Markov 微分博弈模型及其在网络安全中的应用

张恒巍[1,2],黄世锐[1]
(1. 信息工程大学三院,河南郑州 450001;2. 信息保障技术重点实验室,北京 100093)

摘 要: 当前基于博弈理论的网络安全研究成果难以应用于实时、连续、随机对抗的网络攻防过程. 本文针对网络安全防御的实时性和网络状态变化的随机性,基于动态、实时对抗的视角分析攻防行为,在结合微分博弈模型和 Markov 决策方法的基础上进行扩展,构建 Markov 攻防微分博弈模型,分析具有多个阶段且每阶段持续时间较短的攻防过程;提出多阶段博弈均衡解计算方法,设计多阶段最优防御策略选取算法. 仿真实验结果表明,模型和算法有效且可行.

关键词: 网络安全; 网络攻防; 博弈论; 微分博弈; Markov 决策; 网络防御; 攻防行为分析; 最优防御策略

中图分类号: TP309 **文献标识码:** A **文章编号:** 0372-2112 (2019)03-0606-07

电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2019.03.013

Markov Differential Game Model and Its Application in Network Security

ZHANG Heng-wei[1,2], HUANG Shi-rui[1]
(1. *The Third Institute, Information Engineering University, Zhengzhou, Henan* 450001, *China*;
2. *Science and Technology on Information Assurance Laboratory, Beijing* 100093, *China*)

Abstract: The current research of network security based on game theory fails to analyze the real-time, continuous, and random network attack and defense process. For the randomness of security states and the real-time character of network defense decision-making, we analyzed the network attack and defense behaviors from the view of dynamic and real-time confrontation. Then we combined and extended the differential game model and Markov decision-making method. On these basis, a Markov attack-defense differential game model is constructed, which can be adopted to analyze the multi-stage attack and defense process with short duration in each stage. Besides, a multi-stage game equilibrium solution is proposed, and an optimal defense strategy selection algorithm is designed. Finally, the experiments demonstrate that the model and method proposed in this paper are valid.

Key words: network security; network attack and defense; game theory; differential game; Markov decision-making; network defense; attack-defense analysis; optimal defense strategies

1 引言

网络安全的本质在攻防对抗,博弈论作为研究决策主体之间行为直接相互作用时的决策问题的理论,与网络攻防的基本特征高度契合,博弈模型应用于网络攻防分析和防御行动规划已经成为近年来新的研究方法和热点,并取得了部分成果[1,2].

传统动态博弈模型只能分析时间间断、离散的网络攻防对抗过程[3~5],随着技术发展,网络攻防逐渐趋向动态化、连续化、高频化,传统模型方法已不能满足实时防御的更高要求. 微分博弈是时间实时变化情况下描述冲突对抗中连续控制过程的理论方法[6],能够分析动态连续变化的网络攻防过程[7,8]. 另一方面,网络攻防过程和攻防策略的变化以及网络系统运行环境的改变均会导致系统安全状态的动态变化,并具有随机性,因此有学者结合博弈论与 Markov 决策方法,将网络攻防作为多阶段 Markov 过程进行研究[9]. 但是,现有研究成果均是将静态博弈或动态博弈与 Markov 决策相结合[10,11],难以满足实时防御决策的需求. 在借鉴微分博弈的基础上,结合 Markov 决策方法并加以扩展,可以显著提高网络防御规划决策的有效性和适用性. 由于微分博弈是多维相空间中的连续变化过程[12],模型的构建和运用难度大,结合

收稿日期:2017-09-07;修回日期:2018-02-05;责任编辑:梅志强
基金项目:国家自然科学基金(No. 61303074, No. 61309013);河南省科技攻关计划基金(No. 182102210144);信息保障技术重点实验室开放基金(No. KJ-15-110)

Markov 决策后难度进一步增加,目前据我们所知,尚未有公开文献对上述方法予以讨论.

本文主要工作在于,首次以微分博弈理论和 Markov 决策方法为基础,将一定时间内的网络攻防对抗转化为多个阶段且每阶段持续时间较短的连续攻防过程,构建 Markov 攻防微分博弈模型进行研究;在各阶段内利用微分博弈完成实时攻防分析,基于 Markov 转移概率实现对不同阶段可能跳变路径的分析;在求解和分析多阶段博弈均衡的基础上,设计最优防御策略选取算法,并通过仿真实验验证了模型和方法的有效性.

2　Markov 攻防微分博弈模型

对于经典微分博弈,当系统从初始状态开始动态博弈过程时,均衡策略是一个时间路径,双方的博弈结果和系统的状态变化是确定、可度量的. 但是,网络攻防过程一般受到不完全信息限制,并具有随机性. 因此,本文将一定持续时间的网络攻防对抗转化为多个阶段且每阶段时间较短的攻防过程,如图 1 所示. 攻防双方由初始状态开始连续决策、动态对抗,随着时间推移,一方面受到攻防行为的作用,另一方面受到系统环境和博弈要素变化的影响,网络系统以概率 η 跳变到另一个状态. 为此,构建多阶段 Markov 攻防微分博弈模型,引入折现因子 μ,量化计算从初始阶段到最后阶段的攻防双方的期望总收益,并以此作为双方的目标函数,实现博弈均衡求解和防御策略选取.

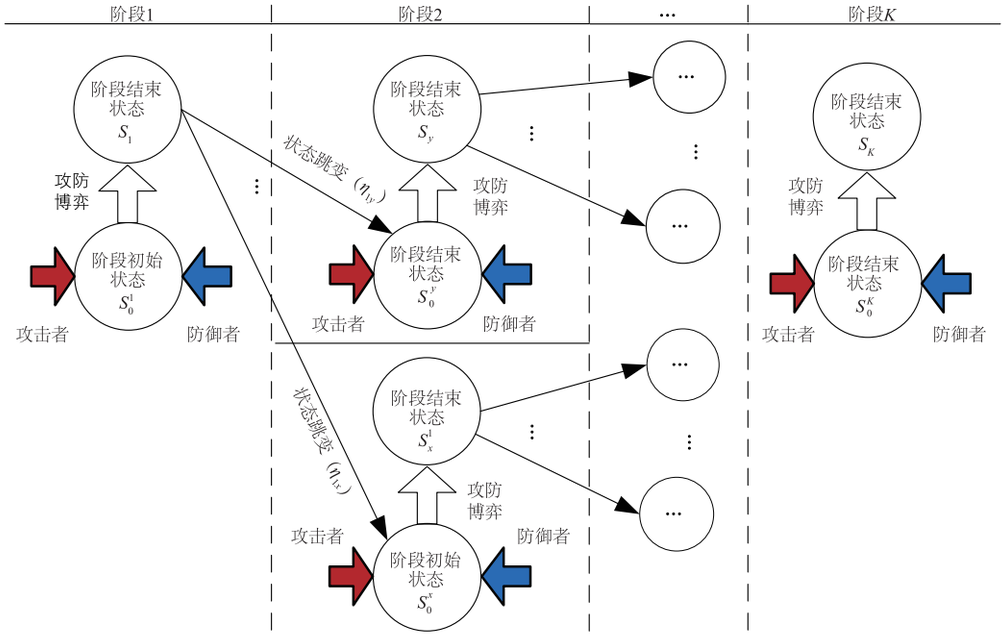


图 1　Markov攻防微分博弈过程示意图

对于由大量节点构成的网络系统,为描述不同节点安全状态的动态改变,借鉴传染病动力学 SIR 模型[13]并加以扩展,把网络系统中的节点类比为 SIR 模型中的个体,描述安全状态演化过程,网络节点具有四个不同的安全等级:正常级 N(Normal)、感染级 I(Infected)、修复级 R(Restored)、受损级 M(Malfunctioned),不同等级间具有四种迁移方式.

$N\to I$:防御策略失败,正常节点被感染,破坏效果处于潜伏期;攻击者能利用该节点攻击邻接节点.

$N\to R$:防御策略成功,正常节点具有对攻击的免疫能力.

$I\to R$:防御策略识别感染节点并清除感染,避免感染节点的损失并且使其具有免疫能力.

$I\to M$:防御策略失败,破坏效果出现,感染节点丧失服务功能.

设网络节点总数为 Q,t 时刻四种等级的节点数量用变量 $N(t)$、$I(t)$、$R(t)$ 和 $M(t)$ 表示. 假设节点密度 θ,r 表示两个节点的网络距离,$r=1$ 时,代表节点直接连接. 对于一个感染节点,与其直接连接的节点数量为 $\pi\theta$. 若假设网络节点数量较大且感染节点相互距离较远,忽略感染节点影响范围的重叠效应,则 t 时刻和感染节点直接连接的正常节点的数量为 $\pi\theta I(t)N(t)/Q$.

依据攻防行为分类方法,采用平均攻击强度 e_A^H、e_A^M、$e_A^L\in[0,1]$,表示强、中、弱三类攻击策略,则攻击效用为 $a(t)=p_A^H(t)e_A^H+p_A^M(t)e_A^M+p_A^L(t)e_A^L$,简记为 a;采用平均防御强度 e_D^H、$e_D^L\in[0,1]$,表示强、弱两类防御策

略,则防御效用可以表示为 $d(t)=p_{\mathrm{D}}^{\mathrm{H}}(t)e_{\mathrm{D}}^{\mathrm{H}}+p_{\mathrm{D}}^{\mathrm{L}}(t)e_{\mathrm{D}}^{\mathrm{L}}$,简记为 d. 攻防效用差值表示为 $\eta(t)=a(t)-d(t)$,由此可得描述安全等级迁移可能性的迁移参数 $\eta_{NI},\eta_{NR},\eta_{IR},\eta_{IM}$.

$$\eta_{NI}=\begin{cases}0, & \eta(t)\leq 0\\ \eta(t), & \eta(t)>0\end{cases},\eta_{NR}=\begin{cases}|\eta(t)|, & \eta(t)\leq 0\\ 0, & \eta(t)>0\end{cases}$$
$$\eta_{IR}=\begin{cases}|\eta(t)|, & \eta(t)\leq 0\\ 0, & \eta(t)>0\end{cases},\eta_{IM}=\begin{cases}0, & \eta(t)\leq 0\\ \eta(t), & \eta(t)>0\end{cases} \tag{1}$$

描述网络节点安全等级变化的微分方程 $f(\mathrm{d}N(t),\mathrm{d}I(t),\mathrm{d}R(t),\mathrm{d}M(t))$ 如下.

$$\begin{cases}\mathrm{d}N(t)=-\eta_{NI}(t)\pi\theta I(t)N(t)/Q-\eta_{NR}(t)N(t)\\ \mathrm{d}I(t)=\eta_{NI}(t)\pi\theta I(t)N(t)/Q-\eta_{IM}(t)I(t)-\eta_{IR}(t)I(t)\\ \mathrm{d}R(t)=\eta_{NR}(t)N(t)+\eta_{IR}(t)I(t)\\ \mathrm{d}M(t)=\eta_{IM}(t)I(t)\end{cases} \tag{2}$$

对于等级迁移 $N\rightarrow I$,设回报系数为 r_1;$N\rightarrow R$ 或 $I\rightarrow R$,回报系数为 r_2;$I\rightarrow M$,回报系数为 r_3. 参考文献[14],采用统计平均值定义回报系数 $r_1、r_2、r_3\in[0,10]$,可得策略回报 $r_{\mathrm{D}}(t)$ 和 $r_{\mathrm{A}}(t)$.

$$\begin{aligned}r_{\mathrm{D}}(t)=&r_2[\eta_{NR}(t)N(t)+\eta_{IR}(t)I(t)]\\&-r_1\eta_{NI}(t)\pi\theta I(t)N(t)/Q-r_3\eta_{IM}(t)I(t)\end{aligned} \tag{3}$$

$$\begin{aligned}r_{\mathrm{A}}(t)=&r_1\eta_{NI}(t)\pi\theta I(t)N(t)/Q+r_3\eta_{IM}(t)I(t)\\&-r_2[\eta_{NR}(t)N(t)+\eta_{IR}(t)I(t)]\end{aligned} \tag{4}$$

设策略执行代价为 $v_{\mathrm{D}}(t)$ 和 $v_{\mathrm{A}}(t)$.

$$\begin{aligned}v_{\mathrm{D}}(t)=&\frac{d^2}{2}c_{\mathrm{D}}[N(t)+I(t)+R(t)+M(t)]\\ v_{\mathrm{A}}(t)=&\frac{a^2}{2}c_{\mathrm{A}}[N(t)+I(t)+R(t)+M(t)]\end{aligned} \tag{5}$$

$c_{\mathrm{D}}、c_{\mathrm{A}}\in[1,10]$ 为策略的成本/效用系数[7]. 综合策略回报和代价,可得双方收益函数.

$$\begin{aligned}U_{\mathrm{D}}(P_{\mathrm{A}}(t),P_{\mathrm{D}}(t))=&\int_{t_0}^{t_0+T}[r_{\mathrm{D}}(t)-v_{\mathrm{D}}(t)]\mathrm{d}t\\ U_{\mathrm{A}}(P_{\mathrm{A}}(t),P_{\mathrm{D}}(t))=&\int_{t_0}^{t_0+T}[r_{\mathrm{A}}(t)-v_{\mathrm{A}}(t)]\mathrm{d}t\end{aligned} \tag{6}$$

定义 1 Markov 攻防微分博弈模型 MADG(Markov Attack-Defense Differential Game)可用十元组表示为 $\mathrm{MADG}=(N,K,S,B,t,x,P,\eta,\mu,U)$

(1)$N=(N_{\mathrm{D}},N_{\mathrm{A}})$ 是 Markov 攻防微分博弈的参与者空间,N_{D} 为防御者,N_{A} 为攻击者.

(2)K 是多阶段博弈的阶段数,某阶段博弈用 $G(k)$ 表示,$k=\{1,2,\cdots,K\}$. 将持续时间 $[t_{\mathrm{bin}},t_{\mathrm{end}}]$ 的攻防过程转化为 K 个阶段且每阶段持续时间 T 的多阶段过程,其中 $G(k)$ 的时间为 $[t_k,t_k+T]$,且 $t_1=t_{\mathrm{bin}}$,$\sum_{k=1}^{K}T=t_{\mathrm{end}}-t_{\mathrm{bin}}$.

(3)$S=\{S_0^k,S_k\mid k=1,\cdots,K\}$ 是安全状态集合. 为便于分析不同阶段博弈 $G(k)$,采用 $\{S_0^1\cdots S_0^k\cdots S_0^K\}$ 代表不同阶段的初始状态,$\{S_1\cdots S_k\cdots S_K\}$ 代表结束状态,且 $S_0^k,S_k\in S$.

(4)$B=(\mathrm{DS},\mathrm{AS})$ 是攻防行动空间.

$\mathrm{DS}=\{\mathrm{DS}_k^j\mid 1\leq k\leq K,1\leq j\leq n\}$,$\mathrm{AS}=\{\mathrm{AS}_k^i\mid 1\leq k\leq K,1\leq i\leq m\}$,$\mathrm{DS}_k^j$ 和 AS_k^i 分别表示双方第 k 阶段的可选策略.

(5)$t\in[t_{\mathrm{bin}},t_{\mathrm{end}}]$ 表示攻防博弈的时刻.

(6)$x_k(t)=\{N_k(t),I_k(t),R_k(t),M_k(t)\}$ 代表阶段 k 内网络节点变量,$N_k(t)+I_k(t)+R_k(t)+M_k(t)=Q$.

(7)$P=\{P_{\mathrm{D}}^k(t),P_{\mathrm{A}}^k(t)\}$ 是双方在 $G(k)$ 的控制策略.

$P_{\mathrm{D}}^k(t)=\{p_{\mathrm{D}}^k(t)_j\mid 1\leq k\leq K,1\leq j\leq n\}$ 表示 $G(k)$ 内防御者在 t 时刻选取的混合策略,$\sum_{j=1}^{n}p_{\mathrm{D}}^k(t)_j=1$;$P_{\mathrm{A}}^k(t)=\{p_{\mathrm{A}}^k(t)_i\mid 1\leq k\leq K,1\leq i\leq m\}$ 是攻击者选取的混合策略,$\sum_{i=1}^{m}p_{\mathrm{A}}^k(t)_i=1$.

(8)$\eta_{ij}=\eta(S_j\mid S_i)$ 代表系统从状态 S_i 跳变至状态 S_j 的概率,且 $i=j$ 时,$\eta_{ij}\equiv 0$.

(9)μ 是折现因子,表示阶段 k 中的收益相较初始阶段的折现比例,$0\leq\mu\leq 1$.

(10)$U^k=\{U_{\mathrm{D}}^k,U_{\mathrm{A}}^k\}$ 是博弈收益集合,U_{A}^k 和 U_{D}^k 是 k 阶段中的攻防收益函数.

设计目标函数 R,在函数中引入折现因子 μ,描述攻防双方收益与博弈阶段之间的折现关系.

$$\begin{cases}R_{\mathrm{D}}^k(S_0^k,S_k)=U_{\mathrm{D}}^k(P_{\mathrm{D}}^k(t),P_{\mathrm{A}}^k(t))\\ \qquad\qquad+\sum_{e,h\in[k,K]}\mu\eta(S_h\mid S_e)R_{\mathrm{D}}^h(S_0^h,S_h)\\ R_{\mathrm{A}}^k(S_0^k,S_k)=U_{\mathrm{A}}^k(P_{\mathrm{D}}^k(t),P_{\mathrm{A}}^k(t))\\ \qquad\qquad+\sum_{e,h\in[k,K]}\mu\eta(S_h\mid S_e)R_{\mathrm{A}}^h(S_0^h,S_h)\end{cases} \tag{7}$$

U_{D}^k 和 U_{A}^k 的函数形式参考式(6),目标函数 R 是博弈阶段 $G(k)$ 的收益与折现收益之和,攻防双方的目标是使各自的目标函数达到最大值.

3 Markov 多阶段博弈均衡求解方法

对于博弈阶段 $G(k)$,根据微分博弈基本理论[6],若 $(P_{\mathrm{A}}^k(t)^*,P_{\mathrm{D}}^k(t)^*)$ 是 k 阶段的最优控制策略,则满足:

$$\begin{cases}\forall P_{\mathrm{A}}^{k}(t),U_{\mathrm{A}}^{k}(P_{\mathrm{A}}^{k}(t)^{*},P_{\mathrm{D}}^{k}(t)^{*})\geq U_{\mathrm{A}}^{k}(P_{\mathrm{A}}^{k}(t),P_{\mathrm{D}}^{k}(t)^{*})\\ \forall P_{\mathrm{D}}^{k}(t),U_{\mathrm{D}}^{k}(P_{\mathrm{A}}^{k}(t)^{*},P_{\mathrm{D}}^{k}(t)^{*})\geq U_{\mathrm{D}}^{k}(P_{\mathrm{A}}^{k}(t)^{*},P_{\mathrm{D}}^{k}(t))\end{cases} \tag{8}$$

在多阶段博弈中,根据 Markov 决策准则,各参与人必有一个 Markov 最优响应策略[15],若该策略为 $\{(P_{\mathrm{A}}^{k}(t)^{*},P_{\mathrm{D}}^{k}(t)^{*})\mid 1\leq k\leq K\}$,则 $(P_{\mathrm{A}}^{k}(t)^{*},P_{\mathrm{D}}^{k}(t)^{*})$ 对阶段 k 满足下列条件:

$$\begin{aligned}P_{\mathrm{A}}^{k}(t)^{*}&=\operatorname{argmax}[R_{\mathrm{D}}^{k}(S_0^{k},S_k)]\\ P_{\mathrm{D}}^{k}(t)^{*}&=\operatorname{argmax}[R_{\mathrm{A}}^{k}(S_0^{k},S_k)]\end{aligned} \tag{9}$$

攻防过程由有限的 k 个阶段博弈构成,同时每个阶段中的策略集 DS_k 和 AS_k 是有限的,因此 MADG 为多阶段-多状态的有限 Markov 微分博弈模型. 根据微分博弈基本定理和文献[9,15]的结论,可以保证多阶段均衡策略的存在性和合理性.

引入贴现因子 μ 计算多阶段攻防收益,将多阶段均衡策略求解问题,转化为以整体收益最高为目标的动态规划问题进行计算.

$$\forall k\in K,t\in[t_{\mathrm{bin}},t_{\mathrm{end}}]$$

$$\begin{cases}\max R_{\mathrm{A}}^{k}(S_0^{k},S_k)=\max[U_{\mathrm{A}}^{k}(P_{\mathrm{A}}^{k}(t),P_{\mathrm{D}}^{k}(t))+\sum\limits_{e,h\in[k,K]}\mu\eta(S_h\mid S_e)R_{\mathrm{A}}^{h}(S_0^{h},S_h)]\\ \max R_{\mathrm{D}}^{k}(S_0^{k},S_k)=\max[U_{\mathrm{D}}^{k}(P_{\mathrm{A}}^{k}(t),P_{\mathrm{D}}^{k}(t))+\sum\limits_{e,h\in[k,K]}\mu\eta(S_h\mid S_e)R_{\mathrm{D}}^{h}(S_0^{h},S_h)]\\ N_k(t)+I_k(t)+R_k(t)+M_k(t)=Q\\ \mathrm{d}N_k(t)=-\eta_{NR}(t)N_k(t)-\eta_{NI}(t)\pi\theta I_k(t)N_k(t)/Q\\ \mathrm{d}R_k(t)=\eta_{NR}(t)N_k(t)+\eta_{IR}(t)I_k(t)\\ \mathrm{d}I_k(t)=-I_k(t)[\eta_{IM}(t)+\eta_{IR}(t)]+\eta_{NI}(t)\pi\theta I_k(t)N_k(t)/Q\\ \mathrm{d}M_k(t)=\eta_{IM}(t)I_k(t)\end{cases} \tag{10}$$

求解方程式(10)可得最优控制策略集合 $\{(P_{\mathrm{A}}^{k}(t)^{*},P_{\mathrm{D}}^{k}(t)^{*})\}$. 依据博弈理论,防御方应将 $P_{\mathrm{D}}^{k}(t)^{*}$ 作为最优防御策略.

4　多阶段最优防御策略选取算法及对比分析

算法 1　多阶段最优防御策略选取算法

输入:Markov 攻防微分博弈模型 MADG
输出:多阶段最优防御策略 $P_{\mathrm{D}}^{k}(t)^{*}$
BEGIN
1. 初始化 MADG = $(N,K,S,B,t,x,P,\eta,\mu,U)$;
2. 构建防御行为空间 DS 和攻击行为空间 AS;
3. 构建各阶段系统状态集合 $\{S_0^1\cdots S_0^k\cdots S_0^K\}$ 和 $\{S_1\cdots S_k\cdots S_K\}$;
4. 初始化状态转移概率 $\eta_{ij}=\eta(S_j\mid S_i)$ 以及常量系数 r_1,r_2,r_3 和 $c_{\mathrm{D}},c_{\mathrm{A}}$;
For ($k=1$;$k\leq K$;$k++$)
{　//计算 k 阶段的博弈收益
5. 依据式(2)构建 k 阶段节点安全等级变化的微分方程 $f(\mathrm{d}N(t),\mathrm{d}I(t),\mathrm{d}R(t),\mathrm{d}M(t))$;
6. 计算 $N_k(t)$、$I_k(t)$、$R_k(t)$、$M_k(t)$;
7. 由式(6)计算收益 $U_{\mathrm{D}}^{k}(P_{\mathrm{D}}^{k}(t),P_{\mathrm{A}}^{k}(t))$ 和 $U_{\mathrm{A}}^{k}(P_{\mathrm{D}}^{k}(t),P_{\mathrm{A}}^{k}(t))$;
}
8. 利用折现因子 μ 和公式(7),计算折扣收益 $\sum\limits_{e,h\in[k,K]}\mu\eta_{eh}R_{\mathrm{D}}^{h}(S_0^{h},S_h)$ 和 $\sum\limits_{e,h\in[k,K]}\mu\eta_{eh}R_{\mathrm{A}}^{h}(S_0^{h},S_h)$;
9. 基于动态规划方法,以 $R_{\mathrm{D}}^{k}(S_0^{k},S_k)$ 和 $R_{\mathrm{A}}^{k}(S_0^{k},S_k)$ 作为目标函数,求解方程式(10),得到最优策略集合 $\{(P_{\mathrm{A}}^{k}(t)^{*},P_{\mathrm{D}}^{k}(t)^{*})\}$;
10. Return $\{P_{\mathrm{D}}^{k}(t)^{*}\}$;
END

将本文提出的方法和其它文献进行对比,结果见表 1. 模型通用性是指博弈类型集和策略集是否可以扩展至 n;若可以,说明模型的通用性较好;反之,通用性差. 均衡求解是指文中是否给出均衡解的计算过程. 决策时效性是指决策结果(最优策略)的有效时间. 由于网络攻防速度快且安全状态改变具有随机性,因此防御决策需要连续化、实时化,否则最优策略可能无法适应攻防节奏而失去效力. 本文方法具备分析多阶段、多状态、连续攻防的能力,可为多阶段攻防过程实时选取最优防御策略,具有更好的理论价值和实用性.

表 1　模型和方法对比

文献	攻防博弈过程	博弈类型	决策时效性	模型通用性	均衡求解
文献[4]	单阶段	动态博弈	未考虑	差	简单
文献[5,6]	离散多阶段	动态博弈	差	较好	简单
文献[8,9]	单阶段连续时间	微分博弈	较好	较好	简单
文献[10,11]	离散多阶段	Markov 动态博弈	差	好	详细
本文	多阶段连续时间	Markov 微分博弈	好	好	详细

5　仿真实验与分析

5.1　实验环境描述

采用仿真工具 Scalable Simulation Framework(SSF-Net)[16],模拟不同规模和初始状态的网络攻防场景. 采用 Route Views Project 中的连接数据集(NetTFData20170616103000)设计实验拓扑结构. 其中,终端集群节点数量 1800,Web 服务器集群节点数量 50,

应用服务器集群节点数量35,数据服务器集群节点数量30. 网络系统安全状态划分为9种,如表2所示. 假设不同阶段间的状态转移概率固定,依据历史数据和专家经验确定具体概率,如表3所示. 基于MIT的攻防行为数据库[17],构建攻防策略集,如表4所示.

表2 系统安全状态

安全状态	状态描述	安全状态	状态描述
S_0^1,S_1	全网节点正常状态	S_0^6	攻击者获取应用服务器guest权限
S_0^2	攻击者获取终端guest权限	S_0^7	攻击者获取应用服务器root权限
S_0^3	攻击者获取终端root权限	S_0^8	攻击者获取数据服务器guest权限
S_0^4	攻击者获取Web服务器guest权限	S_0^9	攻击者获取数据服务器root权限
S_0^5	攻击者获取Web服务器root权限	—	——

表3 各阶段之间的状态转移概率

状态跳变	跳变概率	状态跳变	跳变概率
$S_1\rightarrow S_0^2$	$\eta(2\mid1)=0.9$	$S_4\rightarrow S_0^8$	$\eta(8\mid4)=0.4$
$S_2\rightarrow S_0^3$	$\eta(3\mid2)=0.7$	$S_5\rightarrow S_0^6$	$\eta(6\mid5)=0.9$
$S_2\rightarrow S_0^4$	$\eta(4\mid2)=0.7$	$S_5\rightarrow S_0^7$	$\eta(7\mid5)=0.3$
$S_2\rightarrow S_0^5$	$\eta(5\mid2)=0.3$	$S_5\rightarrow S_0^8$	$\eta(8\mid5)=0.5$
$S_3\rightarrow S_0^4$	$\eta(4\mid3)=0.9$	$S_6\rightarrow S_0^7$	$\eta(7\mid6)=0.7$
$S_3\rightarrow S_0^5$	$\eta(5\mid3)=0.5$	$S_6\rightarrow S_0^8$	$\eta(8\mid6)=0.8$
$S_4\rightarrow S_0^5$	$\eta(5\mid4)=0.8$	$S_7\rightarrow S_0^8$	$\eta(8\mid7)=0.8$
$S_4\rightarrow S_0^6$	$\eta(6\mid4)=0.7$	$S_7\rightarrow S_0^9$	$\eta(9\mid7)=0.3$
$S_4\rightarrow S_0^7$	$\eta(7\mid4)=0.1$	$S_8\rightarrow S_0^9$	$\eta(9\mid8)=0.6$

表4 各博弈阶段的攻防动作集

博弈阶段	AS,e_A	DS,e_D
$S_0^2\rightarrow S_2$	Install Trojan,e_A^H	InstallPW-worm patches,e_D^H
	Attack address blacklist,e_A^M	Limit SYN/ICMP packets,e_D^M
	Install socket analyzer program,e_A^L	Repair Data,e_D^L
$S_0^3\rightarrow S_3$	Steal account and crack it,e_A^H	Reset Listener ports,e_D^H
	Send abnormal data to GIOP,e_A^M	Recorrect homepage,e_D^M
	Web-rhost attack,e_A^L	Renew root data,e_D^L
$S_0^4\rightarrow S_4$	Http LQ-sniffer,e_A^H	Address blacklist,e_D^H
	Attack SSH on Web Sever,e_A^M	Add physical resource,e_D^M
	Sr-Hard blood,e_A^L	Redeploy firewall rule,e_D^L
$S_0^5\rightarrow S_5$	Steal account and crack,e_A^H	Correct homepage,e_D^H
	Homepage attack,e_A^M	Install SGD on Web Sever,e_D^M
	Apache chunk overflow,e_A^L	Restart server process,e_D^L
$S_0^6\rightarrow S_6$	LPC to LSASS process,e_A^H	install delete Trojan patches,e_D^H
	install delete Trojan,e_A^M	Repair database,e_D^M
	SMTPsniffer,e_A^L	Delete suspicious account,e_D^L
$S_0^7\rightarrow S_7$	Shutdown server tenor,e_A^H	Limitdatebase ports,e_D^H
	Ftp rhost attack,e_A^M	Filtrate malicious packets,e_D^M
	installVBW Trojan,e_A^L	Patch SSH on Ftp Sever,e_D^L
$S_0^8\rightarrow S_8$	Oracle TNS Listener,e_A^H	Limit access toMS. SDO_CS,e_D^H
	THS chunk overflow,e_A^M	Install Oracle patches,e_D^M
	Ssh buffer overflow,e_A^L	Delete suspicious account,e_D^L
$S_0^9\rightarrow S_9$	CF-exploit attack,e_A^H	Reset Oracle access authority,e_D^H
	Shutdowndatabase server,e_A^M	Alter data read-write rule,e_D^M
	installSQL Listener program,e_A^L	Redeploy firewall policy,e_D^L

5.2 攻防仿真与分析

结合网络对抗实际,设攻防过程为40m,分为4个阶段,每阶段时间$T=10$m,折现因子$\mu=0.6$. 以数据服务器集群作为攻击目标,存在两条主要攻击路径,具体分析如下:

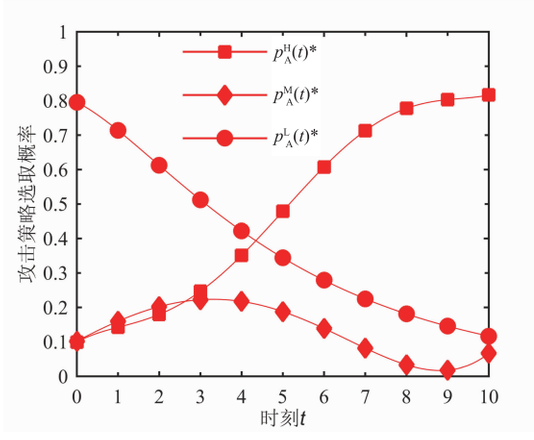
攻击路径(1):

$$S_1\rightarrow S_0^2\rightarrow S_2\rightarrow S_0^5\rightarrow S_5\rightarrow S_0^7\rightarrow S_7\rightarrow S_0^9\rightarrow S_9$$

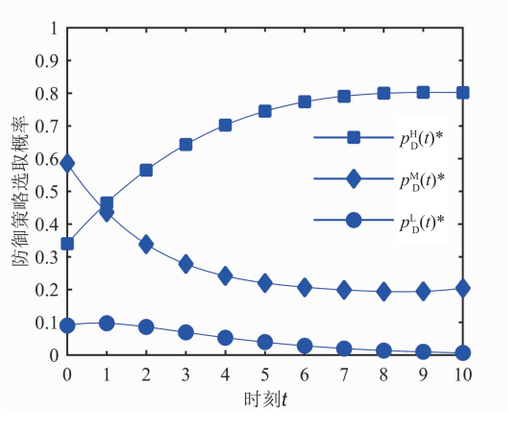
第1阶段$S_0^2\rightarrow S_2$,攻防双方最优策略轨迹如图2所示. 本阶段结束后以概率$\eta(5\mid2)=0.3$从状态S_2跳变至S_0^5. 攻击方前期以低强度策略为主,注重隐匿性和渗透性;防御方前期对攻击感知不足,在考虑代价的情况下采用中强度策略作为平时选择. 后期,随着攻击持续,攻击暴露概率增大,攻击方以高强度策略为主;防御方采用高强度策略应对. 结束时的攻防最优策略分别是(0.81,0.08,0.11)和(0.8,0.2,0).

第2阶段$S_0^5\rightarrow S_5$,攻防双方最优策略轨迹如图3所示,双方的意图逐渐明朗化. 前期,攻防双方均采用高

强度策略对抗;随着时间推移,考虑策略代价,攻击者转而增加中低强度策略选取概率,防御者随之采取中强度策略为主. 结束时的攻防最优策略分别是(0. 11, 0. 28,0. 61)和(0. 2,0. 62,0. 18).



(a)　第1阶段最优攻击策略轨迹



(b)　第1阶段最优防御策略轨迹

图 2　攻击路径(1)的第1阶段最优攻防策略轨迹

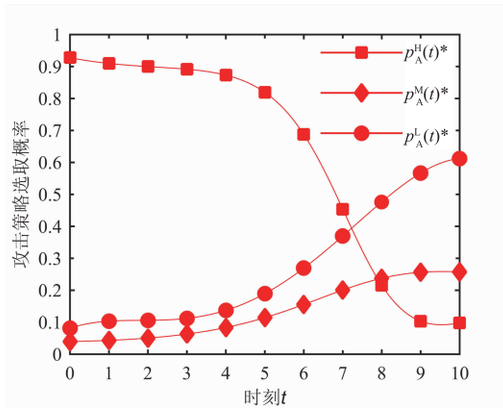
第 3 阶段 $S_0^7 \rightarrow S_7$,攻击目标是增大应用服务器集群中 root 权限节点的数量,为攻击数据服务器集群奠定基础. 前期,攻防双方采取高强度策略对抗;后期,攻击方以中高强度攻击策略为主,由于应用服务器属于核心资产,防御方高度警惕,持续以大概率选取强防御策略. 结束时的攻防最优策略分别是(0. 2,0. 61,0. 19)和(0. 71,0. 02,0. 27).

第 4 阶段 $S_0^9 \rightarrow S_9$,攻击者对最终目标—数据服务器集群实施攻击,中高强度攻击策略概率保持在 0. 7 以上;防御方为保护关键核心资产,高强度策略概率始终在 0. 5 以上. 结束时的攻防最优策略分别是(0. 38,0. 5, 0. 12)和(0. 7,0. 19,0. 11).

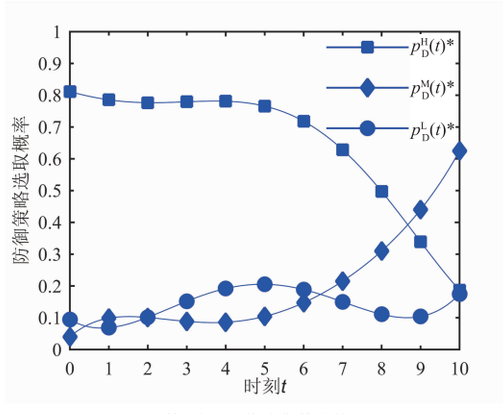
攻击路径(2):

$$S_1 \rightarrow S_0^2 \rightarrow S_2 \rightarrow S_0^4 \rightarrow S_4 \rightarrow S_0^8 \rightarrow S_8 \rightarrow S_0^9 \rightarrow S_9$$

具体分析与上一条路径同理类似. 不同路径中攻



(a)　第2阶段最优攻击策略轨迹



(b)　第2阶段最优防御策略轨迹

图 3　攻击路径(1)的第2阶段最优攻防策略轨迹

防双方各阶段的收益如表 5 所示. 攻击路径(1)中,攻防总收益分别是 114. 6 和 -61. 6;攻击路径(2)中,攻防总收益分别是 172. 3 和 -120. 9.

由数据分析可以看出,两条攻击路径的攻防总收益存在明显差异,从防御角度出发,路径(1)更符合期望,而应尽量避免路径(2). 两条路径的第 1 阶段相同,但在阶段末的状态改变时两条路径产生分化. 为减小路径(2)出现概率,应当降低 $S_2 \rightarrow S_0^4$ 的发生可能. 因此,针对 $S_0^4 \rightarrow S_4$ 对应的 AS,防御方可以利用动态调整网络访问端口、增设白名单、安装反嗅探装置等主动防御策略,降低状态改变到 S_0^4 的概率,减小路径(2)发生可能.

表 5　不同攻击路径下的攻防双方博弈收益

攻击路径(1)			攻击路径(2)		
阶段	攻击收益	防御收益	阶段	攻击收益	防御收益
$S_0^2 \rightarrow S_2$	18. 9	-9. 8	$S_0^2 \rightarrow S_2$	34. 4	-22. 5
$S_0^5 \rightarrow S_5$	25. 8	-13. 4	$S_0^4 \rightarrow S_4$	41	-28. 6
$S_0^7 \rightarrow S_7$	31. 6	-16. 1	$S_0^8 \rightarrow S_8$	44. 8	-32. 1
$S_0^9 \rightarrow S_9$	38. 3	-22. 3	$S_0^9 \rightarrow S_9$	52. 1	-37. 7

6 结束语

当前博弈论在网络安全领域的研究成果,难以满足分析快速变化、连续对抗的网络攻防过程的要求,本文基于微分博弈和 Markov 决策方法,构建了 Markov 攻防微分博弈模型,提出了基于动态规划的均衡解计算方法,设计了多阶段最优防御策略选取算法,通过仿真实验验证了模型和方法的有效性. 与已有工作相比,本文方法兼顾网络防御决策的实时性和安全状态变化的随机性,最优防御策略选取的时效性、针对性和指导意义更强. 未来工作主要包括进一步研究移动目标防御、拟态防御等主动防御技术,分析其作用机理;研究增强状态转移概率准确性的方法.

参考文献

[1] Moore D, Shannon C, Voelker G M. Budgeting process for information security expenditures[J]. Communications of the ACM, 2016, 49(10): 121 – 125.

[2] Drew Fudenberg, Jean Tirole. Game Theory [M]. Boston: Massachusettes Institute of Technology Press, 2015.

[3] White J, Park J S, Kamhoua C A, et al. Game theoretic attack analysis in online social network services[A]. 2016 International Conference on Social Networks Technology [C]. Los Angeles: IEEE Press, 2016. 312 – 319.

[4] Shordon L, Meliyt J. Network survivability analysis based on game model[J]. Multimedia Information Networking and Security, 2016, 55(5): 199 – 204.

[5] 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.
ZHANG Heng-wei, LI Tao. Optimal active defense based on multi-stage attack-defense signaling game [J]. Acta Electronica Sinica, 2017, 45(2): 431 – 439. (in Chinese)

[6] David W K Yeung, Leon A Petrosyan. Differential Games Theory[M]. New York: Springer Press, 2014.

[7] Nilim A, Ghaoui L E. Active defense strategy selection based ondifferential game[J]. Operations Research, 2016, 43(12): 163 – 169.

[8] 张恒巍,黄世锐. 基于攻防微分博弈的网络安全防御决策方法[J]. 电子学报,2017,45(6):887 – 894.
ZHANG Heng-wei, HUANG Shi-rui. Network defense decision-making method based on attack-defense differential game[J]. Acta Electronica Sinica, 2017, 45(6): 887 – 894. (in Chinese)

[9] Erwin A, Alex P. Modeling analysis and control of network security based on Markov games[J]. IEEE Transactions on Automatic Control, 2017, 57(5): 41 – 49.

[10] SUN Wei, KONG Xiangwei, HE Dequan, et al. Research on attack and deference in information security based on stochastic game [J]. ACM Information security Science and Technology, 2015, 27(9), 1408 – 1412.

[11] 王元卓,林闯,程学旗,等. 基于随机博弈模型的网络攻防量化分析方法[J]. 计算机学报,2014,33(9):1748 – 1764.
WANG Yuan-zhuo, LIN Chuang, CHENG Xue-qi, et al. Analysis fornetwork attack-defense based on stochastic game model[J]. Chinese Journal of Computers, 2014, 33(9): 1748 – 1764. (in Chinese)

[12] 范红旗,王胜. 二人微分对策问题信息模式的数学描述[J]. 电子学报,2015,42(2):355 – 361.
FAN Hong-qi, WANG Sheng. Mathematical descriptionfor information pattern of stochastic differential games [J]. Acta Electronica Sinica, 2015, 42(2): 355 – 361. (in Chinese)

[13] Martin A Nowak. Evolutionary Dynamics: Exploring the Equations of Life[M]. Boston: Harvard University Press, 2013.

[14] Richard Lippmann, Joshua W Haines. Analysis and results of the network intrusion detection evaluation[A]. The 19'th International Workshop on Recent Advances in Intrusion Detection[C]. New York: ACM Press, 2016. 162 – 182.

[15] Nilim A, Ghaoui L E. Robust control of Markov decision processes with uncertain transition matrices [J]. Operations Research, 2016, 53(5): 780 – 798.

[16] Valizadeh M, Koch W. Scalable Simulation Framework on Network [DB/OL]. http://www. ssfnet. org, 2012-11-08/2017-06-16.

[17] Gordon L, Loeb M, Lucyshyn W, Richardson R. 2015 CSI/FBI computer crime and security survey[A]. 2015 Computer Security Institute [C]. San Francisco: IEEE Press, 2015. 169 – 179.

作者简介

张恒巍 男,1978 年出生,河南洛阳人,博士,信息工程大学副教授,研究方向为网络安全与攻防对抗、信息安全风险评估.
E-mail: zhw11qd@126.com

黄世锐(通信作者) 男,1994 年出生,广东汕头人,信息工程大学硕士研究生,研究方向为网络安全预警与主动防御.
E-mail: hsrfzhac@qq.com