计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200700108

基于吸收 Markov 链的网络入侵路径分析方法

张 凯[1,2,3] 刘京菊[1,3]
1 国防科技大学电子对抗学院 合肥 230037
2 中国酒泉卫星发射中心 甘肃 酒泉 732750
3 网络空间安全态势感知与评估安徽省重点实验室 合肥 230037
(zkdfbbking@163.com)

摘 要 从攻击者角度对网络进行入侵路径分析对于指导网络安全防御具有重要意义。针对现有的基于吸收 Markov 链的分析方法中存在的对状态转移情形考虑不全面的问题和状态转移概率计算不合理的问题,提出了一种基于吸收 Markov 链的入侵路径分析方法。该方法在生成攻击图的基础上,根据攻击图中实现状态转移所利用的漏洞的可利用性得分,充分考虑了非吸收节点状态转移失败的情况,提出了一种新的状态转移概率计算方法,将攻击图映射到吸收 Markov 链模型;利用吸收 Markov 链的状态转移概率矩阵的性质,计算入侵路径中节点的威胁度排序和入侵路径长度的期望值。实验结果表明,该方法能够有效计算节点威胁度排序和路径长度期望;通过对比分析,该方法的计算结果相比现有方法更符合网络攻防的实际情况。

关键词:网络安全;入侵路径分析;攻击图;吸收 Markov 链;节点威胁度排序;路径长度期望
中图法分类号 TP393.8

Attack Path Analysis Method Based on Absorbing Markov Chain

ZHANG Kai[1,2,3] and LIU Jing-ju[1,3]
1 College of Electronic Engineering, National University of Defense Technology, Hefei 230037, China
2 Jiuquan Satellite Launch Center, Jiuquan, Gansu 732750, China
3 Anhui Province Key Laboratory of Cyberspace Security Situation Awareness and Evaluation, Hefei 230037, China

Abstract The analysis of network attack path from the perspective of attackers is of great significance to guide network security defense. The existing analysis methods based on absorbing Markov chain have some problems, such as incomplete consideration of state transition and unreasonable calculation of state transition probability. In order to solve these problems, this paper proposes an attack path analysis method based on absorbing Markov chain. Based on the generation of attack graph and the exploitability score of vulnerability, the situation that the failure state transition of non-absorbing nodes will be fully considered. In order to map the attack graph to the absorbing Markov chain model, this paper proposes a new method to calculate the state transition probability. Then, by using the properties of the state transition probability matrix of the absorbing Markov chain, it calculates the threat ranking of the nodes in the attack path and the expected length of the attack path. Then, the application feasibility of absorbing Markov chain with multi absorbing states is discussed. The results of the experiment show that the proposed method can effectively calculate the node threat ranking and path length expectation. Through comparative analysis, this method is more in line with the actual situation of network attack and defense than the existing methods.

Keywords Network security, Attack path analysis, Attack graph, Absorbing Markov chain, Node threat ranking, Path length expectation

1 引言

网络中的恶意攻击者能够利用网络主机上存在的漏洞实施网络攻击,导致国家、企业和个人遭受无法挽回的经济损失。当攻击者对网络进行入侵时,通常从利用网络中的漏洞获得对节点的特权开始,然后逐渐入侵到其他节点,最终到达目标节点。从初始节点到目标节点的入侵路径可用于描述攻击者的特定攻击行为。从攻击者的角度对网络入侵路径进行分析能够定位网络中的薄弱节点,找出可能的入侵路径,从而为网络防御提供明确的策略指导。

目前,用于网络入侵路径分析的模型很多。Huang 等[1]利用攻击图评估路径节点在 APT 攻击下的脆弱性。Yu 等[2]将警报关联问题视为推理问题,利用隐彩色 Petri 网从部分警报中发现入侵者的行为,预测入侵者的下一个目标。Wang 等[3]针对现有攻击路径预测方法无法准确反映攻击者的攻击能力对后续攻击路径的影响问题,提出了基于因果知识网络

到稿日期:2020-07-17 返修日期:2020-08-13 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。
通信作者:刘京菊(jingjul@aliyun.com)

的攻击路径预测方法。Liu 等[4]融合攻击方、防护方和网络环境信息,通过时间和空间维度分析各安全态势要素及其相互影响关系对网络安全态势的影响,辅助攻击意图识别和路径分析的研究。

由于网络拓扑本身是基于图的结构,基于图论的攻击图模型已成为研究网络多步攻击的主要方法。Zeng 等[5]基于不确定图模型提出了一种攻击图生成算法,逆向模拟生成攻击图,能够模拟实际攻击情况并找出最可靠的攻击路径。Kerem[6]综述了攻击图的研究现状,指出攻击图的可达性分析和路径研究是攻击图研究的重点之一。Zeng 等[7]将基于攻击图的分析方法总结为 5 类,即图算法、贝叶斯网络、Markov 模型、成本优化算法和不确定性算法,并指出其中基于 Markov 模型的分析方法能够识别出高威胁度节点,评估最可能的入侵路径,具有容易训练和预测效果理想的优势。Sheyner 等[8]最早将概率论与攻击图相结合,基于 Markov 决策过程,利用条件转移概率分析非确定性节点,评估攻击者最可能采用的攻击路径。Wang 等[9]首先提出在隐马尔可夫模型框架下定量分析攻击图的方法,基于提出的模型,通过一系列可观测的值来预测下一个系统状态,从而获得针对特定观察序列的最可能的攻击路径。Miehling 等[10]认为防御者在任何给定时间只能部分观察攻击者的行为,并且需要在信息不完整时做出决定,因此提出将部分可观测 Markov 决策过程用于攻击图分析的优化策略。

Abraham 等[11]将攻击图建模为吸收 Markov 链,Markov 链的转移概率基于通用漏洞评分标准(Common Vulnerability Scoring System,CVSS)来计算,该模型可以对网络进行安全评估,计算其路径长度期望和路径概率。但是,该方法在模型中没有考虑节点状态转移失败的情况,对非吸收节点没有设置指向自身的状态转移边,导致状态转移概率的计算不够合理。

Abraham 等[12]将时间因素引入 Markov 链,并使用漏洞生命周期模型[13]来计算漏洞或补丁在披露后的几天内出现的可能性,指出漏洞的影响将随着时间的推移逐渐减小,应随着时间推移降低从此漏洞进行状态转换的可能性。Hu 等[14]认为现有研究主要集中于理想攻击场景中的路径预测,然而理想攻击路径并不都是入侵者采取的真实路径,因此提出了基于吸收 Markov 链的多步攻击路径预测方法。但对于初始状态,该方法没有考虑状态转移失败的情况,且在计算状态转移概率时将指向自身的状态转移边的得分赋值为同一个数值,不符合网络攻防的实际情况。

本文针对基于吸收 Markov 链的路径分析方法中存在的对状态转移情形考虑不全面的问题和状态转移概率计算方法不合理的问题,提出了一种改进状态转移概率计算方法的基于吸收 Markov 链的入侵路径分析方法,其更加符合网络攻防的实际情况。本文的主要贡献如下:

(1)全面考虑了非吸收节点状态转移失败的情况。对每一个非吸收节点设置指向节点自身的状态转移边,表示状态转移失败的情况。

(2)改进状态转移概率的计算方法。根据每个节点状态转移所对应的漏洞的可利用性得分,在考虑状态转移失败的情况下,为非吸收节点指向自身的状态转移边设置不同的状态转移概率。

2　攻击图与吸收 Markov 链

2.1　攻击图及相关定义

根据攻击图中节点和边表示的不同含义,攻击图可以分为很多类型[15]。本文使用的攻击图模型的定义如下。

定义 1(攻击图(Attack Graph,AG))　攻击图可用四元组表示,$AG=(S,E,V,A)$,其中:

(1)$S=S_I\cup S_T\cup S_G$,S_I,S_T,S_G 互不相交。S_I 表示起始状态节点集合,对于任意 $S_i\in S_I$,不存在指向 S_i 的边;S_T 表示过渡状态节点集合,对于任意 $S_t\in S_T$,存在 $S_i\neq S_j$,$S_i\in S$,$S_j\in S$,有 S_i 指向 S_t 的边和 S_t 指向 S_j 的边。S_G 表示目标状态节点的集合,对于 $S_g\in S_G$,不存在 $S_i\in S$,使得 S_g 指向 S_i。

(2)$E=\{E_{ij}\mid i,j=1,2,\cdots,n\}$,$E_{ij}$ 表示由状态 S_i 指向 S_j 的边,即 E_{ij} 对应一次 S_i 到 S_j 的状态转移。

(3)$V=\{V_i\mid i=1,2,\cdots,m\}$ 表示所有可利用漏洞的集合,对于任意 V_i 有 $EXP(V_i)$ 表示该漏洞的可利用性得分。

(4)$A=\{A_i\mid i=1,2,\cdots,m\}$ 表示原子攻击的集合,A_i 为原子攻击,表示对漏洞的一次利用,每次成功的漏洞利用对应一次状态转移 E_{ij}。原子攻击具有成功概率 $P(A_i)$,有 $P(A_i)=P(E_{ij})$,即原子攻击成功概率等于状态转移概率。若攻击者在状态 S_i 成功发动原子攻击,达到了状态 S_j,则完成了状态转移。

图 1(a)给出了攻击图的示例,其中 S_1 为起始状态节点,S_4 为目标状态节点,状态转移所利用的漏洞和基于 CVSS 的可利用性得分标注于边上。

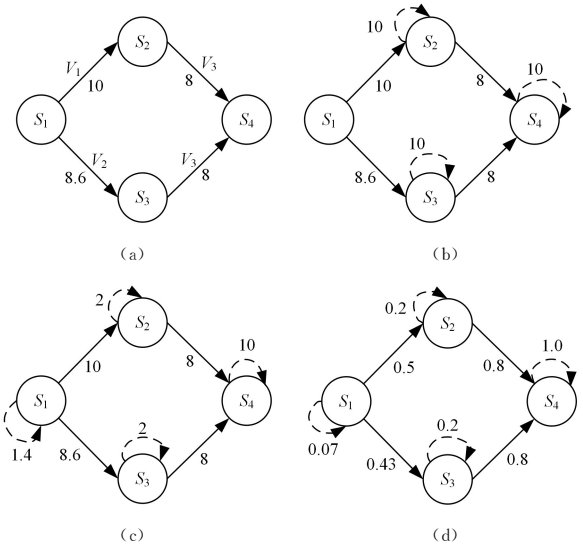

图 1　攻击图示例

Fig. 1　Example of attack graph

根据攻击图的定义,可以定义如下概念:1)入侵路径(Attack Path,AP),表示攻击者从初始状态节点到达目标状态节点的状态转移序列,$AP_1=S_1\rightarrow S_2\rightarrow S_4$ 为一条从 S_1 到 S_4 的入侵路径;2)入侵路径长度(Attack Path Length,APL),表示路径所含有向边的数量,路径 AP_1 的长度为 2;3)路径成功概率 $P(AP)$,表示路径中所有状态转移都成功的概率,$P(AP_1)=$

$P(E_{12})\times P(E_{24})$；4)入侵路径长度期望，表示攻击者为实现其入侵目标需要实施的入侵步骤的期望值；5)状态节点访问次数期望，表示攻击者在实现攻击目标的过程中对某个状态节点访问次数的期望值。根据状态节点访问次数期望值可以对节点的威胁度进行排序，攻击者在实现攻击目标的过程中需要访问某节点的次数越多，该节点对攻击的贡献越大，节点的威胁度越高，排序就越靠前。

2.2 吸收 Markov 链

Markov 链是时间和状态都离散的 Markov 过程[16]，记状态空间为 $S=\{S_1,S_2,\cdots,S_n\}$，该过程从这些状态之一开始，能够从一个状态转移到另一个状态。如果 Markov 链当前处于状态 S_i，则它以概率 P_{ij} 进入下一状态 S_j，该概率不取决于链在当前状态之前所处的状态，仅与当前状态有关，这一性质称为无后效性，概率 P_{ij} 称为状态转移概率。该过程可以保持其所处的状态，并且概率为 P_{ii}。所有的 P_{ij} 按行列组合即可构成矩阵 $\boldsymbol{P}$。

定义 2(吸收 Markov 链[16]) 如果一个 Markov 链含有至少一个吸收状态且从任意状态出发都能最终到达吸收状态，则称之为吸收 Markov 链。若吸收 Markov 链有 r 个吸收状态和 t 个非吸收状态，则其所有状态数 $n=t+r$，将代表吸收状态的行和列置于矩阵右下方，状态转移概率矩阵可表示为：

$$\boldsymbol{P}=\begin{pmatrix} Q & R \\ O & I \end{pmatrix}$$

其中，$\boldsymbol{Q}$ 是 $t\times t$ 的矩阵，表示非吸收状态之间的转移概率矩阵；$\boldsymbol{R}$ 是 $t\times r$ 的矩阵，表示非吸收状态到吸收状态的转移概率；$\boldsymbol{O}$ 表示 $r\times t$ 的全 0 矩阵；$\boldsymbol{I}$ 是 $r\times r$ 的单位矩阵。

在攻击图中，当前状态是否能向下一个状态转移只取决于当前状态是否满足漏洞利用的前置条件，而与当前状态之前的状态无关，若将攻击图中的状态节点集合当作 Markov 链的状态空间，则状态间的转移正好符合 Markov 链的无后效性；攻击图中原子攻击的成功率可以作为 Markov 链的状态转移概率；由于攻击图必然包括目标节点，攻击图的目标状态可以作为 Markov 链的吸收状态。因此，攻击图能够映射到吸收 Markov 链模型，从而利用吸收 Markov 链对入侵路径进行分析。

3 基于吸收 Markov 链的入侵路径分析

基于吸收 Markov 链进行入侵路径分析的流程如图 2 所示。

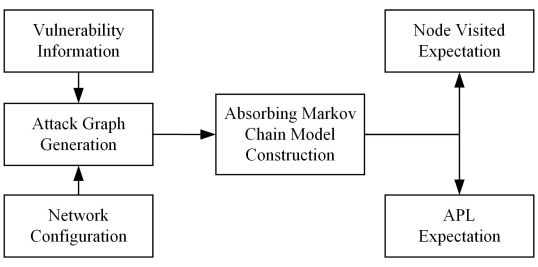

图 2 基于吸收 Markov 链的入侵路径分析流程

Fig. 2 Attack path analysis process based on absorbing Markov chain

首先对目标网络环境生成攻击图；其次以生成的攻击图为基础构建吸收 Markov 链模型，这一步最重要的是状态转移概率矩阵的计算；然后根据状态转移概率矩阵的性质计算节点访问次数的期望和入侵路径长度的期望；最后对所得结果进行分析。

3.1 攻击图到吸收 Markov 链的映射

本文利用 Mulval[17] 工具自动化生成攻击图，Mulval 生成的攻击图为逻辑攻击图，为了将攻击图映射到吸收 Markov 链，需要对生成的攻击图进行化简，文献[18]提供了攻击图的化简方法。

3.1.1 基于 CVSS 的漏洞可利用性评分度量

CVSS 是漏洞评估领域的公开标准，其中包含了漏洞可利用性评分准则，如表 1 所列，这些准则全面评估了漏洞利用的难度，具体的数值可以通过查询美国国家漏洞库(National Vulnerability Database，NVD)得到。本文基于 CVSS 进行漏洞的可利用性评分度量，增强了研究的通用性。

表 1 漏洞可利用性指标

Table 1 Base metrics of vulnerability

Base Metrics	Metric Value	Numerical Value
Access Vector (AV)	Local (L)	0.395
	Adjacent Network (A)	0.646
	Network(N)	1.0
Access Complexity (AC)	High (H)	0.35
	Medium(M)	0.61
	Low (L)	0.71
Authentication (Au)	Multiple(M)	0.45
	Single (S)	0.56
	None (N)	0.704

在 NVD 漏洞库中查询得到漏洞的具体信息，根据各指标的取值，通过式(1)计算得到可利用性得分(Exploitability Score，ES)。

$$ES=20\times AV\times AC\times Au \tag{1}$$

3.1.2 状态转移概率的计算

在理想的情况下，每次实施的原子攻击都能够成功，从而达成状态的转移，但在实际的网络攻防对抗中，由于攻击者自身的技能水平不同、网络环境的复杂多变、目标网络的防御措施等因素的影响，原子攻击的成功实施存在一定的概率，可以将原子攻击失败看作是状态节点到自身的状态转移。

对于状态转移概率 P_{ij} 的计算，文献[11]提出的方法没有考虑漏洞利用失败的情形，直接将节点的所有边对应的漏洞的可利用得分进行归一化，从而得到状态转移概率；文献[14]中的方法考虑了漏洞利用失败的情况，但是没有考虑初始节点，且在计算向自身转移的概率时，对指向自身的边设置的可利用性得分均为 10，如图 1(b)所示，然后对得分进行归一化得到状态转移概率。针对文献[11,14]中的问题，本文给出如下改进方法：

(1)充分考虑漏洞利用失败的情形，对所有的非吸收节点设置指向自身的状态转移边，包括初始状态节点。

(2)改进状态转移边的概率计算方式。节点指向自身的状态转移边的概率计算需要综合以该节点为起始的边，所有以该节点为起始的边对应的可利用漏洞的可利用性得分的平均值越大，说明向其他状态转移成功的概率越大，节点保持自身状态的概率就越小。

改进算法的具体步骤如算法 1 所示。

算法 1　状态转移概率归一化度量算法

输入：攻击图 AG=(S,E,V,A)

输出：状态转移概率矩阵 **P**

步骤 1　取集合 S 中的未遍历的节点 S_i。

步骤 2　计算 S_i 的出度为 k，对于 S_i 指向的 k 个节点集合 $S_{end}=\{S_j \mid j=1,2,\cdots,k\}$，计算边 E_{ij} 对应的原子攻击所利用的漏洞的可利用性得分总和 $Score_i$。

步骤 3　计算状态 S_i 到自身的状态转移概率，即 P_{ii}，若 $Score_i=10*k$，则 $P_{ii}=1/(1+10*k)$；否则 $P_{ii}=(10*k\text{-}Score_i)/10*k$。

步骤 4　对于集合 S_{end} 中的 S_j，若 $Score_i=10*k$，则 $P_{ij}=EXP(v_j)/(1+10*k)$；否则 $P_{ij}=EXP(v_j)/10*k$。

步骤 5　重复步骤 1 直至遍历 S 中的所有状态。

步骤 6　算法结束。

攻击图为有向图，不妨假设以邻接矩阵 $\boldsymbol{A}$ 存储攻击图，$\boldsymbol{A}$ 的维度为攻击图中的状态数，记为 n；如果 $\boldsymbol{A}$ 中的元素 A_{ij} 非零则表示存在状态节点 S_i 到 S_j 的边，A_{ij} 的值为 S_i 到 S_j 所利用漏洞的可利用性评分。在矩阵 $\boldsymbol{A}$ 上运行算法 1，每个状态节点都需要遍历 1 次，共有 n 个节点。对于每一个节点，首先需要遍历一遍其他节点以求和（步骤 2 中的 $Score_i$），共 n 次；然后计算概率也需要遍历所有节点，共 n 次。即对于每一个节点需要遍历 $2n$ 次，则共需要 $2n^2$ 次，因此时间复杂度为 $O(n^2)$。在计算过程中需要维护 $n\times n$ 的矩阵 $\boldsymbol{P}$，因此空间复杂度为 $O(n^2)$。

当算法 1 应用于一些极端情况，如出度较大的节点时，计算的概率可能会失真，例如某个节点有 10 条流出边，且每条边对应的漏洞得分都为 5，则指向自身的边的得分就高达 50，归一化后概率为 0.5，其他边的状态转移概率仅为 0.05。

以图 1(a)中的攻击图为例运行算法 1，对各条指向节点自身的状态转移边的赋值如图 1(c)所示，则相比文献[14]中的方法，本文算法计算得到的得分差别较大，对得分进行归一化，即可得到状态转移概率，如图 1(d)所示。

3.2　入侵路径分析

基于吸收 Markov 链的入侵路径分析主要利用状态转移概率矩阵相关的定理，通过算法 1 计算状态转移概率矩阵，然后根据相关定理来计算对状态的访问次数期望和到达吸收状态时的状态转移次数的期望。状态的访问次数的期望对应攻击图中节点的威胁度，状态转移次数的期望则对应入侵路径长度的期望，下面结合定理进行阐述。

定理 1　对于一个吸收 Markov 链的状态转移概率矩阵 $\boldsymbol{P}$ 中的矩阵 $\boldsymbol{Q}$，矩阵 $\boldsymbol{I}\text{-}\boldsymbol{Q}$ 的逆矩阵记为矩阵 $\boldsymbol{N}$，则 $\boldsymbol{N}$ 中的元素 N_{ij} 表示从状态 S_i 出发在到达吸收状态时经过状态 S_j 的次数。

$$N=(I-Q)^{-1} \tag{2}$$

由于吸收 Markov 链是以攻击图为基础构建的，那么对链中的状态访问次数的期望就对应攻击图中状态节点的访问次数期望。对于攻击者来说，在实现攻击目标过程中需要逐步获得权限从而到达不同的状态节点，对某节点的访问次数越高，说明该节点对攻击者越重要，节点威胁度就越大。

定理 2　假设吸收 Markov 链从状态 S_i 开始，设 t_i 为到达吸收状态之前的状态转移次数的期望，设 t 为第 i 个元素为 t_i 的列向量，有：

$$t=N\times c \tag{3}$$

其中，c 为全为 1 的列向量。

根据第 2.1 节中的定义，入侵路径中的边表示状态的转移，那么根据定理 2 求得的从不同状态出发到达吸收状态时需要的状态转移次数的期望，就对应攻击图中从起始状态到目标状态的路径长度的期望。

定理 3　设 B_{ij} 表示从状态 S_i 出发，最终到达吸收状态 S_j 的概率，$\boldsymbol{B}$ 是由 B_{ij} 组成的矩阵，则：

$$\boldsymbol{B}=N\times R \tag{4}$$

多攻击目标的攻击图可以映射为多吸收状态 Markov 链，根据定理 3 可以计算从初始状态到达不同吸收状态的概率，即表示在攻击图中，攻击者从初始状态节点开始，最终到达不同攻击目标的概率，使得网络安全管理员有针对性地对到达概率较高的节点采取防御措施。

4　实验与分析

4.1　实验环境与攻击图生成

本文构建的网络环境拓扑如图 3 所示。网络中包括外部防火墙 Firewall-1、内部防火墙 Firewall-2、攻击者主机、服务器主机。

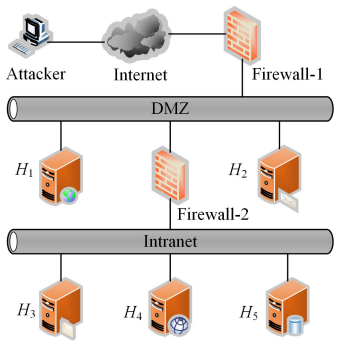

图 3　实验网络拓扑

Fig. 3　Topology of experiment network

外网防火墙将内部网络与 Internet 隔离开来；内网防火墙进一步将内部网络分为 DMZ 区和 Internal 区；Firewall-1 仅允许 Internet 中的主机访问 DMZ 区域的主机 H_1 和 H_2；Firewall-2 仅允许 H_1 和 H_2 访问 H_3 和 H_4；在 Intranet 中，H_3 和 H_4 可以访问 H_5。网络中各主机运行的服务信息和漏洞信息如表 2 所列。

表 2　网络主机配置及漏洞信息

Table 2　Host configuration and vulnerability information

Host	Service	Vulnerability	No.	ES
H_1	Web	CVE-2014-0098	V_1	10
H_2	Email	CVE-2016-0037	V_2	8
H_3	Linux	CVE-2018-2773	V_3	3.4
	Office	CVE-2018-8247	V_4	8.6
H_4	Bmc	CVE-2013-4782	V_5	10
H_5	Radius	CVE-2014-1878	V_6	10

根据网络拓扑和防火墙策略及采集的漏洞信息，利用工具 Mulval 生成攻击图，并将攻击图化简，得到图 4 所示的攻击图，不同状态间转移所依赖的漏洞已经标注于边上，边值表示状态转移所依赖漏洞的可利用性得分，具体的状

态描述信息如表 3 所列。

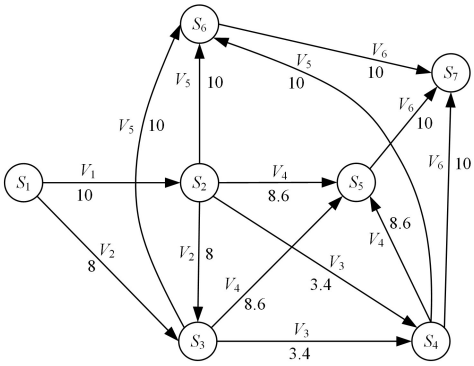

图 4 化简后的攻击图

Fig. 4 Simplified attack graph

根据算法 1 计算状态转移概率，得到吸收 Markov 链的状态转移概率矩阵。对于非目标节点添加指向自身的状态转移边，并将攻击图中的可利用性得分替换为对应的状态转移概率，得到吸收 Markov 链攻击图，如图 5 所示，图中虚线为非目标节点指向自身的状态转移边，状态转移概率标注于边上。

表 3 状态节点信息

Table 3 State node information

State	Description	State	Description
S_1	Attacker	S_5	(H_3, root)
S_2	(H_1, root)	S_6	(H_4, root)
S_3	(H_2, root)	S_7	(H_5, root)
S_4	(H_3, user)		

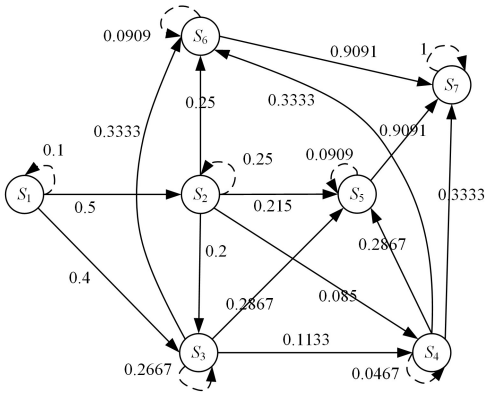

图 5 吸收 Markov 链攻击图

Fig. 5 Attack graph based on absorbing Markov chain

图 5 对应的状态转移概率矩阵 $\boldsymbol{P}$ 为：

$$\boldsymbol{P}=\begin{pmatrix}0.1 & 0.5 & 0.4 & 0 & 0 & 0 & 0\\ 0 & 0.25 & 0.2 & 0.085 & 0.215 & 0.25 & 0\\ 0 & 0 & 0.2667 & 0.1133 & 0.2867 & 0.3333 & 0\\ 0 & 0 & 0 & 0.0476 & 0.2867 & 0.3333 & 0.3333\\ 0 & 0 & 0 & 0 & 0.0909 & 0 & 0.9091\\ 0 & 0 & 0 & 0 & 0 & 0.0909 & 0.9091\\ 0 & 0 & 0 & 0 & 0 & 0 & 1\end{pmatrix}$$

4.2 节点威胁度排序与分析

4.2.1 节点威胁度排序

根据状态转移概率矩阵 $\boldsymbol{P}$，利用式(2)和式(3)计算得到期望矩阵 $\boldsymbol{N}$ 和 $\boldsymbol{t}$。

$$\boldsymbol{N}=\begin{pmatrix}1.1111 & 0.7407 & 0.8081 & 0.1622 & 0.4812 & 0.5595\\ 0 & 1.3333 & 0.3637 & 0.1623 & 0.4812 & 0.5595\\ 0 & 0 & 1.3637 & 0.1622 & 0.4812 & 0.5594\\ 0 & 0 & 0 & 1.0500 & 0.3311 & 0.3849\\ 0 & 0 & 0 & 0 & 1.1000 & 0\\ 0 & 0 & 0 & 0 & 0 & 1.1000\end{pmatrix}$$

$$t=\begin{pmatrix}3.8629\\ 2.8999\\ 2.5666\\ 1.7661\\ 1.1000\\ 1.1000\end{pmatrix}$$

期望矩阵 $\boldsymbol{N}$ 的不同行代表从不同节点出发到达吸收节点时经过各节点次数的期望值。节点的访问次数越高，该节点的威胁程度就越大。矩阵 $\boldsymbol{N}$ 的第 1 行，表示从状态节点 S_1 出发，若要到达最终的吸收状态，则需要经过其他状态节点的次数的期望；访问 S_2, S_3, S_4, S_5, S_6 状态节点的期望次数分别为 0.7407, 0.8081, 0.1622, 0.4812, 0.5595，据此得到节点威胁度排序 $S_3>S_2>S_6>S_5>S_4$，因此在采取防御策略时应优先考虑修复状态 S_3（H_2 上的漏洞）。

根据文献[14]中的方法计算状态转移概率矩阵 $\boldsymbol{P}$，并根据式(2)和式(3)计算矩阵 $\boldsymbol{N}$ 和 $\boldsymbol{t}$。

$$P=\begin{pmatrix}0 & 0.556 & 0.444 & 0 & 0 & 0 & 0\\ 0 & 0.25 & 0.2 & 0.085 & 0.215 & 0.25 & 0\\ 0 & 0 & 0.3125 & 0.1060 & 0.2687 & 0.3125 & 0\\ 0 & 0 & 0 & 0.2591 & 0.2228 & 0.2591 & 0.2591\\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5\\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5\\ 0 & 0 & 0 & 0 & 0 & 0 & 1\end{pmatrix}$$

$$N=\begin{pmatrix}1.0000 & 0.7413 & 0.8615 & 0.2083 & 0.8746 & 1.0170\\ 0 & 1.3333 & 0.3879 & 0.2085 & 0.8747 & 1.0171\\ 0 & 0 & 1.4545 & 0.2081 & 0.8744 & 1.0169\\ 0 & 0 & 0 & 1.3479 & 0.6041 & 0.6994\\ 0 & 0 & 0 & 0 & 2.0000 & 0\\ 0 & 0 & 0 & 0 & 0 & 2.0000\end{pmatrix}$$

$$t=\begin{pmatrix}4.7027\\ 3.8215\\ 3.5540\\ 2.6506\\ 2.0000\\ 2.0000\end{pmatrix}$$

从矩阵 $\boldsymbol{N}$ 的第 1 行可以看出，从 S_1 状态节点出发，在到达吸收状态 S_7 时访问 S_2, S_3, S_4, S_5, S_6 节点的期望值分别为：0.7413, 0.8615, 0.2083, 0.8746, 1.0170。据此得到的节点威胁度排序为 $S_6>S_5>S_3>S_2>S_4$。这与本文方法得出的结果相差较大，主要差别在于节点 S_6 和 S_5 与节点 S_3 和 S_2 的威胁排序上，下面从 4 个方面对两种方法的结果进行分析。

4.2.2 分析

实验网络生成的攻击图共包含 15 条入侵路径，具体路径

信息和概率分布如表 4 所列。

表 4　入侵路径信息和概率分布

Table 4　Attack path information and probability distribution

Attack Path	Length	Success Probability
$S_1\rightarrow S_2\rightarrow S_4\rightarrow S_7$	3	0.014 1
$S_1\rightarrow S_2\rightarrow S_5\rightarrow S_7$	3	0.097 7
$S_1\rightarrow S_2\rightarrow S_6\rightarrow S_7$	3	0.113 6
$S_1\rightarrow S_3\rightarrow S_4\rightarrow S_7$	3	0.015 1
$S_1\rightarrow S_3\rightarrow S_5\rightarrow S_7$	3	0.104 3
$S_1\rightarrow S_3\rightarrow S_6\rightarrow S_7$	3	0.121 2
$S_1\rightarrow S_2\rightarrow S_3\rightarrow S_4\rightarrow S_7$	4	0.003 8
$S_1\rightarrow S_2\rightarrow S_3\rightarrow S_5\rightarrow S_7$	4	0.026 1
$S_1\rightarrow S_2\rightarrow S_3\rightarrow S_6\rightarrow S_7$	4	0.030 3
$S_1\rightarrow S_2\rightarrow S_4\rightarrow S_5\rightarrow S_7$	4	0.011 1
$S_1\rightarrow S_2\rightarrow S_4\rightarrow S_6\rightarrow S_7$	4	0.012 9
$S_1\rightarrow S_3\rightarrow S_4\rightarrow S_5\rightarrow S_7$	4	0.011 8
$S_1\rightarrow S_3\rightarrow S_4\rightarrow S_6\rightarrow S_7$	4	0.013 7
$S_1\rightarrow S_2\rightarrow S_3\rightarrow S_4\rightarrow S_5\rightarrow S_7$	5	0.003 0
$S_1\rightarrow S_2\rightarrow S_3\rightarrow S_4\rightarrow S_6\rightarrow S_7$	5	0.003 4

从最短路径方面进行分析，在 6 条长度为 3 的最短路径中，统计经过各状态节点的最短路径数分布，如图 6 所示。从图中可以看出，最短路径中经过 S_2 和 S_3 节点的路径多于经过 S_5 和 S_6 节点的路径数。

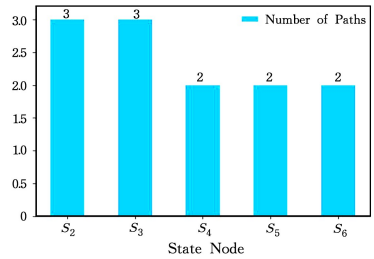

图 6　经过节点的最短路径数

Fig. 6　Number of shortest paths through nodes

从路径的概率分布进行方面，对两种方法得到的路径中成功概率最大的前 6 条路径进行分析，统计路径中各状态的访问次数分布。表 5 列出了文献[14]中成功概率最大的前 6 条路径。两种方法经过节点的路径数分布如图 7 所示。

表 5　成功概率最大的前 6 条路径

Table 5　First six paths with the highest probability of success

Path	Success Probability
$S_1\rightarrow S_2\rightarrow S_6\rightarrow S_7$	0.069 5
$S_1\rightarrow S_3\rightarrow S_6\rightarrow S_7$	0.069 4
$S_1\rightarrow S_2\rightarrow S_5\rightarrow S_7$	0.059 8
$S_1\rightarrow S_3\rightarrow S_5\rightarrow S_7$	0.059 7
$S_1\rightarrow S_2\rightarrow S_3\rightarrow S_6\rightarrow S_7$	0.017 4
$S_1\rightarrow S_2\rightarrow S_3\rightarrow S_5\rightarrow S_7$	0.014 9

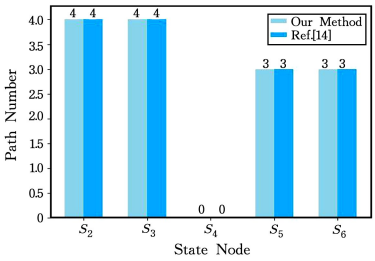

图 7　经过节点的最大概率路径数

Fig. 7　Number of paths passing through nodes

从节点的度分布方面进行，两种方法得到的路径是相同的，因此节点的度也是相同的。各状态节点的度分布如图 8 所示，从图中可以看出节点 S_2 和 S_3 的度分布高于节点 S_5 和 S_6。

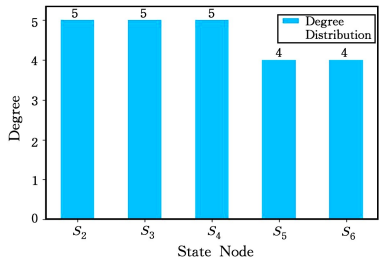

图 8　节点度分布

Fig. 8　Degree distribution of nodes

从节点修复效果方面进行分析，对节点对应的漏洞进行修复，即在攻击图中删除与该节点相关的所有边，对修复后的网络生成攻击图，并分析所得路径数，如图 9 所示。可以看出，在修复了节点 S_2 或 S_3 后，攻击图中剩余可达路径仅剩 5 条，而修复 S_5 或 S_6 后仍有 9 条路径可用。

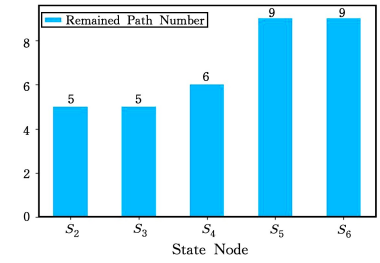

图 9　修复节点后剩余路径数

Fig. 9　Number of paths remaining after node repair

综上分析可得，本文方法在评估节点威胁度方面具有更高的准确性，更加符合实际情况。

4.3　入侵路径分析

4.3.1　威胁路径长度的期望值

根据状态转移概率矩阵求得的矩阵 $\boldsymbol{t}$ 可以得到从不同节点出发到达吸收状态时的状态转移次数期望值，即入侵路径长度的期望值，本文方法与文献[14]中的方法求得的期望值分布如图 10 所示，可以看到两种方法的计算结果在趋势上是一致的。

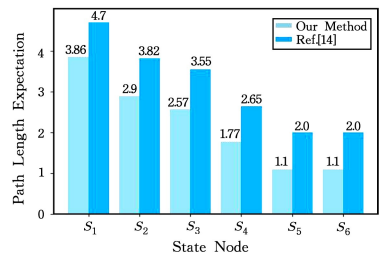

图 10　入侵路径长度期望

Fig. 10　Expectation of path length

4.3.2　分析

本文方法得到的期望值比文献[14]中的方法得到的期望值小，这是因为文献[14]中的方法将节点保持状态的得分统一设置为 10，这导致节点保持自身状态不变的概率变大，从

而使得路径长度的期望值增大;本文方法中的得分设置得相对偏小,路径长度期望值也偏小。考虑到从状态节点 S_6 出发,要到达吸收状态节点 S_7 只需要一步,即利用主机 H_5 的漏洞,而该漏洞的可利用得分为 10,从实际情况进行分析,攻击者对于利用一个可利用性得分为 10 的漏洞应该具有极高的成功率,按照本文的概率计算方法,利用成功的概率为 0.9091,这是合理的,而文献[14]中的方法计算的概率为 0.5,是偏小的,与实际情况不符。

结束语 本文针对基于吸收 Markov 链的入侵路径分析方法中存在的状态转移情形考虑不全面和状态转移概率计算不合理的问题,通过对所有非吸收节点设置指向自身的状态转移边,提出了一种改进的状态转移概率计算方法,使得状态转移概率值更符合实际。通过实验结果分析,本文方法在节点威胁度排序和入侵路径长度期望的计算上都更加符合网络攻防对抗的实际情况。

网络环境复杂且多变,下一步的工作将重点研究入侵路径的动态分析,建立更加完善的入侵路径分析模型。

参 考 文 献

[1] HUANG Y H, WU Y F, YANG H P, 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.

[2] YU D, FRINCKE D. Improving the quality of alerts and predicting intruder's next goal with Hidden Colored Petri-Net[J]. Computer Networks, 2007, 51(3): 632-654.

[3] WANG S, TANG G M, KOU G, et al. Attack path prediction method based on causal knowledge net[J]. Journal on Communications, 2016(10): 198.

[4] LIU Y L, FENG D G, LIAN Y F, et al. Network Situation Prediction Method Based on Spatial-Time Dimension Analysis[J]. Journal of Computer Research and Development, 2014, 51(8): 1681-1694.

[5] ZENG S W, WEN Z H, DAI L W, et al. Analysis of Network Security Based on Uncertain Attack Graph Path[J]. Computer Science, 2017, 44(S1): 361-365.

[6] KAYNAR K. A taxonomy for attack graph generation and usage in network security[J]. Journal of Information Security and Applications, 2016, 29: 27-56.

[7] ZENG J, WU S, CHEN Y, et al. Survey of Attack Graph Analysis Methods from the Perspective of Data and Knowledge Processing[J]. Security & Communication Networks, 2019, 2019(11): 1-16.

[8] SHEYNER O, HAINES J, JHA S, et al. Automated Generation and Analysis of Attack Graphs [C]// Proceedings 2002 IEEE Symposium on Security and Privacy. 2004.

[9] WANG S, ZHANG Z, KADOBAYASHI Y. Exploring attack graph for cost-benefit security hardening: A probabilistic approach[J]. Computers & Security, 2013, 32: 158-169.

[10] 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.

[11] ABRAHAM S, NAIR S. Cyber security analytics: a stochastic model for security quantification using absorbing markov chains [J]. Journal of Communications, 2014, 9(12): 899-907.

[12] ABRAHAM S, NAIR S. A Predictive Framework for Cyber Security Analytics using Attack Graphs[J]. International Journal of Computer Networks & Communications, 2015, 7(1).

[13] FREI S. Security econometrics : The dynamics of (in) security [M]. BookSurge Publishing, 2009.

[14] HU H, LIU Y L, ZHANG H Q, et al. Route Prediction Method for Network Intrusion Using Absorbing Markov Chain[J]. Journal of Computer Research and Development, 2018, 55(4): 831-845.

[15] 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.

[16] GRINSTEAD C M, SNELL J L. Introduction to probability [M]. American Mathematical Soc., 2012.

[17] OU X, GOVINDAVAJHALA S, APPEL A W. MulVAL: A Logic-based Network Security Analyzer[C]// USENIX security symposium. 2005, 8: 113-128.

[18] YOUSEFI M, MTETWA N, ZHANG Y, et al. A novel approach for analysis of attack graph[C]// IEEE International Conference on Intelligence and Security Informatics (ISI). IEEE, 2017: 7-12.

ZHANG Kai, born in 1992, postgraduate. His main research interests include network security situation awareness and so on.

LIU Jing-ju, born in 1974, professor. Her main research interests include network security situation awareness and network security detection.