第 46 卷 第 10 期 计 算 机 工 程 2020 年 10 月
Vol. 46 No. 10 Computer Engineering October 2020

·网络空间安全· 文章编号：1000-3428(2020)10-0143-08 文献标志码：A 中图分类号：TP301.6

基于 SQAG 模型的攻击熵优化算法

张 俊，张安康，王 辉

（河南理工大学 计算机科学与技术学院，河南 焦作 454000）

摘 要： 为降低网络安全风险，更好地实现网络攻击路径的优化，在现有网络攻击图的基础上构建 SQAG 模型对网络攻击进行建模。该模型将攻击过程离散化，每一时刻的攻击图包含攻击者在当前时刻已经占据的节点。同时利用攻击熵优化算法对子攻击路径进行成本收益分析，从而消除冗余路径。通过对攻击过程进行合理的推演，将精确推理的联结树算法应用到时序网络攻击图中，实时得到任意时刻攻击图的节点置信度。实验结果表明，在防火墙收紧访问尺度情况下，该模型网络攻击节点置信度随时间延长而降低，利用攻击熵优化算法消除冗余路径，可得到更准确的节点置信度。

关键词： SQAG 模型；置信度；攻击熵；成本收益分析；联结树算法

开放科学（资源服务）标志码（OSID）：

中文引用格式： 张俊，张安康，王辉. 基于 SQAG 模型的攻击熵优化算法[J]. 计算机工程，2020，46(10)：143-150.

英文引用格式： ZHANG Jun，ZHANG Ankang，WANG Hui. Attack entropy optimization algorithm based on SQAG model[J]. Computer Engineering，2020，46(10)：143-150.

Attack Entropy Optimization Algorithm Based on SQAG Model

ZHANG Jun，ZHANG Ankang，WANG Hui

（School of Computer Science and Technology，Henan Polytechnic University，Jiaozuo，Henan 454000，China）

【Abstract】 In order to reduce network security risks and better realize the optimization of network attack paths，this paper constructs a SQAG model for network attacks based on the existing network attack graphs. The model discretizes the attack process，in which the attack graph at each moment contains the nodes occupied by the attacker at that time. The attack entropy optimization algorithm is used to implement cost-benefit analysis of sub-attack paths，so as to reasonably eliminate redundant paths. Through reasonable deduction of the attack process，the joint tree algorithm that performs precise reasoning is applied to the sequential network attack graph to obtain the node confidence degree of the attack graph at any moment in real time. Experimental results show that when the firewall tightens the access scale，the confidence degree of each node in the proposed model decreases with time in the attack process. The redundant paths are eliminated by using the attack entropy optimization algorithm to obtain a more accurate confidence degree of nodes.

【Key words】 SQAG model；confidence degree；attack entropy；cost-benefit analysis；joint tree algorithm

DOI： 10.19678/j.issn.1000-3428.0056289

0 概述

互联网的快速发展能够满足社会各方面的现实需求，但也引发了越来越多的安全问题。根据国家互联网应急中心发布的《2018 年中国互联网网络安全报告》[1]，截至 2018 年底，国家互联网应急中心 CNCERT/CC 共接收到境内外报告的网络安全事件 106 700 起。网络安全事件的频发给人们的正常工作与生活带来很大的影响[2]。

为降低网络空间面临的安全风险，研究人员对网络攻击进行了建模。文献[3]将贝叶斯网与攻击图相结合，利用近似贝叶斯推理预测大规模网络路径。文献[4]利用贝叶斯推理对网络进行动态风险评估。文献[5]在利用贝叶斯预测攻击路径时引入时间增益补偿率。文献[6]通过贝叶斯推理直接求出所有节点的置信度。文献[7]在假定目标节点被占据的条件下求出了其余节点的置信度。文献[8]通过攻击图的置信度对网络安全进行了合理度量。

基金项目： 国家自然科学基金(61300216)。

作者简介： 张 俊(1975—)，男，副教授、博士，主研方向为网络安全；张安康，硕士研究生；王 辉(通信作者)，副教授、博士。

收稿日期： 2019-10-14 **修回日期：** 2019-11-08 **E-mail：** 710832453@qq.com

文献[9]通过加固置信度较高的节点,降低了网络系统被攻击的风险,使其更加可靠。

本文通过 SQAG 模型对网络攻击进行建模,利用攻击熵算法消除冗余路径,运用联结树算法进行精确推理,从而获得理想情况下节点置信度的预测值及消除冗余路径后节点置信度的合理值。

1 相关研究

近年来,研究人员在网络攻击图的基础上利用贝叶斯推理对攻击路径进行预测。文献[10]指出攻击图的路径生成和路径分析是攻击图的研究重点之一,明确了在当前的网络攻击图研究中,当节点数增加时路径数目的增加方式是一个 NP 难问题,但没有给出减少冗余路径的具体方法。

文献[11]提出一个面向内部攻击意图推断的概率攻击图模型,在模型中引入条件概率表来描述单步攻击检测结果的不确定,并在模型基础上给出计算攻击意图的算法以及利用累积概率来预测最有可能发生的攻击路径的方法。该文中给出的起始节点概率是静态的,如果对该模型引入时间维度,那么当防火墙收缩访问尺度时,起始节点概率值将随时间变化,模型可实现对整个攻击过程中攻击意图的预测。

文献[12]提出一种成本收益分析方法,其中风险系数随时间呈指数级降低。将这种风险系数应用在一个包含多次试探攻击的场景中时,攻击时间会变长,但在没有新的节点被占据的情况下,攻击者对网络结构的了解程度不会加深,攻击经验也不会增加,因此风险成本并不会随时间以指数级的速度降低。

文献[13]提出在满足证据变量时间偏序关系的情况下,利用贝叶斯推理计算攻击路径节点置信度的方法。该方法采用似然加权法来近似计算节点置信度,但是在把一个攻击过程分解为多个时间片段后,当每一个时间片段中的网络攻击图对节点置信度的精度要求都比较高时,似然加权法将需要较大的样本量,极大地增加了计算量。

为优化攻击路径,本文提出 SQAG 模型和攻击熵优化算法,在现有资源状态攻击图的基础上引入时间和当前时刻已经占据的节点,转换成时序贝叶斯网络攻击图模型,根据联结树算法计算出某时刻各个节点的置信度。通过计算某一时刻的攻击熵,进而计算出该时刻的攻击风险成本,对基于深度优先搜索算法得出的攻击路径进行筛选,以实现对现有的攻击路径进行优化的目的。

2 时序网络攻击图

在一个攻击者为获取目标节点而反复进行试探性攻击操作的背景下,本文构造的时序网络攻击图模型主要有 3 个目标:1)利用该模型对攻击过程进行建模,预测任意时刻的攻击路径;2)根据当前时刻的时序网络攻击图,对子攻击路径进行成本收益分析;3)使模型能够体现出攻击者在该时刻之前已经获取的网络结构、攻击经验等信息,并利用攻击熵对其进行度量。为实现这些目标,在已有模型的基础上进行扩展,添加时间和每一时刻已经占据的资源节点。本文构造的时序网络攻击图模型及其参数定义如下:

定义 1(时序网络攻击图模型) SQAG = (T,R,A,E,W,P,Π) 是与发生时刻对应的一组有向无环图集。其中:

$T=\{t_1,t_2,\cdots,t_n,n\in\mathbb{N}^+\}$ 表示攻击过程中由各个时刻组成的集合。在集合 T 中,任意 2 个时刻间的时间差相等。

$R=\{r_i\mid i=1,2,\cdots,n_r,n_r\in\mathbb{N}^+\}$ 为资源节点集合,资源节点 r_i 为布尔变量,r_i = true 表示当前攻击者占据资源节点 r_i,r_i = false 表示当前攻击者未占据资源节点 r_i,若 $R_0\subseteq R$ 为初始资源节点集合,则它是一组以一定函数来改变被攻击者占据概率的资源节点;若 $R_c\subseteq R$ 为已占据资源节点集合,则它是攻击者在多次攻击过程中曾成功占据的资源节点;若 $R_g\subseteq R$ 为目标资源节点集合,则它是攻击者试图去占据的资源节点。

$A=\{a_j\mid j=1,2,\cdots,n_a,n_a\in\mathbb{N}^+\}$ 为攻击行为节点集合,攻击节点 a_j 为布尔变量,a_j = true 表示当前攻击行为 a_j 已经发生,a_j = false 表示当前攻击行为 a_j 未发生。

$E=E_1\cup E_2$ 为关联攻击节点与资源节点的有向边集合,$E_1\subseteq R\times A$ 表示只有当攻击者成功占据资源节点之后攻击行为才能发生,$E_2\subseteq R\times A$ 表示只有当攻击行为发生之后才可以占据资源节点。

W 为攻击权集合,$\forall w\in\mathbb{W}$均与 $e\in\mathbb{E}$相关联,由二元组(c,g)表示,c 表示攻击行为 a_j 向资源状态节点 r_i 发动攻击时付出的成本,g 表示攻击成功时可以获得的收益。

$P=P_1\cup P_2\cup P_3$,其中,P_1 为初始节点发生的概率,$P_2=P(a_i=\text{true}\mid\text{Pre}(a_i)=\text{true})$,$\text{Pre}(a_i)$表示攻击事件之前的资源节点取值,$P_3=P(\text{Con}(a_j)=\text{true}\mid a_j=\text{true})$,$\text{Con}(a_j)$表示攻击事件之后的资源节点取值。

Π 表示某一时刻攻击图中某一节点的置信度,$\Pi(r_i,t_i)$表示攻击者在 t_i 时刻成功占据节点 r_i 的概率,$\Pi(a_j,t_i)$表示攻击行为 a_j 在 t_i 时刻发生的概率。

根据以上定义,可以得出 SQAG 模型。用有向边连接资源节点与攻击行为节点,用深色标注攻击者在该时刻曾经占据的资源节点,并综合考虑 and 节点和 or 节点的几种排列组合情况,生成时序网络攻击图的主干结构。根据经验定义起始节点的概率

P_1 以及条件概率 P_2 和 P_3,得到如图 1 所示的时序网络攻击图。

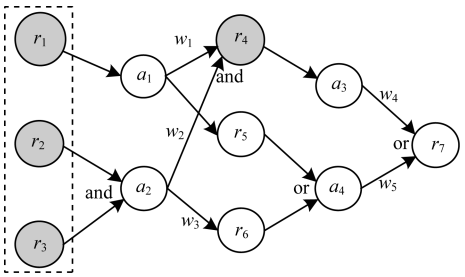


图 1　时序网络攻击图

Fig. 1　Sequential network attack graph

如图 1 所示,在 t 时刻,攻击者已经成功占据的资源节点为 $R_c=\{r_1,r_2,r_3,r_4\}$,新一轮的攻击者以概率 $P_{t_1}(R_0)$ 分别占据起始资源节点 r_1、r_2、r_3,攻击行为 a_1、a_2 以条件概率 P_2 发生,并以条件概率 P_3 分别占据相对应的资源节点 r_4、r_5、r_6,最终向目标节点 r_7 靠进。

3　基于 SQAG 模型的攻击路径

攻击者在占据系统内全部目标资源节点前,从起始节点开始持续进行攻击,若攻击被阻断,则回到起始节点开始新一轮的攻击。为了能够直观地描述攻击演进过程,在时序网络攻击图的基础上定义攻击路径。

定义 2(攻击路径)　在 t 时刻,对于任意一个目标资源节点 $r_n\in R_g$,从起始节点 R_0 开始,存在一组由资源节点和攻击节点交替排列的序列 $r_0,a_1,r_1,a_2,\cdots,a_j,r_i,\cdots,r_n$,使得节点序列中任意 2 个相邻节点间总有 $<r_i,a_j>\in E_1$ 或 $<a_j,r_i>\in E_2(0<i,j<n)$,则记节点序列 $a_1,r_1,a_2,\cdots,a_j,r_i,\cdots,r_n$ 为对应时刻的攻击路径,记为STEP_{t-x}。

以图 1 为例,攻击者可以从起始节点 r_2 和 r_3 出发,经过 a_2、r_6、a_4 最终到达目标节点 r_7,其中由 $r_2\cap r_3$、a_2、r_6、a_4、r_7 按发生顺序组成的节点序列即为一条攻击路径(符号 $\cap$ 表示节点间的 and 关系)。为便于表示攻击路径所包含的元素及其之间的关系,用某时刻的线序关系集来表示攻击路径,即 $\mathrm{STEP}_{t-x}=\{<r_0,a_1>,<a_1,r_1>,\cdots,<a_j,r_i>,\cdots,<a_n,r_n>\}$,其中任意一个线序关系表示一条子攻击路径。在图 1 中,t 时刻通向节点 $R_g=\{r_7\}$ 的 3 条攻击路径如下:

$$\mathrm{STEP}_{t-1}=\{<r_1,a_1>,<a_1,r_5><r_5,a_4>,<a_4,r_7>\}$$

$$\mathrm{STEP}_{t-2}=\{<r_2\cap r_3,a_2>,<a_2,r_6>,<r_6,a_4>,<a_4,r_7>\}$$

$$\mathrm{STEP}_{t-3}=\{<r_2\cap r_3,a_2>,<a_2\cap<r_1,a_1>,r_4>,<r_4,a_3>,<a_3,r_7>\}$$

4　攻击熵优化算法

攻击者在时序网络攻击图中任意时刻有多种占据目标节点的路径方式。当某一时刻时序网络攻击图的节点增多时,通向目标节点的攻击路径数量会呈指数级增长。为有效预测实际情况下可能出现的攻击路径,本文定义了攻击熵,用以度量攻击者对网络结构的了解程度和所获取的攻击经验。利用攻击熵计算风险成本,进而将某一时刻下子攻击路径的成本值和收益值进行对比,达到优化攻击路径的目的。

4.1　攻击熵的定义

定义 3(攻击熵)　攻击熵是用来描述攻击者对网络结构了解程度和攻击经验平均不确定性的变量,结合时序网络攻击图模型,推导其数学表达式如下:

设 $P(\mathrm{Attack}_{t-r})$ 是当前时刻未占据节点的联合概率分布,表示为:

$$P(\mathrm{Attack}_{t-r})=P_r(X=\mathrm{Attack}_{t-r})=\frac{1}{n_r}$$

其中,Attack_{t-r} 是取值于离散联合分布集 $R-R_c$ 上的随机变量,$r\in\{1,2,\cdots,n_r\}$,n_r 为该时刻未占据节点的数目。

在此基础上,定义攻击熵如下:

$$H(X)=-\sum_{n=1}^{n_r}P(\mathrm{Attack}_{t-r})\times\mathrm{In}(P(\mathrm{Attack}_{t-r}))$$

随着攻击的演进,攻击图中未被占领的资源节点数目减少,而攻击者对网络结构了解加深,获得的攻击经验增加,这使得信息的不确定性降低,表现为攻击熵的数值降低,计算得出的风险成本也降低。

4.2　子攻击路径的成本分析

子攻击路径的成本由风险成本和操作成本两部分构成。文献[14]通过对子攻击路径的发生原因进行分析,提出操作成本 $\cos t(e)$ 由元操作成本 $\cos t$(meta-operation)和操作序列成本 $\cos t$(sequence)两部分组成,表示如下:

$$\cos t(e)=\alpha\times\cos t(\text{meta-operation})+\beta\times\cos t(\text{sequence})\quad(1)$$

其中,e 表示子攻击路径 $<a_j,r_i>$。

除了攻击操作需要花费成本外,攻击者在实施网络攻击的过程中还要面临被安全管理人员发现的风险,即需要承担相应的风险成本。因此,在分析子攻击路径成本时,也要考虑风险因素对子攻击路径是否发生影响。

风险系数(用 θ 表示)是用来衡量某次攻击行为在占领资源节点时被发现的可能性大小的一种度量,它一方面取决于攻击目标对攻击行为的影响系数(用 M 表示),攻击目标越重要,管理者对其越重视,攻击行为就越容易被检测到;另一方面取决于攻

击行为自身的影响系数(用 Γ 表示),攻击行为越复杂,被发现的可能性就越高;另外还取决于攻击熵,攻击熵越低,攻击者规避风险的能力越强。因此,可以把风险系数 θ 定义为:

$$\theta(e)=\Gamma(a_j)\times M(r_i)\times\frac{H(X)}{H_{\max}(X)} \tag{2}$$

归一后的攻击熵 $\frac{H(X)}{H_{\max}(X)}$ 值在 0 ~ 1 之间,由式(2)易知引入攻击熵后风险系数变小,符合攻击者在多次试探性攻击之后对网络结构了解加深,攻击经验增加的实际情况。

被攻击目标 r_i 越重要,攻击者的操作 a_j 复杂度越高,攻击熵越大,从而暴露的风险也越大。根据以上分析,风险成本 $\cos t(r)$ 可表示如下:

$$\cos t(r)=\cos t(e)\times\theta(e) \tag{3}$$

从式(3)可以看出,风险成本随着风险系数的增长而增长,这符合实际攻击中风险成本的增长趋势。进而子攻击路径的攻击成本可用下式计算:

$$\cos t(e_{ji})=\varepsilon\times\cos t(e)+\mu\times\cos t(r) \tag{4}$$

其中,ε 和 μ 分别表示两种成本的权重。

在某一时刻,当子攻击路径上获得的收益大于成本时,认为该条子攻击路径是可行的,否则是不可行的。

4.3 攻击熵优化算法

为解决时序网络攻击图中攻击路径随资源节点个数增加呈指数级增长的问题,本文运用攻击熵优化算法对子路径进行成本收益分析,以减少冗余路径,缩小网络攻击图的规模,具体算法如下:

算法 1 攻击熵优化算法

输入 TSAG

输出 $STEP_{t-x}$

1. n = n_r + n_a
2. Init;matric;G. arc[n,n];G. rel[n,n];G. visited[n];
3. for each $r_i \in Rora_i \in A$
4. if a_i,r_j 之间存在有向边 e
5. G. arc[i,j] = 1
6. if r_i,r_j 之间或者 a_i,a_j 之间存在 and 关系
7. G. rel[i,j] = 1
8. if r_i,r_j 之间或者 a_i,a_j 之间存在 or 关系
9. G. rel[i,j] = 2
10. end for
11. //用矩阵存储 SQAG 模型的攻击关系与参数
12. for each $r_i \in R_0$
13. DFS(G,i)
14. G. visited[i] = true;
15. 输出 G. visited[i];
16. for each j < n
17. if G. rel[i,j] = 1
18. 输出 G. visited[j];
19. //输出当前访问节点
20. Run ACCA
21. c = cost(e_{ji})
22. if(G. arc[i,j] = =1 && ! G. visited[j]&&(c < g))
23. //进行成本收益分析
24. i = j goto 第 11 行
25. end for
26. end for

在消除冗余路径的过程中,为利用攻击熵合理计算出风险成本,本文提出一种攻击成本计算算法,具体步骤如下:

算法 2 攻击成本计算算法

输入 SQAG

输出 $\cos t(e_{ji})$

1. Init a = 0;b = 0;
2. for each $r_i \in R$
3. if $r_i \notin R_c$
4. a = a + 1;
5. else b = b + 1;
6. end for
7. //计算当前未占据节点和已占据节点
8. $P(Attack_{t-1;a})=\frac{1}{a}$

   $P_{max}(Attack_{t-1;a+b})=\frac{1}{a+b}$
9. $H(X)=-\sum_{n=1}^{a}P(Attack_{t-r})\times In(P(Attack_{t-r}))$

   $H_{max}(X)=-\sum_{n=1}^{a+b}P_{max}(Attack_{t-r})\times In(P_{max}(Attack_{t-r}))$
10. //计算当前时刻攻击熵和最大攻击熵
11. $\theta(e)=\Gamma(a_j)\times M(r_i)\times\frac{H(X)}{H_{max}(X)}$
12. $\cos t(r)=cost(e)\times\theta(e)$
13. $\cos t(e_{ji})=\varepsilon\times cost(e)+\mu\times cost(r)$

算法 1 存储了当前时刻攻击图的信息之后,计算出未占据的节点数,通过调用算法 2 计算出攻击熵,对其进行归一化,将资源重要性程度影响系数、攻击行为自身影响系数、归一化后的攻击熵累乘计算出风险系数,进而计算出合理的风险成本,加上操作成本后得出攻击成本,进而在算法 1 中进行成本收益分析,消除冗余路径。攻击熵优化算法利用攻击熵这一参数将攻击者在攻击过程中已经获取的网络结构信息和攻击经验包含在内,使得风险成本的计算比较合理。

5 时序网络攻击图中节点置信度的计算

在时序网络攻击图模型的基础上,利用贝叶斯推理计算节点置信度。相关学者采用似然加权法抽样得出节点置信度的近似值[15-16],但是精度要求高时,似然加权法需要大量的抽样。为能够快速地得到每一时刻各个攻击图各条路径上的节点置信度,本文选择了精确推理中时间效率最高的算法,即联结树算法[17]。

定义 4(团集) 团集(Clique,C)是由 SQAG 模型中某一时刻攻击图中 3 个节点组成集合的集合,其中 3 个节点组成的集合称为团。

定义 5(联结树)　联结树(Joint Tree,JT)是由 SQAG 模型的团集 C 中的元素 C_i和有向边集合 S 中的元素 S_j连接成的树状结构。其中,团集和边集满足任何 2 个团 C_i和 C_j之间路径上的每个边包含于边集 S 中,相邻 2 个点之间的边 $S_{ij}=C_i\cap C_j$。令团$C_1=\{a_3,a_4,r_7\}$,团 $C_2=\{a_3,a_4,a_5\}$,则 S_{12}为边 $<a_3,a_4>$。

联结树算法步骤如下:

步骤 1　将时序网络攻击图转换为联结树

为将 TSAG 模型上的推理转换成 JT 中的推理,需要将对应的时序网络攻击图 2(a)转变为联结树 JT = (C,S),将转换过程分解为以下 3 步:

1)建立 Moral 图

(1)找出时序网络攻击图中每一个资源节点和攻击节点的父节点。

(2)用无向边将父节点两两相连。

(3)将有向边改为无向边。

如图 2(a)所示的时序网络攻击图,其 Moral 图为图 2(b)。

2)三角化 Moral 图

将 Moral 图中每一个大于或等于 4 的环的非相邻节点用无向边连接起来,实现 Moral 图的三角化。

如图 2(b)所示的 Moral 图,其三角化后的 Moral 图为图 2(c)。

3)建立联结树

(1)确定所有的团集,其中团集是三角化后的 Moral 图中最大全连通图组成的集合,团集中每对不同的团都由无向边连接。

(2)在团集中添加一些边和分割节点构造出一棵联结树 T。

图 2(c)为三角化后的 Moral 图,其联结树如图 2(d)所示。

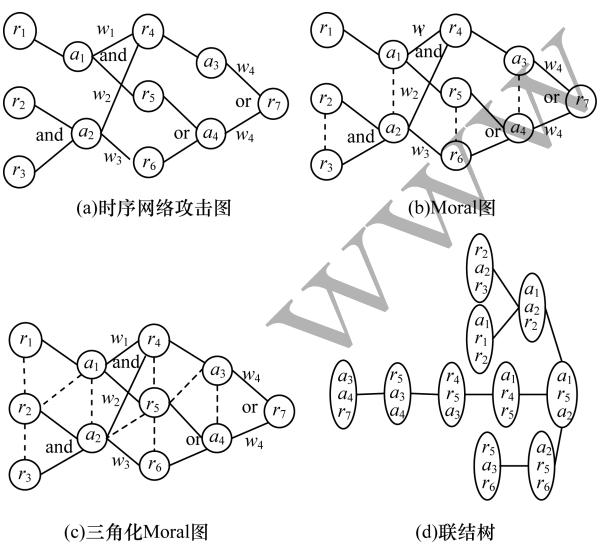


图 2　时序网络攻击图转换成联接树的过程

Fig. 2　Process of converting sequential network attack graph into joint tree

步骤 2　初始化联结树

联结树的初始化是对所有节点给定发生与否的概率,进而得到每个团的分布函数。由时序网络攻击图构建的联结树中的团 C_i由 3 个节点组成,其中每一个节点有 2 种状态,则共有 8 种状态组合。令 Φ_i 表示团 C_i的分布函数,Φ_{ij}代表团 C_i第 j 个状态组合的分布函数,初始化如下:

for each $\Phi_{ij}=1$

步骤 3　消息传递

消息传递通过更新任意 2 个团之间的联合概率分布,使得联结树达到全局一致状态。如图 3 所示是相邻团的节点间进行的一次消息传递[18]。

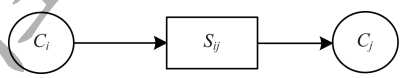


图 3　一次消息传递过程

Fig. 3　Process of a message delivery

从团 C_i到团 C_j进行一次消息传递包括以下 3 步:

1)产生消息:

$$\Phi_S^{\text{new}}=\sum_{C_i\setminus S}\Phi_{C_i} \tag{5}$$

2)吸收信息,更新团节点的分布函数 Φ_{C_j}:

$$\Phi_{C_j}=\frac{\Phi_S^{\text{new}}}{\Phi_S}\Phi_{C_j} \tag{6}$$

3)更新分隔节点的分布函数:

$$\Phi_S=\Phi_S^{\text{new}} \tag{7}$$

步骤 4　概率计算

概率计算即是计算 SQAG 模型中任意节点发生的概率。通过概率计算,可以得到当前节点的置信度。在任意一个包含节点 V 的团 C_i中,通过 $P(V)=\sum_{C\setminus(V)}\Phi_C$ 可计算出节点 V 的分布[19]。

步骤 5　条件概率计算

条件概率计算即是在 SQAG 模型中某些观测到的节点值为 True 的条件下,计算另外的某个节点发生的条件概率。通过对联结树进行消息传递得到全局一致的联结树[20]。当联结树加入条件 e 后达到全局一致时,对任意的团节点 C 有 $\Phi_C=P(C,e)$,e 表示加入的条件。

$$P(V\mid e)=\frac{P(V,e)}{P(e)}=\frac{P(V,e)}{\sum_V P(V,e)} \tag{8}$$

将某时刻的时序网络攻击图转换为联结树后,对联结树进行赋值,得到一个带有状态组合分布函数的联结树。通过团节点之间的消息传递,联结树达到全局一致。在这种状态下,可得出任意节点的置信度。

6 算法仿真与验证

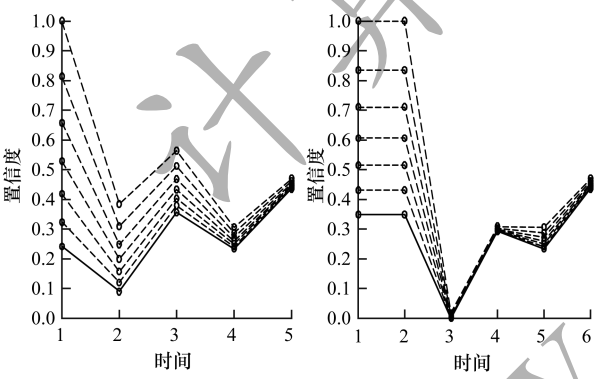
在一个攻击者为占据所有目标节点而不断反复从起始节点进行攻击的场景中,为证明该模型能够较好地模拟出网络攻击节点置信度随时间变化,本文设计如下实验:首先在 matlab 软件上实现联结树算法,建立防火墙随时间收紧情况下的模型,结合联结树算法得出各条路径上节点置信度随时间的变化;其次在假设子攻击路径参数(收益、操作成本系数、风险成本系数)为定值的情况下,计算出某时刻优化后的攻击路径及各节点置信度。

以图 1 的时序网络攻击图为例,利用联结树算法推理计算节点的置信度,其中,条件概率 $P_2=0.7$, $P_3=0.5$。随着攻击的进行,防火墙会收缩访问尺度,故可根据经验设定起始节点概率:

$$P(v_i \in R_0)=1-0.1\times t_i,$$

$$i=1,2,3,t_i=0,1,2,3,4,5,6$$

以 STEP_{t-1} 和 STEP_{t-2} 为例,画出在满足上述假设的情况下 7 个时刻各条攻击路径上节点发生的置信度,如图 4 所示,其中图 4(a)和图 4(b)攻击路径节点分别为 r_1、a_1、r_5、a_4、r_7 和 r_2、r_3、a_2、r_6、r_7。



(a)STEP_{t-1}的节点置信度变化 (b)STEP_{t-2}的节点置信度变化

图 4 节点置信度随时间的变化过程

Fig. 4 Change process of node confidence degree with time

在图 4(a)和图 4(b)中,随着防火墙收缩访问尺度,各条折线上从高到低的圆点分别代表 STEP_{t-1} 和 STEP_{t-2} 上 7 个时刻各节点置信度随时间均呈下降趋势。可以看出,随着攻击的演进,防火墙收缩访问尺度,起始节点发生概率降低,导致该时刻此条路径上的所有节点置信度都相应地降低。以上分析没有考虑随着攻击熵的增加,某些子攻击路径的风险成本增加,导致攻击路径变化。

计算各个攻击阶段下对应 Attack_{t-r} 的攻击熵,如图 5 所示。

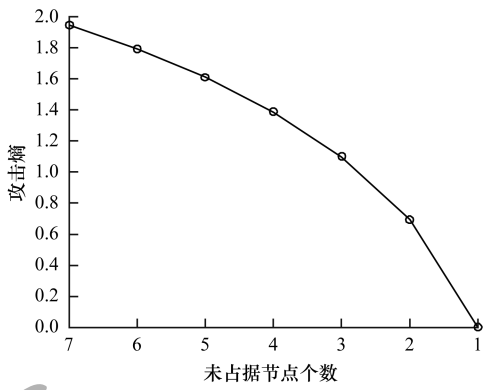


图 5 攻击熵与未占据节点个数的关系

Fig. 5 Relationship of attack entropy and number unoccupied nodes

经过长时间的攻击之后,攻击者未占据的节点随时间逐渐减少,攻击者对网络系统不确定性的了解程度降低,攻击经验增加,攻击熵也随之降低。

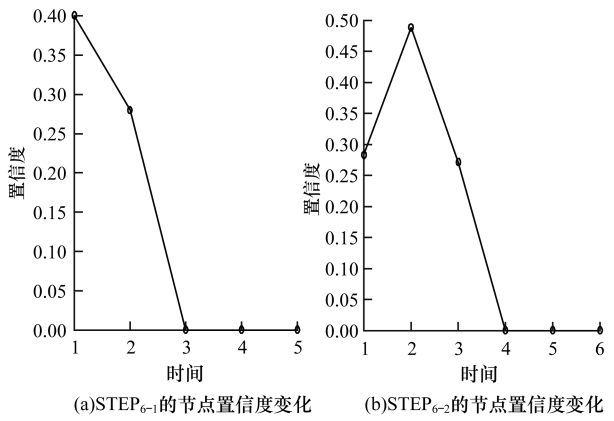
根据经验,设定图 1 中的网络系统某时刻攻击权对应的子攻击路径上的攻击参数如表 1 所示。运行算法 1 和算法 2 的代码可得,攻击权 w_1 对应的子攻击路径上的成本小于收益,而攻击权 w_2、w_3、w_4、w_5 对应的子攻击路径上的成本大于收益,从而可知利用算法 1 和算法 2 可合理删除攻击者认为不可能发生的子攻击路径,证明了该算法可以实现路径优化。

表 1 $t_i=6$ 时子攻击路径上的攻击参数

Table 1 Attack parameters on sub-attack path at $t_i=6$

参数	w_1	w_2	w_3	w_4	w_5
攻击收益 g	2	2	2	2	2
节点影响系数 M	0.5	0.6	0.7	0.9	0.9
自身影响系数 Γ	0.8	0.7	0.6	0.5	0.4
操作成本 $\cos t(e)$	1.3	1.4	1.5	1.6	1.7
未占据节点 n_r	7	6	5	4	3
操作成本系数 ε	1	1	1	1	1
风险成本系数 μ	1	1	1	1	1
c 是否大于 g	N	Y	Y	Y	Y

为进一步描述优化后的攻击路径,利用联结树算法进行推理,计算 STEP_{6-1} 和 STEP_{6-2} 上各节点置信度如图 6 所示,其中图 6(a)和图 6(b)攻击路径节点分别为 r_1、a_1、r_5、a_4、r_7 和 r_2、r_3、a_2、r_6、r_7。由图 6 可以看出,STEP_{6-1} 可行的攻击路径有 $\{<r_1,a_1>\}$,STEP_{6-2} 可行的攻击路径有 $\{<r_2\cap r_3,a_2>\}$,与图 4 推演出的攻击路径对比可知,利用攻击熵优化算法消除了冗余子攻击路径,进而利用联结树算法推理优化后的节点置信度更加接近实际情况。

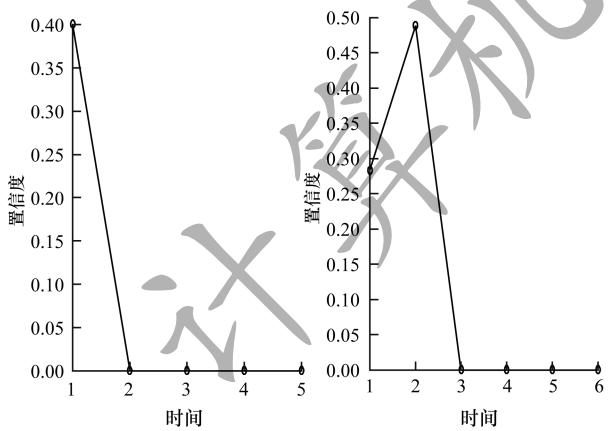


(a)$STEP_{6-1}$的节点置信度变化　(b)$STEP_{6-2}$的节点置信度变化

图 6　$STEP_{6-1}$和 $STEP_{6-2}$上的节点置信度

Fig. 6　Node confidence degree on $STEP_{6-1}$ and $STEP_{6-2}$

对于未使用攻击熵优化算法的情况,即此时风险系数的计算方式为 $\theta(e)=\Gamma(a_j)\times M(r_i)$,进行优化后的攻击路径及其节点置信度如图 7 所示 ,其中图 7(a)和图 7(b)攻击路径节点分别为 r_1、a_1、r_5、a_4、r_7 和 r_2、r_3、a_2、r_6、r_7。



(a)$STEP_{6-1}$的节点置信度变化　(b)$STEP_{6-2}$的节点置信度变化

图 7　未运用攻击熵时 $STEP_{6-1}$和 $STEP_{6-2}$上的节点置信度

Fig. 7　Node confidence degree on $STEP_{6-1}$ and $STEP_{6-2}$ without attack entropy

对比图 6 和图 7 可知,图 6 中 $STEP_{6-1}$上 a_1 和 $STEP_{6-2}$上 a_2 置信度都不为 0,而在图 7 中它们的置信度为 0。这是由于攻击熵的引入使得在计算风险成本时,将攻击者在多次试探性攻击的过程中所获取的攻击经验和对网络结构的了解程度包含在内,导致攻击者的风险成本变低,故可能发生的子攻击路径变多。由图 6 和图 7 对比可以看出,引入攻击熵算法可合理消除冗余路径。

7　结束语

对网络攻击进行合理建模并实时高效地预测攻击路径是网络安全领域研究的热点。本文在已有研究基础上,提出时序攻击图模型和攻击熵优化算法,通过定义在时序网络攻击图中的攻击熵概念,描述攻击者对网络结构了解程度的加深与攻击经验的增加对风险成本的影响,利用攻击熵计算风险成本,合理地消除时序网络攻击图中的冗余路径。在使用攻击熵优化算法合理减少冗余路径后,应用精确推理中的联结树算法得到节点置信度的精确值,进而对攻击路径进行实时预测。本文所提算法没有考虑防御者对攻击方式的了解程序,因此在进行成本收益分析时,不能充分体现双方的博弈过程,下一步将对此进行改进。

参考文献

[1] State Internet Center. 2018 China Internet network security report[EB/OL]. [2019-09-01]. https://www.cert.org.cn/publish/main/46/index.html. (in Chinese)
国家互联网中心.《2018 年中国互联网网络安全报告》[EB/OL]. [2019-09-01]. https://www.cert.org.cn/publish/main/46/index.html.

[2] ABRAHAM S, NAIR S. A predictive framework for cyber security analytics using attackgraphs[J]. International Journal of Computer Networks & Communications, 2015, 7(1):1-17.

[3] JIA Wei, LIAN Yifeng, FENG Dengguo, et al. Bayesian-network-approximate-reasoning-based method for network vulnerabilities evaluation[J]. Journal on Communications, 2008, 29(10):191-198. (in Chinese)
贾炜,连一峰,冯登国,等. 基于贝叶斯网络近似推理的网络脆弱性评估方法[J]. 通信学报,2008,29(10):191-198.

[4] POOLSAPPASIT N, DEWRI R, RAY I. Dynamic security risk management using Bayesian attack graphs[M]. [S.1.]: IEEE Computer Society Press, 2012.

[5] WANG Hui, WANG Yincheng, LU Shikai. Attack path optimization algorithm based on time gain compensation rate[J]. Computer Engineering, 2018, 44(8):190-197, 204. (in Chinese)
王辉,王银城,鹿士凯. 基于时间增益补偿率的攻击路径优化算法[J]. 计算机工程,2018,44(8):190-197,204.

[6] XIE P, LI J H, OU X, et al. Using Bayesian networks for cyber security analysis[C]//Proceedings of 2010 IEEE/IFIP International Conference on Dependable Systems and Networks. Washington D. C., USA: IEEE Press, 2010:211-220.

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

[8] XIAO Daoju, YANG Sujuan, ZHOU Kaifeng, et al. A study of evaluation model for network security[J]. Journal of Huazhong University of Science and Technology (Nature Science), 2002, 30(4):37-39. (in Chinese)

肖道举,杨素娟,周开锋,等. 网络安全评估模型研究[J]. 华中科技大学学报(自然科学版),2002,30(4):37-39.

[9] JIANG Wei,FANG Binxing,TIAN Zhihong,et al. Evaluating network security and optimal active defense based on attack-defense game model[J]. Chinese Journal of Computers,2009,32(4):817-827. (in Chinese)
姜伟,方滨兴,田志宏,等. 基于攻防博弈模型的网络安全测评和最优主动防御[J]. 计算机学报,2009,32(4):817-827.

[10] KAYNAR K. A taxonomy for attack graph generation and usage in network security[J]. Journal of Information Security and Applications,2016,29(4):27-56.

[11] CHEN Xiaojun, FANG Binxing, TAN Qingfeng, et al. Inferring attack intent of malicious insider based on probabilistic attack graph model[J]. Chinese Journal of Computers,2014,37(1):62-72. (in Chinese)
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.

[12] WANG Hui, WANG Yunfeng, WANG Kunfu. Research on predicting attack path based on Bayesian inference[J]. Application Research of Computers,2015,32(1):226-231. (in Chinese)
王辉,王云峰,王坤福. 基于贝叶斯推理的攻击路径预测研究[J]. 计算机应用研究,2015,32(1):226-231.

[13] ZHANG Shaojun, LI Jianhua, SONG Shanshan, et al. Using Bayesian inference for computing attack graph node beliefs[J]. Journal of Software, 2010, 21(9): 2376-2386. (in Chinese)
张少俊,李建华,宋珊珊,等. 贝叶斯推理在攻击图节点置信度计算中的应用[J]. 软件学报,2010,21(9):2376-2386.

[14] WANG Hui, LIU Shufen. A scalable predicting model for insider threat[J]. Chinese Journal of Computers, 2006,29(8):1346-1355. (in Chinese)
王辉,刘淑芬. 一种可扩展的内部威胁预测模型[J]. 计算机学报,2006,29(8):1346-1355.

[15] GHASEMIGOL M, GHAEMI-BAFGHI A, TAKABI H. A comprehensive approach for network attack forecasting[J]. Computers & Security,2015,58:83-105.

[16] LI Heng, WANG Yongjun, CAO Yuan. Searching forward complete attack graph generation algorithm based on hypergraph partitioning[J]. Procedia Computer Science,2017,107:27-38.

[17] HU Xiaojian, YANG Shanlin, MA Shanxi. Inference structure and construction algorithms of Bayesian network based on junction tree[J]. Journal of System Simulation,2004,16(11):2559-2563. (in Chinese)
胡小建,杨善林,马溪骏. 基于联结树的贝叶斯网的推理结构及构造算法[J]. 系统仿真学报,2004,16(11):2559-2563.

[18] LIU Junna. Research on inference algorithm in Bayesian network[D]. Hefei: Hefei University of Technology, 2007. (in Chinese)
刘俊娜. 贝叶斯网络推理算法研究[D]. 合肥:合肥工业大学,2007.

[19] ZHU Mingmin. Research on Bayesian network structure learning and reasoning[D]. Xi'an: Xidian University, 2013. (in Chinese)
朱明敏. 贝叶斯网络结构学习与推理研究[D]. 西安:西安电子科技大学,2013.

[20] LIU Zhen, TAN Liang, ZHOU Mingtian. Bayesian inference based on global message propagation[J]. Computer Science,2006,33(9):166-168. (in Chinese)
刘震,谭良,周明天. 基于全局消息传播的贝叶斯推理[J]. 计算机科学,2006,33(9):166-168.

编辑 索书志