doi:10.11835/j.issn.1000-582X.2017.12.005

信息系统脆弱性被利用概率计算方法

柴继文[1],王　胜[1],梁晖辉[1],胡　兵[2],向　宏[2]
(1.国网四川省电力公司电力科学研究院,成都 610072;
2.重庆大学 信息物理社会可信服务计算教育部重点实验室,重庆 400030)

摘　要:针对现有信息系统风险评估工作中对脆弱性的评估未考虑各脆弱性间的相关性,评估结果受到较多人为主观因素的影响,提出"被利用难易程度"和"被选择概率"两个指标将现有对脆弱性的"被利用难易程度"评价转换为更为科学的"被利用概率"评价,并用贝叶斯网络的正向推理计算脆弱性节点的累积"被选择概率"。通过理论和实验分析,与相关的研究成果相比,提出的脆弱性被利用概率计算方法更准确、合理。

关键词:风险评估;脆弱性;贝叶斯网络;被利用概率
中图分类号:TP309　　**文献标志码**:A　　**文章编号**:1000-582X(2017)12-035-08

A computing approach of information system vulnerability's exploited probability

CHAI Jiwen[1], *WANG Sheng*[1], *LIANG Huihui*[1], *HU Bing*[2], *XIANG Hong*[2]
(1. State Gid Sichuan Electric Power Research Institute, Chengdu 610072, P.R.China;
2. Key Laboratory of Dependable Service Computing in Cyber Physical Society, Ministry of Education, Chongqing University, Chongqing 400030, P.R.China)

Abstract: The evaluation results are impacted by many subjective factors since the existing risk assessment for information systems does not take the correlation of vulnerabilities into account. By combining two assessment vectors, i.e. access complexity and chosen probability, we transfer the so called "accessed complexity" evaluation method into an "exploited probability" evaluation approach, and use Bayesian networks' forward inference to accumulation each of vulnerability's chosen probability. Theoretical and experimental analysis show that the proposed "exploited probability" evaluation method is more accurate and reasonable than associated existing research work.

Keywords: risk assessment; vulnerability; Bayesian network; exploited probability

对信息系统脆弱性进行科学量化的评价是实施信息安全风险评估的一项重要基础性工作。现有风险评估标准[1-2]对脆弱性评价包括"对资产的损害程度"和"被利用难易程度"两个方面,其量化计算基本参照国际主流标准—通用脆弱性评分系统(CVSS)的计算方法[3]。但随着信息系统日趋复杂,CVSS 最让人诟病的是其计算

收稿日期:2017-07-14
基金项目:国网四川省电力公司科技项目(5219991351VR);国家自然科学基金资助项目(61472054)。
Supported by Science and Technology Program of State Grid Sichuan Electric Power Company(5219991351VR) and National Natural Science Foundation of China(61472054).
作者简介:柴继文(1963—),男,国网四川省电力公司电力科学研究院副总工程师,高级工程师。
向宏(联系人),男,重庆大学教授,博士生导师,(E-mail)xianghong@cqu.edu.cn。

过程未考虑脆弱性在信息系统中相互关联对脆弱性被选择利用的影响,使得对脆弱性的"被利用难易程度"评价不够准确,影响对"安全事件发生可能性"的评价和最终对信息系统的风险分析。

Phillips 等人[4]在 1998 年首次提出基于攻击图的网络安全分析方法,基于攻击图的分析方法可以发现不同脆弱性之间的关联关系,攻击图模型与 CVSS 相结合已成为评价脆弱性的有效途径之一。在面对基于攻击图的脆弱性被选择概率计算中各脆弱性节点之间的相关性问题时,目前有 3 种技术途径:一是假设攻击图中任意节点之间都相互独立,如文献[5-6]假设攻击图中任意节点相互独立,通过宽度优先搜索算法计算各节点的概率,但在攻击图中不同子节点可能具有相同的父节点,因此,任意节点的独立性假设会导致各个阶段概率计算的错误;二是通过计算各个节点的最大可达概率,回避各节点之间的相关性问题,如文献[7]提出了有效攻击路径分析技术,采取前向搜索方式和深度优先搜索策略寻找各个节点的有效攻击路径,文献[8]提出了适用于大规模网络的最大可达概率的概念和计算方法,通过删除攻击图中的不可达路径简化攻击图,解决了攻击图中循环路径导致的攻击图难以理解和概率重复计算问题,但没有解决相关性问题;三是结合攻击图模型和 CVSS 构建贝叶斯网络,利用贝叶斯网络正向推理解决各个节点的相关性问题,如文献[9-10]基于攻击图建立贝叶斯网络,利用贝叶斯条件概率公式计算各个节点的概率。

通过上述 3 种方法的对比,第 3 种方法所计算出的攻击图中各个节点的概率是相对准确的,但是由于基于攻击图的贝叶斯网络异常复杂,在贝叶斯网络中的精确推理是 NP 难题,文献[9-10]所建立的概率模型仅在理论上解决了概率的计算问题。

在现有研究工作的基础上,将对脆弱性的"被利用难易程度"评价转换为"被利用概率"这一更为科学的评价,包括"被利用难易程度"和"被选择概率"两个评价要素。通过使用攻击图模型和 CVSS 构建信息系统脆弱性的贝叶斯网络,并利用贝叶斯网络的正向推理计算网络中各脆弱性节点的累积"被选择概率",解决各脆弱性节点的相关性对最终风险评估结果的影响。

1 漏洞被利用概率计算

1.1 被利用概率计算模型

所提出的脆弱性被利用概率计算方法包括以下 5 个步骤:1)生成攻击图;2)根据 CVSS 评价体系计算各个脆弱性被利用难易程度;3)将步骤 2 的计算结果赋值于攻击图中相应的节点,构建贝叶斯网络;4)计算贝叶斯网络中各个节点的被选择概率;5)利用步骤 2 和步骤 4 的计算结果计算脆弱性的被利用概率。参与计算的各个要素进行定义如下:

定义 1 脆弱性被利用难易程度是指脆弱性被攻击者成功利用的可能性大小,是脆弱性的固有属性,记为:$E(v)$。

定义 2 脆弱性被选择概率是指攻击图中脆弱性节点被攻击者选择利用的可能性大小,由节点之间的关联性和节点被利用难易程度所确定,记为:$S(v)$。

定义 3 脆弱性被利用概率是指攻击者选择攻击图中脆弱性节点并成功利用的可能性大小,记为:$P(v)=E(v)\times S(v)$。

1.2 被利用难易程度计算

现有研究文献对脆弱性被利用难易程度的评价主要依据 CVSS 进行计算,或者选取 CVSS 中的"Access Complexity"指标进行赋值,如文献[8];或者直接以 CVSS 最终评价结果进行赋值,如文献[5,12]。但 CVSS 对脆弱性的评价包括基本度量(base metrics)、生命周期度量(temporal metrics)和环境度量(environmental metrics)3 个方面[3],各个方面的评价指标中既有对脆弱性执行难易程度的评价,又有对脆弱性执行效果的评价,因此,上述文献对计算不够准确和全面。

根据定义 1,从最新 CVSS 评价标准 3.0 版中提取与评价脆弱性利用难易程度相关的指标(参见表 1),利用公式(1)计算脆弱性被利用难易程度。

$$E(v)=AV\times AC\times PR\times UI\times E\times RC。\tag{1}$$

表 1　脆弱性难易程度评价指标

Table 1　The complexity evaluation for use vulnerability

度量因素	类型	赋值
攻击向量(*AV*)	本地	0.395
	邻近网络	0.646
	远程	1.0
攻击复杂度(*AC*)	高	0.35
	中	0.61
	低	0.71
所需权限(*PR*)	需要多重身份验证	0.45
	需要单重身份验证	0.56
	不需要	0.74
用户交互(*UI*)	需要	0.45
	不需要	1.0
利用可能性(*E*)	未证明	0.85
	理论证明	0.9
	实际可以	0.95
	高利用性	1.0
	未定义	1.0
报告确认程度(*RC*)	未确认	0.9
	未证实	0.95
	已确认	1.0
	未定义	1.0

1.3　被选择概率计算

为了避免贝叶斯网络中精确推理的 NP 难题，在文献[5,8,14]的研究基础上，对文献[5]的条件概率计算方法和文献[8]的攻击图节点概率计算算法进行改进，提出节点被选择概率算法。根据定义 2，在计算攻击图中各节点的被选择概率时，需要首先确定各节点间的相关性，借鉴文献[13]的研究成果将节点间的关系分为“直接、或、与、混合”4 种类型，如图 1 所示。

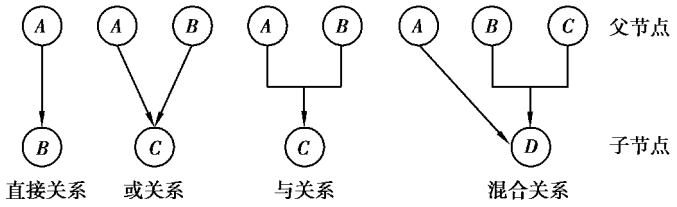

图 1　脆弱性间关系类型

Fig.1　The type of relationship between vulnerability

1)直接关系，子节点存在唯一一个父节点；

2)或关系，子节点存在多个父节点，且父节点间相互独立，任意父节点被成功利用，就可导致子节点被利用，直接关系为或关系的一种特殊形式；

3)与关系，子节点存在多个父节点，在所有父节点被成功利用的前提下，子节点才能被利用；

4)混合关系，子节点的多个父节点间同时存在或和与关系。

在考虑攻击图中各父节点的相互关系对子节点被选择的影响时，需要在计算各子节点独立被选择概率的基础上，根据父节点间的关系类型计算各子节点的关联被选择概率。下述定理阐述了如何计算攻击图中各节点被选择概率。

定理 1：对于攻击图中任意节点 $v_1 \in V$，$V=(v_1|v_2|\cdots|v_1|\cdots|v_n)$，$V$ 中的元素为同一父节点 T 的子节点，且相互独立，则 v_i 的独立被选择概率为

$$O(v_i)=\frac{E(v_i)}{\sum_{j=1}^{n}E(v_j)},\begin{pmatrix}1\leqslant i\leqslant n\\1\leqslant j\leqslant n\end{pmatrix}。$$

证明 1。若 $n=1$，则 v_i 是父节点 T 的唯一个子节点，在 T 被成功利用后，v_i 是执行下一步攻击的唯一选择，可知上述定理成立；若 $n>1$，因为 V 中元素相互独立，在 T 被成功利用后，v_i 被选择执行下一步攻击的概率取决于 v_i 自身被利用的难易程度在集合 V 中所占比重，由公式(1)可知 $E(v)$越高，被利用难度越低，被选择概率就越大。

定理 2。对于攻击图中任意节点 v，其独立被选择概率为 $c(v)$，父节点为 $T=(t_1|\cdots|t_i|t_{i+1}|\cdots|t_n)$，各父节点被选择概率为 $S(t)=\{S(t_1)|S(t_2)|\cdots|S(t_n)\}$，设 $t_1\rightarrow t_i$ 节点间为或关系，$t_{i+1}\rightarrow t_i$ 节点间为与关系，则节点 v 的被选择概率为

$$S(v)=O(v)\times\max\left\{S(t_1)\cdots S(t_i)\prod_{i+1}^{n}S(t)\,|\,t\in\tau\right\}。$$

证明 2。根据文献[8]的定义 1、定义 2 和定理 1 可知，若节点 v 的父节点间为或关系，则父节点对 v 的关联影响为所有父节点被选择概率中的最大值 $\max\{S(t)|c\in T\}$；若父节点间为与关系，则父节点对 v 的关联影响为 $\prod_{t\in T}S(t)$，因此，若节点 v 的 n 个父节点间既有或关系，又有与关系，则节点 v 的被选择概率为其独立被选择概率与所有父节点被选择概率中最大值的乘积。

2 对比实验分析

2.1 攻击图获取

为验证本文漏洞被利用概率计算方法的科学性，采用文献[5]中生成的攻击图，如图 2 所示，并利用文献[5,8]的算法和本文的算法进行对比分析。图中节点为攻击过程中可被选择利用的脆弱性，节点间的连线表示各节点的关联关系。各节点通过节点标识、节点被利用时源和目标主机、节点的 CVE 和 bugtraq 编号进行描述。

2.2 被利用难易程度计算

根据表 1 所确定的脆弱性难易程度评价指标，利用公式(1)计算 $E(v)$(参见表 2)，并将 $E(v)$值赋予攻击图中对应的节点，与文献[5,8]的计算结果对比参见表 3。

表 2 脆弱性被利用难易程度值

Table 2 The values for complexity to exploit vulnerability

Bugtraq	*AV*	*AC*	*PR*	*UI*	*E*	*RC*	$E(v)$
9 751	1.0	0.71	0.74	1.0	1.0	1.0	0.53
10 181	1.0	0.71	0.74	1.0	0.9	1.0	0.47
10 108	1.0	0.71	0.74	1.0	1.0	1.0	0.53
10 201	0.40	0.71	0.74	1.0	0.95	1.0	0.20
6 410	1.0	0.71	0.74	1.0	1.0	1.0	0.53
10 212	1.0	0.71	0.74	1.0	0.9	0.95	0.45

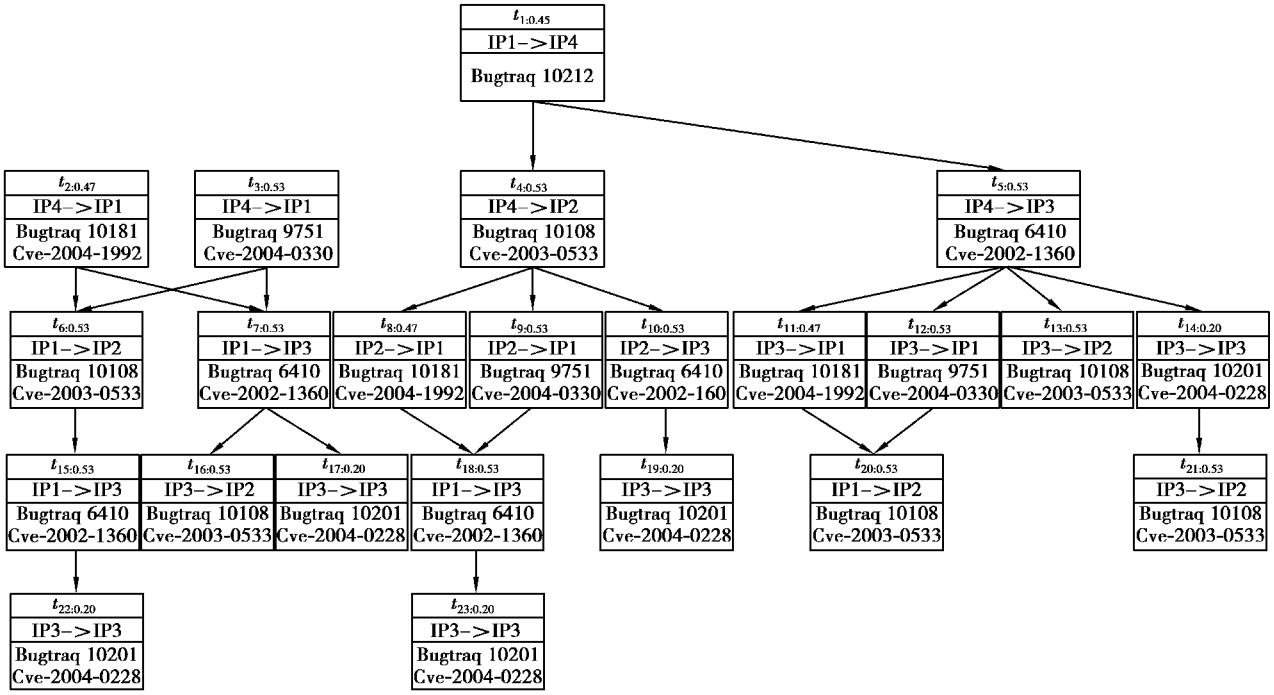

图 2　攻击图

Fig.2　Attack graph

表 3　各节点被利用难易程度值对比

Table 3　Compared the values of complexity for exploited nodes

节点编号	$E(v)$	文献[5]取值	文献[8]取值
t_1	0.45	0.95	0.9
t_2	0.47	0.64	0.9
t_3	0.53	0.98	1.0
t_4	0.53	0.86	1.0
t_5	0.53	0.98	1.0
t_6	0.53	0.86	1.0
t_7	0.53	0.98	1.0
t_8	0.47	0.64	0.9
t_9	0.53	0.98	1.0
t_{10}	0.53	0.98	1.0
t_{11}	0.47	0.64	0.9
t_{12}	0.53	0.98	1.0
t_{13}	0.53	0.86	1.0
t_{14}	0.20	0.70	0.95
t_{15}	0.53	0.98	1.0
t_{16}	0.53	0.86	1.0
t_{17}	0.20	0.70	0.95
t_{18}	0.53	0.98	1.0
t_{19}	0.20	0.70	0.95
t_{20}	0.53	0.86	1.0
t_{21}	0.53	0.86	1.0
t_{22}	0.20	0.70	0.95
t_{23}	0.20	0.70	0.95

2.3 被选择概率计算

应用第2节提出的节点被选择概率算法,计算图2中各节点的被选择概率$S(v)$,参见表4。

表4 各节点被选择概率值

Table 4 Each node is selected in probability

节点编号	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
$O(v)$	1.0	0.228	0.257	0.257	0.257	1.0	1.0	0.307
$S(v)$	1.0	0.228	0.257	0.257	0.257	0.257	0.257	0.079
节点编号	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}	t_{16}
$O(v)$	0.346	0.346	0.272	0.306	0.306	0.116	1.0	0.726
$S(v)$	0.089	0.089	0.070	0.079	0.079	0.030	0.257	0.187
节点编号	t_{17}	t_{18}	t_{19}	t_{20}	t_{21}	t_{22}	t_{23}	
$O(v)$	0.274	1.0	1.0	1.0	1.0	1.0	1.0	
$S(v)$	0.070	0.089	0.089	0.079	0.030	0.257	0.089	

2.4 被利用概率计算

应用第2节提出的节点被选择概率值的计算方法,计算图2中各脆弱性节点的被利用概率$P(v)$,并与文献[5,8]的计算结果进行对比,参见表5。

表5 各节点被利用概率值

Table 5 Each node is exploited in probability

节点编号	$E(v)$	文献[5]取值	文献[8]取值
t_1	0.45	0.946	0.9
t_2	0.107	0.113	0.81
t_3	0.136	0.260	0.9
t_4	0.136	0.201	0.9
t_5	0.136	0.264	0.9
t_6	0.136	0.131	0.9
t_7	0.136	0.174	0.9
t_8	0.037	0.032	0.81
t_9	0.047	0.073	0.9
t_{10}	0.047	0.075	0.9
t_{11}	0.033	0.034	0.81
t_{12}	0.042	0.078	0.9
t_{13}	0.042	0.061	0.9
t_{14}	0.006	0.041	0.86
t_{15}	0.136	0.129	0.9
t_{16}	0.099	0.082	0.9
t_{17}	0.014	0.055	0.86
t_{18}	0.047	0.101	0.9
t_{19}	0.018	0.052	0.86
t_{20}	0.042	0.093	0.9
t_{21}	0.016	0.035	0.86
t_{22}	0.051	0.090	0.86
t_{23}	0.018	0.071	0.86

2.5　实验结果对比分析

对比图 3 中 3 条曲线，由于文献[8]在计算脆弱性被利用难易程度时，直接使用 CVSS 中可利用性(E)指标的值，忽略了其他对脆弱性被利用难易程度的影响指标，因此，最终对脆弱性被利用概率的计算结果普遍高于文献[5]和本文的计算结果；但由于文献[5]在计算各个节点被利用概率时未考虑各个节点的父节点间相关性对子节点的关联影响，属于相同父节点的子节点被利用概率分布曲线存在较大波动，如图 3 中红色曲线上的 t_3、t_4、t_5 节点，文献[8]与本文的算法对上述节点的计算结果走向趋势较为一致。

通过上述对比，本文在评价信息系统脆弱性被利用概率时采用的算法更为科学、合理。

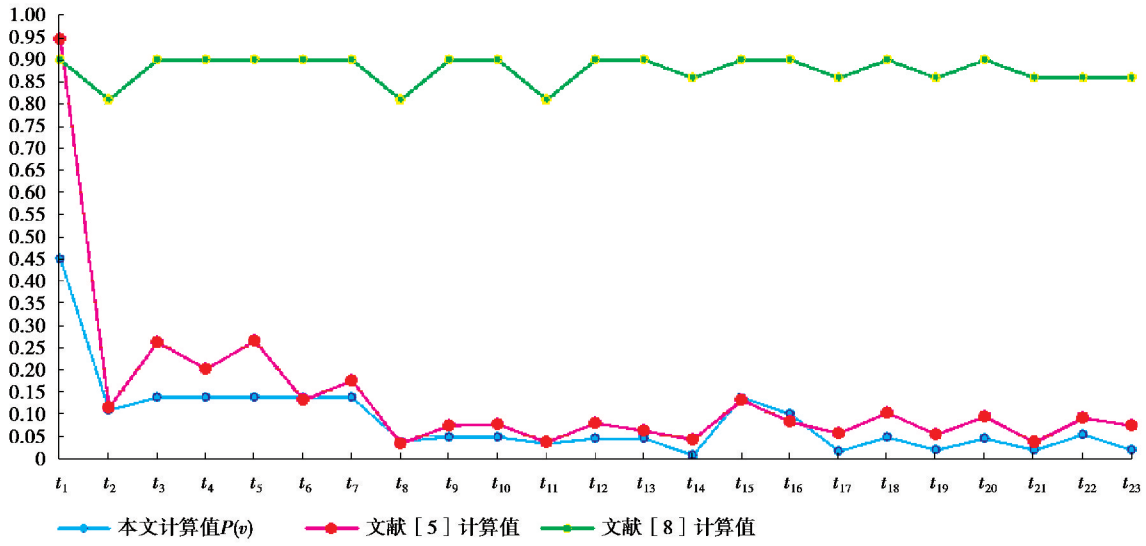

图 3　被利用概率对比分析结果

Fig.3　Comparative analysis in the probability of being exploited

3　结　语

为提高信息安全风险评估中对安全事件发生可能性评估结果的准确性，提出了一种基于贝叶斯网络的脆弱性被利用概率的算法。该算法优化了脆弱性被利用难易程度的计算指标，并从脆弱性在贝叶斯网络中独立性选择概率和关联被选择概率两个方面评价脆弱性的被选择概率。在尽可能客观、准确评价脆弱性的“被利用难易程度”和“被选择概率”基础上计算脆弱性的“被利用概率”。通过理论和实验对比分析，验证了本文所提方法可更加准确地评估脆弱性节点的被利用概率。

参考文献：

[1] GB/T 20984—2007，信息安全技术信息安全风险评估规范[S].

[2] NIST SP 800—30，Risk Management Guide for Information Technology Systems [S].

[3] Peter Mell，Karen Scarfone，Sasha Romanosky. A Complete Guide to the Common Vulnerability Scoring System Version 2.0.www. first. org/cvss.

[4] Phillips C，Laura S P. A graph-based system for network vulnerability analysis[C]// Proc of Workshop on New Security Paradigms. New York：ACM Press，1998：71-79.

[5] 谢丽霞，江典盛，张利，等.漏洞威胁的关联评估方法[J].计算机应用，2012，32(3)：679-682.
XIE Liixia，JIANG Diansheng，ZHANG Li，et al. Vulnerability threat correlation assessment method[J]. Journal of Computer Applications，2012，32(3)：679-682.(in Chinese)

[6] 黄永洪,吴一凡,杨豪璞,等.基于攻击图的 APT 脆弱节点评估方法[J].重庆邮电大学学报(自然科学版),2017,29(4):535-541.
HUANG Yonghong, WU Yifan, YANG Haopu, et al. Graph-based vulnerability assessment for APT attack[J]. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition),2017,29(4):535-541.(in Chinese)

[7] 陈锋.基于多目标攻击图的层次化网络安全风险评估方法研究[D].长沙:国防科技大学,2009.
CHEN Feng. A Hierarchical Network Security Risk Evaluation Framework Based on Multi-Goal Attack Graphs[D]. Changsha: National University of Defense Technology,2009.(in Chinese)

[8] 叶云,徐锡山,齐治昌.大规模网络中攻击图的节点概率计算方法[J].计算机应用与软件,2011,28(11):137-192.
YE Yun, XU Xishan, QI Zhichang. Attack graph's nodes probabilistic computing approach in a large-scale network[J].Computer Applications and Software,2011,28(11):137-192.(in Chinese)

[9] Frigault M, Wang L. Measuring network security using Bayesian network-based attack graphs[C]// Proceedings of the 3rd IEEE International Workshop on Security, Trust, and Privacy for Software Applications. Turku, Finland,2008:698-703.

[10] Frigault M, Wang L. Measuring network security using dynamic bayesian network[C]// Proc.4th ACM Workshop on Quality of Protection. Alexandria VA, USA,2008:23-30.

[11] 冯月进,张凤斌.最大相关最小冗余限定性贝叶斯网络分类器学习算法[J].重庆大学学报,2014,37(6):71-77.
FENG Yuejin, ZHANG Fengbin. Max-relevance min-redundancy restrictive BAN classifier learning algorithm[J]. Journal of Chongqing University,2014,37(6):71-77.(in Chinese)

[12] 张凤荔,冯波.基于关联性的漏洞评估方法[J].计算机应用研究,2014,31(3):812-814.
ZHANG Fengli, FENG Bo. Vulnerability assessment based on correlation[J]. Application Research of Computers,2014,31(3):812-814.(in Chinese)

[13] 张玺,黄曙光,夏阳,等.一种基于攻击图的漏洞风险评估方法[J].计算机应用研究,2010,27(1):284-286.
ZHANG Xi, HUANG Shuguang, XIA Yang, et al. Attack graph-based method for vulnerability risk evaluation[J]. Application Research of Computers,2010,27(1):284-286.(in Chinese)

[14] Ghosh N, Ghosh S K. An approach for security assessment of network configurations using attack graph[C]// Proc of the 1st International Conference on Networks & Communications. Washington DC: IEEE Computer Society,2009:283-288.

(编辑 王维朗)