基于贝叶斯攻击图的网络攻击面风险评估方法

周余阳[1,2,3]，程光[1,2,3]，郭春生[1,2,3]

（1. 东南大学网络空间安全学院，江苏 南京 211189；
2. 东南大学计算机科学与工程学院，江苏 南京 211189；
3. 教育部计算机网络和信息集成重点实验室（东南大学），江苏 南京 211189）

摘　要：针对移动目标防御中网络攻击面缺少客观风险评估的不足，为了有效地实现网络系统的安全风险评估，实现对潜在的攻击路径进行推算，提出一种基于贝叶斯攻击图的网络攻击面风险评估方法。通过对网络系统中资源、脆弱性漏洞及其依赖关系建立贝叶斯攻击图，考量节点之间的依赖关系、资源利用之间的相关性以及攻击行为对攻击路径的影响，推断攻击者到达各个状态的概率以及最大概率的攻击路径。实验结果表明了所提网络攻击面风险评估方法的可行性和有效性，能够为攻击面动态防御措施的选择提供很好的支撑。

关键词：移动目标防御；安全风险评估；贝叶斯攻击图；攻击面；攻击路径

中图分类号：TP393

文献标识码：A

doi: 10.11959/j.issn.2096-109x.2018053

Risk assessment method for network attack surface based on Bayesian attack graph

ZHOU Yuyang[1,2,3], CHENG Guang[1,2,3], GUO Chunsheng[1,2,3]

1. School of Cyber Science and Technology, Southeast University, Nanjing 211189, China
2. School of Computer Science and Engineering, Southeast University, Nanjing 211189, China
3. Key Laboratory of Computer Network and Information Integration of Ministry of Education (Southeast University), Nanjing 211189, China

Abstract: Aiming at the lack of objective risk assessment for the network attack surface on moving target defense, in order to realize the security risk assessment for the network system, and calculate the potential attack paths, a risk assessment method for network attack surface based on Bayesian attack graph was proposed. The network system resources, vulnerability and dependencies between them were used to establish Bayesian attack graph. Considering dependencies between nodes, the correlation between the resource and the influence of attacks on the attack path, the probability of each state that attackers can reach and the maximum probability attack path can be inferred. The experimental results prove the feasibility and effectiveness of the proposed network attack surface risk assessment method, which can provide a good support for the selection of dynamic defensive measures of attack surface.

Key words: moving target defense, security risk assessment, Bayesian attack graph, attack surface, attack path

---

收稿日期：2018-05-03；**修回日期**：2018-06-02

通信作者：程光，gcheng@njnet.edu.cn

基金项目：国家自然科学基金资助项目（No.61602114）；国家重点研发计划基金资助项目（No. 2017YFB0801703）

Foundation Items: The National Natural Science Foundation of China (No.61602114), The National Key R&D Plan Program of China (No.2017YFB0801703)

1 引言

随着信息技术的飞速发展，互联网开始在社会的各行各业有所应用，国家电力、交通、金融、能源等领域的有效运作都离不开互联网的支撑，人们的生产生活方式也与互联网息息相关。互联网一方面给人们带来了诸多的便利和好处，而另一方面，互联网存在着巨大的安全隐患。社会对互联网的依赖性越强，网络攻击带来的危害就越严重，因此，网络空间的安全尤为重要。

有研究表明，互联网安全的最大问题在于整体态势的易攻难守[1]。攻击者具备足够的时间进行攻击准备并组织进攻，能够长期针对攻击目标（网络设备、通信链路、基础协议等）的固有脆弱性进行反复的漏洞分析和渗透测试，直至达成最终目标[2]。移动目标防御（MTD, moving target defense）是美国提出的“改变游戏规则”防御研究方向，通过不断、持续地转移攻击面，减少系统的静态性、同构性和确定性，迷惑或误导攻击者，增加攻击者实施攻击的成本和难度，以此挫败攻击者的攻击。

而作为移动目标防御中的重要一环，网络攻击面的动态转移一直是研究者持续关注的研究重点。传统的网络防御通常采用防火墙、入侵检测系统、用户认证、数据加密和解密、漏洞扫描、防病毒软件等，但单一的安全防护技术已经不能确保网络和系统的安全，而且大部分安全防御技术是被动、滞后的。由于网络架构和配置的静态性，这些方法在检测攻击时依赖先验知识，通过静态特征对攻击进行匹配，而攻击者可以持续地收集和分析网络和系统的信息，使其有充分的时间找到目标中存在的缺陷和漏洞，从而对目标发动攻击。基于网络攻击面动态转移的移动目标防御方法，主要目的是切断攻击者实施网络侦查和探测目标漏洞这一环节，迫使攻击者不断追逐攻击目标，阻止攻击者连接到目标系统或引导攻击者连接到虚假、错误的目标，从而消除攻击者攻击时间、空间上的优势。

然而，现有网络攻击面的动态转移大多为基于时间驱动的网络资源随机转移或基于简单事件的动态转移方式，由于没有对安全状态的整体认知，故其并不具有广泛的针对性和目的性；同时，对于网络攻击面动态转移前后的安全状况以及风险状况，目前还不存在客观的风险评估分析方法。在目前众多网攻击的模型描述方法中，最常见的就是攻击树和攻击图方法，其通过模拟不同节点之间的因果依赖关系，研究复杂的多步攻击行为。攻击树具有直观性、可视化等特点，能够很好地用图形化语言描述网络攻击。但由于树状模型存在扩展性不强的缺陷，因此，攻击树在描述复杂的网络攻击时，效果将大打折扣，同时，复杂网络系统建模时的复杂度和工作量也在很大程度上制约了攻击树方法的使用[3-4]。而相比于攻击树，攻击图方法由于基于单调性假设，具有良好的可扩展性，能够很好地对复杂的网络攻击进行描述[5-6]。

Poolsappasit 等[7]于 2012 年提出了基于贝叶斯攻击图的动态安全风险管理方法，评估安全管控方法的收益与开销，奠定了基于攻击图的安全评估与管控的研究方向；Miehling 等[8]在攻防双方信息部分可见的博弈情况下，基于贝叶斯攻击图利用开销函数制定最优的防御策略；类似地，Nguyen 等[9]在贝叶斯攻击图的基础上提出一种多阶攻击图，定量对比与评估不同防御策略的开销；Bopche 等[10]在动态网络中构建贝叶斯攻击图，基于图相似性理论在时域上定量测量攻击面的转移，从而实现网络安全风险的评估。

在国内的研究中，陈小军等[11]利用概率攻击图，结合节点置信概率，对内部攻击者的攻击意图和攻击路径进行推算；高妮等[12]采用攻击图描述攻击间因果关系，结合通用漏洞评分和局部条件概率分布，推理并动态更新单步攻击的后验概率；刘威歆等[13]针对现有攻击图方法存在冗余路径误报的问题，提出基于攻击图的多源告警关联分析方法，结合图关系与阈值进行联动预测，借此提升关联分析的有效性；雷程等[14]提出的移动目标防御效能评估方法中，以攻击图为基础建立分层网络资源图，结合图中变点检测与标准化度量方法，对移动目标防御的成本与收益进行了动态度量。

在上述研究的基础上，本文提出一种基于贝叶斯攻击图的网络攻击面风险评估方法，通过动

态构建贝叶斯攻击图，综合考量节点之间的依赖关系、相关性与可达概率分布情况，能够有效地评估网络系统的安全风险情况，并对潜在的攻击路径进行推算，能够为攻击面动态防御措施的选择提供很好的支撑。

2　网络攻击建模

2.1　攻击面及其相关定义

目前，学术界对攻击面尚无一个明确的、统一的定义。Howard 等[15]最早将攻击面描述为系统的承受攻击能力，其中包含攻击目标与促成因素、通道与协议、访问权限 3 个维度；Manadhata 等[16]正式提出了系统攻击面的概念，其针对大量攻击案例分析后，总结归纳出攻击者可以利用系统函数、通道和系统环境中的数据项对系统进行攻击，并据此将系统攻击面定义为系统函数、通道和数据三元组的子集；Peng 等[17]在云服务安全的研究中，将虚拟机实例的攻击面归结为所有外部可访问资源的总和；而在网络攻击面的相关研究[18-20]中，研究者将网络攻击面定义为暴露在攻击者面前的网络资源（端口、IP 地址等）以及已被攻破、可利用的脆弱性漏洞。

结合上述研究可以发现，攻击面实质上属于系统资源，是系统的重要组成部分。然而，并不是全部的系统资源都可以称为攻击面，只有攻击者能够利用某种资源攻击系统时，该资源才成为攻击面的一部分。于是，本文将攻击面归结为系统中可被利用、遭受攻击的资源集合，攻击者可通过攻击面实施攻击，达到窃取系统资源、破坏系统的目的。

虽然现有的移动目标防御研究已经广泛使用了攻击面的概念，但目前在攻击面定义及其外延概念的表述上还存在精确性与通俗性的不足。因此，为了进一步对攻击面的特性进行刻画以及后续阐述网络攻击的建模方法，本文在此基础上给出了以下若干定义。

定义 1　（资源）对于系统 *Sys*，将可访问、可接入的系统组成统称为系统资源 *res*，并定义 *Res* 为全体系统资源的全集。对任意资源 $res \in Res$，其可配置参数集为 C_r；对任意可配置参数 $c \in C_r$，其取值集合为 *Vc*。

定义 2　（攻击面）对于系统 *Sys*，定义系统的攻击面 *As* 为攻击者当前可用于发动攻击的系统资源 *res* 的集合。若攻击者当前时刻能够利用 *n* 种资源 $res_1,res_2,\cdots,res_n$ 实施攻击，则 $res_1, res_2,\cdots, res_n$ 构成攻击面 *As*，即 $As=\{res_1,res_2,\cdots,res_n\}$，并满足 *As* 是全体系统资源 *Res* 的子集，即 $As \subseteq Res$。

定义 3　（攻击面转移）对于攻击面 *As*，若攻击面上资源数量发生变化或资源可配置参数取值发生变化，则表示攻击面发生了转移。即 t_1 时刻 $As_1=\{res_1,res_2,\cdots, res_m\}$，$C_{r1}=\{c_1,c_2,\cdots,c_m\}$；$t_2$ 时刻 $As_2=\{res_1,res_2,\cdots,res_n\}$，$C_{r2}=\{c_1,c_2,\cdots,c_n\}$。若 $m \neq n$ 或 $C_{r1} \neq C_{r2}$，表明攻击面 $As_1 \neq As_2$，则称该系统从 t_1 时刻到 t_2 时刻发生了攻击面转移。

定义 4　（风险系数）对于系统资源 *res*，若攻击者成功攻陷该资源的概率为 *p*，则称 *p* 为资源 *res* 的风险系数。

2.2　贝叶斯攻击图定义

攻击图将网络资源形式化，反映了网络内可能存在的攻击路径，并能根据图中所示情况对攻击者攻击时更有可能采用的路径进行评估与判断。作为一种非常有效的概率推理模型，贝叶斯网络由 Pearl[21]于 1988 年率先提出。在贝叶斯网络中，初始节点将会赋予初始概率值，而有向边则反映了节点之间的因果关系，从而可以根据初始节点概率以及节点间的因果关系，对后续所有节点的条件概率进行相应推导。

因此，目前的学者也大多基于贝叶斯网络进行网络资源上攻击路径发生概率和节点被攻陷概率计算的研究。通常，会将基于贝叶斯网络的攻击图称作贝叶斯攻击图。类似于 Poolsappasit 等[7]提出的方法，本文假定贝叶斯攻击图中各个节点采用伯努利随机变量表示当前的属性状态，属性主要包含以下 4 种实例：系统漏洞、系统特性、网络特性和访问权限。

定义 5　（节点属性状态）对任意节点的属性 *S*，其包含 *S*=1/True 与 *S*=0/False 这 2 个状态。*S*=1/True 时，表示属性 *S* 当前已被攻击者攻破，节点到达该状态；相应地，*S*=0/False 时表示属性 *S* 当前尚未被攻击者攻破。

攻击图中代表攻击者攻击行为的最小单位称为原子攻击。根据攻击图的种类和应用场景的不

同，原子攻击可以是一次漏洞利用、一次社会工程攻击或一次未授权的登录行为，或仅表示网络状态发生了改变而没有呈现出具体攻击行为的细节。在贝叶斯攻击图中，原子攻击以有向边表示，根据初始状态，有向边以及节点间的依赖关系和状态概率形成的四元组集合构成了贝叶斯攻击图，其具体定义如下。

定义 6 （贝叶斯攻击图）给定节点属性状态集合 S，有向边集合（原子攻击集合）A，依赖关系集合ε，概率集合 Pr，定义贝叶斯攻击图为一个四元组 $BAG=(S,A,\varepsilon,P_r)$，需要满足以下 4 个条件。

1) $A\in S\times S$，$\forall a\in A, a=pre(a)\rightarrow post(a)$。其中，$pre(a)$ 表示 a 的起始状态节点，$post(a)$ 表示 a 的结果状态节点。

2) $S=S_{\text{ini}}\bigcup S_{\text{mid}}\bigcup S_{\text{fin}}$，其中，$S_{\text{ini}}$、$S_{\text{mid}}$、$S_{\text{fin}}$ 分别代表初始状态、中间状态以及最终状态。其中

① $\forall S_i\in S_{\text{ini}}$，不存在 $a\in A$ 使 S_i=$post(a)$。

② $\forall S_j\in S_{\text{mid}}$，存在 $a_1,a_2\in A$ 使 S_j=$pre(a_1)$ =$post(a_2)$。

③ $\forall S_k\in S_{\text{fin}}$，不存在 $a\in A$ 使 S_k=$pre(a)$。

3) $\forall S_i\in S_{mid}\bigcup S_{fin}$，存在 $S_j\in S$ 使 $(S_j,S_i)\in\varepsilon$，称 S_j 为 S_i 的父节点，记作 $Par[S_i]=\{S_j\in S\mid(S_j,S_i)\in\varepsilon\}$。

4) $\forall S_i\in S$，$Pr(S_i)$ 代表状态 S_i=1 的概率，即状态 S_i 的可到达概率，同理，状态 S_i=0 的概率可表示为 $Pr(\neg S_i)$=1 – $Pr(S_i)$。

2.3 贝叶斯攻击图构建

首先，为了生成贝叶斯攻击图，需要分别输入 n 个初始节点的初始状态 S_{ini} 以及其直接后续的有向边集合 A_0、依赖关系集合ε_0 和概率集合 P_{r0}，应用贝叶斯攻击图生成算法，遍历完整的节点状态以及攻击路径，实现攻击图的构建，如算法 1 所示。

算法 1 贝叶斯攻击图生成（Bayesian attack graph generate）算法

输入 初始状态序列 $S_{\text{ini}}=\{s_{\text{ini}1},s_{\text{ini}2},\cdots,s_{\text{ini}n}\}$ 及对应的后续有向边集合 A_0、依赖关系集合ε_0 和概率集合 Pr_0

输出 攻击图 $BAG(S, A, \varepsilon, Pr)$

1) $BAG\ (S, A, \varepsilon, Pr)\leftarrow(S_{\text{ini}}, A_0, \varepsilon_0, Pr_0)$

2) $n\leftarrow0$

3) WHILE $S_n\neq S_{\text{fin}}$ DO //判断是否到达最终态

4) FOR EACH A_n, ε_n, Pr_n IN $BAG\ (S, A, \varepsilon, Pr)$

5) Calculate S_{n+1} //计算下一阶段节点状态

6) $n\leftarrow n$+1

7) Infer A_n, ε_n, Pr_n//推断

8) IF A_n, ε_n, Pr_n 不为空 THEN

9) Add $(S_n, A_n, \varepsilon_n, Pr_n)$ to BAG //更新后续一阶段攻击图

10) ELSE

11) Add$(S_n, 0, 0, 0)$ to BAG//到达最终态

12) END WHILE

13) RETURN $BAG\ (S, A, \varepsilon, Pr)$

其次，由于单调性假设，攻击者在攻击过程中不断地提高自己的能力且不会失去已有的能力，那么攻击者重复索取已拥有的权限将不满足利益最大的原则。同样地，由于贝叶斯攻击图属于有向无环图，所以需要对攻击图中出现的环路即含圈攻击路径进行剔除。将算法 1 生成的攻击图作为输入，遍历全图，剔除环路后，生成新图，如算法 2 所示。

算法 2 环路剔除（loop remove）算法

输入 原始攻击图 BAG

输出 新攻击图 BAG'

1) $(S, A, \varepsilon, Pr)\leftarrow$ Traversal(BAG) //遍历攻击图

2) $n\leftarrow$Stage(BAG)

3) FOR EACH $S_i\in S, a_i\in A, \varepsilon_i\in\varepsilon, Pr_i\in Pr$

4) FOR EACH $S_j\in S, a_j\in A, \varepsilon_j\in\varepsilon, Pr_j\in Pr$

5) IF EXIST $S_j=post(a_i)$ IN ε_i WITH Pr_i AND $S_i=post(a_j)$ IN ε_j WITH Pr_j THEN //判断是否属于环路

6) IF $i<j$ THEN

7) Remove(a_j,ε_j,Pr_j)FROM(S, A, ε, Pr)

8) ELSE

9) Remove(a_i,ε_i,Pr_i)FROM(S, A, ε, Pr)

10) $BAG'\leftarrow$Update (S, A, ε, Pr) //更新攻击图

11) RETURN BAG'

2.4 节点条件概率计算

为了能够计量各个节点的条件概率分布情况，需要首先评估攻击者成功利用漏洞，攻破某个资源的成功概率。因此，本文引入了通用脆弱性评分系统[22]（CVSS, common vulnerability scoring system），CVSS 能够提供完整的评分参

数，开放评分框架，使评分者直接了解总体评分的由来；采用相同框架对全体脆弱性进行评分，评分标准规范统一；采用动态评估与脆弱性所在资源之间依赖关系相结合的方式，量化资源脆弱性利用的难易程度。

本文参照 Zangeneh 等[23]提出的方法，基于攻击途径（access vector）、攻击复杂度（access complexity）、身份认证（authentication instances）这 3 个测度，对漏洞利用的成功率进行测算。在一次原子攻击过程中，节点状态由 S_i 转移到 S_j，假定其中某一脆弱性漏洞与该状态转移过程相关，将其记作 e_i，那么该漏洞的成功利用率为

$$Pr(e_i) = AV \cdot AC \cdot AU \tag{1}$$

$Pr(e_i)$仅代表该漏洞的成功利用率，而在攻击图所描述的多步攻击中，漏洞能否成功利用、结果状态是否可达，还依赖于其起始状态节点的可达概率。为此还需计算节点的条件概率，即父节点可达的前提下，子节点的可到达率 $Pr(S_j \mid Par[S_j])$。

由于父节点与子节点之间存在“与”关系和“或”关系，即$\varepsilon \in$ {AND,OR}，故在此分以下 2 种情况进行讨论。

1) 当ε=AND，表示任意父节点的状态均为 1 时，子节点才有概率可达；否则，子节点不可达。

即若 S_j=1，那么 $\forall S_i \in Par[S_j], S_i = 1$。由此，可计算条件概率为

$$Pr(S_j \mid Par[S_j]) = \begin{cases} 0, & \exists S_i \in Par[S_j], S_i = 0 \\ \prod\limits_{S_i=1} Pr(e_i), & \text{其他} \end{cases} \tag{2}$$

2) 当ε=OR，表示仅存在一个父节点的状态为 1 时，该子节点便有概率可到达；而当任意父节点状态均为 0 时，则表示该子节点必定不可到达。

即若 S_j=1，那么 $\exists S_i \in Par[S_j], S_i = 1$。由此，可计算条件概率为

$$Pr(S_j \mid Par[S_j]) = \begin{cases} 0, & \forall S_i \in Par[S_j], S_i = 0 \\ 1 - \prod\limits_{S_i=1} [1 - Pr(e_i)], & \text{其他} \end{cases} \tag{3}$$

3　安全风险评估

3.1　风险评估框架

在设计风险评估框架时，需考虑到网络风险的三要素，即资源、脆弱性、威胁。本文基于此建立了如图 1 所示的风险评估框架。

首先，需要对系统资源、系统脆弱性以及系统所受威胁有一个整体认知，鉴于贝叶斯攻击图的组成元素包含资源节点状态以及节点之间的因果关系，本文所构建框架将通过网络系统的资源信息及专家知识对网络系统中所存在资源的状态属性进行识别，对资源之间的状态转换以及依赖关系进行确认。

其次，在贝叶斯攻击图的构建过程中，节点在状态转移过程中需要条件概率的测算才能对后续状态进行推导，故图 1 框架中引入了脆弱性检测组件，从而检测出整个网络系统的脆弱性漏洞，

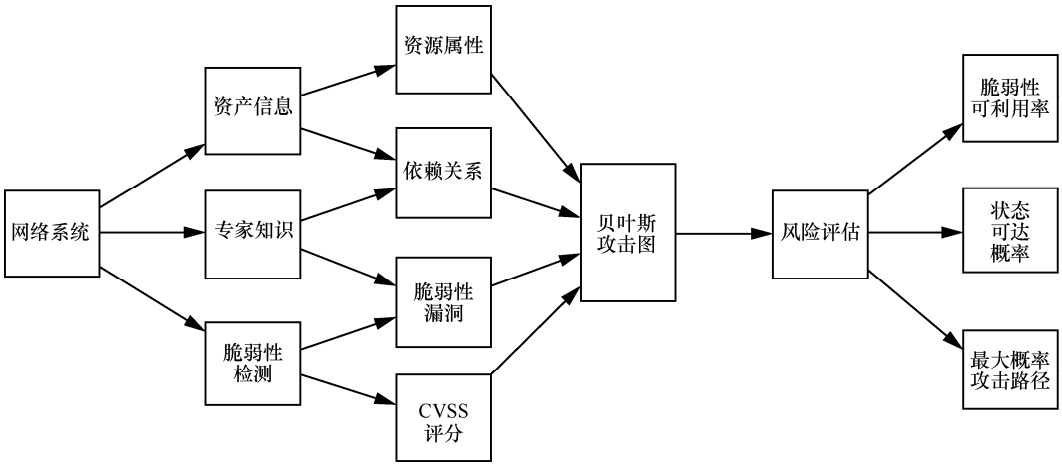


图 1　风险评估框架

并结合文献[7]提出的方法引入对脆弱性漏洞的可利用率进行度量的测算，此外还额外结合父子节点之间存在的依赖关系对计算到达后续状态的条件概率进行完善，使攻击图更加精确和有效地反映网络系统各资源的实际状态转移情况。

最后，在构建的贝叶斯攻击图基础之上，为了进一步细化风险评估，使评估结果更加简洁、直观，本框架参照文献[14,23]所提出的方法，对脆弱性可利用率进行定量度量，并根据本文提出的状态可达概率的概念，获取各个节点状态的可达情况，实现对每个节点状态总体可达概率的测算，并基于此由最终状态对前向状态进行反向查找，从而推导得出最大概率的攻击路径。

该风险评估框架综合考虑了系统资源属性与脆弱性漏洞利用之间的关系，以状态节点和有向转移概率的形式构成了贝叶斯攻击图。既考虑了节点之间的依赖关系，又引入测量标准对脆弱性漏洞利用率进行数学计算。同时，为了直观清晰地描述系统当前的威胁情况，通过脆弱性可利用率、状态可达概率以及最大概率攻击路径的形式表示整体的安全风险状况，能够为防御者后续的攻击面动态转移策略的制订与执行提供基础与支撑。

3.2 风险的相关性

为了更精确地有效评估网络系统的安全风险，本文在前述 CVSS 的基础上，考量资源相互之间的关系对资源脆弱性的影响因素，以保证评估数据的有效性。资源之间的相互关系主要体现在资源脆弱性的依赖关系和关联性对攻击者实施攻击的影响，即不同类型资源间的依赖关系和相同类型资源间的关联关系对攻击者实施攻击、漏洞利用的影响。其中，不同类型资源之间的依赖关系对攻击实施的影响已经在 2.4 节条件概率计算中进行了讨论，故本节重点讨论相同类型资源之间的相关性。

由于不同网络资源节点之间可能存在相同的资源类型，如不同网络资源可能都采用了 TCP，如果攻击者在某一节点上成功通过 TCP 的某个脆弱性漏洞将该节点攻破，那么攻击者在后续的攻击路径上需要再次利用该类资源脆弱性时，将很大程度上提高脆弱性利用的成功率。可以认为，对于相同类型的资源脆弱性，虽然在不同网络资源中的配置参数可能并不完全一致，但是其脆弱性具有相似的属性，存在一定的相关性。因此，攻击者一旦成功利用了某类资源的脆弱性，那么在其后续的攻击行为中，利用同类资源脆弱性的成功率将提高。为此，本文参照雷程等[14]提出的方法，对同类资源的相关性进行了定义，其形式化描述如下。

定义 7 （相关性）给定系统 Sys，对于$\forall res_1$，$res_2 \in Res$，若 res_1，res_2 属于相同类型的资源，该类型可配置参数集为 C_r，那么定义 res_1，res_2 的相关性为 $Cor(res_1, res_2)$: $C_r \times C_r \to [0,1]$。

当 Cor 为 0 时，表示 res_1，res_2 完全不相关，属于两类资源，此时 $Pr(e_{res2} \mid e_{res1}) = Pr(e_{res2})$；而当 Cor 为 1 时，表示 res_1，res_2 完全相同，属于同一资源，那么攻击者在已能成功利用 res_1 的前提下必然能攻破 res_2。由此，可以衍生出相同类型资源脆弱性漏洞利用的条件概率为

$$Pr(e_{res_2} \mid e_{res_1}) = Pr(e_{res_2})^{1-Cor} \tag{4}$$

3.3 攻击路径推断

攻击图的生成和分析主要服务于准确地预测攻击以及推测攻击者的后续攻击行为。于是，能否推断出攻击者的攻击意图以及攻击路径成了攻击图技术是否具有现实意义的重要评估内容。近年来，如何根据攻击图高效地对攻击路径进行推测，并及时采取相应的安全防护措施，减少因网络攻击造成的危害，也已成为网络安全防护领域的一项研究热点。

因此，本文首先定义了总体可达概率，其次给出了全节点概率推导算法，最后给出了最大概率攻击路径推导算法。

3.3.1 总体可达概率

在给定攻击图以及脆弱性可利用概率的前提下，将状态的总体可达概率定义为发生一系列攻击行为到达当前节点状态的总体可能性。其具体定义如下。

定义 8 （总体可达概率）给定贝叶斯攻击图 BAG 以及脆弱性漏洞可利用概率集合 $\{Pr(e_1), Pr(e_2), \cdots, Pr(e_n)\}$，总体可达概率($TAP$)定义如下。

1) 初始状态必然可达，即 $TAP(S_{\text{ini}})$=1。

2) $\forall S_i \in S_{\text{mid}} \bigcup S_{\text{fin}}$， $\forall Par[S_i] \in S$：

① 若 $\varepsilon(Par[S_i],S_i)$=AND，$Pr(e_{\_Par[S_i]\to S_i})$ 表示 S_i 父节点转移到其状态所需要脆弱性漏洞的可利用率，则 $TAP(S_i)=\prod TAP(Par[S_i])\ Pr(e_{\_Par[S_i]\to S_i})$；

② 若 $\varepsilon(Par[S_i],S_i)$=OR，则 TAP 可表示为 $TAP(S_i)=1-\prod\{1-TAP(Par[S_i])Pr(e_{\_Par[S_i]\to S_i})\}$。

3.3.2 全节点概率推导算法

根据 3.3.1 节中关于总体可达概率的定义，给出如下全节点概率推导算法，实现了对攻击图整体全部节点的可达概率评估，较为全面地评估攻击者的攻击意图，为后续攻击路径的推导提供数据基础。

算法 3 全节点概率推导（whole node probability derivation）算法

输入 攻击图 $BAG(S, A, \varepsilon, Pr)$

输出 攻击图上各个节点的总体可达概率

1) InitQueue(Q)　　//初始化队列
2) PushQueue(Q,S_{ini})　　//压入初始节点
3) WHILE (EmptyQueue(Q)) DO
4)　s = PopQueue(Q)
5)　FOR EACH $S_i \in S$
6)　　IF $s \in Par[S_i]$　THEN //判断是否父节点
7)　　PushQueue(Q,S_i)　//压入其子节点
8)　　IF NOT EXIST $Par[s]$ THEN
9)　　$WNPD(s)$=1
10)　　ELSE
11)　　$WNPD(s)$=$TAP(s)$
12) END WHILE
13) RETURN 各个节点的总体可达概率

3.3.3 最大概率攻击路径推导算法

根据 3.3.2 节中全节点概率推导算法上，以最终状态为起点，反向查找，寻找总体可达概率最大的节点，依次将其添加至攻击路径上，直到寻找到初始节点，则能推导出一条最大概率的攻击路径，具体算法如下。

算法 4 最大概率攻击路径推导（maximum probability attack path derivation）算法

输入 攻击图 $BAG(S, A, \varepsilon, Pr)$；各个节点的总体可达概率集合 TAP

输出 最大概率攻击路径 *Attack Path*

1) InitQueue(Q)　　//初始化队列
2) F=argmax($TAP(S_{fin})$) //在最终状态中查找总体可达概率最大的节点
3) Add F to *Attack Path*　//加入攻击路径
4) PushQueue(Q,F)
5) WHILE (EmptyQueue(Q)) DO
6)　s = PopQueue(Q)
7)　IF NOT EXIST $Par[s]$ THEN //是否初始节点
8)　　Add s to *Attack Path*
9)　ELSE IF $\varepsilon(Par[s],s)\in$AND THEN
10)　　PushQueue(Q, $Par[s]$) //将 s 的全部父节点压入队列
11)　　Add $Par[s]$ to *Attack Path*
12)　ELSE IF $\varepsilon(Par[s],s)\in$OR THEN
13)　　M=argmax($TAP(Par[s])$) //查找父节点中总体可达概率最大节点
14)　　PushQueue(Q, M)
15)　　Add M to *Attack Path*
16) END WHILE
17) RETURN *Attack Path*

4 实验验证分析

为了验证基于贝叶斯攻击图的网络攻击面风险评估方法的可行性与有效性，本节首先利用如图 2 所示的网络拓扑，构建了小型的实验网络环境，并对此网络环境中的安全漏洞和潜在攻击手段进行了检测。然后利用第 2 节介绍的网络攻击建模方法实现攻击图的构建和对应的有向边之间的条件概率表。最后通过第 3 节介绍的安全风险评估方法，实现风险的前后关联，相应调整对应资源的可利用率，从而计算出全节点的总体可达概率，最终实现最大概率攻击路径的推测。

4.1 实验网络环境

在图 2 所示的实验网络拓扑中，网络被划分为 2 个部分：外网部分和内部工作网络。外网与内网之间用防火墙进行逻辑隔离，在内部工作网络中有 3 台主机 H_1、H_2、H_3，其中，H_1 为一台互联网信息服务（IIS, Internet information services）服务器，H_2 为一台文件服务器，H_3 为一台数据服务器，具体的配置信息如表 1 所示。

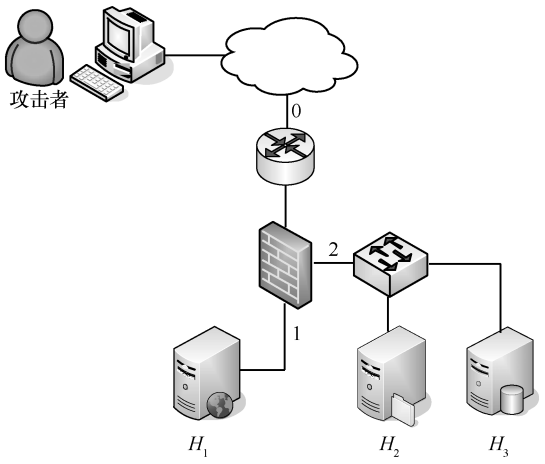


图 2 实验网络拓扑示意图

表 1 主机节点配

节点编号	节点名称	系统配置信息
H_1	网络服务器	Windows NT 4.0
H_2	文件服务器	Windows Server 2003 SP1
H_3	数据服务器	Red Hat 7.0

同时，防火墙将整体实验网络环境划分为 3 个网段，其中，外网为网段 0，H_1 属于网段 1，H_2、H_3 同属于网段 2，防火墙具体实施策略如表 2 所示，对于同网段主机之间的访问遵循该网段所开放端口的访问，而对于不在防火墙策略中的其他访问则视为非法访问，一律禁止。

4.2 实验准备工作

首先，在实验网络环境中运用 Nessus 对实验网络中的主机进行脆弱性检测，再对检测到的节点资源脆弱性信息进行汇总，同时结合表 2 防火墙的配置规则，提取出可能被攻击者利用的脆弱性资源，如表 3 所示。由于本文主要考察资源脆弱性对网络服务造成的影响，所以对扫描得到的脆弱性所对应的网络服务资产，按照如表 3 所示的 CVSS 基本度量组指标分值，结合如表 4 所示的节点资源脆弱性的信息列表，再通过式(1)对攻击途径、攻击复杂度、身份认证进行测算，可得出相应的脆弱性成功利用率，其结果如表 5 所示。

表 3 CVSS 基本度量指标分值

度量指标	度量等级	等级分值
AV	网络	0.85
	邻近网络	0.62
	本地	0.55
	物理	0.20
AC	低	0.78
	中	0.56
	高	0.24
AU	无	0.85
	低	0.62
	高	0.27

表 4 节点资源脆弱性信息

编号/节点	资源	脆弱性名称	CVE 编号
A / H_1	IIS	IIS buffer overflow	CVE-2009-1012
B / H_2	FTP	FTP rhost overwrite	CVE-2011-4800
C / H_2	RSH	RSH login	CVE-2006-0408
D / H_2	SSH	SSH buffer overflow	CVE-1999-1455
E / H_3	Squid proxy	Squid port scan	CVE-2001-1030
F / H_3	LICQ	LICQ-remote-to-user	CVE-2001-0439
G / H_3	MySQL DB	Local-setuid-bof	CVE-2006-3368

4.3 攻击图构建与总体可达概率分布

根据实验网络环境的拓扑结构、利用 Nessus 扫描获得的节点脆弱性结果，得到如图 3 所示的贝叶斯攻击图。其中，椭圆形标记代表节点状态，数字代表主机标号，而无圈文字代表攻击，需要特别说明的是，在攻击图中仅有攻击 LICQ_ remote_to_user(1,3) 与攻击 LICQ_remote_to_ user(2,3)指向状态 User(3)的两条有向边属于“或”关系，其余的有向边指向同一状态的依赖关系均属于“与”关系。

根据 2.4 节中的相关定义以及表 5 中脆弱性利用成功率的情况，可以得出各个状态之间的条件概率分布情况；再根据 3.2 节中风险相关性的

表 2 防火墙策略

网段	可访问主机	端口号	服务
0	H_1	80	IIS
1	H_2	21、22、514	FTP、SSH、RSH
	H_3	80、5190、3306	Squid proxy、LICQ、MySQL DB
2	H_1	80	IIS

表 5　　脆弱性成功利用率

节点	CVE 编号	*AV*	*AC*	*AU*	$Pr(e_i)$
H_1	CVE-2009-1012	0.85	0.56	0.27	0.129
H_2	CVE-2011-4800	0.62	0.56	0.62	0.215
H_2	CVE-2006-0408	0.62	0.78	0.62	0.300
H_2	CVE-1999-1455	0.62	0.78	0.27	0.131
H_3	CVE-2001-1030	0.85	0.78	0.85	0.564
H_3	CVE-2001-0439	0.85	0.78	0.62	0.411
H_3	CVE-2006-3368	0.55	0.56	0.62	0.191

定义以及 3.3 节中关于总体可达概率的定义，最终得出了如表 6 所示的总体可达概率分布情况。值得说明的是，对于图 3 中的通信交互状态，如 ftp(1,2)、ssh(1,2)等，由于其属于表 2 防火墙策略中所允许的正常访问，故将其状态可达概率认定为 1，即表明其必定可达；而对于脆弱性状态，如 CVE-2009-1012、CVE-2011-4800 等，由于已在表 5 中详细说明了其可利用成功率，故也不在总体可达概率情况表中额外说明。此外，对于攻击图中的攻击 squid_port_scan(1,3) 与 squid_port_scan(2,3)，LICQ_remote_to_user (1, 3)与 LICQ_remote_to_user (2, 3)，发动攻击的资源与状态存在相关性，于是根据 3.2 节中相关内容对其后状态可达概率进行相应修正。

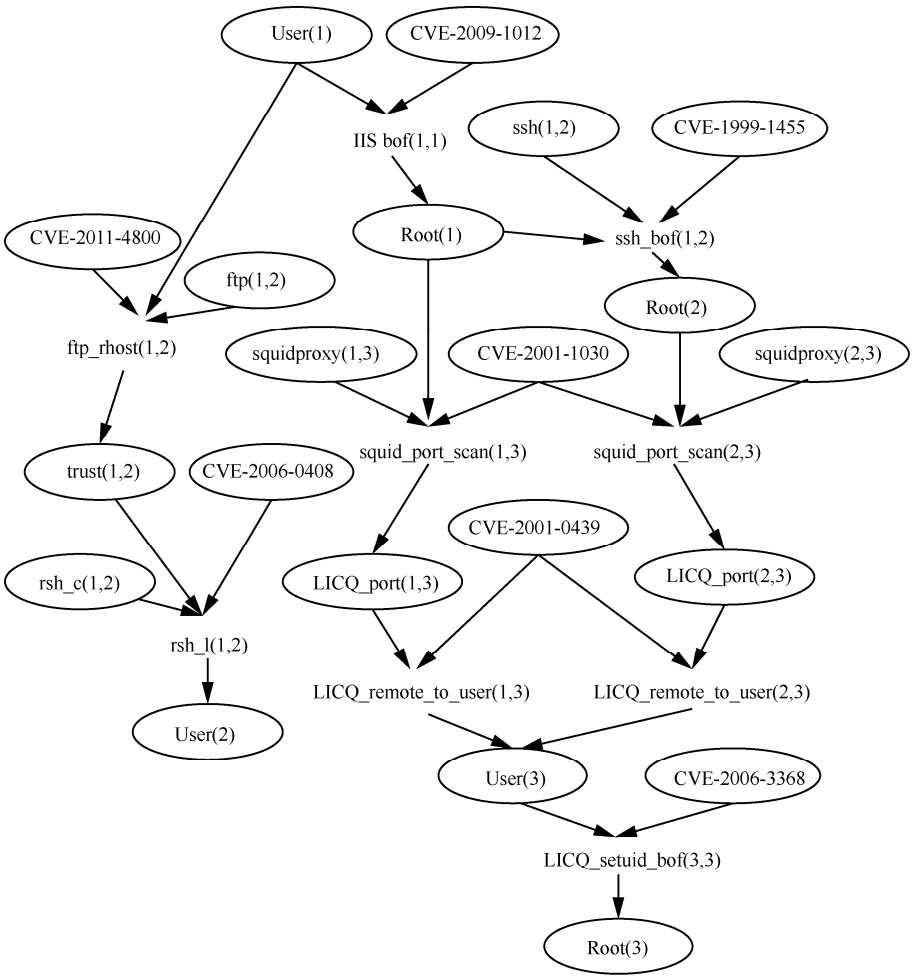


图 3　实验网络构建的贝叶斯攻击图

表 6 总体可达概率情况

状态	*TAP*	状态	*TAP*
User(1)	1	LICQ_port(1,3)	0.109
Root(1)	0.129	LICQ_port(2,3)	0.014
Trust(1,2)	0.215	User(3)	0.067
User(2)	0.065	Root(3)	0.013
Root(2)	0.017	—	—

4.4 最大概率攻击路径的推导

根据前述的贝叶斯攻击图以及总体可达概率分布情况，计算得到的攻击路径以及其可到达率如表 7 所示。

根据表 7，可以发现路径 3 的可到达率显著高于路径 1 与路径 2，说明攻击者的状态可到达率随着攻击路径的增长、脆弱性数量的增加而降低。但是，值得注意的是，路径 3 的最终状态仅仅到达 User(2)，并不是一条成功的攻击路径。同时，当对攻击路径 2 进行拆解后，可以发现攻击 User(1)→A→Root(1)→D→Root(2)的可到达率为 0.017，也从侧面说明 Root 权限获取的攻击难度明显高于 User 权限。

上述的攻击路径以及可到达率结果建立在实验环境中无观测的情况下，可发现最终的可到达率均相对较低，而在实验中为了实现安全风险的监控，部署了 Snort 进行各类脆弱性利用攻击行为或网络异常行为事件的捕获，根据不同的监测结果对最大概率攻击路径以及其成功率进行相应修正和调整，即监测到某一脆弱性被利用后，表示该脆弱性利用后的状态必定可达，具体情况如表 8 所示。

根据表 8，同样可以发现，由于路径 3 的最终状态并没有到达 Root(3)，所以即使监测到该路径上发生了脆弱性利用，也可以认为攻击者实行了无效攻击，导致攻击成功率为 0；而对于相同的攻击路径，观测到的脆弱性利用越接近最终状态，其攻击的成功率越高，尤其当监测到编号 G 的脆弱性利用时，表明攻击者已经成功利用 Local-setuid-bof 对 H_3 实施攻击，那么攻击者已经顺利获取 Root(3)的权限，表明攻击成功。

表 8 最大概率攻击路径及其成功率

脆弱性利用监测	最大概率攻击路径	攻击成功率
A	1	0.09
B	3	0
C	3	0
D	2	0.09
E	1 或 2	0.105
F	1 或 2	0.191
G	1 或 2	1

4.5 实验分析

为验证本文风险评估方法的准确性与先进性，同样依据图 3 所示实验网络的贝叶斯攻击图，采用文献[23]提出的风险评估方法得出了总体可达概率与不同攻击路径的可到达率。同时，为了更直观地显示本文方法的准确性，绘制出了图 4 与图 5 的总体可达概率对比图与攻击路径可到达率对比图。

由图 4 与图 5 可以看出，本文提出的风险评估方法，节点状态 LICQ_port(1,3)、LICQ_port(2,3)、User(3)与 Root(3)的总体可达概率显著高于文献[23]所提方法，这是由于 3.2 节中对风险的相关性进行了较深入的讨论与分析，认为攻击者一旦成功利用了某类资源的脆弱性，那么在其后续的攻击行为中，利用同类资源脆弱性的成功率将会提高。因此，在上述节点状态时，由于其受到攻击时所

表 7 攻击路径及其可到达率

编号	攻击路径	脆弱性数	可到达率
1	User(1)→A→Root(1)→E→LICQ_port(1,3)→F→User(3)→G→Root(3)	4	0.011 5
2	User(1)→A→Root(1)→D→Root(2)→E→LICQ_port(2,3)→F→User(3)→G→Root(3)	5	0.001 5
3	User(1)→B→trust(1,2)→C→User(2)	2	0.064 5

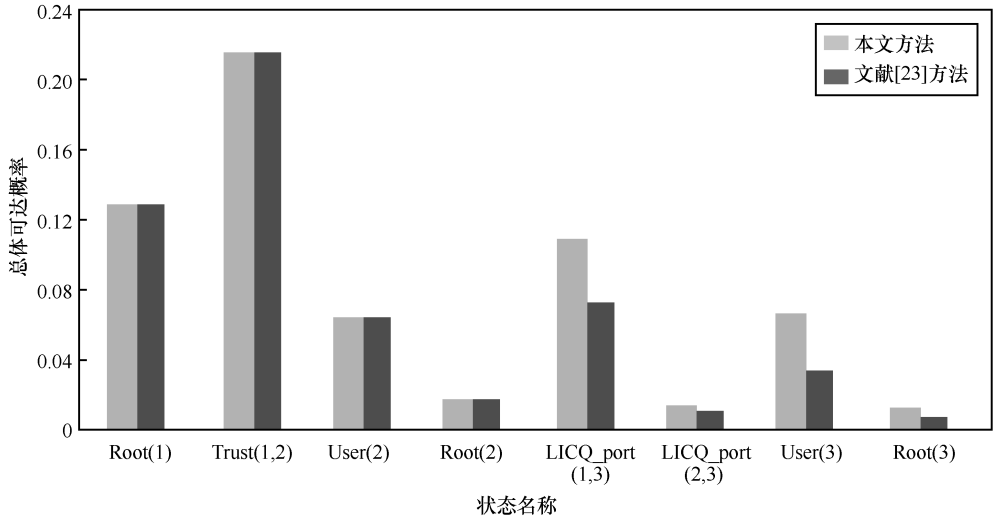


图 4　总体可达概率对比

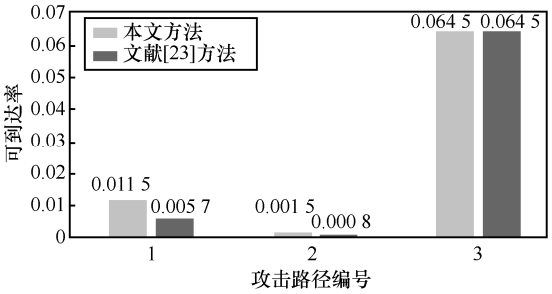


图 5　攻击路径可到达率对比

利用的脆弱性漏洞之间存在相关性，那么总体可达概率也将会相应升高；相应地，由于总体可达概率的提升，导致本文方法中这些状态所在的攻击路径 1 与攻击路径 2 的可到达率也高于文献[23]方法所得结果。由此可以看出，本文的风险评估方法能够更加精确地对网络系统的整体风险情况进行有效评估，能够更加可靠地为网络攻击面的转移提供有力的依据。

除此以外，图 6 还对 2 种风险评估方法中攻击路径的推导效率进行了对比，从图中可以看出，本文方法攻击路径的推导时间明显少于文献[23]

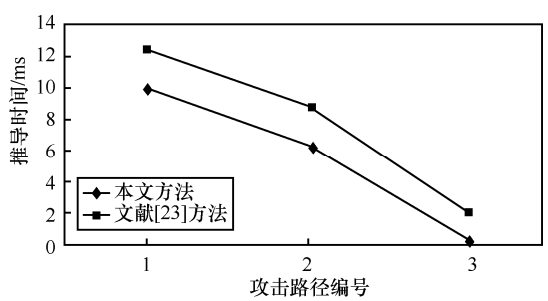


图 6　攻击路径推导时间对比

的推导时间，这是由于本文 3.3.3 节中提出的最大概率攻击路径推导算法中利用反向查找的方式，优先从最终状态入手，可以有效地避开冗余路径和无效路径（如攻击路径 3），能够很大程度地提升攻击路径的推导效率。

5　结束语

为了有效地评估网络系统的安全风险情况，并能够对潜在的攻击路径进行推算，本文提出一种基于贝叶斯攻击图的网络攻击面风险评估方法，通过对网络系统中资源、脆弱性漏洞及其依赖关系建立贝叶斯攻击图，进行安全风险的评估，推断攻击者到达各个状态的概率以及最大概率的攻击路径。本文提出的模型和算法考虑了节点之间的依赖关系、资源利用之间的相关性以及攻击行为对攻击路径的影响，提高了风险评估的准确性。实验结果表明，本文的工作可以有效地推导出各个状态的可达概率，并给出了不同情况下攻击者的最大概率攻击路径，能够为防御者实施攻击面的动态转移提供很好的支撑。

对于未来可开展的工作，一方面目前的建模工作主要还是基于小型的实验网络，在大规模网络中推广时还存在问题，如何实现大规模的自动攻击图构建将是后续研究的一个方向；另一方面，本文的风险评估基于脆弱性利用率、资源可被攻破概率以及攻击者攻击成功率，而在实际的攻击中，攻击者也会根据系统防护情

况、攻击开销等其他因素进行攻击选择，而防御者在进行攻击面转移方案的选择时，也会考虑防御性能与防御开销的权衡，这些都将成为未来的研究重点。

参考文献：

[1] JAJODIA S, GHOSH A K, SWARUP V, et al. Moving target defense: creating asymmetric uncertainty for cyber threats[J]. Springer Ebooks, 2011: 54.

[2] 蔡桂林, 王宝生, 王天佐, 等. 移动目标防御技术研究进展[J]. 计算机研究与发展, 2016, 53(5): 968-987.
CAI G L, WANG B S, WANG T Z, et al. Research and development of moving target defense technology[J]. Journal of Computer Research and Development, 2016, 53(5): 968-987.

[3] WHITLEY J N, PHAN R C W, WANG J, et al. Attribution of attack trees[J]. Computers & Electrical Engineering, 2011, 37(4): 624-628.

[4] DALTON II G C, EDGE K S, MILLS R F, et al. Analysing security risks in computer and radio frequency identification (RFID) networks using attack and protection trees[J]. International Journal of Security and Networks, 2010, 5(2/3): 87-95.

[5] AMMANN P, PAMULA J, RITCHEY R, et al. A host-based approach to network attack chaining analysis[C]// Computer Security Applications Conference. 2005: 72-84.

[6] 叶子维, 郭渊博, 王宸东, 等. 攻击图技术应用研究综述[J]. 通信学报, 2017, 38(11): 121-132.
YE Z W, GUO Y B, WANG C D, et al. Survey on application of attack graph technology[J]. Journal on Communications, 2017, 38(11): 121-132.

[7] POOLSAPPASIT N, DEWRI R, RAY I. Dynamic security risk management using Bayesian attack graphs[J]. IEEE Transactions on Dependable and Secure Computing, 2012, 9(1): 61-74.

[8] MIEHLING E, RASOULI M, TENEKETZIS D. Optimal defense policies for partially observable spreading processes on bayesian attack graphs[C]// ACM Workshop on Moving Target Defense. 2015: 67-76.

[9] NGUYEN T H, WRIGHT M, WELLMAN M P, et al. Multi-stage attack graph security games: heuristic strategies, with empirical game-theoretic analysis[C]//ACM Workshop on Moving Target Defense. 2017: 87-97.

[10] BOPCHE G S, MEHTRE B M. Graph similarity metrics for assessing temporal changes in attack surface of dynamic networks[J]. Computers & Security, 2017, 64: 16-43.

[11] 陈小军, 方滨兴, 谭庆丰, 等. 基于概率攻击图的内部攻击意图推断算法研究[J]. 计算机学报, 2014, 37(1): 62-72.
CHEN X J, FANG B X, TAN Q F, et al. Inferring attack intent of malicious insider based on probabilistic attack graph model[J]. Chinese Journal of Computers, 2014, 37(1): 62-72.

[12] 高妮, 高岭, 贺毅岳, 等. 基于贝叶斯攻击图的动态安全风险评估模型[J]. 四川大学学报(工程科学版), 2016, 48(1): 111-118.
GAO N, GAO L, HE Y Y, et al. Dynamic security risk assessment model based on Bayesian attack graph[J]. Journal of Sichuan University(Engineering Science Editon), 2016, 48(1): 111-118.

[13] 刘威歆, 郑康锋, 武斌, 等. 基于攻击图的多源告警关联分析方法[J]. 通信学报, 2017, 36(9): 135-144.
LIU W X, ZHENG K F, WU B, et al. Alert processing based on attack graph and multi-source analyzing[J]. Journal on Communications, 2017, 36(9): 135-144.

[14] 雷程, 马多贺, 张红旗, 等. 基于变点检测的网络移动目标防御效能评估方法[J]. 通信学报, 2017, 38(1): 126-140.
LEI C, MA D H, ZHANG H Q, et al. Performance assessment approach based on change-point detection for network moving target defense[J]. Journal on Communications, 2017, 38(1): 126-140.

[15] HOWARD M, PINCUS J, WING J M. Measuring relative attack surfaces[J]. Computer Security in the 21st Century, 2003: 109-137.

[16] MANADHATA P K, WING J M. An attack surface metric[J]. IEEE Transactions on Software Engineering, 2011, 37(3): 371-386.

[17] PENG W, LI F, HUANG C T, et al. A moving-target defense strategy for cloud-based services with heterogeneous and dynamic attack surfaces[C]// IEEE International Conference on Communications. 2014: 804-809.

[18] SUN K, JAJODIA S. Protecting enterprise networks through attack surface expansion[C]//Workshop on Cyber Security Analytics, Intelligence and Automation. 2014: 29-32.

[19] CYBENKO G, JAJODIA S, WELLMAN M P, et al. Adversarial and uncertain reasoning for adaptive cyber defense: building the scientific foundation[C]//International Conference on Information Systems Security. 2014: 1-8.

[20] FOREMAN J C, GURUGUBELLI D. Identifying the cyber attack surface of the advanced metering infrastructure[J]. The Electricity Journal, 2015, 28(1): 94-103.

[21] PEARL J. Probabilistic reasoning in intelligent system[M]. Morgan Kaufinann: Network of Plausible Inference, 1988: 1-86.

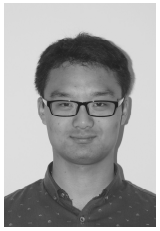
[22] MELL P M, KENT K A, ROMANOSKY S. The common vulnerability scoring system (CVSS) and its applicability to federal agency systems[J]. NIST Interagency/Internal Report (NISTIR) - 7435, 2007, 1(1): 119-127.

[23] ZANGENEH V, SHAJARI M. A cost-sensitive move selection strategy for moving target defense[J]. Computers & Security, 2018(75): 72-91.

[作者简介]

周余阳（1994-），男，江苏泰州人，东南大学博士生，主要研究方向为网络安全、移动目标防御。

程光（1973-），男，安徽黄山人，博士，东南大学教授、博士生导师，主要研究方向为网络测量、网络安全和网络管理。



郭春生（1994-），男，河南南阳人，东南大学硕士生，主要研究方向为网络安全。