基于多阶段 Markov 信号博弈的移动目标防御最优决策方法

蒋 侶[1],张恒巍[1,2],王晋东[1]
(1. 战略支援部队信息工程大学三院,河南郑州 450001; 2. 河南省信息安全重点实验室,河南郑州 450001)

摘 要: 随着移动目标防御技术研究的不断深入,移动目标防御策略选取问题成为当前研究的热点问题之一,本文提出一种基于多阶段 Markov 信号博弈模型的移动目标防御最优策略选取方法. 首先,结合攻防实际,提出实施攻击所需构建的攻击链模型. 其次,在考虑状态随机跳变的基础上,将多阶段信号博弈模型与 Markov 决策过程相结合,构建基于多阶段 Markov 信号博弈的移动目标防御模型. 同时,引入 Logistic 映射刻画攻防博弈系统中可能造成概率更新过程失真的随机干扰因素. 在形式化建模的基础上,设计折扣收益目标函数,并提出均衡求解算法,给出最优防御策略选取算法. 最后,通过仿真实验验证模型和方法的有效性.

关键词: 移动目标防御; Markov 决策; 多阶段信号博弈; 最优策略选取; logistic 映射
中图分类号: TP309 **文献标识码:** A **文章编号:** 0372-2112 (2021)03-0527-09
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.12263/DZXB.20191070

A Markov Signaling Game-Theoretic Approach to Moving Target Defense Strategy Selection

JIANG Lü[1], ZHANG Heng-wei[1,2], WANG Jin-dong[1]
(1. *The Third Institute, PLA SSF Information Engineering University, Zhengzhou, Henan* 450001, *China*;
2. *Henan Key Laboratory of Information Security, Zhengzhou, Henan* 450001, *China*)

Abstract: With the development of the research on moving target defense (MTD) technique, how to effectively select the optimal strategy of moving target defense has become an urgent issue in the current research. To solve this problem, we propose a MTD optimal strategy selection method based on multi-stage Markov signaling game model. Firstly, combined with the actual attack-defense process, we construct an attack chain model that attackers need to build to carry out the attack. Secondly, due to the random jump between states, we combine multi-stage signaling with markov decision process (MDP) to construct the corresponding MTD model. Meanwhile, we adopt Logistic mapping to characterize the stochastic interference factors that may cause the distortion of the probability updating in the attack-defense process. Additionally, on the basis of formally modeling, we design an objective function with discounted total payoff. Besides, we give a solution method for multi-stage signaling game equilibrium and design an optimal defense strategy selection algorithm for MTD. Finally, the simulation demonstrates the effectiveness and feasibility of the proposed model and method.

Key words: moving target defense; Markov decision; multi-stage signaling game; optimal strategy selection; logistic mapping

1 引言

近年来,随着网络技术的飞速发展,网络安全事件频发,对国家重要领域造成了巨大的安全损失[1]. 网络安全的本质在攻防对抗,当前攻防过程中“易攻难守”的特点十分突出. 首先,攻击者有时间优势,能够对目标系统内核等关键信息进行渗透测试和脆弱性分析;其次,攻击者有能力优势,目标系统架构的确定性、静态性和同构性为攻击者提供了较好静态分析条件;最后,攻击者有成本优势,攻击者只需花费少量代价找到破坏

收稿日期:2019-09-19;修回日期:2020-10-26;责任编辑:梅志强
基金项目:河南省科技攻关计划基金(No. 182102210144)

系统的某个“点”[2]. 网络安全现阶段亟需一种主动防御技术.

移动目标防御(Moving Target Defense,MTD)作为一种主动防御技术,能够有效提升防御效能,增强目标系统防御能力[3]. 如何高效利用 MTD 技术进行网络攻防对抗,增强目标系统的防御效能成为 MTD 领域的重要研究热点问题. 网络攻防对抗过程中存在的目标对立性、关系非合作性以及策略依存性与博弈论的基本特征相符[4]. 将博弈论用于 MTD 防御决策问题逐渐成为近年来的研究热点.

目前,利用博弈理论研究移动目标防御决策已经取得了部分成果. 文献[5]将 MTD 攻防对抗抽象成一种单阶段完全信息静态博弈模型,将最优防御策略转化成目标系统安全性和可用性之间的平衡,但完全信息假设与实际攻防过程不符. 文献[6]提出一种基于不完全信息静态博弈的 MTD 模型,该模型中攻击者需攻破 k 个平台漏洞,将最优防御策略抽象为最大化平台差异性,但实际攻防对抗过程是动态变化的,后行动者往往能够观察到先行动者的行为. 文献[7]提出一种基于不完全信息动态博弈的 MTD 模型,采用贝叶斯攻击图量化每个攻击面的安全等级,将最优防御策略抽象为防御策略有效性与防御成本之间的均衡. 文献[8]针对 Web 应用环境下的 MTD 决策问题,提出基于 Bayesian Stackelberg 博弈的 MTD 模型,分析了目标系统漏洞重要程度与不同敏感程度的攻击者对攻防收益的影响,并据此寻找最优策略. 以上研究成果均未考虑攻防信息对攻防对抗过程的影响,忽略了攻防信息可能起到的主动防御效果.

文献[9]针对 DDoS(Distributed Denial of Service Attack,拒绝式服务攻击)的防御决策问题,采用信号博弈模型研究攻防行为和信号作用机理,在均衡分析的基础上设计了防御决策算法. 文献[10]采用多阶段信号博弈建模攻防场景,分析攻防双方的博弈收益,设计了一种信息安全防御机制. 文献[11]提出了一种重复信号博弈模型,将多阶段信号博弈视为单阶段博弈的重复. 文献[12]采用信号博弈模型研究 MTD 对抗过程,并给出精炼贝叶斯均衡的求解算法,同时设计了一种最优策略选取算法. 文献[13]提出了一种基于信号博弈的 MTD 模型,将攻击策略视为攻击者的一种信号,防御者据此选取最优防御策略,但该模型未能体现信号机制的主动防御效果. 以上研究成果未考虑到实际 MTD 攻防过程中,攻防双方策略的改变以及网络系统运行环境的改变均会导致攻防系统状态的动态变化,并具有随机性.

针对上述问题,本文以信号博弈理论为基础,结合 MTD 攻防实际,提出基于多阶段 Markov 信号博弈的移动目标防御决策方法,主要工作如下:

(1) 在分析动态攻防博弈的基础上,将 Markov 决策过程与多阶段信号博弈模型相结合,构建移动目标防御多阶段 Markov 信号博弈模型,用于多阶段攻防过程分析.

(2) 在考虑状态随机跳变的基础上,同时考虑博弈系统中存在的随机干扰因素对阶段间概率更新的影响,提出了 Logistic 映射,用于刻画概率更新过程中受到的随机干扰影响.

(3) 在给出的攻防策略收益量化方法基础上,给出了精炼贝叶斯均衡的求解步骤,并设计了移动目标防御多阶段 Markov 信号博弈的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可行性.

2 多阶段 Markov 信号博弈模型构建

2.1 多阶段 Markov 信号博弈过程分析

移动目标防御技术以防御方可控的方法通过改变目标系统的资源属性增加攻击者的攻击难度[14]. 信号博弈(signaling game)是研究信号如何影响博弈均衡及博弈收益的一种博弈理论.

在多阶段信号博弈模型中,从博弈初始状态开始,博弈系统经过动态攻防对抗达到某个相对稳定状态,由于攻防策略是随时间变化的,攻击目的、攻防双方偏好以及目标系统运行环境可能发生改变,直接导致稳定状态无法维持,并随机转移到新状态,从而开始下一阶段的信号博弈. 与此同时,相邻阶段间,上一阶段后验概率更新成下一阶段先验概率的过程受到随机干扰因素(Random Interference,RI)的影响,如攻击者观测能力有限以及系统环境发生变化等,导致后验概率在更新过程中受到干扰发生失真. 但由于影响攻防双方可行策略集和目标系统状态改变的因素较多,作用原理复杂,并具有一定的随机性. 本文借鉴 MDP(Markov Decision Process,MDP,马尔科夫决策过程)刻画不同阶段间状态随机跳变,采用 Logistic 映射描述后验概率更新过程中的失真现象,将多阶段 Markov 信号博弈和 Logistic 映射相结合,构建带 Logistic 映射的多阶段 Markov 信号博弈模型对 MTD 攻防行为和防御决策进行分析.

2.2 攻防策略收益量化

根据文献[14,15]中关于攻防成本与收益计算的定义,本文从探测面扩展和攻击面转换的角度对攻防策略收益进行量化分析.

防御成本通常由 MTD 攻击面转换成本、探测面扩展成本、MTD 系统负面成本和信号发送成本. 攻击成本是指攻击者为构建攻击链所需的时间、专业知识以及软硬件资源等. 防御策略有效性(effectiveness of defense strategies)是指当攻击者采用攻击策略 a_k,防御者采用

MTD 策略 d_s 的有效性,记为 $\varepsilon(a_k,d_s)$,$\varepsilon(a_k,d_s)$取值可根据 CNNVD(China National Vulnerability Database of Information Security,中国国家信息安全漏洞库)得到. 如果防御策略能够完全阻止攻击,$\varepsilon(a_k,d_s)=1$;当防御策略完全无效时,$\varepsilon(a_k,d_s)=0$,其他情况时,$0<\varepsilon(a_k,d_s)<1$. 系统损失代价(System Damage Cost,SDC)是指攻防双方在对抗导致系统功能发生故障或敏感信息丢失所带来的损失,通常取正值,其值大小通常由资源重要程度,攻击致命度,资源属性损害进行描述[14]. 系统资源重要程度由防御者进行定义并赋值,攻击致命度由防御策略有效性参数给出,资源属性伤害可根据攻击导致的性能降低给出赋值,通常与该资源的 CPU 占用率和占用时间成反比.

根据文献[16],攻击收益(Attack Payoff,AP)是指攻击者通过攻击链所获得总收益,防御收益(Defense Payoff,DP)是指防御者通过采取 MTD 策略保护目标信息系统所获得的收益.

$$\mathrm{AP}=(1-\varepsilon)\mathrm{SDC}(a_k,d_s)-\mathrm{AC}(a_k) \tag{1}$$

$$\mathrm{DP}=-(1-\varepsilon)\mathrm{SDC}(a_k,d_s)-\mathrm{DC}(\theta_i,m_j,d_s) \tag{2}$$

2.3　移动目标防御多阶段 Markov 信号博弈模型

定义 1　多阶段 Markov 信号博弈模型(Moving Target Defense Multi-Stage Markov Signaling Game Model,MTD_M^{2}SGM)可表示为十二元组$(N,\Theta,T,B,M,p_k,\tilde{p}_k,S_0,S,\xi,\eta,U)$. 其中

(1)$N=(N_a,N_d)$是信号博弈模型局中人集合,N_a为攻击者,N_d为 MTD 系统.

(2)$\Theta=(\Theta_a,\Theta_d)$分别是局中人 N_a 和 N_d 的类型空间.

(3)T 是多阶段信号博弈的阶段总数,$G(k)$表示第 k 个阶段博弈过程,$k\in[1,T]$,$k\in N^+$.

(4)$B=(D,A)$是攻防双方的策略空间. $D=\{d_i^k\mid 1\leqslant k\leqslant T,1\leqslant i\leqslant g\}$,$d_i^k$ 为防御者在阶段 $G(k)$中第 i 个策略;$A=\{a_1,a_2,\cdots,a_h\}$为攻击策略,$g,h\geqslant 1$.

(5)$M=\{m_1,m_2,\cdots,m_n\}$是防御者的信号空间,防御信号名称与防御类型相对应,防御者可以自主选择发送真实或虚假信号.

(6)p_k 是攻击者在阶段 k 的先验概率,它是在随机干扰因素作用下由上阶段后验概率更新得到,可记为 $p_k=(p_1^k,p_2^k,\cdots,p_n^k)$,满足 $p_i{}^k=p_k(\theta_i)\geqslant 0$,$\sum_{i=1}^{n}p_i{}^k=1$. 采用 Logistic 映射刻画随机干扰因素对概率更新过程的影响,即 $x_{n+1}=\mu x_n(1-x_n)$,$x_0=\tilde{p}_{k-1}$,$x_n=p_k$,迭代次数为 $n=40$,$\mu=3.9$.

(7)$\tilde{p}_k$ 是攻击者在 k 阶段的后验概率,$\tilde{p}_k(\theta_i\mid m_j)$表示攻击者接收到防御信号 m_j 时,利用贝叶斯法则修正先验概率 p_k 得到的防御类型 θ_i 的后验概率.

(8)$S_0=\{S_0^1,S_0^2,\cdots,S_0^T\}$是攻防过程中 MTD 系统的各阶段的初始状态集合.

(9)$S=\{S_1,S_2,\cdots,S_T\}$是攻防对抗过程中 MTD 系统的相对稳定状态集合.

(10)ξ 是折扣因子,ξ^k 表示在博弈阶段 k 中的收益相较初始阶段的折现比例,$0\leqslant\xi^k\leqslant 1$.

(11)η 是 MTD 系统状态转移概率,其中 $\eta_{ij}=\eta(S_0^j\mid S_i)$表示系统从状态 S_i 跳变至状态 S_0^j 的概率. 转移概率 η_{ij}取决于攻防双方的策略和 MTD 系统环境. 通过 Markov 概率矩阵进行刻画,当 $i=j$,$\eta_{ij}=0$.

(12)$U=(U_A^k,U_D^k)$表示攻防双方在第 k 个阶段的收益函数集合.

设计目标准则函数 R,用于判断攻防双方策略的优劣. 常用目标准则函数有折扣期望准则函数和平均回报准则函数[16]. 由于移动目标防御对抗过程中,攻防收益与时间有关,本文引入折现因子 ξ,采用折扣期望准则函数,即

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

3　均衡求解与算法设计

3.1　均衡分析

当攻防博弈阶段处于 $G(k)$时,攻防策略和信号策略分别为 $D^k=\{d_1^k,d_2^k,\cdots,d_g^k\}$,$A^k=\{a_1^k,a_2^k,\cdots,a_h^k\}$和 $M=\{m_1,m_2,\cdots,m_n\}$. 根据多阶段信号博弈均衡定理[17],若(d^{k*},a^{k*},m^{k*})为博弈阶段 $G(k)$的精炼贝叶斯均衡,则满足下列条件:

$$\begin{cases}\forall d_i^k,U_X^k(d_i^k,a^{k*},m^{k*})\geqslant U_X^k(d^{k*},a^{k*},m^{k*})\\ \forall a_i^k,U_X^k(d^{k*},a_i^k,m^{k*})\geqslant U_X^k(d^{k*},a^{k*},m^{k*})\quad X=\{A,D\}\\ \forall m^k,U_X^k(d^{k*},a^{k*},m^k)\geqslant U_X^k(d^{k*},a^{k*},m^{k*})\end{cases} \tag{4}$$

由于攻防对抗过程由多个博弈阶段构成,且每个阶段会受到上一阶段攻防策略的影响. 根据 Markov 决策过程理论[18],若(d^{k*},a^{k*},m^{k*})是 MTD_M^{2}SGM 的 Markov 最优响应策略,则(d^{k*},a^{k*},m^{k*})使目标准则函数 $R_D^k(S_0^k,S_k)$和 $R_A^k(S_0^k,S_k)$达到最大值,当 $k\in[1,T]$,满足下列条件:

$$\begin{aligned}(d^{k*},m^{k*})\in\operatorname{argmax}R_D^k(S_0^k,S_k)&=U_D^k(d^{k*},a^{k*},m^{k*})\\&+\sum_{h\in[k,T]}\xi^h\eta_{kh}(S_0^h\mid S_k)R_D^h(S_0^h,S_h)\\ a^{k*}\in\operatorname{argmax}R_A^k(S_0^k,S_k)&=U_A^k(d^{k*},a^{k*},m^{k*})\\&+\sum_{h\in[k,T]}\xi^h\eta_{kh}(S_0^h\mid S_k)R_A^h(S_0^h,S_h)\end{aligned} \tag{5}$$

MTD_M²SGM 为多阶段-多状态的有限 Markov 信号博弈模型. 根据信号博弈基本定理[19]和文献[20,21]的结论可证明多阶段均衡策略存在性.

3.2 均衡求解

定义 2 单阶段博弈的精炼贝叶斯均衡可表示为 $(m^*(\theta,d^*),a^*(m),\tilde{p}^*(\theta|m))$,其中 $m^*(\theta,d^*)$ 表示防御类型为 θ 的防御者释放信号 m^* 且选择防御策略 $d^*(m^*)$,简记为 $m^*(\theta)$;$a^*(m)$ 为攻击者的信号依存策略;$\tilde{p}^*(\theta|m)$ 为攻击者对防御者的后验概率. 精炼贝叶斯均衡满足以下条件:

$$(1)\ a^*(m)\in\underset{a\in A}{\operatorname{argmax}}\sum\tilde{p}(\theta|m)U_A(\theta,m,a)$$

$$(2)\ m^*(\theta)\in\underset{m\in M}{\operatorname{argmax}}U_D(\theta,m,a^*(m))$$

$\tilde{p}^*(\theta|m)$ 是攻击者根据先验概率 p、防御信号 m 以及最优策略组合 $(m^*(\theta),a^*(m))$ 通过贝叶斯法则得到的. 具体求解方法详见文献[19].

针对移动目标防御多阶段攻防对抗的收益计算问题,引入折扣因子 ξ,将未来阶段的收益折算成基于初始阶段的折扣收益;同时引入 Logistic 映射,刻画阶段间影响后验概率更新的随机干扰因素. 本文将 MTD_M² SGM 的均衡求解问题,转化为以最大化整体收益的动态规划问题.

$$\forall k\in[1,T],d_i^k\in D,a_i^k\in A$$

$$\begin{cases}\max R_D^k(S_0^k,S_k)=\max\Big[U_D^k(d^{k*},a^{k*},m^{k*})+\sum\limits_{h\in[k,T]}\xi^h\eta_{kh}(S_0^h\mid S_k)R_D^h(S_0^h,S_h)\Big]\\ \max R_A^k(S_0^k,S_k)=\max\Big[U_A^k(d^{k*},a^{k*},m^{k*})+\sum\limits_{h\in[k,T]}\xi^h\eta_{kh}(S_0^h\mid S_k)R_A^h(S_0^h,S_h)\Big]\\ x_0=\tilde{p}_{k-1},x_n=p_k,x_{n+1}=\mu x_n(1-x_n)\\ \mu=3.9,n=40,k=2,3,\cdots,T\end{cases}\tag{6}$$

其中,d^k,m^k 和 a^k 分别表示在博弈阶段 $G(k)$,防御方防御策略、信号策略和攻击策略. 通过求解方程式(6)即可得到攻防双方最优策略 $(m^{k*}(\theta,d^{k*}),a^{k*}(m))$. 攻防双方在此策略组合下双方收益达到最大值,故防御方应将 $m^{k*}(\theta,d^{k*})$ 作为最优防御策略.

3.3 最优防御策略选取算法设计与对比分析

基于上述分析,设计基于多阶段 Markov 信号博弈的移动目标防御最优决策算法.

算法 1 基于多阶段 Markov 信号博弈的移动目标防御最优决策算法

Input:移动目标防御多阶段 Markov 信号博弈模型 MTD_M²SGM

Output:多阶段最优防御策略 $m^{k*}(\theta,d^{k*})$

BEGIN

1. 初始化博弈模型 MTD_M²SGM;
2. 构建防御策略空间 D 和攻击策略空间 A;
3. 构建各阶段安全状态集合 $S_0=\{S_0^1,S_0^2,\cdots,S_0^T\}$ 和 $S=\{S_1,S_2,\cdots,S_T\}$;
4. 初始化 Logistic 映射:$x_{n+1}=\mu x_n(1-x_n)$,$x_0=\tilde{p}_{k-1}$,$x_n=p_k$,迭代次数为 $n=40$,$\mu=3.9$,$k=2,3,\cdots,T$;
5. 初始化状态转移概率 $\eta_{ij}=\eta(S_0^j|S_i)$,阶段 $k=1$;;
6. While($k\leqslant T$);

{ //计算不同阶段的单阶段信号博弈收益

(1) 构建 $p_k=(p_1^k,p_2^k,\cdots,p_n^k)$,满足 $\sum\limits_{i=1}^{n}p_i^k=1$;

(2) 分别计算 $G(k)$ 阶段攻防双方收益函数 $U_A^k(\theta,m,a)$ 和 $U_D^k(\theta,m,a)$;

(3) 利用折扣因子 ξ,计算攻防双方折扣收益 $\sum\limits_{h\in[k,T]}\xi^h\eta_{kh}(S_0^h|S_k)R_X^h(S_0^h,S_h)$;

(4) 基于 3.2 节动态规划算法,以 $\max R_D^k(S_0^k,S_k)$ 和 $\max R_A^k(S_0^k,S_k)$ 作为目标函数,求解 $(m^{k*}(\theta,d^{k*}),a^{k*}(m))$;

(5) 利用贝叶斯公式 $p_k(\theta|m)=\dfrac{p_k(m|\theta)p_k(\theta)}{\sum\limits_{\theta'}p_k(m|\theta')p_k(\theta')}$ 计算后验概率 $\tilde{p}_k^*(\theta|m)$;

(6) For($i=1;i<n;i++$)

{$x_0=\tilde{p}_k^*(\theta|m)$; $x_{i+1}=\mu x_i(1-x_i)$;}

(7) $p_{k+1}(\theta|m)=x_{40}$;

(8) Output($(m^{k*}(\theta,d^{k*}))$; //输出当前阶段最优防御策略

(9) $k=k+1$;

}

END

令 Logistic 映射迭代次数为 r,平均时间复杂度为 $o(k(u^3+n^2+2n+r))$,平均空间复杂度为 $o(k(un+r))$. 将本文提出的模型及方法和其它文献进行对比,如表 1 所示.

表 1 模型与算法对比分析

文献	博弈过程	博弈类型	信号发送	均衡求解	具体应用
文献[11]	单阶段	完全信息静态	防御者	详细	策略选取
文献[13]	单阶段	单阶段信号博弈	防御者	详细	攻防分析
文献[16]	多阶段	信号博弈	攻击者	一般	策略选取
文献[22]	多阶段	马尔科夫博弈	—	一般	均衡求解
文献[21]	多阶段	马尔科夫博弈	—	详细	策略选取
本文	多阶段	马尔科夫信号博弈	防御者	详细	策略选取

4 仿真实验及分析

4.1 仿真环境描述及过程分析

为验证 MTD_M²SGM 模型及最优防御策略选取算法的可行性与有效性,构建如图 1 所示的系统拓扑结构. 实验网络系统主要由业务网、接入网和外部互联网构成,主要包括网络防御设备、Web 服务器、文件服务器、数据库服务器和客户端.

采用文献[16,23]的方法分析仿真实验结构,将移动目标防御攻防对抗过程分为八个阶段,各状态如

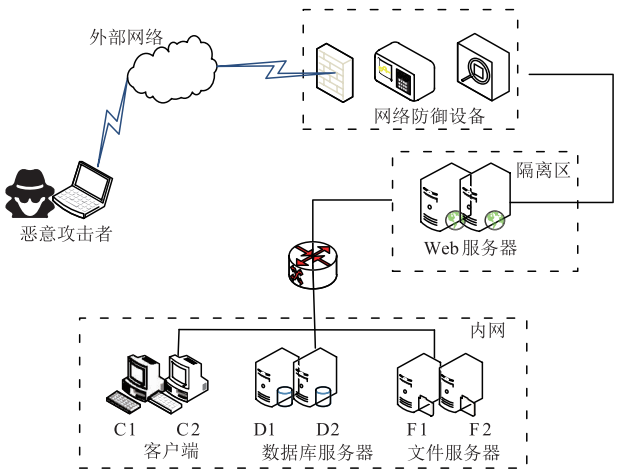

图1　仿真实验系统拓扑示意图

表 2 所示.

表 2　仿真实验系统不同阶段下状态表

攻击链阶段	状态	状态描述
扫描检测	S_0^1	实验系统各节点处于正常状态
	S_1	获取网络安全防御设备 root 权限
	S_0^2	获取 Web 服务器 access 权限
	S_2	获取 Web 服务器 user 权限
脆弱性利用	S_0^3	获取文件服务器 F1 user 权限
	S_3	获取客户端 C2 user 权限
	S_0^4	获取文件服务器 F2 user 权限
	S_4	获取数据库服务器 D1 access 权限
攻击植入	S_0^5	获取客户端 C1 root 权限
	S_5	获取数据库服务器 D2 user 权限
	S_0^6	获取文件服务 F2 root 权限
	S_6	获取客户端 C2 root 权限
攻击维持	S_0^7	获取数据库服务器 D1 root 权限
	S_7	获取敏感信息并破坏数据库服务器 D1
	S_0^8	获取数据库服务器 D2 root 权限
	S_8	数据库服务器 D2 被植入木马

其中 S_0^k 是阶段 $G(k)$ 的初始状态,S_k 是阶段 $G(k)$ 的结束状态. 通过历史数据和专家经验[24] 确定 η_{ij}. 各阶段更新后的概率矩阵如表 3 所示.

表 3　状态转移概率及概率更新

状态跳变	概率更新	状态跳变	概率更新
$S_1\to S_0^2$ $\eta(2\mid 1)=0.7$	$\begin{bmatrix}0.5 & 0.5\\ 0.4 & 0.6\end{bmatrix}$	$S_2\to S_0^4$ $\eta(4\mid 2)=0.8$	$\begin{bmatrix}0.3 & 0.7\\ 0.8 & 0.2\end{bmatrix}$
$S_1\to S_0^4$ $\eta(4\mid 1)=0.4$	$\begin{bmatrix}0.4 & 0.6\\ 0.6 & 0.4\end{bmatrix}$	$S_4\to S_0^5$ $\eta(5\mid 4)=0.8$	$\begin{bmatrix}0.5 & 0.5\\ 0.9 & 0.1\end{bmatrix}$
$S_4\to S_0^8$ $\eta(8\mid 4)=0.4$	$\begin{bmatrix}0.4 & 0.6\\ 0.5 & 0.5\end{bmatrix}$	$S_5\to S_0^8$ $\eta(8\mid 5)=0.5$	$\begin{bmatrix}0.6 & 0.4\\ 0.5 & 0.5\end{bmatrix}$
$S_5\to S_0^7$ $\eta(7\mid 5)=0.9$	$\begin{bmatrix}0.6 & 0.4\\ 0.3 & 0.7\end{bmatrix}$	$S_2\to S_0^3$ $\eta(3\mid 2)=0.7$	$\begin{bmatrix}0.6 & 0.4\\ 0.3 & 0.7\end{bmatrix}$
$S_3\to S_0^4$ $\eta(4\mid 3)=0.8$	$\begin{bmatrix}0.4 & 0.6\\ 0.1 & 0.9\end{bmatrix}$	$S_3\to S_0^5$ $\eta(5\mid 3)=0.3$	$\begin{bmatrix}0.3 & 0.7\\ 0.3 & 0.7\end{bmatrix}$
$S_4\to S_0^6$ $\eta(6\mid 4)=0.7$	$\begin{bmatrix}0.7 & 0.3\\ 0.8 & 0.2\end{bmatrix}$	$S_5\to S_0^6$ $\eta(6\mid 5)=0.6$	$\begin{bmatrix}0.2 & 0.8\\ 0.7 & 0.3\end{bmatrix}$
$S_6\to S_0^7$ $\eta(7\mid 6)=0.9$	$\begin{bmatrix}0.9 & 0.1\\ 0.3 & 0.7\end{bmatrix}$	$S_7\to S_0^8$ $\eta(8\mid 7)=0.7$	$\begin{bmatrix}0.8 & 0.2\\ 0.4 & 0.6\end{bmatrix}$

通过扫描工具扫描整个仿真系统,根据国家信息安全漏洞库(CNNVD)数据,利用文献[21,23]的分析方法将防御类型分为高等级防御类型 θ_{H} 和低等级防御类型 θ_{L},构建攻防策略集,给出防御策略有效性参数 $\varepsilon_{ij}=(a_i,d_j)$,如表 4 所示. 针对各阶段攻防策略 $\{a_i^k,d_j^k\}$,参考文献[11,13]的计算方法,计算各阶段的攻防收益,如表 5 所示. 设定折扣因子 $\xi=0.4$,利用 Matlab2015 工具实现计算各阶段目标准则函数值,得到防御者最优信号策略 $m^*(\theta)$、最优防御策略 $d^*(m)$、攻击者最优攻击策略 $a^*(m)$,如表 6 所示.

表 4　各阶段攻防策略及防御策略有效性

博弈阶段	攻击策略 A	防御类型	防御策略 D	防御策略有效性 ε
$S_0^1\to S_1$	Steal account and crack it	θ_{H}	Delete account + Fixed Frequency	$\begin{bmatrix}0.8 & 0.6 & 0.3\\ 0.3 & 0.6 & 1\\ 0.2 & 0.5 & 0.9\end{bmatrix}$
	Oracle TNS listener		IP Enlarging + Random Frequency	
	install Web Listener program	θ_{L}	Reinstall Listener + Port Enlarging	
$S_0^2\to S_2$	install delete Trojan	θ_{H}	Protocol Changing + Random Frequency	$\begin{bmatrix}0.5 & 0.6 & 1\\ 0.3 & 0.7 & 0.2\\ 1 & 0.8 & 0.2\end{bmatrix}$
	LPC to LSASS process		IP Enlarging + Route Enlarging	
	SMTP sniffer	θ_{L}	Uninstall delete Trojan	
$S_0^3\to S_3$	LPC to LSASS process	θ_{H}	Fingerprint switch + Fixed Frequency	$\begin{bmatrix}0.4 & 0.8 & 0.8\\ 0.4 & 0.6 & 0.3\\ 0.2 & 0.3 & 0.6\end{bmatrix}$
	Install socket analyzer		Limit packets + Fixed Frequency	
	Attack Address blacklist	θ_{L}	Blacklist Enlarging	

续表

博弈阶段	攻击策略 A	防御类型	防御策略 D	防御策略有效性 ε
$S_0^4 \rightarrow S_4$	Shutdown server tenor install DLI Trojan install VBW Trojan	θ_H θ_L	Restart database IP Hopping + port Changing Delete Trojan + Random Frequency	$\begin{bmatrix} 0.9 & 0.1 & 0.3 \\ 0.2 & 0.3 & 1 \\ 0.2 & 0.3 & 1 \end{bmatrix}$
$S_0^5 \rightarrow S_5$	Oracle TNS Listener THS chunk overflow install delete Trojan	θ_H θ_L	Storage Enlarging Delete trojan + Random Frequency Port Changing + Fixed Frequency	$\begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.8 & 0.1 & 0.2 \\ 0.1 & 1 & 0.3 \end{bmatrix}$
$S_0^6 \rightarrow S_6$	Ssh buffer overflow Wu-FtpSockprintf install SQL Listener program	θ_H θ_L	IP Enlarging + IP Hopping StorageEnlaring + Fixed Frequency Protocol changing + Random Frequency	$\begin{bmatrix} 0.3 & 0.9 & 0.3 \\ 0.8 & 0.2 & 0.7 \\ 0.4 & 0.1 & 0.6 \end{bmatrix}$
$S_0^7 \rightarrow S_7$	Shutdown database server Steal account and crack Oracle TNS Listener	θ_H θ_L	IP Hopping + port Changing Port Changing + Fixed Frequency Route changing + Random Frequency	$\begin{bmatrix} 0.3 & 0.2 & 0.1 \\ 0.8 & 0.8 & 0.8 \\ 0.7 & 0.8 & 0.6 \end{bmatrix}$
$S_0^8 \rightarrow S_8$	LPC to LSASS process install SQL Listener Oracle TNS Listener	θ_H θ_L	IP Enlarging + Random Frequency Port Changing + Random Frequency Protocol changing + Fixed Frequency	$\begin{bmatrix} 0.8 & 0.9 & 0.7 \\ 0.5 & 0.6 & 0.5 \\ 0.5 & 0.8 & 0.4 \end{bmatrix}$

表 5 各阶段攻防收益

博弈阶段	θ_H^k	θ_L^k
$S_0^1 \rightarrow S_1$	$\begin{bmatrix} (5,-8) & (12,-20) & (18,-25) \\ (15,-20) & (10,-18) & (-4,-4) \\ (22,-30) & (18,-22) & (-1,-2) \end{bmatrix}$	$\begin{bmatrix} (8,-10) & (15,-20) & (12,-15) \\ (15,-10) & (12,-12) & (-4,-2) \\ (18,-22) & (20,-22) & (6,-3) \end{bmatrix}$
$S_0^3 \rightarrow S_3$	$\begin{bmatrix} (20,-10) & (8,-10) & (5,-5) \\ (10,-20) & (5,-5) & (10,-10) \\ (20,-10) & (5,-6) & (12,-8) \end{bmatrix}$	$\begin{bmatrix} (10,-4) & (6,0) & (5,-10) \\ (10,-5) & (15,-10) & (2,-10) \\ (5,0) & (0,-5) & (3,-5) \end{bmatrix}$
$S_0^5 \rightarrow S_5$	$\begin{bmatrix} (10,-10) & (8,-15) & (10,-10) \\ (10,-4) & (4,-8) & (8,-4) \\ (4,-8) & (10,-5) & (15,-8) \end{bmatrix}$	$\begin{bmatrix} (12,-4) & (8,-15) & (10,-10) \\ (4,-6) & (5,-5) & (4,-8) \\ (8,-6) & (10,-5) & (6,-10) \end{bmatrix}$
$S_0^7 \rightarrow S_7$	$\begin{bmatrix} (12,-5) & (8,-12) & (10,-10) \\ (12,-5) & (8,-12) & (4,0) \\ (4,0) & (10,-14) & (2,-2) \end{bmatrix}$	$\begin{bmatrix} (20,-25) & (0,-15) & (20,-10) \\ (12,-5) & (5,-8) & (4,-4) \\ (12,-8) & (8,-8) & (4,-8) \end{bmatrix}$
$S_0^2 \rightarrow S_2$	$\begin{bmatrix} (12,-4) & (8,-15) & (10,-10) \\ (5,-5) & (0,-12) & (8,-8) \\ (4,-0) & (10,-10) & (10,-8) \end{bmatrix}$	$\begin{bmatrix} (6,-10) & (8,-10) & (10,-4) \\ (5,-5) & (20,-12) & (10,-2) \\ (5,-8) & (8,-6) & (4,-8) \end{bmatrix}$
$S_0^4 \rightarrow S_4$	$\begin{bmatrix} (5,-10) & (15,-18) & (8,-10) \\ (6,-10) & (10,-8) & (15,-6) \\ (10,0) & (5,-5) & (10,0) \end{bmatrix}$	$\begin{bmatrix} (12,-4) & (8,-5) & (10,-10) \\ (6,-2) & (0,-8) & (12,-2) \\ (12,-4) & (15,-5) & (5,-5) \end{bmatrix}$

续表

博弈阶段	θ_H^k	θ_L^k
$S_0^6 \rightarrow S_6$	$\begin{bmatrix} (5,-5) & (12,-8) & (10,-20) \\ (5,-6) & (10,-10) & (4,-8) \\ (10,-2) & (10,-4) & (5,-8) \end{bmatrix}$	$\begin{bmatrix} (10,-2) & (8,-6) & (10,-2) \\ (4,-8) & (10,-20) & (4,-6) \\ (15,-20) & (10,-6) & (10,-8) \end{bmatrix}$
$S_0^8 \rightarrow S_8$	$\begin{bmatrix} (8,-12) & (6,-10) & (12,-10) \\ (14,-6) & (6,0) & (14,-12) \\ (12,-10) & (4,-10) & (12,-10) \end{bmatrix}$	$\begin{bmatrix} (2,-4) & (0,-4) & (2,-6) \\ (14,-16) & (12,-14) & (10,-8) \\ (8,-10) & (4,-4) & (16,-12) \end{bmatrix}$

表 6　各阶段均衡策略

博弈阶段	防御策略 $d^*(m)$	信号策略 $m^*(\theta)$	攻击策略 $a^*(m)$	攻击收益 R_A^k	防御收益 R_D^k
$S_0^1 \rightarrow S_1$	d_3	m_{H}	a_3	19.3	-70.2
$S_0^2 \rightarrow S_2$	d_3	m_{L}	a_1	12.4	-51.6
$S_0^3 \rightarrow S_3$	d_2	m_{L}	a_1	20.8	-44.8
$S_0^4 \rightarrow S_4$	d_1	m_{L}	a_2	18.6	-55.4
$S_0^5 \rightarrow S_5$	d_2	m_{H}	a_3	21.9	-40.2
$S_0^6 \rightarrow S_6$	d_1	m_{L}	a_1	28.4	-60.6
$S_0^7 \rightarrow S_7$	d_3	m_{H}	a_2	34.2	-50.5
$S_0^8 \rightarrow S_8$	d_2	m_{H}	a_1	14.6	-20.8

4.2　仿真结果分析

实验中的攻击链主要有两条,具体分析如下:

(1)攻击链 1:$S_0^1 \rightarrow S_1 \rightarrow S_0^2 \rightarrow S_2 \rightarrow S_0^4 \rightarrow S_4 \rightarrow S_0^6 \rightarrow S_6 \rightarrow S_0^7 \rightarrow S_7$

第 1 阶段 $S_0^1 \rightarrow S_1$,攻击者通过被动监听网络安全防御设备,分析并获取系统脆弱性;第 2 阶段 $S_0^2 \rightarrow S_2$,系统以概率 $\eta_{12}=0.7$ 从 S_1 跳变到 S_0^2,攻击者扫描并分析 Web 服务器可能存在的脆弱性;第 3 阶段 $S_0^4 \rightarrow S_4$,系统以概率 $\eta_{24}=0.8$ 从 S_2 跳变到 S_0^4,攻击者利用 Web 服务器存在的脆弱性并以之为跳板,获取 F2 的 user 权限和 D1 的 access 权限;第 4 阶段 $S_0^6 \rightarrow S_6$,系统以概率 $\eta_{46}=0.7$ 从 S_4 跳变到 S_0^6,获取 F2 的 root 权限和 C2 的 root 权限;第 5 阶段 $S_0^7 \rightarrow S_7$,系统以概率 $\eta_{67}=0.9$ 从 S_6 跳变到 S_0^7,获取 D1 的 root 权限并窃取 D1 中的敏感信息并破坏数据库 D1. 通过表 5 可得,攻击链 1 的攻击总收益 $R_{AT1}=112.9$,防御总收益 $R_{DT1}=-288.3$.

(2)攻击链 2:$S_0^1 \rightarrow S_1 \rightarrow S_0^2 \rightarrow S_2 \rightarrow S_0^3 \rightarrow S_3 \rightarrow S_0^5 \rightarrow S_5 \rightarrow S_0^8 \rightarrow S_8$

前两个阶段与攻击链 1 一致;第 3 阶段 $S_0^3 \rightarrow S_3$,系统以概率 $\eta_{23}=0.7$ 从 S_2 跳变到 S_0^3,攻击者主要利用 Web 服务器存在的脆弱性并以之为跳板,获取 F1 的 user 权限和 C2 的 user 权限;第 4 阶段 $S_0^5 \rightarrow S_5$,系统以概率 $\eta_{35}=0.3$ 从 S_3 跳变到 S_0^5,获取 C1 的 root 权限和 D2 的 user 权限;第 5 阶段 $S_0^8 \rightarrow S_8$,系统以概率 $\eta_{58}=0.5$ 从 S_5 跳变到 S_0^8,获取 D2 的 root 权限并向 D2 植入木马便于下次攻击. 通过表 5 可得,攻击链 2 的攻击总收益 $R_{AT2}=89$,防御总收益 $R_{DT2}=-227.6$.

通过分析上述两条攻击链可知,从防御者的角度来看,因为 $R_{DT1}<R_{DT2}$,$R_{AT1}>R_{AT2}$,攻击链 2 显然更符合防御者的期望,防御者应尽量避免攻击链 1 的形成. 对比分析攻击链 1 和攻击链 2 可以发现,第 1、2 阶段两者相同;在第 3 阶段,攻击链 1 跳变到 S_0^4,攻击链 2 跳变到 S_0^3,为降低攻击链 1 形成的概率,需要减小 $\eta(S_2 \mid S_0^4)$,分析 $S_0^4 \rightarrow S_4$ 过程中的攻击策略 A^4,防御者可以利用随机改变端口信息、增设黑名单等方式改变访问控制策略或增加策略(delete DLI Torjan)的频率以及其他针对性强的防御策略,降低 $\eta(S_2 \mid S_0^4)$ 的值,削弱攻击链 1 的可能性.

5　结束语

移动目标防御是一种改变攻防态势地位的主动防御技术,在实际应用中具有很好的效果和巨大的潜力,如何进行最优策略选取是移动目标防御研究领域的关键问题之一. 本文提出基于多阶段 Markov 信号博弈的移动目标防御决策方法,解决了在多阶段攻防博弈过

程中,随机因素影响状态转换和后验概率更新过程,并给出了求解精炼贝叶斯均衡解的求解步骤,设计了最优防御策略求解算法.下一步工作主要将从实时连续的角度研究移动目标防御对抗过程,提出时效性强的防御决策方法.

参考文献

[1] Bin-xing F. A hierarchy model on the research fields of cyberspace security technology[J]. Chinese Journal of Network & Information Security,2015,1(01):2 -7.

[2] Jajodia S. Moving Target Defense: Creating Asymmetric Uncertainty for Cyber Threats[M]. New York: Springer-Verlag,2011. 42 -45.

[3] 蔡桂林,王宝生,王天佐,罗跃斌,王小峰,崔新武. 移动目标防御技术研究进展[J]. 计算机研究与发展,2016,53(05):968 -987.
CAI G L,WANG B S,WANG T Y,et al. Research on development ofmoving target defense technology[J]. Journal of Computer Research and Development,2016,53(05):968 -987. (in Chinese)

[4] 刘效武,王慧强,吕宏武,禹继国,张淑雯. 网络安全态势认知融合感控模型[J]. 软件学报,2016,27(08):2099 -2114.
LIU X W,WANG H Q,LV H W,et al. Fusion-based cognitive awareness-control model for network security situation[J]. Journal of Software,2016,27(08):2099 -2114. (in Chinese)

[5] 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(01):43 -49.
ZHU J M,WANG Q. Analysis of cyberspace security based on game theory[J]. Chinese Journal of Network and Information Security,2015,1(01):43 -49. (in Chinese)

[6] Pratyusa K. Manadhata. Game theoretic approaches to attack surface shifting [J]. ACM Transactions on Information and System Security,2017,23(2):145 -153.

[7] Carter K M,Riordan J F,Okhravi H. A game theoretic approach to strategy determination for dynamic platform defenses[A]. Proceedings of the First ACM Workshop on Moving Target Defense[C]. New York: ACM,2014. 21 -30

[8] Kambhampati S,et al. Moving target defense for web applications using bayesian stackelberg games:(extended abstract)[A]. 2016 International Conference on Autonomous Agents and Multiagent Systems[C]. Singapore: AAMAS,2016. 1377 -1378.

[9] Filler T,Judas J,Fridrich J. Signaling game model:DDOS defense analysis [J]. Journal of Security Engineering,2016,39(3):414 -417.

[10] Zhang H,Dingkun Y U,Han J,et al. Network security threat assessment based on the signaling game[J]. Journal of Xidian University,2016,43(3):137 -143. .

[11] 张恒巍,余定坤,韩继红,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(05):51 -61.
ZHANG H W,YU D K,HAN J H,et al. Network security threat assessment based on the signaling game[J]. Journal on Communications,2016,37(05):51 -61. (in Chinese)

[12] Okhravi H,Comella A,Robinson E,et al. Creating a cyber moving target for critical infrastructure applications using platform diversity [J]. International Journal of Critical Infrastructure Protection,2014,5(1):30 -39.

[13] Feng X,Zheng Z,Cansever D. A signaling game model for moving target defense[A]. INFOCOM 2017-IEEE Conference on Computer Communications[C]. Piscataway: IEEE,2017. 1 -9.

[14] Huang S R,Zhang H W,et al. Markov differential game for network defense decision-making method [J]. IEEE Access,2018,6:39621 -39634

[15] Lei C,Zhang H Q,Wan L M,et al. Incomplete information Markov game theoretic approach to strategy generation for moving target defense[J]. Computer Communications,2018,116:184 -199.

[16] 刘江,张红旗,刘艺. 基于不完全信息动态博弈的动态目标防御最优策略选取研究[J]. 电子学报,2018,46(1):82 -89.
LIU J,ZHANG H Q,LIU Y. Research on optimal selection of moving target defense policy based on dynamic game with incomplete information[J]. Acta Electronica Sinica,2018,46(1):82 -89. (in Chinese)

[17] Zhu Q,Başar T. Game-theoretic approach to feedback-driven multi-stage moving target defense[A]. International Conference on Decision and Game Theory for Security[C]. Berlin: Springer,2013. 246 -263.

[18] Manadhata P K,Wing J M. An attack surface metric[J]. IEEE Transactions on Software Engineering,2011,37(3):371 -386.

[19] Drew Fudenberg,Jean Tirole. Game Theory [M]. Boston: Massachusetts Institute of Technology Press,2012. 270 -295.

[20] Gao X,Zhu Y F. DDos Defense mechanism analysis based on signaling game model[A]. Proceedings of the 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics[C]. Piscataway: IEEE Press,2013. 414 -417.

[21] Zangeneh V,Shajari M. A cost-sensitive move selection strategy for moving target defense[J]. Computers & Security,2018,75(JUN.):72 -91.

[22] Sengupta S,Chowdhary A,Huang D,et al. General sum

Markov games for strategic detection of advanced persistent threats using moving target defense in cloud networks [A]. International Conference on Decision and Game Theory for Security[C]. Berlin:Springer,2019. 492 – 512

[23] Maleki H, Valizadeh S, Koch W, et al. Markov modeling of moving target defense games[A]. Proceedings of the 2016 ACM Workshop on Moving Target Defense[C]. New York:ACM,2016. 81 – 92.

[24] Gordon L, Loeb M, Lucyshyn W, Richardson R. 2014 CSI/FBI computer crime and security survey[A]. Proceedings of the Computer Security Institute[C]. San Francisco, California:IEEE Press,2014. 11 – 34.

作者简介

蒋　侣　男,1995 年出生,四川广安人,战略支援部队信息工程大学硕士,助理工程师,主要研究方向为移动目标防御、网络安全与攻防对抗.

张恒巍(通信作者)　男,1978 年出生,河南洛阳人,博士,战略支援部队信息工程大学副教授、研究生导师,研究方向为网络安全博弈、信息安全风险评估、智能系统安全测评.
E-mail:zhw11qd@126. com

王晋东　男,1966 年出生,山西洪桐人,战略支援部队信息工程大学教授,主要研究方向为网络与信息安全、云资源管理.