基于 Markov 演化博弈的网络防御策略选取方法

张恒巍[1,2],黄健明[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: 当前运用博弈理论的网络安全研究大多采用完全理性假设,本文针对现实社会中攻防双方的有限理性限制条件和攻防过程的动态变化特征,基于非合作演化博弈理论,从有限理性约束出发,将演化博弈模型与 Markov 决策相结合,构建多阶段 Markov 攻防演化博弈模型,实现对多阶段、多状态攻防对抗的动态分析推演;依据博弈的折扣总收益设计目标函数,提出多阶段博弈均衡的求解方法,给出最优防御策略选取算法. 通过实验验证了模型和方法的有效性.

关键词: 网络安全;网络攻防;博弈论;有限理性;演化博弈;网络防御;Markov 决策;多阶段最优防御
中图分类号: TP309 **文献标识码:** A **文章编号:** 0372-2112 (2018)06-1503-07
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2018.06.033

Network Defense Strategy Selection Method Based on Markov Evolutionary Game

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

Abstract: Most research of network security based on the traditional game theory use completely rational assumption. For the condition of bounded rationality and the characteristic of dynamic changing process, we combined the evolutionary game model with Markov decision-making process based on the theory of non-cooperative evolutionary game, which is on the restraint of bounded rationality. Thus we constructed a multi-stage Markov attack-defense evolutionary game model to achieve multi-stage and multi-state dynamic analysis and evolution. Besides, on the basis of the sum discount payoffs to design objective function, the method to solve the evolutionary stable equilibrium was proposed, and the optimal defense strategy selection algorithm was provided. Finally, the validity of the model and method is validated by numerical simulation.

Key words: network security; network attack-defense; game theory; bounded rationality; evolutionary game; network defense; Markov decision-making; multi-stage optimal defense

1 引言

确保网络空间安全是当前亟待解决的问题[1]. 网络安全的本质在于攻防对抗[2]. 博弈论是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[3]. 网络攻防具有的目标对立性、关系非合作性、策略依存性特点均与博弈论的基本特征吻合[4],将博弈模型应用于网络攻防分析已经成为当前的研究热点.

传统的攻防博弈模型遵循博弈人完全理性的基本假设,针对网络攻防过程中攻防双方同时行动或者非同时行动两种场景,可以分别采用静态博弈模型[5,6]或者动态博弈模型[7]研究网络防御决策问题. 当进一步考虑攻防行为信息对博弈过程的影响时,可在引入信号博弈模型[8,9]的基础上,研究有限信息条件下的动态防御决策. 面对现实社会中攻防双方只具

收稿日期:2017-01-16;修回日期:2017-03-23;责任编辑:覃怀银
基金项目:国家自然科学基金(No. 61303074, No. 61309013);信息保障技术重点实验室开放基金(No. KJ-15-110)

备有限理性的现实情况,可以借鉴演化博弈理论,突破博弈人完全理性假设的限制,通过演化博弈模型[10~13]分析网络群体行为,研究网络系统安全状态的演化规律和相应的防御决策方法.另一方面,考虑到网络系统的环境变化以及安全状态转移具有动态性和随机性,研究者在结合博弈论与 Markov 决策方法的基础上,进一步提出攻防随机博弈模型[14,15],研究非确定性条件下的防御决策方法,具有更好的实用性.但是,目前的研究成果均是将矩阵博弈与 Markov 决策方法相结合,难以满足攻防双方的有限理性限制条件.针对上述不足,本文提出多阶段 Markov 攻防演化博弈模型,引入贴现因子 ξ 对不同阶段的博弈收益进行折扣处理,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 多阶段 Markov 攻防演化博弈模型构建

2.1 多阶段 Markov 攻防演化博弈过程分析

在攻防博弈过程中,依据演化博弈基本理论[16],网络攻防系统能够在一定时间内达到演化稳定状态.但是,由于攻防双方的目标、偏好、可行策略集都可能发生变化;系统运行环境也可能发生改变,因此演化稳定状态并非长期稳定、一成不变.当演化稳定状态被打破后,系统以概率 η 从稳定状态跳变到另一个非稳定状态,进而开始下一阶段的演化博弈.从全局视角出发,网络攻防系统处于"演化—跳变—演化"的动态过程中,如图 1 所示.本文在借鉴 Markov 过程[17]的基础上,将各个演化阶段之间的状态跳变描述为随机过程,将多阶段演化博弈与 Markov 决策方法相结合,构建多阶段 Markov 演化博弈.

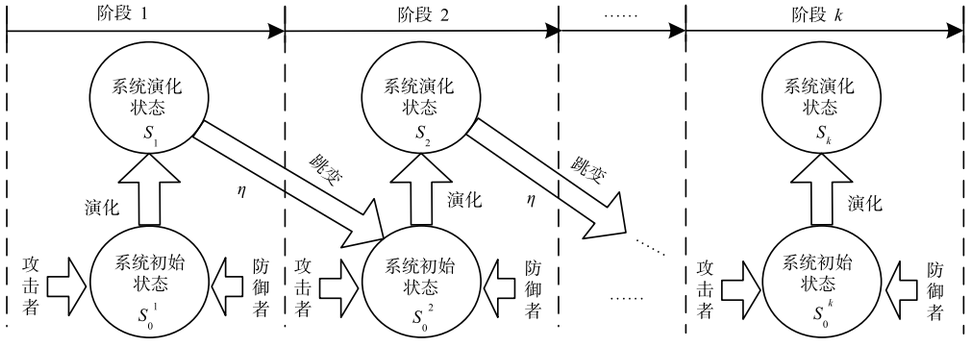

图1 多阶段Markov演化博弈示意图

2.2 单阶段攻防演化博弈模型

首先对单阶段攻防演化博弈进行分析,然后以此为基础构建多阶段 Markov 攻防演化博弈模型.

定义 1 单阶段攻防演化博弈模型 ADE(Attack-Defense Evolutionary game model)表示为四元组,ADE = (N,B,P,U).

① $N=(N_D,N_A)$ 是博弈者空间,N_D 为防御方,N_A 为攻击方.

② $B=(\mathrm{DS},\mathrm{AS})$ 是博弈行动空间,DS 和 AS 分别表示防御者和攻击者的行动策略集.

③ $P=\{p_i,q_j\}$ 是博弈信念集合,p_i 表示攻击策略 AS_i 的选择概率,$p_i\in[0,1]$ 且 $\sum_{i=1}^{m}p_i=1$;q_j 表示防御策略 DS_j 的选择概率,$q_j\in[0,1]$ 且 $\sum_{j=1}^{n}q_j=1$.

④ $U=\{U_D,U_A\}$ 是收益函数集合,U_D 和 U_A 分别表示防御者和攻击者的收益函数.

在攻防博弈中,当防御者以概率 q_j 选取策略 DS_j,攻击者以概率 p_i 选取策略 AS_i 时,用 a_{ij} 和 b_{ij} 分别表示攻击者和防御者的策略收益,则可以采用公式(1)计算不同防御策略的期望收益 U_{DS_i} 和平均防御收益 $\overline{U}_D$.

$$\begin{aligned}U_{\mathrm{DS}_1}&=p_1b_{11}+p_2b_{21}+\cdots+p_mb_{m1}\\U_{\mathrm{DS}_n}&=p_1b_{1n}+p_2b_{2n}+\cdots+p_mb_{mn}\\\overline{U}_D&=q_1U_{\mathrm{DS}_1}+q_2U_{\mathrm{DS}_2}+\cdots+q_nU_{\mathrm{DS}_n}\end{aligned}\tag{1}$$

用 $q_j(t)$ 和 $p_i(t)$ 分别表示在时间 t,防御方和攻击方中选择策略 DS_j 和策略 AS_i 的人数比例,则 $q_j(t)$ 和 $p_i(t)$ 的动态变化速率可用复制动态方程[16]表示为

$$\begin{aligned}D(q_j)&=\frac{\mathrm{d}q_j(t)}{\mathrm{d}t}=q_j(U_{\mathrm{DS}_j}-\overline{U}_D)\\A(p_i)&=\frac{\mathrm{d}p_i(t)}{\mathrm{d}t}=p_i(U_{\mathrm{AS}_i}-\overline{U}_A)\end{aligned}\tag{2}$$

2.3 多阶段 Markov 攻防演化博弈模型

定义 2 多阶段 Markov 攻防演化博弈模型 M^2ADE(Multi-stage Markov Attack-Defense Evolutionary game model)表示为 $\mathrm{M}^2\mathrm{ADE}=(N,T,B,P,\xi,S_0,S,\eta,U)$.

① $N=(N_D,N_A)$ 是博弈者空间,N_D 为防御方,N_A 为攻击方.

② T 是多阶段博弈的阶段总数,当前博弈阶段表

示为 $G(k)$,$k=\{1,\cdots,T\}$,$T\in\mathbb{N}$.

③ $B=(DS,AS)$是攻防行动空间,$AS=\{AS_k^i\mid 1\leqslant k\leqslant T,1\leqslant i\leqslant m\}$,$AS_k^i$ 表示攻击方在第 k 个阶段的可选策略;$DS=\{DS_k^j\mid 1\leqslant k\leqslant T,1\leqslant j\leqslant n\}$,$DS_k^j$ 表示防御方在第 k 个阶段的可选策略.

④ $P=\{p_k^i,q_k^j\}$是博弈信念集合,在博弈阶段 k,p_k^i 表示选择攻击策略 AS_k^i 的概率,$p_k^i\in[0,1]$且 $\sum_{i=1}^{m}p_k^i=1$;q_k^j 表示选择防御策略 DS_k^j 的概率,$q_k^j\in[0,1]$且 $\sum_{j=1}^{n}q_k^j=1$.

⑤ ξ 是折现因子,表示博弈阶段 k 中的收益相较初始阶段的折现比例,$0\leqslant\xi\leqslant 1$.

⑥ $S_0=\{S_0^1\cdots S_0^k\cdots S_0^T\}$是网络系统的初始安全状态集合.

⑦ $S=\{S_1\cdots S_k\cdots S_T\}$是网络系统的安全状态集合.

S_0 和 S 中的状态与博弈阶段对应,在初始状态为 S_0^k 的博弈阶段 $G(k)$中,演化博弈最终达到状态 S_k.

⑧ η 表示安全状态转移概率,$\eta_{ij}=\eta(S_j\mid S_i)$表示系统从状态 S_i 跳变至状态 S_j 的概率.

⑨ $U=\{U_D^k,U_A^k\}$是收益函数集合,U_D^k 和 U_A^k 代表第 k 个博弈阶段中防御者和攻击者的收益函数.

设计目标函数 R,用于判断攻防双方策略的优劣[18]. 由于博弈收益与时间有关,本文采用折扣期望准则设计目标函数,引入折现因子 ξ 计算未来折扣收益值,攻防双方的目标是使各自的目标函数达到最大值.

$$\begin{cases}R_D^k(S_0^k,S_k)=U_D^k+\sum\limits_{e,h\in[k,T]}\xi^h\eta(S_h\mid S_e)R_D^h(S_0^h,S_h)\\ R_A^k(S_0^k,S_k)=U_A^k+\sum\limits_{e,h\in[k,T]}\xi^h\eta(S_h\mid S_e)R_A^h(S_0^h,S_h)\end{cases}\tag{3}$$

3　博弈均衡求解与防御策略选取算法设计

3.1　博弈均衡分析

在演化博弈阶段 $G(k)$,攻防策略分别为 $DS_k=\{DS_k^1,\cdots,DS_k^n\}$和 $AS_k=\{AS_k^1,\cdots,AS_k^m\}$,若$(DS_k^*,AS_k^*)$为第 k 阶段的演化稳定策略,则对于任意攻防策略 DS_k^j、AS_k^i 满足:

$$\begin{cases}U_D^k(DS_k^*,AS_k^*)\geqslant U_D^k(DS_k^j,AS_k^*)\\ U_A^k(DS_k^*,AS_k^*)\geqslant U_A^k(DS_k^*,AS_k^i)\end{cases}\tag{4}$$

由于每个博弈阶段都会受之前阶段攻防策略的影响,根据 Markov 决策准则,参与人必有一个 Markov 最优响应策略[19]. 因此,若$\{(DS_k^*,AS_k^*)\mid 1\leqslant k\leqslant T\}$为 Markov 最优响应策略,则$(DS_k^*,AS_k^*)$使目标函数 $R_D^k(S_0^k,S_k)$和 $R_A^k(S_0^k,S_k)$对任意阶段 k 均满足下列条件:

$$\begin{aligned}DS_k^*\in\operatorname{argmax}R_D^k(S_0^k,S_k)&=\operatorname{argmax}\Big[U_D^k(DS_k^*,AS_k^*)+\sum_{e,h\in[k,T]}\xi^h\eta(S_h\mid S_e)R_D^h(S_0^h,S_h)\Big]\\ AS_k^*\in\operatorname{argmax}R_A^k(S_0^k,S_k)&=\operatorname{argmax}\Big[U_A^k(DS_k^*,AS_k^*)+\sum_{e,h\in[k,T]}\xi^h\eta(S_h\mid S_e)R_A^h(S_0^h,S_h)\Big]\end{aligned}\tag{5}$$

定理 1　多阶段 Markov 攻防演化博弈 M^2DE 存在混合策略下的纳什均衡.

证明　M^2ADE 博弈由多个独立且相似的单阶段演化博弈构成. 一方面,由于每个独立的单阶段演化博弈均属于有限博弈,因此,必定存在混合策略下的纳什均衡[16]. 另一方面,由多阶段 Markov 演化博弈模型的定义,依据转移概率和收益函数可知,存在与 M^2ADE 等价的有限随机博弈,且收益函数为凸函数. 根据有限随机博弈的均衡策略存在性定理[20,21]可知,该有限随机博弈存在混合策略下的纳什均衡. 综上,定理得证.

3.2　博弈均衡求解

3.2.1　单阶段演化博弈均衡求解

首先给出单阶段演化博弈均衡的求解过程和步骤,为计算多阶段博弈均衡提供支持.

(1) 在攻防双方的可选策略集上建立概率推断 p_i 和 q_j;

(2) 计算攻击方和防御方的复制动态方程;

计算攻击策略的期望收益和平均收益

$$U_{AS_i}=\sum_{j=1}^{n}q_ja_{ij},\ \overline{U}_A=\sum_{i=1}^{m}p_iU_{AS_i}\tag{6}$$

得到攻击方的复制动态方程

$$A(p_i)=\frac{\mathrm{d}p_i(t)}{\mathrm{d}t}=p_i(U_{AS_i}-\overline{U}_A)\tag{7}$$

对于防御方,同理可得

$$\begin{aligned}U_{DS_j}&=\sum_{i=1}^{m}p_ib_{ij},\ \overline{U}_D=\sum_{j=1}^{n}q_jU_{DS_j},\\ D(q_j)&=\frac{dq_j(t)}{dt}=q_j(U_{DS_j}-\overline{U}_D)\end{aligned}\tag{8}$$

(3) 计算演化稳定均衡解

联立双方的复制动态方程,得到方程组 $Y=\begin{bmatrix}D(q)\\A(p)\end{bmatrix}=0$,计算可得演化稳定均衡解.

3.2.2　多阶段博弈均衡求解

引入贴现因子 ξ,将未来收益折算成基于初始阶段的折扣收益. 在此基础上,采用动态规划法求解多阶段均衡策略.

对于 $k=\{1,\cdots,T\}$,$AS_k^i\in AS$,$DS_k^j\in DS$

$$\begin{cases}\max R_D^k(S_0^k,S_k) = \max\left[U_D(\mathrm{DS}_k^j,\mathrm{AS}_k^i) + \sum_{e,h\in[k,T]}\xi^h\eta(S_h\mid S_e)R_D^h(S_0^h,S_h)\right]\\ \max R_A^k(S_0^k,S_k) = \max\left[U_A(\mathrm{DS}_k^j,\mathrm{AS}_k^i) + \sum_{e,h\in[k,T]}\xi^h\eta(S_h\mid S_e)R_A^h(S_0^h,S_h)\right]\\ D_k(q_k^j) = \dfrac{\mathrm{d}q_k^j(t)}{\mathrm{d}t} = q_k^j(U_{\mathrm{DS}_k^j} - \overline{U_{D_k}}) = 0\\ A_k(p_k^i) = \dfrac{\mathrm{d}p_k^i(t)}{\mathrm{d}t} = p_k^i(U_{\mathrm{AS}_k^i} - \overline{U}_{A_k}) = 0\\ \sum_{j=1}^n q_k^j = 1,\sum_{i=1}^m p_k^i = 1;q_k^j\in[0,1],p_k^i\in[0,1]\end{cases} \tag{9}$$

求解上述方程可以得到最优解集合$\{(\mathrm{DS}_k^*,\mathrm{AS}_k^*)\}$,依据博弈理论,混合策略$(\mathrm{DS}_k^*,\mathrm{AS}_k^*)$是第$k$阶段攻防双方的最优选择,因此$\mathrm{DS}_k^*$即为最优防御策略.

3.3 最优防御策略选取算法设计与分析

算法1 多阶段 Markov 攻防演化博弈的最优防御策略选取算法

输入:博弈模型 $\mathrm{M}^2\mathrm{ADE}$

输出:多阶段最优防御策略

BEGIN

1. 初始化 $\mathrm{M}^2\mathrm{ADE}=(N,T,B,P,\xi,S_0,S,\eta,U)$;
2. 构建防御行为空间 DS 和攻击行为空间 AS;
3. 构建安全状态集合 $S_0=\{S_0^1\cdots S_0^k\cdots S_0^T\}$ 和 $S=\{S_1\cdots S_k\cdots S_T\}$;
4. 初始化状态转移概率 $\eta_{ij}=\eta(S_j\mid S_i)$;
5. *For* $(k=1;k\leqslant T;k++)$
   {
6. 构建博弈信念集合 $P=\{p_k^i,q_k^j\}$,$p_k^i\in[0,1]$且$\sum_{i=1}^m p_k^i=1$,$q_k^j\in[0,1]$且$\sum_{j=1}^n q_k^j=1$;
7. 针对攻防策略对$(\mathrm{AS}_k^i,\mathrm{DS}_k^j)$,计算攻防收益值$a_k^{ij}$、$b_k^{ij}$;
8. 计算策略的期望收益 $U_{\mathrm{DS}_k^j}=p_k^1b_k^{1j}+p_k^2b_k^{2j}+\cdots+p_k^mb_k^{mj}$,$U_{\mathrm{AS}_k^i}=q_k^1a_k^{i1}+q_k^2a_k^{i2}+\cdots+q_k^na_k^{in}$;
9. 计算攻防双方的平均收益 $\overline{U}_{D_k}=\sum_{j=1}^n q_k^jU_{\mathrm{DS}_k^j}$,$\overline{U}_{A_k}=\sum_{i=1}^m p_k^iU_{\mathrm{AS}_k^i}$;
   }
10. 计算折扣收益 $\sum_{e,h\in[k,T]}\xi^h\eta(S_h\mid S_e)R_h(S_0^h,S_h)$;
11. 基于3.2.2节的动态规划法,以$\max R_D^k(S_0^k,S_k)$和$\max R_A^k(S_0^k,S_k)$为目标函数,求解方程(9),得到博弈均衡解$(\mathrm{DS}_k^*,\mathrm{AS}_k^*)$;
12. Return (DS_k^*);

END

算法的时间复杂度为$O(k(m+n)^2)$,空间复杂度为$O(knm)$.将本文模型与方法同现有研究成果进行比较,如表1所示.考虑到现实社会中的决策行为不可能达到完全理性的状态,因此基于非完全理性限制条件开展攻防博弈研究更加贴合实际.博弈过程是指博弈模型是否具备分析多阶段攻防过程的能力,具备这一能力的模型能够分析更加复杂、多回合的攻防过程,对于防御决策的指导作用更强.均衡求解是指文献中是否给出均衡解的计算过程,由于动态多阶段博弈求解相比单阶段博弈更加困难,尤其是引入 Markov 过程后更加复杂,如果没有详细的计算方法和过程步骤会降低实际应用能力.

表1 对比分析

文献	行为理性	博弈过程	博弈类型	均衡求解	具体应用
[5,6]	完全理性	—	静态博弈	简单	防御策略选取
[7,8]	完全理性	单阶段	动态博弈	详细	动态攻防分析
[11,12,13]	非完全理性	单阶段	演化博弈	简单	防御机制分析
[14,15]	完全理性	多阶段	Markov 矩阵博弈	详细	安全效能评估
本文	非完全理性	多阶段	Markov 演化博弈	详细	防御策略选取

4 仿真实验与分析

4.1 攻防过程仿真

构建如图2所示的信息系统进行实验,系统主要由安全防御设备、Web 服务器、文件服务器、数据库服务器和客户终端组成.

基于文献[20,21]的方法,设定网络攻防过程分为八个阶段,如图3所示.其中,S_0^k为初始状态,S_k为演化状态;实线表示单阶段内的攻防博弈过程,虚线表示安全状态跳变.

参考文献[18,22],假设状态转移概率固定,依据历

史数据和专家经验确定取值,如表 2 所示. $\eta_{ij}=\eta(S_j|S_i)$　　表示从状态 S_i 跳变至状态 S_j 的概率.

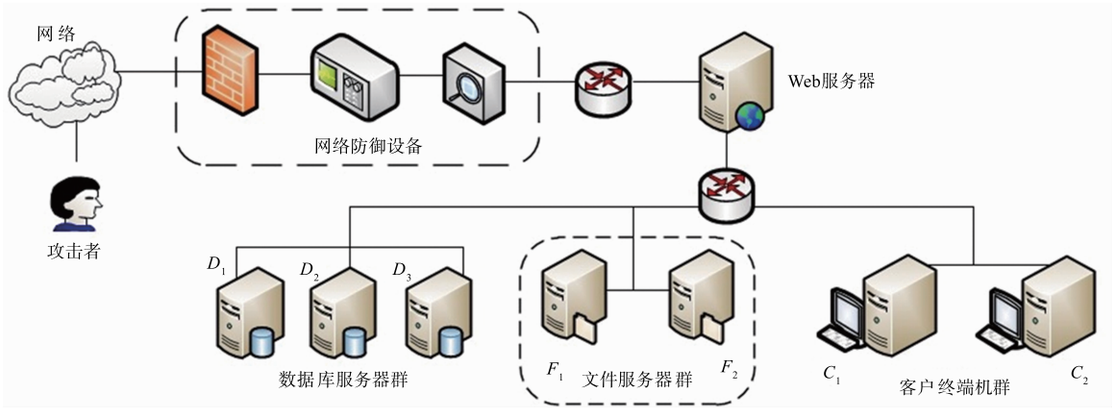

图2　实验系统结构示意图

表 2　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	状态跳变	跳变概率
$S_1\rightarrow S_0^2$	$\eta(2\|1)=0.8$	$S_1\rightarrow S_0^6$	$\eta(6\|1)=0.6$	$S_2\rightarrow S_0^3$	$\eta(3\|2)=0.7$
$S_2\rightarrow S_0^8$	$\eta(8\|2)=0.4$	$S_3\rightarrow S_0^4$	$\eta(4\|3)=0.6$	$S_3\rightarrow S_0^5$	$\eta(5\|3)=0.9$
$S_3\rightarrow S_0^8$	$\eta(8\|3)=0.6$	$S_4\rightarrow S_0^7$	$\eta(7\|4)=0.8$	$S_5\rightarrow S_0^7$	$\eta(7\|5)=0.9$
$S_6\rightarrow S_0^1$	$\eta(1\|6)=0.5$	$S_6\rightarrow S_0^3$	$\eta(3\|6)=0.8$	$S_6\rightarrow S_0^7$	$\eta(7\|6)=0.8$
$S_7\rightarrow S_0^4$	$\eta(4\|7)=0.6$	$S_8\rightarrow S_0^1$	$\eta(1\|8)=0.9$	$S_8\rightarrow S_0^4$	$\eta(4\|8)=0.8$

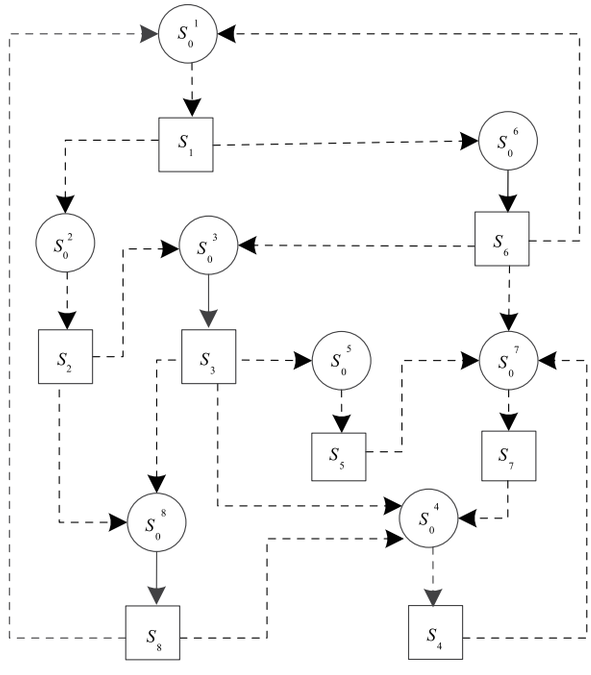

图3　网络攻防状态转移图

通过 Nessus 扫描实验信息系统,结合国家信息安全漏洞库(CNNVD)信息[22],利用文献[9,14,24]的方法在分析路由文件、漏洞信息的基础上,构建各阶段的攻防策略集,如表 3 所示.

表 3　不同博弈阶段的攻防行动可选集

博弈阶段	AS	DS
$S_0^1\rightarrow S_1$	Remote buffer overflow	Limit packets from ports
	Wu-Ftp Sockprintf	Install Oracle patches
	install Web Listener program	Reinstall Listener program
$S_0^2\rightarrow S_2$	install delete Trojan	Uninstall delete Trojan
	Steal account and crack it	Limit access to MDSYS. SDO _ CS
	Remote buffer overflow	Renew data(root)
$S_0^3\rightarrow S_3$	LPC to LSASS process	Restart Database server
	Homepage attack	Limit SYN/ICMP packets
	Attack Address blacklist	Add physical resource
$S_0^4\rightarrow S_4$	Attack SSH on Ftp Sever	Repair database
	install DLI Trojan	Correct homepage
	Steal account and crack it	Delete suspicious account
$S_0^5\rightarrow S_5$	Oracle TNS Listener	Address blacklist
	LPC to LSASS process	Patch SSH on Ftp Sever
	install delete Trojan	Restart Database server
$S_0^6\rightarrow S_6$	Oracle TNS Listener	Restart Database server
	Wu-Ftp Sockprintf	Repair database
	install SQL Listener program	Correct homepage
$S_0^7\rightarrow S_7$	Homepage attack	Reinstall Listener program
	install VBW Trojan	Limit access to MDSYS. SDO _ CS
	Oracle TNS Listener	Add physical resource
$S_0^8\rightarrow S_8$	LPC to LSASS process	Uninstall delete Trojan
	Attack Address blacklist	Repair database
	Remote buffer overflow	Restart Database server

参考文献[7,9,11]的方法,设折现因子 $\xi=0.5$,采

用 Matlab2012 软件实现 3.3 节的策略选取算法,算法运行时间为 32.7s,各阶段的均衡策略如表 4 所示,其中 DS_k^* 为各阶段的最优防御策略.

表 4 各阶段博弈均衡策略

博弈阶段	防御策略 DS_k^*	攻击策略 AS_k^*	攻击收益	防御收益
$S_0^1 \rightarrow S_1$	{0.5,0.5,0}	{0.45,0.4,0.15}	27.7	-30.5
$S_0^2 \rightarrow S_2$	{0.45,0.5,0.05}	{0.47,0.23,0.3}	18.6	-20.4
$S_0^3 \rightarrow S_3$	{0.4,0.55,0.05}	{0.3,0.7,0}	34.8	-43.2
$S_0^4 \rightarrow S_4$	{0.23,0.22,0.55}	{0.8,0.1,0.1}	50.3	-35.9
$S_0^5 \rightarrow S_5$	{0.4,0.5,0.1}	{0.2,0.5,0.3}	23.5	-19.5
$S_0^6 \rightarrow S_6$	{0.4,0.6,0}	{1,0,0}	56.3	-40.3
$S_0^7 \rightarrow S_7$	{1,0,0}	{0.3,0.7,0}	23.9	-34.2
$S_0^8 \rightarrow S_8$	{0.2,0,0.8}	{0.3,0.2,0.5}	32.8	-27.5

4.2 实验分析

以文件服务器 F_1、F_2 为攻击目标,分析仿真过程可知,存在两条主要攻击路径,路径①:$S_0^1 \rightarrow S_1 \rightarrow S_0^2 \rightarrow S_2 \rightarrow S_0^8 \rightarrow S_8$ 和路径②:$S_0^1 \rightarrow S_1 \rightarrow S_0^6 \rightarrow S_6 \rightarrow S_0^3 \rightarrow S_3 \rightarrow S_0^5 \rightarrow S_5$,其中路径①可以获取服务器 F_2 的 root 权限,路径②可以获取服务器 F_1 的 root 权限.

路径①的攻击总收益为 $U_{AZ}^1 = 79.1$,防御总收益为 $U_{DZ}^1 = -78.4$;对于路径②有 $U_{AZ}^2 = 142.3$,$U_{DZ}^2 = -133.5$. 可知 $U_{AZ}^1 < U_{AZ}^2$ 且 $U_{DZ}^1 > U_{DZ}^2$,显然路径①更加符合防御方的愿望. 对比两条路径可以发现,其第 1 阶段相同,但在第 2 阶段,路径①从状态 S_1 跳变至 S_0^2,而路径②从状态 S_1 跳变至 S_0^6. 为降低路径②的发生可能,应当尽量减小概率 $\eta(6|1)$,若能使 $\eta(6|1)=0$,则路径②将不会实现,可以满足防御方的愿望. 进一步分析攻击动作集可知,针对 S_0^6 对应的 AS = {Oracle TNS Listener, Wu-Ftp Sockprintf, install SQL Listener program},防御者可以利用动态调整网络访问端口、增设白名单等方式改变访问控制规则或者增加新的针对性防御策略,降低该攻击动作集的实施可能性,减小跳变至 S_0^6 的概率,降低路径②的发生可能.

5 结论

本文基于多阶段 Markov 攻防演化博弈模型研究了网络防御策略选取问题,主要工作包括:(1) 在分析动态攻防博弈过程的基础上,构建了 Markov 演化博弈模型,具备分析多智能体-多状态攻防行为的能力. (2)基于折扣总收益设计了攻防博弈的目标函数,实现了对多阶段攻防博弈的量化分析. (3)提出了基于动态规划的多阶段博弈均衡计算方法,设计了多阶段最优防御策略选取算法. 研究成果对于在有限理性条件的多阶段动态攻防中实施网络防御决策具有指导意义,能够为开展网络空间攻防对抗研究提供模型和方法支持.

当前攻防手段越来越多样化,网络系统规模不断增大,本文的实验中系统状态集与攻防策略集均规模有限,分析和研究较大规模、复杂环境下网络攻防过程的能力不足. 此外,假设状态转移概率固定,并且主要依赖历史数据和专家经验,准确性和动态性存在不足,需要对状态转移概率从多属性角度开展更加精确的研究. 上述不足是下一步工作的主要关注点.

参考文献

[1] 方滨兴. 从层次角度看网络空间安全技术的覆盖领域[J]. 网络与信息安全学报,2015,1(1):1-6.
FANG Bing-xing. A hierarchy model on the research fields of cyberspace security technology[J]. Chinese Journal of Network and Information Security,2015,1(1):1-6. (in Chinese)

[2] Gordon L, Loeb M. Budgeting process for information security expenditures[J]. Communications of ACM,2016,49(9):121-125.

[3] Drew Fudenberg, Jean Tirole. Game Theory[M]. Boston: Massachusetts Institute of Technology Press,2012.

[4] 朱建明,王秦. 基于博弈论的网络空间安全若干问题分析[J]. 网络与信息安全学报,2015,1(1):43-49.
ZHU Jian-ming, WANG Qin. Analysis of cyberspace security based on game theory[J]. Chinese Journal of Network and Information Security,2015,1(1):43-49. (in Chinese)

[5] Lye K W, Jeannette W. Markov game strategies in network security[J]. Information Security,2005,24(1):71-86.

[6] 余定坤,王晋东,张恒巍. 静态贝叶斯博弈主动防御策略选取方法[J]. 西安电子科技大学学报,2016,43(1):163-169.
YU Ding-kun, WANG Jin-dong, ZHANG Heng-wei. Active defense strategy selection based on static Bayesian game[J]. Journal of Xidian University,2016,43(1):163-169. (in Chinese)

[7] 林旺群,王慧,刘家红. 基于非合作动态博弈的网络安全主动防御技术研究[J]. 计算机研究与发展,2013,48(2):306-316.
LIN Wang-qun, WANG Hui, LIU Jia-hong. Research on active defense technology in network security based on non-cooperative dynamic game theory[J]. Journal of Computer Research and Development,2013,48(2):306-316. (in Chinese)

[8] 张恒巍,王晋东,李涛. 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):32-43.
ZHANG Heng-wei, WANG Jin-dong, Li Tao. Defense policies selection method based on attack-defense signaling

game model[J]. Journal on Communications, 2016, 37(5):32 – 43. (in Chinese)

[9] 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)

[10] 王元卓,于建业,邱雯. 网络群体行为的演化博弈模型与分析方法[J]. 计算机学报,2015,38(2):282 – 300.
WANG Yuan-zhuo, YU Jian-ye, QIU Wen. Evolutionary game model and analysis methods for network group behavior[J]. Chinese Journal of Computers,2015,38(2):282 – 300. (in Chinese)

[11] 朱建明,宋彪,黄启发. 基于系统动力学的网络安全攻防演化博弈模型[J]. 通信学报,2014,35(1):54 – 61.
ZHU Jian-ming, SONG Biao, HUANG Qi-fa. Evolution game model of offense-defense for network security based on system dynamics[J]. Journal on Communications, 2014,35(1):54 – 61. (in Chinese)

[12] SUN Wei, KONG Xiangwei, HE Dequan. Research on attack and deference in information security based on evolutionary game[J]. Information Science,2016,27(9),1408 – 1412.

[13] LIU Fengming, DING Yongsheng. Dynamics analysis of evolutionary game-based trust computing for P2P networks[J]. Application Research of Computers, 2016, 33(8):2460 – 2463.

[14] 姜伟,方滨兴,田志宏. 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,2013,47(10):1714 – 1723.
JIANG Wei, FANG Bing-xing, TIAN Zhi-hong. Research on defense strategies selection based on attack-defense stochastic game model[J]. Journal of Computer Research and Development, 2013, 47(10):1714 – 1723. (in Chinese)

[15] 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,2013,33(9):1748 – 1764.
WANG Yuanzhuo, LIN Chuang, CHENG Xueqi, et al. Analysis for network attack-defense based on stochastic game model[J]. Chinese Journal of Computers, 2013, 33(9):1748 – 1764. (in Chinese)

[16] Herbert Gintis. Game Theory Evolving[M]. Boston: Priceton University Press,2015.

[17] 方兆本,廖伯其. 随机过程[M]. 北京:科学出版社,2015.
FANG Zhao-ben, LIAO Bo-qi. Stochastic Process[M]. Beijing: China Science Press,2015. (in Chinese)

[18] 张勇. 基于 Markov 博弈模型的网络安全态势感知方法[J]. 软件学报,2016,22(3):495 – 508.
ZHANG Yong. Network security situation awareness approach based on Markov game model[J]. Journal of Software,2016,22(3):495 – 508. (in Chinese)

[19] Borkovsky R N, Doraszelski U, Kryukov Y. A user's guide to solving dynamic stochastic games using the homotopy method[J]. Operation Research, 2015, 58(4): 1116 – 1132.

[20] Doraszelski U, Escobar J F. A theory of regular Markov perfect equilibria in dynamic stochastic games genericity, stability and purification[J]. Theoretical Economics, 2015,5(2):369 – 402.

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

[22] China National Vulnerability Database of Information Security[DB/OL]. http://www.cnnvd.org.cn,2015 – 05 – 23/2016 – 10 – 26.

[23] Maleki H, Valizadeh M H, Koch W, et al. Markov modeling of moving target defense games[J]. Cryptology and Security Engineering,2016,29(10):47 – 83.

[24] 张恒巍,余定坤. 信号博弈网络安全威胁评估方法[J],西安电子科技大学学报,2016,43(3):137 – 143.
ZHANG Heng-wei, YU Ding-kun. Network security threat assessment based on signaling game[J]. Journal of Xidian University,2016,43(3):137 – 143. (in Chinese)

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,主要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

黄健明(通信作者)　男,1992 年出生,湖南张家界人,信息工程大学硕士研究生,主要研究方向为网络安全主动防御.
E-mail:hjm-i-jbb@126.com