文章编号:1671-6833(2022)01-0090-07

基于非零和信号博弈的主动防御模型

黄万伟[1], 袁 博[1], 王苏南[2], 张校辉[3]

(1.郑州轻工业大学 软件学院,河南 郑州 450001; 2.深圳职业技术学院 电子与通信工程学院,广东 深圳 518005; 3.河南信安通信技术股份有限公司,河南 郑州 450001)

摘 要: 近几年以 APT 为代表的网络攻击危害日趋严重,现有的基于信号博弈理论的研究虽然可以在一定程度上模拟 APT 攻防过程,但忽略了攻防过程中双方收益不对等的现象以及多阶段的对抗过程,导致模型和方法缺乏普适性。为此提出一个基于非零和信号博弈的主动防御模型,依据信号博弈理论,结合分析网络攻防多阶段的对抗过程建立攻防博弈树;收益量化过程中基于收益不对等的情况,采用非零和方法与贴现因子对攻防过程中多阶段的收益情况进行建模;提出了符合网络攻防特征的量化方法,并通过分析模型中存在的纳什均衡和精炼贝叶斯均衡,得出当前最优防御策略算法。通过仿真实验对该模型和方法进行验证,结果证实了所提模型和方法的可行性和有效性。

关键词: 非零和; 信号博弈; 贴现因子; 最优防御策略

中图分类号: TP309 **文献标志码**: A **doi**:10.13705/j.issn.1671-6833.2021.05.010

0 引言

在过去几年里,以高级持续性威胁(advanced persistent threat, APT)为代表的网络攻击对网络信息安全造成了严重的威胁,国家和企业面临着严峻的考验[1-2]。APT 攻击带有明确的目的性,攻击方式复杂多变,对目标系统实施有计划、全方位的入侵,以达到最终目的[3]。传统的被动防御技术只能在破坏活动发生后采取降低破坏程度的补救措施,无法应对当前愈来愈灵活的 APT 攻击[4],因此,亟须一种有一定预测能力的主动防御方法,在 APT 造成重大损失之前将损失降到最低。

主动防御技术具有灵活多变的特点,是解决当前网络安全问题的有效方法之一[5],其中应用于主动防御技术中的博弈论是解决 APT 问题的有效指导理论[6]。博弈建模需要解决两个问题[7]:如何准确描述攻防双方的策略选择和行动顺序;如何对攻防过程进行量化。针对以上问题,研究者已经做出了一些研究。孙文君等[8]提出一种基于 FlipIt 模型的攻防博弈模型;Fang 等[9]提出了一种面向 APT 的攻击路径预测模型,能在一定程度上预测下一步攻击行动。但以上博弈过程是在信息可靠的基础上提出来的,在实际应用过程中有很大的局限性。姜伟等[10]提出了攻防策略分类及其成本量化方法,随后在文献[11]中构建了非合作两角色零和随机博弈模型,然而在攻防收益量化的过程中采用零和方法[12],并不符合实际攻防情况。针对零和方法存在的问题,孙骞等[13]和李静轩等[14]构建了基于非零和的博弈模型,并给出了策略选取方法,而该方法没有考虑到双方信息不完全对称的问题。Yang[15]提出基于攻防信号的博弈模型,在建模时选择攻击者作为信号的释放者,但实际攻防过程中防御者很难获取该信号,缺乏可行性。针对文献[15]存在的问题,张恒巍等[16-17]建立了基于信号博弈的主动防御模型,并选取了防御方作为信号的释放者。同时,张为等[18]从防御角度构建博弈模型,在一定程度上可以增强系统的防御能力,却不能很好地体现出多阶段信号博弈的收益。

针对上述问题,本文基于信号博弈理论与非零和思想提出了一种基于非零和信号博弈的主动防御模型,选择防御者作为信号的释放者,通过释放混合信号来提高防御能力,提出了贴现因子来对多阶段收益的情况进行量化描述,通过求解当

收稿日期:2021-02-24;**修订日期**:2021-05-24

基金项目:河南省高等学校重点科研项目(20B520039)

作者简介:黄万伟(1979—),男,江苏盐城人,郑州轻工业大学副教授,博士,主要从事智能硬件、网络安全和大数据技术研究,E-mail: huangww79@163.com。

前最优策略算法得出本阶段最优防御策略。

1　非零和信号博弈模型

1.1　攻防博弈过程分析

网络安全包括 3 个主题:攻击者、网络系统和合法用户。合法用户只是获得网络系统的服务[19],因此本文主要考虑攻击者和网络系统间的信号博弈。在 APT 信号博弈过程中,首先由防御者释放信号,由于防御者的信息影响着攻击者的决策,因此防御者释放夹杂虚假信息的信号可以增加攻击者的攻击代价,并根据已知的攻击方式构建防御策略。攻击者在攻击前会先对目标系统进行检测,收集目标在外网中的详细信息以及防御方发出的信号,找到目标系统防御的薄弱点,生成先验概率集合,根据接收到的信号,遵循贝叶斯法则生成后验概率集合,并选择攻击策略。在下一阶段,防御者根据所预测的攻击者的攻击构建防御策略,并继续释放夹杂虚假信息的信号,攻击者根据收到的信号生成新的后验概率集合,并选择攻击策略。后面阶段的攻防过程以此类推。多阶段对抗过程如图 1 所示。

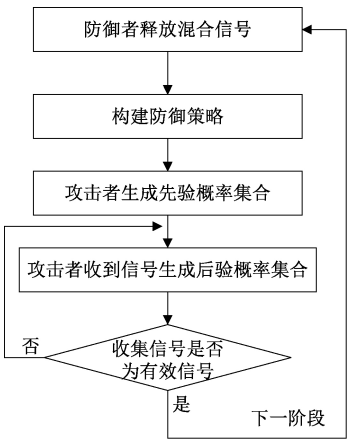

图 1　多阶段对抗过程

Figure 1　Multistage confrontation process

根据以上叙述,将其所释放的夹杂虚假信息的信号定义为混合信号。借鉴信号博弈的基本理论,其对抗过程是一个多阶段策略选择不断变化的过程:攻防双方在每阶段都会从对方释放的信号中来预测出对方在下一阶段采取的行动,并从己方最大利益出发选择最优策略,其攻防博弈树如图 2 所示。

从以上分析可以得出,双方当前的策略选择主要与上一时刻双方所释放的信号有关,其状态转移具有一定的 Markov 性。

1.2　模型的假设

通过分析网络攻防过程,本模型假设如下。

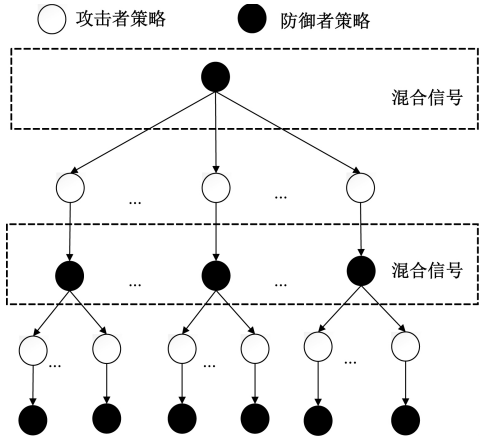

图 2　攻防博弈树

Figure 2　Attack and defense game tree

假设 1　利益最大化假设,即博弈参与双方所选择的策略都是基于实现自身收益最大化,其策略选择受其操作代价和实际条件的影响。

假设 2　状态有限性假设,即 APT 攻防过程的步骤是有限的,因为攻防双方对资源的消耗是不可逆的,双方不可能无休止地进行行动,因此整个博弈过程不存在死锁现象。

假设 3　信息可知性假设,即攻防双方基本了解对方可能进行的行动手段及其相应结果。APT 攻击是过去攻击的结合,而不是新的攻击[9],所以防御方会对攻击者可能采取的攻击行为及其结果有大概了解,而攻击者同样会对防御者可能采取的防御行为及结果基本了解。

1.3　模型的定义

基于上述 3 个假设,定义非零和信号博弈模型 $NSG=(N,S,\theta,\omega,M,P,C,R,\tau,U)$,其中:

(1) 参与者空间 $N=\{N_D,N_A\}$。N_D 代表防御者,N_A 代表攻击者,多阶段网络攻防过程也可以看作是两者之间的博弈过程。

(2) 状态空间 $S=\{S_1,S_2,\cdots,S_t \mid t\in N^+\}$。每个状态都是表示在上一阶段双方博弈后系统所处的状态,该状态只与上一状态双方选择的策略有关。

(3)类型空间 θ。θ 所代表的是防御者的类型空间,其现实意义是代表防御者的资源类型,按照防御能力可分为多种类型,即 $\theta=\{\theta_1,\theta_2,\cdots,\theta_n \mid n\in N^+\}$。

(4) 攻防策略空间 $\omega\{D,A\}$。防御者的策略空间为 $D=\{D_1,D_2,\cdots,D_i\}$,其中 D_i 表示某一防御策略,为防御动作的集合 $D_i=\{d_1,d_2,\cdots,d_m \mid m\in N^+\}$,$d_m$ 表示某一具体的防御动作;同理,攻击者的策略空间为 $A=\{A_1,A_2,\cdots,A_j\}$,其中 $A_j=\{a_1,a_2,\cdots,a_n \mid n\in N^+\}$,$a_n$ 表示某一具体的

攻击动作。

(5) 信号空间 $M=\{m_1,m_2,\cdots,m_k \mid k\in N^+\}$。信号空间是所有信号的集合,其中包含着真实信号和伪装信号,设 k 为防御者释放的信号数量,当 $k=1$ 时,代表释放的为真实信号,当 $k=2$ 时代表为1个真实信号和1个伪装信号的混合信号,当 $k=3$ 时表示为1个真实信号和2个伪装信号的混合信号,$k=4,5,\cdots$,以此类推。

(6) 概率 P,$0\leqslant P\leqslant 1$。$P(m_k)$ 为释放信号 m_k 的概率,其值是由攻击者在收集目标系统的信息时确定的。$P(A_j)$ 表示攻击方采用策略 A_j 的概率,其值与释放信号的概率有关,根据贝叶斯法则可以得出下式[18]:

$$\begin{cases} P(A_j)=\sum\limits_{n=1}^{k}P(A_j\mid m_n)\cdot P(m_n);\\ P(m_n)\geqslant 0;\\ \sum\limits_{n=1}^{N}P(m_n)=1。\end{cases}\tag{1}$$

$P(D_i)$ 表示防御方采用 D_i 的概率,其值与攻击者采取的攻击策略的概率有关,根据贝叶斯法则可以得出以下公式:

$$\begin{cases} P(D_i)=\sum\limits_{n=1}^{j}P(D_i\mid A_n)\cdot P(A_n);\\ P(A_j)\geqslant 0;\\ \sum\limits_{n=1}^{j}P(A_n)=1。\end{cases}\tag{2}$$

$P(D_i,A_j)$ 表示防御者采用防御策略 D_i,攻击者采用攻击策略 A_j 后,攻击者攻击成功的概率。设防御者释放的混合信号量为 k,因此攻击者命中真实信号的概率为 $1/k$,根据假设条件3,防御者针对攻击策略 A_j 选择防御策略 D_i 的概率为 $P_D(D_i)$,所以针对该攻击防御失败的概率即为 $1-P_D(D_i)$,可得出攻击者攻击成功概率为

$$P(D_i,A_j)=\sum P_A(A_j)(1/k)\cdot(1-P_D(D_i))。\tag{3}$$

(7)代价 C。C_D 表示防御代价,C_A 表示攻击代价,C_m 表示释放信号的代价。同文献[11]的方法,防御代价 C_D 的求解过程如式(4)所示:

$$C_D=O_{\text{cost}}+N_{\text{cost}}+R_{\text{cost}}。\tag{4}$$

式中:O_{cost} 为操作代价,代表执行该防御策略所付出的代价;N_{cost} 为负面代价,代表执行该防御策略导致提供的服务质量下降等带来的损失;R_{cost} 为残余损失,代表执行该防御策略后攻击者对系统造成的未被消除的损失。

(8)即时收益 R。R_D 表示防御者获得的即时收益,R_A 表示攻击者的即时收益。其中,$R_D(D_i,A_j)+R_A(D_i,A_j)\neq 0$,该公式表示在防御者采取防御策略 D_i 和攻击者采取攻击策略 A_j 时所获得的即时收益之和,其值不为零,此处体验了博弈论中非零和的思想,也更符合网络攻防实际收益情况。

防御者的即时收益 $R_D(D_i,A_j)$ 可由式(5)来表示:

$$R_D(D_i,A_j)=V\cdot AL\cdot(1-P(D_i,A_j))-\sum C_D-kC_m。\tag{5}$$

式中:V 表示资源的价值;AL 表示该资源对于系统的重要程度;$1-P(D_i,A_j)$ 表示防御成功的概率;$\sum C_D$ 代表执行防御策略时的防御代价。

攻击者的即时收益 $R_A(D_i,A_j)$ 可由式(6)来表示:

$$R_A(D_i,A_j)=V\cdot AL\cdot P(D_i,A_j)-\sum C_A。\tag{6}$$

式中:$P(D_i,A_j)$ 为攻击成功的概率;$\sum C_A$ 为执行攻击策略时的攻击代价。

(9)贴现因子 τ,$0\leqslant\tau\leqslant 1$,表示先前阶段收益对之后收益的影响,此处体现了多阶段博弈的收益情况。

(10)收益函数 $U=(U_D,U_A)$,为即时收益与以往阶段收益之和。U_D 所代表的是防御者的收益,由式(7)来表示;U_A 所代表的是攻击者的收益,由式(8)来表示。

$$U_D(D_i,A_j)=\sum_{l=1}^{t}(\tau^{l-1}R_D(D_i,A_j));\tag{7}$$

$$U_A(D_i,A_j)=\sum_{l=1}^{t}(\tau^{l-1}R_A(D_i,A_j))。\tag{8}$$

2 策略求解

2.1 均衡分析

根据1.2节中的假设条件,攻击者和防御者采取的策略都是保证自身利益的最大化,攻防双方存在以下均衡。

(1)纳什均衡。由1.1节的博弈过程进行分析,其状态转移具有 Markov 性,且每一个阶段都代表着一个博弈过程,给定一个 $NSG=(N,S,\theta,\omega,M,P,C,R,\tau,U)$,其状态空间 S、防御者和攻击者的策略空间 $\omega\{D,A\}$ 为有限集合,则一定存在一个稳定的纳什均衡[11]。即在模型 NSG 中存在策略 (D_i^*,A_j^*) 是一个纳什均衡,$\forall D,\exists D^*$ 使 $U_D(D^*)\geqslant U_D(D)$;$\forall A,\exists A^*$ 使 $U_A(A^*)\geqslant U_D(A)$。

(2)精炼贝叶斯均衡。攻击者通过对防御者释放信号的收集,对信号空间 $M=\{m_1,m_2,\cdots,m_k\}$ 和防御策略空间 $D=\{d_1,d_2,\cdots,d_i\}$ 进行分析,得到后验概率集合 $P_A(A_j \mid m_t)$,经过计算总能得到最优策略 A^* 以满足条件 $A^* \in \max U_A(D,A)$,对于防御者来说同理可以得到最优策略 D^* 使得 $D^* \in \max U_A(D,A^*)$。

2.2　求解当前最优策略

通过上述模型的描述以及对存在均衡解的分析,设计了一个求解当前最优防御策略的算法,如下所示。

算法 1　当前最优策略算法。

输入:非零和信号模型 NSG、V、AL;

输出:当前最优防御策略 D^*。

①**Initialize** $(N,S,\theta,\omega,M,P,C,R,\tau,U)$;

②**Initialize** $\theta=\{\theta_1,\theta_2,\cdots,\theta_n\}$;

③**Initialize** $D=\{d_1,d_2,\cdots,d_i\}$, $A=\{a_1,a_2,\cdots,a_j\}$;

④**Initialize** $M=\{m_1,m_2,\cdots,m_k\}$;

⑤**Initialize** $P_m=\{P(m_1),P(m_2),\cdots,P(m_k)\}$;

⑥**While** $S_t \in S$ **and** $m_k \in M$

⑦　$P(A_j)=\sum_{n=1}^{k} P(A_j \mid m_n) \times P(m_n)$;

⑧　$P(D_i)=\sum_{n=1}^{j} P(D_i \mid A_n) \times P(A_n)$;

⑨　**Foreach** $d_i \in D$ **and** $m_k \in M$

⑩　$R_D(D_i,A_j)=V\times AL\times(1-P(D_i,A_j))-\sum C_D-kC_m$;

⑪　$R_A(D_i,A_j)=V\times AL\times P(D_i,A_j)-\sum C_A$;

⑫　$U_D(D_i,A_j)=\sum_{l=1}^{t}(\tau^{l-1}R_D(D_i,A_j))$;

⑬　$U_A(D_i,A_j)=\sum_{l=1}^{t}(\tau^{l-1}R_A(D_i,A_j))$;

⑭　**End Foreach**

⑮　$U_D(D^*,A^*)=\max(\{U_D(D_i,A_j) \mid i,j \in N^+\})$;

⑯　$U_A(D^*,A^*)=\max(\{U_A(D_i,A_j) \mid i,j \in N^+\})$;

⑰　**Return** D^*

⑱**End while**

根据 1.1 节描述的博弈过程和 2.1 节中对两个均衡的分析,可以得出该算法的复杂度主要在于计算攻防策略的后验概率以及精炼贝叶斯的过程,其时间复杂度为 $O(N^2)$,其空间复杂度主要是用于存储各策略收益和求解的中间结果,其空间复杂度为 $O(mn)$,与其他文献的模型对比如表 1所示。

3　模型验证与分析

3.1　实验环境

为了模型的可行性和有效性,本文按照典型的信息系统进行仿真,采用了如图 3 所示的网络拓扑结构来模拟攻防情景。攻击者想要获取目标的私密信息,首先在发动攻击前对目标系统暴露在外网的信息进行收集分析,例如目标 IP 地址的范围、域名以及对外网提供的应用服务等。通过收集信息,攻击者确定其防御薄弱点以及攻击方法。

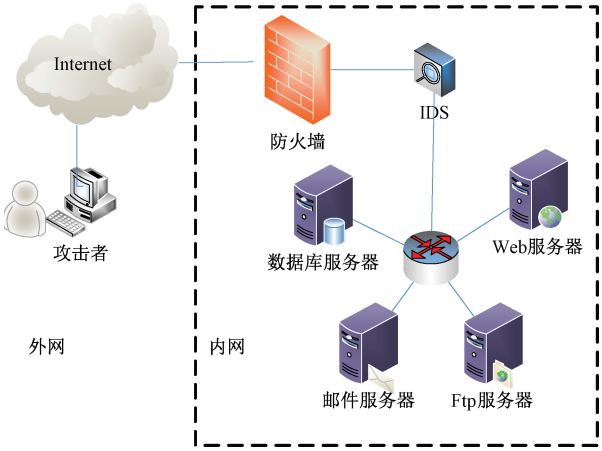

图 3　攻防网络拓扑图

Figure 3　Attack and defense network topology

在此次实验中,攻击者处于外部网络,目标网络由入侵检测设备、Web 服务器集群、数据库服务器、邮件服务器和 Ftp 服务器组成,其访问规则如表 2 所示。根据文献[20]给出模拟实验网络脆弱性信息如表 3 所示。根据其脆弱性进行攻

表 1　算法对比分析

Table 1　Comparative analysis of algorithms

模型出处	行为理性	博弈类型	博弈过程	均衡求解	具体应用
文献[10-11]	完全理性	静态博弈		一般	网络安全评测
文献[13]	非完全理性	动态博弈	单阶段	一般	防御策略选取
文献[16]	完全理性	静态博弈	单阶段	简单	防御策略选取
文献[17]	完全理性	动态博弈	多阶段	详细	动态攻防分析
文献[18]	非完全理性	动态博弈	单阶段	简单	动态攻防分析
本文	非完全理性	动态博弈	多阶段	一般	防御策略选取

击,将 AL 分为 10 个等级,数值越大表示行为越严重,表 4 列出了进行该攻击的 AL 和攻击代价。针对存在的漏洞可以进行相应的防御动作,假设残余损失为 0,其操作代价及其负面代价如表 5 所示。

表 2 防火墙规则

Table 2 Firewall rules

源地址	目的地址	服务类型	访问策略
All	Web 服务器	Http	Allow
All	Web 服务器	Database	Allow
Web 服务	数据库服务器	Oracle	Allow
File 服务	Ftp 服务器	Ftp	Allow
Smtp 服务	邮件服务器	Smtp	Allow

表 3 仿真实验网络脆弱性信息

Table 3 Vulnerability information of simulation experiment network

设备	类型	脆弱性	类别
内部用户	None	Weak password	User
ftp 服务器	Ftp	Ftp.rhost	User
Web 服务器	Http	Apache chunk overflow	Root
Web 服务器	Database	SQL injection attack	Root
数据库服务器	Oracle	Oracle TNS listener	Root
邮件服务器	Smtp	Remote buffer overflow	Root

表 4 攻击动作和代价

Table 4 Attack actions and cost

攻击动作	AL	攻击代价
a_1: Steal account	5	50
a_2: Ftp.rhost	6	90
a_3: Apache chunk overflow	10	70
a_4: SQL injection attack	9	60
a_5: Oracle TNS listener	10	80
a_6: Remove buffer overflow	8	110

表 5 防御动作和代价

Table 5 Defense actions and cost

防御动作	操作代价	负面代价
d_1: 删除可疑账号	5	40
d_2: 限制用户操作	25	0
d_3: 设置 ThreadLimit	25	0
d_4: Injection 检测工具	20	5
d_5: 修改端口访问控制权限	25	0
d_6: Renew root data	20	35

3.2 实验分析

在本实验中,预设数据的总价值 $V=1\ 000$,贴现因子 $\tau=0.3$,释放信号的代价为 $C_m=40$,根据对近年来 APT 案例结合以上列出的网络拓扑结构的网络漏洞进行分析,得到攻击者可能进行的攻击策略。为了简化实验过程,本实验只给出单阶段和两阶段的过程,多阶段情况以此类推。

根据文献[20]中所给出的方法对输入的路由器、漏洞数据、防火墙和防御策略进行建模和分析,得到可能采用的攻击策略 $A_1=\{a_1,a_3\}$,$A_2=\{a_2,a_6\}$,$A_3=\{a_4,a_5\}$,根据以上假设采用对应的防御策略为 $D_1=\{d_1,d_3\}$,$D_2=\{d_2,d_6\}$,$D_3=\{d_4,d_6\}$。

实例 1 博弈开始时,防御方率先发出信号,其概率为 $(P(m_1),P(m_2),P(m_3))=(0.2,0.4,0.4)$,通过式(1)可以计算得到攻击者采用的攻击策略为$(P(A_1),P(A_2),P(A_3))=(0.4,0.3,0.3)$;采用的防御策略通过式(2)可以得到防御策略概率为$(P(D_1),P(D_2),P(D_3))=(0.4,0.3,0.3)$。

实例 2 博弈开始时,由防御方率先发出信号,其概率为$(P(m_1),P(m_2),P(m_3))=(0.3,0.35,0.35)$,通过式(1)可以计算得到攻击者采用的攻击策略为$(P(A_1),P(A_2),P(A_3))=(0.35,0.2,0.45)$;采用的防御策略通过式(2)可以得到防御策略概率为$(P(D_1),P(D_2),P(D_3))=(0.35,0.2,0.45)$。

实例 3 博弈开始时,由防御方率先发出信号,其概率为$(P(m_1),P(m_2),P(m_3))=(0.5,0.3,0.2)$,通过式(1)可以计算得到攻击者采用的攻击策略为$(P(A_1),P(A_2),P(A_3))=(0.25,0.4,0.35)$;采用的防御策略通过式(2)可以得到防御策略概率为$(P(D_1),P(D_2),P(D_3))=(0.25,0.4,0.35)$。

图 4(a)表示随着混合信号数量 k 的增加,各实例单阶段攻击策略收益的变化。从中可以看到,随着释放信号量 k 值的增加,3 个实例的攻击策略收益都是呈现下降趋势。当 $k=1$ 时,即释放的信号均为真实信号时,3 个实例的攻击方的收益均达到最大;当 $k=3$ 时,实例 1 和实例 2 攻击策略收益趋近于 0;当 $k>3$ 时,实例 1 和实例 2 攻击策略收益均小于 0;当 $k=4$ 时,实例 3 攻击策略收益趋近于 0;当 $k>4$ 时,实例 3 的收益小于 0;当 $k>12$ 时,实例 1 的攻击策略收益开始小于实例 2。图 4(b)表示随着混合信号数量 k 的增加,各实例单阶段防御策略收益的变化。从图 4 中可以看到,随着释放信号量 k 值的增加,3 个实例的防御策略收益趋势均为先增加后减少。当 $k=3$ 时,实例 1 和实例 2 的防御策略收益达到最大值;当 $k>3$ 时,实例 1 和实例 2 的防御策略收益开始减少;当 $k=4$ 时,实例 3 的防御策略收益达到最大值;

当 $k>4$ 时,实例 3 的防御策略收益开始减少。

图 5(a)和图 5(b)分别表示随着混合信号数量 k 的增加,各实例两阶段攻击策略和防御策略收益的变化。从图中可以看出,两阶段攻防策略收益分布情况类似于单阶段,两阶段收益=第 2 阶段的即时收益+第 1 阶段收益×贴现因子。通过对比单阶段和两阶段的攻防收益图能够体现出本文所提出的多阶段收益量化思想。

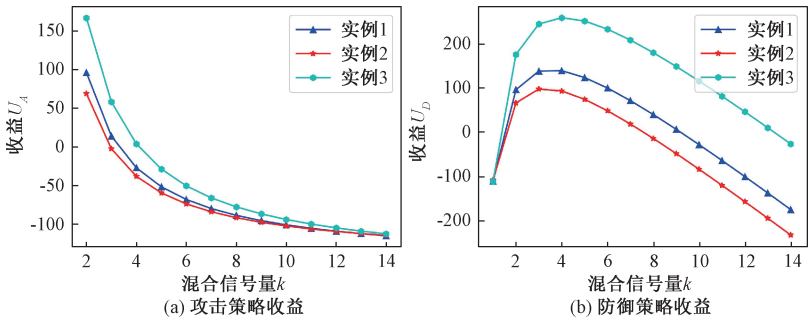

图 4　单阶段策略收益图

Figure 4　Single-stage strategy gains

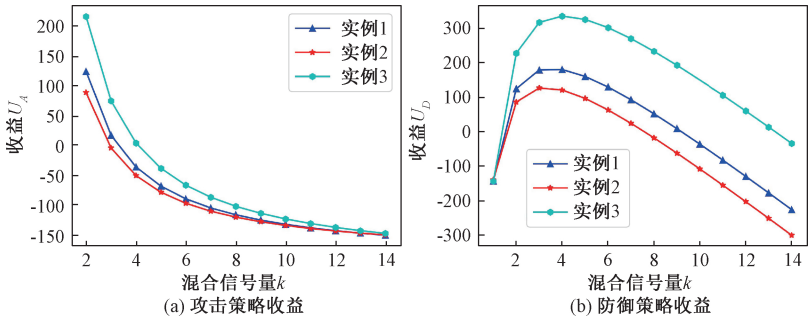

图 5　两阶段策略收益图

Figure 5　Two-stage strategy gains

4　结论

本文以典型的信息系统进行网络攻防仿真实验,分析实验结果得出以下结论:①混合信号的数量与攻击者的收益有着负相关的联系;②防御方不断增加混合信号数量,不一定能获得较高的安全增益;③网络攻防存在着攻防收益的不对称。

本文基于攻防信号博弈理论并结合了博弈理论中非零和的思想,在 3 个假设条件的基础上,将攻防双方的博弈过程进行建模,将攻防策略收益进行量化,根据多阶段信号博弈过程中收益情况引入贴现因子,并通过仿真实例验证了本文所提模型和方法的有效性。

参考文献:

[1] 付钰,李洪成,吴晓平,等.基于大数据分析的 APT 攻击检测研究综述[J].通信学报,2015,36(11):1-14.

[2] IQBAL Z,ANWAR Z.SCERM:A novel framework for automated management of cyber threat response activities[J]. Future generation computer systems,2020,108:687-708.

[3] 王耀光,陈伟权,吴镇邦,等.基于混合差分演化的网络入侵检测算法[J].郑州大学学报(工学版),2017,38(6):29-32,49.

[4] XU D J,LI Y D,XIAO L,et al.Prospect theoretic study of cloud storage defense against advanced persistent threats[C]//2016 IEEE Global Communications Conference. Piscataway:IEEE,2016:1-6.

[5] 陈福才,扈红超,刘文彦.网络空间主动防御技术[M].北京:科学出版社,2018:308-310.

[6] DIJK M,JUELS A,OPREA A,et al.FlipIt:the game of "stealthy takeover"[J].Journal of cryptology,2013,26(4):655-713.

[7] 糜旗,朱杰,徐超,等.基于 APT 网络攻击的技术研究[J].计算机与现代化,2014(10):92-94,122.

[8] 孙文君,苏旸,曹镇.非对称信息条件下 APT 攻防博弈模型[J].计算机应用,2017,37(9):2557-2562.

[9] FANG X P,ZHAI L D,JIA Z P,et al.A game model for predicting the attack path of APT[C]//2014 IEEE 12th International Conference on Dependable, Autonomic and Secure Computing. Piscataway: IEEE, 2014:491-495.

[10] 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.
[11] 姜伟,方滨兴,田志宏,等.基于攻防随机博弈模型的防御策略选取研究[J].计算机研究与发展,2010,47(10):1714-1723.
[12] 杨义先,钮心忻.安全通论:刷新网络空间安全观[M].北京:电子工业出版社,2018.
[13] 孙骞,高岭,刘涛,等.基于非零和博弈的多路径组合攻击防御决策方法[J].西北大学学报(自然科学版),2019,49(3):343-350.
[14] 李静轩,朱俊虎,邱菡,等.基于非零和随机博弈的APT攻击主动防御策略选取[J].计算机应用研究,2020,37(10):3071-3076,3111.
[15] YANG H P. Method for behavior-prediction of APT attack based on dynamic Bayesian game[C]//2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA). Piscataway: IEEE,2016:177-182.
[16] 张恒巍,余定坤,韩继红,等.基于攻防信号博弈模型的防御策略选取方法[J].通信学报,2016,37(5):51-61.
[17] 张恒巍,李涛.基于多阶段攻防信号博弈的最优主动防御[J].电子学报,2017,45(2):431-439.
[18] 张为,苏旸,陈文武.基于信号博弈的主动防御模型[J].计算机工程与应用,2018,54(17):77-82.
[19] 陈永强,付钰,吴晓平.基于非零和攻防博弈模型的主动防御策略选取方法[J].计算机应用,2013,33(5):1347-1349,1352.
[20] ZHANG H W, WANG J D, YU D K, et al. Active defense strategy selection based on static Bayesian game[C]//Third International Conference on Cyberspace Technology (CCT 2015). London: IET, 2015:1-7.

Proactive Defense Model Based on Non-Zero-Sum Signal Game

HUANG Wanwei[1], YUAN Bo[1], WANG Sunan[2], ZHANG Xiaohui[3]

(1.College of Software Engineering, Zhengzhou University of Light Industry, Zhengzhou 450001, China; 2.School of Electronic and Communication Engineering, Shen Zhen Polytechnic, Shenzhen 518005, China; 3.Henan Xin'an Communication Technology Co., Ltd., Zhengzhou 450001, China)

Abstract: In recent years, the damages of network attacks such as launched by APT has become more and more serious. Although existing studies based on signal game theory could simulate the APT attack and defense process to a certain extent, they ignored the phenomenon of unequal benefits between the two sides in the process of attack and defense and the multi-stage confrontation process, resulting in the lack of universality of the model and method. In this paper, a proactive defense model based on non-zero-sum signal game was proposed. First of all, the attack and defense game tree was built based on the signal game theory and the analysis of network attack and defense multi-stage confrontation process. Secondly, the non-zero-sum method and discount factor were used to build the multi-stage income of model in the process of offensive and defensive based on the situation of unequal income. On this basis, a quantitative method was proposed for network attack and defense characteristics, and the current optimal defense strategy algorithm was obtained based on the Nash equilibrium and refined Bayesian equilibrium existing in the analysis model. Finally, the model and method were verified by simulation experiments. The results showed the feasibility and effectiveness of the proposed model and method.

Keywords: non-zero-sum; signal game; discount factor; optimal defense strategy