\_to(S', S_k)$;
        If $e_i \notin E'$ then
            $Add\_to(A', a_i)$;   // push the atomic nodes in $A'$ when satisfied
```

The converting from traditional attack graphs to NAG in Fig.2 shows that the nodes in NAG are apparently fewer than that in the traditional one which helps to analysis network security easier.

Risk assessment using NAG

In NAG, $\forall S_i \in So$， S_i is the initial attribute which attackers possess and currently $P(S_i)=1$; $\forall S_i \in Sin \cup Sf$, according to Bayesian Theory：

$$P(S_i)=P(S_i \mid Pa(S_i), Pa(S_i) \rightarrow S_i) \tag{1}$$

The probability of the attribute nodes associates with its parents and the atomic attacks which come from its parents to it. Considering the relationship of its parents, we define $P(S_i)$ as Eq. 2.

$$P(S_i)=\begin{cases} \prod\limits_{S_j \in Pa(S_i)} P(S_j)\cdot P(Pa(S_i) \rightarrow S_i), & e_i = AND \\ 1-\prod\limits_{S_j \in Pa(S_i)} (1-P(S_j)\cdot P(S_j \rightarrow S_i)), & e_i = OR \\ 1-\prod\limits_{k=1}^{K}(1-\prod\limits_{S_j \in Pa(S_i)_k} P(S_j)\cdot P(Pa(S_i) \rightarrow S_i)), & e_i = MIX \end{cases} \tag{2}$$

When evaluating the network security, its complicated and dynamic characteristics cannot be ignored. The probability can change due to some attack incidents which may be observed by the network security equipments. Thus, this paper introduces the concept of attack evidence(AE) that express the attackers' exploring incidents or the attributes which the attackers already got. Assume that $AE=\{S_1', S_2' \ldots S_n'\}$, $\forall S_i' \in AE$, $S_i'=1$; $\forall S_i \in S\text{-}AE$, the Bayesian posterior probability equation can be used to update the probability:

$$P(S_i|AE)=P(AE|S_i)\times P(S_i)/P(AE) \tag{3}$$

Based on the factors of network risk assessment: asset, vulnerability and threat, this paper defines the risk as: $Risk=f(V, M, P)$. V denotes the value of the asset; M denotes the influence brought by vulnerabilities; P denotes the property of the threat's exploring.

The value of the assets should be educed by the level of confidentiality, integrity and availability. This paper proposed five levels: very low, low, medium, high, very high that was replaced by 1, 2, 3, 4 and 5. The level of confidentiality, integrity and availability was confirmed by experts.

For the complexity and diversity of the vulnerabilities and threats, we use the CVSS to define the two factors. The CVSS is an open vulnerability scoring framework which was put forward by the NIAC. The CVSS is composed of three metric groups: Base, Temporal, and Environmental, each consisting of a set of metrics [9]. In this paper the Temporal and Environmental group are ignored for we are interested in defining and communicating the fundamental characteristics of vulnerability which is the purpose of the Base group. The Base group contains six metrics: access vector (AV), access complexity (AC), authentication (AU), confidentiality impact (CI), integrity impact (II) and availability impact (AI). Their scoring standard is listed in Table 1.

Table 1 Base Group Scoring Evaluation

metric	level 1	level 2	level 3
AV	Local (0.395)	Adjacent Network (0.646)	Network (1.0)
AU	Multiple (0.45)	Single (0.56)	None (0.704)
AC	High (0.35)	Medium (0.61)	Low (0.71)
$CI/II/AI$	None (0)	Partial (0.275)	Complete (0.66)

As shown in Table 1, the Access Vector, Access Complexity, and Authentication metrics capture how the vulnerability is accessed and whether or not extra conditions are required to exploit it. In this study, we use the three metrics to define $P(a)$ in NAG:

$$P(a)=AV\times AU\times AC \tag{4}$$

$P(a)$ only reflects the probability of the atomic attack itself, however in NAG the atomic attack also rely on the attribute nodes' probability($P(S)$). Considering the relationship of AND, the success probability of atomic attack was defined as $P'(a)$:

$$P'(a)=\begin{cases} P(a)\cdot\prod\limits_{S_i\in pre(a)} P(S_i), & \exists S_i\notin So \\ P(post(a)), & pre(a)\subset So \end{cases} \tag{5}$$

The three impact metrics of confidentiality impact, integrity impact and availability impact measure how a vulnerability will directly affect the asset if exploited. This paper combined the three metrics with the score of assets to define the function ($In(a)$) that denotes the impact brought by atomic attack:

$$In(a)=\log_2(\alpha\cdot 2^{CI\cdot C}+\beta\cdot 2^{II\cdot I}+\gamma\cdot 2^{AI\cdot A})/3 \tag{6}$$

C、I、A are the score of assets' confidentiality, integrity and availability that will be defined by experts. α、β、γ are the weights of confidentiality, integrity and availability, and $\alpha+\beta+\gamma=3$。

The risk score of the atomic attack is defined as Eq. 7:

$$Risk(a)=P'(a)\times In(a) \tag{7}$$

In order to evaluate the entire network security, the risk score of the network is expressed as the sum of all the atomic attacks' risk scores in NAG:

$$Risk_{all}=\sum_{a\in A}(P'(a)\cdot In(a)) \tag{8}$$

Case study

An experiment was done to test the assessment method proposed in this paper. The experimental circumstance is shown in Fig. 3. The topology in the experiment is composed of attack and target network.

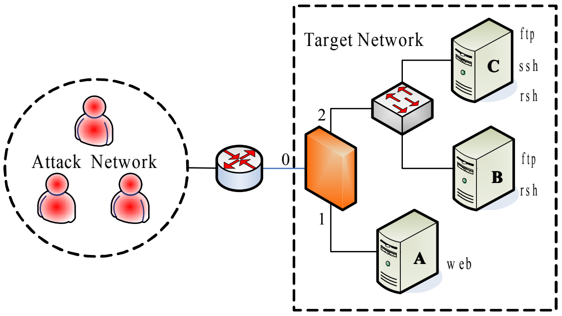


Fig. 3 topology of the testing network

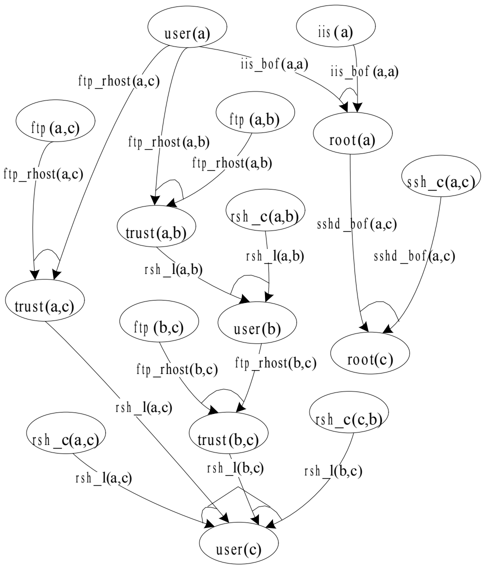


Fig. 4 NAG of the testing network

Machine A is a Web server; Machine B is an database server that offers ftp and rsh services; Machine C is a file server that offers ftp, ssh, and rsh services. The available services are presented by Table 2 which is controlled by the firewall.

Table 2 available services

traffic direction	available ports	available services
0→1	80	web
2→1	80	web
1→2	21, 22, 514	ftp, rsh, ssh

In the experiment, we first collected the vulnerability in the experiment network which can be obtained by using the scanner software Nessus, then marked them with CVE[10] number through which the CVSS scores can be checked. The experts scored the value of victim services by confidentiality, integrity and availability , and in this experiment we assumed that α=1.1、β=1.0、γ=0.9. Then $P(a)$ and $In(a)$ can be computed by Eq. 4 and 6. The summary of vulnerability is presented as Table 3.

Table 3 summary of vulnerability

vulnerability	CVE number	victim host	service	*C*	*I*	*A*
IIS buffer overflow	CVE-2009-1012	A	web	2	3	3
Rsh login	CVE-2006-0408	B	rsh	3	3	4
Ftp rhost overflow	CVE-2011-4800	B	ftp	4	3	4
Ssh buffer overflow	CVE-1999-1455	C	ssh	4	4	3
Rsh login	CVE-2006-0408	C	rsh	4	4	5
Ftp rhost overflow	CVE-2011-4800	C	ftp	3	4	4

Based on the topology, the firewall's security device and vulnerabilities of the network, the NAG of the experiment was drawn in Fig. 4.

Based on the relationships of vulnerabilities shown in NAG, we can get $P(S_i)$ and $P'(a)$ without attack evidence by Eq. 1、2 and 5. Then the attack network started to exploit the target network, and later the target network got some attack evidence which was shown in Table 4.

Table 4 attack evidence

time	victim host	*AE*
1	A	root(a)=1
2	B	user(b)=1
3	C	user(c)=1

The probability of the attribute nodes in NAG then can be updated by Eq. 3 and 5 with attack evidence and then the risk score can be computed in NAG. The assessment results were shown in Table 5.

Table 5 risk scores of the target network

atomic attack	t(0)	t(1)	t(2)	t(3)
iis_bof(a,a)	0.886	1.771	1.771	1.771
rsh_c(a,b)	0.272	0.272	2.115	2.115
ftp_rhost(a,b)	0.975	0.975	2.452	2.452
rsh_c(b,c)	0.040	0.040	0.312	0.532
ftp_rhost(b,c)	0.159	0.159	1.237	2.110
sshd_bof(a,c)	0.256	0.512	0.512	0.512
rsh_c(a,c)	0.312	0.312	0.312	0.532
ftp_rhost(a,c)	1.237	1.237	1.237	2.110
network	4.137	5.278	9.948	12.134

As a web server, machine A's 80 port can be visited by attack and target network. Its asset is not very valuable, so when the evidence shows that the attackers get the root, the risk score doesn't rise quickly. As machine B and C were not open to attack network, when finding the evidence that they got attacked, it shows that the attackers use the relationships of the vulnerabilities in NAG. For the high value of their assets, the risk scores highly rise.

Experimental results which are consisted in the theory of the assessment method show that it not only can evaluate the threat of the network but also dynamically reflect the changing of security situation.

Conclusion

The wide application network technology bring about more and more destructive attacks. To better protect the network, defense method faces the transition from passive defense to active defense. As an important part of the active defense system, real-time risk assessment can analysis the tendency of the network security situation which benefits early warning and strategy implementation. However, previous researches paid more attention to static risk measures which cannot satisfy the requirement of active defense system.

In this paper, a new type of attack graphs model was proposed by combing Bayesian theory, based on which we put forward a risk assessment method by importing CVSS and attack evidence. Through a case in testing network, we proved that unlike previous approaches on attack graphs, the method can dynamically reflect the network security situation which helps to provide active defense with decision-making information. In future work, we focus on the research of finding minimum-cost network hardening measure.

References

[1] R. Deraison. Nessus Vulnerability Scanner. http://www.nessus.org/. 2013

[2] S. Noel, S. Jajodia, L. Wang et al. Measuring Security Risk of Networks Using Attack Graphs. International Journal of Next-Generation Computing, 2010, 1(1):135-147

[3] M. Frigault, L. Wang. Measuring Network Security Using Dynamic Bayesian Network. Conference on Computer and Communications Security Proceedings of the 4th ACM Workshop on Quality of Protection. New York: ACM, 2008, pp.23-30

[4] N. Poolsappasit, R. Dewri, I. Ray. Dynamic Security Risk Management Using Bayesian Attack Graphs. IEEE Transactions on Dependable and Secure Computing, 2012, 9(1): 61-74

[5] Y. Ye, X. Xu, Y. Jia et al. An Attack Graph-Based Probabilistic Computing Approach of Network Security. Chinese Journal of Computers，2010，33(10):1987-1996

[6] X. Zhang, S. Huang, Y. Xia et al. Attack graph-based method for vulnerability risk evaluation. Application Research of Computers. 2010, 27(1):278-280

[7] Common vulnerability scoring system. http://www.first.org/cvss/, 2013.

[8] P. Ammann, D. Wijesekera, S. Kaushik. Scalable, graph-based network vulnerability analysis: Proc. of the 9th ACM Conf. on Computer and Communications Security, New York: ACM, 2002, pp. 217-224

[9] P. Mell, K. Scarfone, S. Romanosky. The Common Vulnerability Scoring System (CVSS) and Its Applicability to Federal Agency Systems, NIST IR 7435. USA: Department of Commerce, 2007

[10] CVE. Common Vulnerabilities and Exposures. http://www.cve.mitre.org/, 2013