基于 Markov 时间博弈的移动目标防御最优策略选取方法

谭晶磊[1,2]，张恒巍[1]，张红旗[1,2]，金辉[1,2]，雷程[1,2]

（1. 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：针对现有博弈模型难以有效建模网络攻防对抗动态连续特性的问题，提出了一种基于 Markov 时间博弈的移动目标防御最优策略选取方法。在分析移动目标攻防对抗过程的基础上，构建了移动目标攻防策略集合，利用时间博弈刻画了单阶段移动目标防御过程的动态性，利用 Markov 决策过程描述了多阶段移动目标防御状态转化的随机性。同时，将攻防双方对资源脆弱性抽象为对攻击面控制权的交替，从而有效保证了博弈模型的通用性。在此基础上，分析并证明了均衡的存在性，设计了最优策略选取算法。最后，通过应用实例验证了所提模型的实用性和算法的有效性。

关键词：时间博弈；移动目标攻击；移动目标防御；最优策略选取；Markov 决策

中图分类号：TN918.1

文献标识码：A

doi: 10.11959/j.issn.1000−436x.2020003

Optimal strategy selection approach of moving target defense based on Markov time game

TAN Jinglei[1,2], ZHANG Hengwei[1], ZHANG Hongqi[1,2], JIN Hui[1,2], LEI Cheng[1,2]

1. Department of Three, Information Engineering University, Zhengzhou 450001, China
2. Henan Key Laboratory of Information Security, Zhengzhou 450001,China

Abstract: For the problem that the existed game model was challenging to model the dynamic continuous characteristics of network attack and defense confrontation effectively, a method based on Markov time game was proposed to select the optimal strategy for moving target defense. Based on the analysis of the attack and defense confrontation process of moving targets, the set of moving target attack and defense strategies was constructed. The dynamics of the single-stage moving target defense process was described by time game. The randomness of multi-stage moving target defense state transformation was described by Markov decision process. At the same time, by abstracting the use of resource vulnerability by attack-defense participants as the alternation of the control of the attack surface, the versatility of the game model was effectively guaranteed. On this basis, the existence of equilibrium was analyzed and proved, and the optimal strategy selection algorithm was designed. Finally, the practicality of the constructed model and the effectiveness of the algorithm are verified by an application example.

Key words: time game, moving target attack, moving target defense, optimal strategy selection, Markov decision

1 引言

全球性网络安全攻防竞赛[1]已经达到前所未有的强度，各类网络攻击事件愈演愈烈，网络攻击者不断制定新的攻击策略。其中，移动目标攻击（MTA, moving target attack）技术是最受攻击者欢迎的攻击

收稿日期：2019–08–02；**修回日期**：2019–09–21

基金项目：国家重点研发计划基金资助项目（No.2016YFF0204002, No.2016YFF0204003）；国家自然科学基金资助项目（No.61902427）

Foundation Items: The National Key Research and Development Program of China (No.2016YFF0204002, No.2016YFF0204003), The National Natural Science Foundation of China (No.61902427)

方法之一，它利用各种不确定的攻击手段隐藏攻击意图，并试图逃避传统网络防御的检测机制。由于传统网络防御机制无法准确预知攻击者下一步攻击行动，MTA 技术在网络攻防博弈中逐渐获得竞争优势，这不仅对网络空间造成了很大的安全威胁，而且产生了高昂的防御成本。

近年来，网络安全战略经历了从被动防御到主动防御的演化升级，新兴的移动目标防御（MTD, moving target defense）技术[2]已经成为平衡网络安全竞争环境的新方法，它通过引入动态性、随机性以及异构性来保护网络空间，旨在利用攻击面的动态变换打破网络系统的静态特性，向攻击者呈现一个不可预测的网络状态，以阻止攻击者的恶意行为，增加攻击者攻击成功的难度。

MTA 与 MTD[3]依据攻防成本和收益选取最优策略攻防收益最大化，具有关系非合作性。在移动目标攻防对抗过程中，MTA 试图通过各种攻击手段控制系统攻击面，将攻击面的暴露范围不断扩大，为后续持续性攻击做好准备；而 MTD 则通过动态化、随机化和多样化的方法控制系统攻击面，转移或者减少系统攻击面，以拒止 MTA 的攻击行动，因而移动目标攻防双方具有目标对立性。移动目标攻防双方对于最优策略的选取不仅仅取决于自身，同时也取决于对手，因此移动目标攻防双方具有策略依存性。移动目标攻防过程所具有的关系非合作性、目标对立性和策略依存性与博弈论的理论特性相契合，博弈论可在选取移动目标防御最优策略的研究中发挥重要作用。

姜伟等[4]提出了一种基于完全信息博弈的最优防御策略选取算法，通过构建攻防随机博弈模型，预测攻击行为，并由此制定最优防御策略。林旺群等[5]提出了基于完全信息动态博弈的最优策略，通过引入“虚拟节点”将攻击图转换为博弈树，采用非合作动态博弈求解最优防御策略，但是该模型并没有给出详细的策略选取算法。Manadhata 等[6]则提出了基于随机博弈的最优攻击面变换方法，为了权衡安全性和可用性，将移动目标防御形式化为二人随机博弈。然而，单阶段博弈难以有效刻画移动目标防御持续动态变化的特性，因此 Vadlamudi 等[7]提出了基于贝叶斯攻击图的移动目标防御最优策略选取方法，它利用贝叶斯攻击图描述了攻击方利用的脆弱性间的关联关系，以及防御方可观测到的攻击行为和网络安全状态，但是仍然难以表征攻防对抗的动态性。为了刻画 MTD 攻防对抗的动态连续特性，Lei 等[8]将攻防双方对资源脆弱性的利用抽象为攻击面和探测面的变化，并在收益函数的计算中考虑了跳变的性能消耗。由于攻防双方的行为策略会导致网络系统状态的改变，且状态转移具有 Markov 性，Maleki[9]提出了基于 Markov 的移动目标防御博弈模型，通过将 Markov 决策过程与博弈模型相结合，对单目标 IP 跳变和多目标 IP 跳变策略进行分析，证明多元素跳变可以有效提高防御的收益，但是基于 Markov 的博弈收益量化仍然依赖攻防对抗的历史数据和专家经验。

虽然现有的研究取得了一定成果，但在模型构建和收益量化方面仍存在不足。一方面，现有的研究工作大多基于随机博弈、贝叶斯博弈等博弈模型，难以有效刻画 MTD 攻防的动态连续特性；另一方面，现有的收益量化方法大都基于历史数据与专家经验表征刻画，难以保证决策结果的客观准确性。基于此，本文引入时间博弈进行博弈的动态性刻画，并利用时间博弈隐蔽对抗的特性构建 MTD 攻防模型，基于 Markov 决策过程表征 MTD 状态的随机迁移特性，通过攻防双方对攻击面的控制时间量化攻防收益。

2　移动目标攻防集合策略构建

2.1　移动目标攻击策略

移动目标攻击体系已经逐步发展并不断完善，常见的移动目标攻击技术如表 1 所示。

表 1　　移动目标攻击技术分类

移动目标攻击策略	具体方法
多态 MTA	变换恶意软件签名
自修改 MTA	动态变换恶意软件代码
混淆 MTA	隐藏恶意活动
自加密 MTA	变换恶意软件签名，并隐藏恶意代码和数据
反虚拟机/反沙箱 MTA	变换追踪环境中的行为，规避自动取证分析
反调试 MTA	变换追踪环境中的行为，规避自动/手动调查
目标漏洞利用 MTA	变换参数和签名，规避自动/手动调查
行为改变 MTA	执行前等待真实的用户活动

多态 MTA 可以有效规避防御者入侵检测系统的特征检测。一方面，多态 MTA 使用多个加密密钥生成相同恶意软件的不同实例，由于新实例具有新的未知静态签名，使基于签名的反恶意软件防御无效。另一方面，多态 MTA 有效载荷（代码和数据）是加密的，可以绕过防御者的深层静态分析。

多态 MTA 通过更改内存中的代码使防御者的攻击检测过程复杂化。

与多态 MTA 类似，自修改 MTA 可以有效规避文件和内存的自动扫描，而混淆 MTA 则可以有效逃避手动检查代码。混淆 MTA 所创建的具有混淆性的代码通常难以被传统检测手段发现，它可以创建带有模糊字符串的有效负载、虚拟代码和复杂的函数调用图，并随机生成恶意软件实例。自加密 MTD 则通过变换恶意软件签名来隐藏恶意代码和数据。

反虚拟机/反沙箱 MTA 是另一种移动目标攻击方法，恶意软件分析通常利用虚拟机或沙箱环境检测恶意软件的运行活动，如果检测到虚拟机或沙箱，则反虚拟机/反沙箱 MTA 会改变其行为并避免任何恶意活动。一旦在真实系统上执行并被标记为良性之后，它就会开始其恶意行为。

反调试 MTA 可以避免调试和运行时的检测分析。如果反调试 MTA 在运行时检测到调试工具，则会更改其执行流程保持良性操作。如果它未被调试工具检测到，则会启动恶意行为。

目标漏洞利用 MTA 可以更改统一资源定位符（URL, uniform resource locator）模式、主机服务器、加密密钥和文件名，还可以通过限制来自相同 IP 地址的漏洞访问次数来规避蜜罐防御。

行为改变 MTA 通常在真实用户交互后发动攻击，因而它可以确保在真实机器上执行攻击。

这些有效的移动目标攻击方法为攻击者赢得了不对称的攻击优势，使传统防御技术处于被动不利的局面。攻击者明确自己的攻击对象、攻击时间、攻击目标和攻击方式，而防御者则处于不确定状态，只能利用大量的成本、时间和资源来规避攻击者可能发起的任何攻击探测和入侵活动。因此，防御者和攻击者之间不存在理论上的对称性。

2.2 移动目标防御策略

防止移动目标攻击的最佳方法是使用基于移动目标防御的新安全解决方案。2009 年，美国国家赛博跨越式发展年会首先提出了移动目标防御这一概念，提出移动目标防御通过持续变换系统呈现给攻击方的攻击面，从而有效增加攻击方探测目标节点脆弱性的代价[10]。2012 年，美国白宫国防安全委员会在赛博空间安全研究进展报告[11]中明确了移动目标的概念，即移动目标是可在多个维度上通过移动来降低攻击方优势并增加系统弹性的技术手段。2014 年，《可改变游戏规则的赛博空间安全研究与发展建议》中则将移动目标防御定义为一种创建、分析、评估和部署多样化、持续时变的机制和策略，以增加攻击实施的复杂度与成本，限制和降低系统脆弱性曝光度和被攻击的概率，提高系统弹性的防御手段[12]。

移动目标防御是一种新的主动防御思想，它通过移动或伪装攻击者探测的资源以扰乱应用程序存储器。当恶意软件获得对移动目标防御保护系统的访问权限时，它无法找到所需的易受攻击的资源以造成损害。就其本质而言，移动目标防御与攻击无关，因此可以有效抵御已知和未知攻击的多种变化。之前的研究[13]已经总结概述了它的基本理论框架，如图 1 所示。

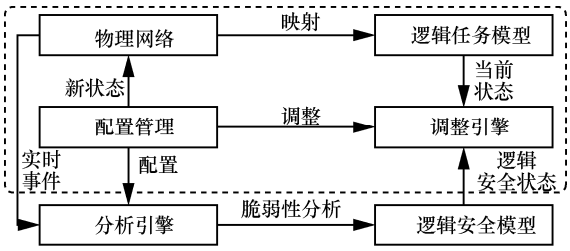

图 1 MTD 理论框架

将物理网络映射到逻辑任务模型，由调整引擎获取逻辑任务模型的当前状态，并由配置管理调整产生新状态进行适应，分析引擎获取物理网络的实时事件，利用传统防御中入侵检测、防火墙等检测机制进行脆弱性分析，由逻辑安全模型产生逻辑安全状态发送给调整引擎，形成一个闭合自反馈的动态调整系统。

移动目标防御技术研究是针对系统不同要素、安全威胁和应用场景设计的可行防御策略，分为系统层 MTD 和网络层 MTD 这 2 个层面，其中，系统层 MTD 包括硬件 MTD 和软件 MTD，网络层 MTD 包括 MAC、IP、协议、路径、操作系统、指纹以及端口 MTD，具体如表 2 所示。本文所采用的移动目标防御策略为网络层 MTD。

3 移动目标防御模型构建

3.1 移动目标防御时间博弈过程分析

2013 年，针对 APT，美国 RSA 实验室的 Dijk [14] 首次提出了时间博弈，与现有的大多数博弈模型不同，时间博弈由防御者和攻击者这 2 个局中人以及公共资源构成，它允许局中人在任意时刻采取行动来控制资源。然而，在局中人实际移动之前，不会显示资源控制权，因此隐蔽性是时间博弈的最大特

点。每个局中人的目标是最大化控制资源时间，同时最小化移动成本。在移动目标攻防过程中，根据时间博弈基本理论，网络攻防系统中的局中人共同争夺对公共资源（攻击面）的控制权，尽可能地最大化自身的收益，图 2 显示了随着时间变化，移动目标攻击者（灰色）和移动目标防御者（黑色）之间的公共资源控制权的切换。

表 2　　移动目标防御策略集合分类

序号	分类名称		具体方法
1	系统层 MTD	软件 MTD	变换应用程序、操作系统、数据
		硬件 MTD	变换处理器
2	网络层 MTD	MAC-MTD	变换 MAC 地址
		IP-MTD	变换 IP 地址
		Procotol -MTD	变换协议
		路径 MTD	变换路径
		OS-MTD	变换操作系统
		Finger-MTD	变换指纹
		Port-MTD	变换端口

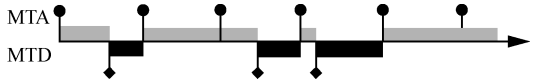

图 2　移动目标防御时间博弈说明示例

移动目标攻防策略的实施都需要付出一定的成本，其中，移动目标攻击者的目标是破坏网络关键服务，并尽可能降低攻击成本。移动目标防御者的目标是增加安全防御预算，减缓或阻止攻击行为，以最大化移动目标攻击者的攻击成本。移动目标攻防双方都需要对系统攻击面进行控制，不同的是，攻击者是利用攻击面可用的脆弱性资源发起攻击，而防御者则是改变或减少攻击面脆弱性资源来提高攻击者的攻击难度，攻击面的控制权会随着局中人的行动发生变化。因此，利用时间博弈刻画单阶段移动目标防御过程更符合真实网络攻防场景。

本文首先利用时间博弈模型刻画单阶段移动目标防御过程，接着从全局视角出发，借鉴 Markov 决策过程[15]，将各博弈阶段之间的状态迁移描述为随机过程，将多阶段时间博弈与 Markov 决策方法相结合，构建多阶段 Markov 时间博弈并进行均衡求解。

3.2　Markov 时间博弈移动目标防御模型构建

首先，对单阶段时间博弈进行分析，如定义 1 所示。

定义 1　单阶段时间博弈模型（STG-MTD）。STG-MTD 表示为六元组(N,B,R,η,U,T)，具体如下。

1) $N=\{N_{\mathrm{MTA}},N_{\mathrm{MTD}}\}$ 是攻防博弈的局中人集合，其中，N_{MTA} 代表移动目标攻击方，N_{MTD} 代表移动目标防御方。

2) $B=\{P_{\mathrm{MTA}},P_{\mathrm{MTD}}\}$ 是攻防博弈可行动作空间，其中，P_{MTA} 和 P_{MTD} 分别代表移动目标攻击者和防御者的移动策略集。

3) R 是移动目标攻防双方所竞争的公共资源，本文将网络中的攻击面视为公共资源。

4) η 是博弈信念集合，η_{MTA_i} 表示移动目标攻击方选择 MTA 策略 P_{MTA_i} $(0\leqslant i\leqslant m)$ 的概率，满足 $\sum_{i=1}^{m}\eta_{\mathrm{MTA}_i}=1$；$\eta_{\mathrm{MTD}_j}$ 表示移动目标防御方选择 MTD 策略 P_{MTD_j} $(0\leqslant j\leqslant l)$ 的概率，满足 $\sum_{j=1}^{l}\eta_{\mathrm{MTD}_j}=1$。

5) $U=\{U_{\mathrm{MTA}},U_{\mathrm{MTD}}\}$ 是移动目标攻防双方的收益函数集合，它由所有局中人对攻击面的控制时间 T_N 和策略实施所需成本 C_N 共同决定，分别为 $U_{\mathrm{MTD}}(C_{\mathrm{MTD}_i},T_{\mathrm{MTD}_j})$ 和 $U_{\mathrm{MTA}}(C_{\mathrm{MTA}_i},T_{\mathrm{MTA}_j})$，$1\leqslant i\leqslant m$，$1\leqslant j\leqslant l$。

6) T 是博弈的总时间，$T=T_{\mathrm{MTD}}+T_{\mathrm{MTA}}$。

以单阶段时间博弈为基础，构建多阶段 Markov 时间博弈模型。

1) 博弈模型定义

定义 2　Markov 时间博弈移动目标防御模型（MTG-MTD）。MTG-MTD 可以表示为十元组$(N,K,R,S,f,B,\eta,U,\beta,T)$，具体如下。

① $N=\{N_{\mathrm{MTA}},N_{\mathrm{MTD}}\}$ 是攻防博弈的局中人集合，其中，N_{MTA} 代表移动目标攻击方，N_{MTD} 代表移动目标防御方。

② K 是多阶段攻防博弈的阶段数，$G(K)$ 代表当前攻防博弈阶段，其中 $K=\{1,\cdots,n\}$，$n\in\mathbb{N}$。

③R 是移动目标攻防双方竞争的公共资源，本文将网络中的攻击面视为公共资源。

④ $S=\{S_1,S_2,\cdots,S_K\}$ 是不同网络攻防阶段安全状态集合。

⑤ f 表示状态迁移概率，$f_{ij}=f(S_j\mid S_i)$ 表示系统从状态 S_i 迁移至状态 S_j 的概率，攻防双方的对抗行为是影响安全状态转换的关键因素，由于攻防双方的可行策略集和网络系统运行环境可能发生改

变，因此状态转换具有一定随机性。

⑥ $B=\{P_{\mathrm{MTA}}^k,P_{\mathrm{MTD}}^k\}$ 是攻防博弈动作空间，其中，P_{MTA} 和 P_{MTD} 分别代表移动目标攻击者和防御者的策略集，$P_{\mathrm{MTA}}^k=\{P_{\mathrm{MTA}_i}^k \mid 1\leqslant k\leqslant K,1\leqslant i\leqslant m\}$，$P_{\mathrm{MTA}_i}^k$ 表示移动目标攻击方在第 k 个阶段的可选 MTA 策略；$P_{\mathrm{MTD}}^k=\{P_{\mathrm{MTD}_i}^k \mid 1\leqslant k\leqslant K,1\leqslant i\leqslant m\}$，$P_{\mathrm{MTD}_i}^k$ 表示移动目标防御方在第 k 个阶段的可选 MTD 策略。

⑦ η 是博弈信念集合，在第 k 阶段，$\eta_{\mathrm{MTA}_i}^k$ 表示移动目标攻击方选择MTA策略 $P_{\mathrm{MTA}_i}^k$ $(0\leqslant i\leqslant m)$ 的概率，满足 $\sum_{i=1}^{m}\eta_{\mathrm{MTA}_i}^k=1$；$\eta_{\mathrm{MTD}_j}^k$ 表示移动目标防御方选择 MTD 策略 $P_{\mathrm{MTD}_j}^k$ $(0\leqslant j\leqslant l)$ 的概率，满足 $\sum_{j=1}^{l}\eta_{\mathrm{MTD}_j}^k=1$。

⑧ $U=\{U_{\mathrm{MTA}}^k,U_{\mathrm{MTD}}^k\}$ 是移动目标攻防双方的收益函数集合，它由所有局中人对攻击面的控制时间 T_N 和策略实施所需成本 C_N 共同决定，$U_{\mathrm{MTD}}(C_{\mathrm{MTD}_i},T_{\mathrm{MTD}_j})$ 和 $U_{\mathrm{MTA}}(C_{\mathrm{MTA}_i},T_{\mathrm{MTA}_j})$，$1\leqslant i\leqslant m$，$1\leqslant j\leqslant l$。移动目标防御者的目的是最小化移动目标攻击者的收益。

⑨ β 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，$0<\beta\leqslant 1$。

⑩ T 是单阶段博弈所需的总时间。

2) 移动目标攻防收益量化

移动目标攻防收益量化是最优防御策略选取的基础，在文献[16]的研究基础上，本文从移动目标攻防双方对攻击面的控制出发，结合移动目标攻防策略特点，对移动目标攻防策略收益进行全面分析量化。

定义 3 防御成本（DC, defense cost）。DC 由移动目标防御者控制攻击面的时间成本 T_{CASC} 和变换攻击面的时间成本 T_{HASC} 两部分组成，$\mathrm{DC}=T_{\mathrm{CASC}}+T_{\mathrm{HASC}}$。

定义 4 攻击成本（AC, attack cost）。AC 指移动目标攻击者发现系统漏洞并采取 MTA 策略时所产生的时间成本。

定义 5 防御有效性（DE, defense effectiveness）。DE 是移动目标防御者实施 MTD 策略对攻击面的控制时间。

定义 6 攻击有效性（AE, attack effectiveness）。AE 是移动目标攻击者实施MTA策略对攻击面的控制时间。

定义 7 防御收益。防御收益指移动目标防御者控制攻击面获得的收益。

$$U_{\mathrm{MTD}}(C_{\mathrm{MTD}_i}^k,T_{\mathrm{MTD}_j}^k)=\mathrm{DE}+\mathrm{AC}-\mathrm{DC}$$

定义 8 攻击收益。攻击收益指移动目标攻击者控制攻击面获得的收益。

$$U_{\mathrm{MTA}}(C_{\mathrm{MTA}_i}^k,T_{\mathrm{MTA}_j}^k)=\mathrm{AE}+\mathrm{DC}-\mathrm{AC}$$

移动目标攻防收益矩阵 $\boldsymbol{M}$ 如下，$U_{\mathrm{MTA}}(C_{\mathrm{MTA}_i}^k,T_{\mathrm{MTA}_j}^k)$ 和 $U_{\mathrm{MTD}}(C_{\mathrm{MTD}_i}^k,T_{\mathrm{MTD}_j}^k)$ 分别表示策略组合 $(P_{\mathrm{MTA}_i}^k,P_{\mathrm{MTD}_i}^k)$ 下的攻击收益值和防御收益值，满足定义 7 和定义 8。

$$\boldsymbol{M}=\begin{bmatrix} P_{\mathrm{MTA}_{11}},P_{\mathrm{MTD}_{11}} & P_{\mathrm{MTA}_{12}},P_{\mathrm{MTD}_{12}} & \cdots & P_{\mathrm{MTA}_{1m}},P_{\mathrm{MTD}_{1m}} \\ P_{\mathrm{MTA}_{21}},P_{\mathrm{MTD}_{21}} & P_{\mathrm{MTA}_{22}},P_{\mathrm{MTD}_{22}} & \cdots & P_{\mathrm{MTA}_{2m}},P_{\mathrm{MTD}_{2m}} \\ \cdots & \cdots & \ddots & \cdots \\ P_{\mathrm{MTA}_{n1}},P_{\mathrm{MTD}_{n1}} & P_{\mathrm{MTA}_{n2}},P_{\mathrm{MTD}_{n2}} & \cdots & P_{\mathrm{MTA}_{nm}},P_{\mathrm{MTD}_{nm}} \end{bmatrix}$$

令 R 为目标准则函数，用于判断移动目标攻防双方策略选取的优劣。常用的准则函数[17]主要有折现期望回报准则函数和平均回报准则函数。在移动目标攻防对抗过程中，由于网络系统信息的价值与时间相关，因此采用折现期望回报准则函数作为博弈双方的目标函数，其中，$\beta\sum_{S}f(S,P_{\mathrm{MTA}},P_{\mathrm{MTD}},S')R_{S'}$ 表示攻防双方分别采取策略 P_{MTA} 和 P_{MTD} 时相较于初始阶段的折现收益值，S 为初始阶段状态，S' 为未来阶段状态，U_S 为初始阶段状态下的攻防收益值。

$$R_S(P_{\mathrm{MTA}}^k,P_{\mathrm{MTD}}^k)=U_S(P_{\mathrm{MTA}}^k,P_{\mathrm{MTD}}^k)+\beta\sum_{S'}f(S,P_{\mathrm{MTA}}^k,P_{\mathrm{MTD}}^k,S')R_{S'}$$

移动目标攻击方通过侦察网络攻击面，发现并利用系统资源脆弱性，进而导致系统性能开销增大或系统功能不可用。移动目标防御方通过选取 MTD 策略从而增大或转换攻击面，进而在保证网络功能正常安全运行的前提下提高系统的安全性。

由以上定义可知，经过有限次博弈后，系统在不同状态间进行迁移，可用攻防博弈树表示。在 TG-MTD 模型构建的基础上，第 4 节给出了模型的均衡策略分析求解和具体的最优防御选取算法。

4 博弈均衡求解与防御策略选取算法设计

根据第 2 节的分析，不同博弈阶段中攻防双方对攻击面的控制顺序动态变化。因此，本节首先提出时间博弈的子博弈精炼纳什均衡求解方法，然后

分析多阶段攻防博弈的求解过程。

4.1　博弈均衡分析

在时间博弈阶段 $G(K)$，移动目标攻防策略分别为 $P^k_{\mathrm{MTA}}=\{P^k_{\mathrm{MTA}_1},\cdots,P^k_{\mathrm{MTA}_n}\}$ 和 $P^k_{\mathrm{MTD}}=\{P^k_{\mathrm{MTD}_1},\cdots,P^k_{\mathrm{MTD}_n}\}$，若 $(P^k_{\mathrm{MTA}^*},P^k_{\mathrm{MTD}^*})$ 为第 k 阶段的时间稳定策略，则对于任意攻防策略 $P^k_{\mathrm{MTA}_j}$ 和 $P^k_{\mathrm{MTA}_i}$ 满足

$$\begin{cases}(U^k_{\mathrm{MTA}^*},U^k_{\mathrm{MTD}^*})\geqslant(U^k_{\mathrm{MTA}_j},U^k_{\mathrm{MTD}^*})\\(U^k_{\mathrm{MTA}^*},U^k_{\mathrm{MTD}^*})\geqslant(U^k_{\mathrm{MTA}^*},U^k_{\mathrm{MTD}_i})\end{cases}\tag{1}$$

$(P^k_{\mathrm{MTA}^*},P^k_{\mathrm{MTD}^*})$ 为某一阶段时间博弈 $G(K)$ 的子博弈精炼纳什均衡，该策略组合为 $G(K)$ 的一个纳什均衡，且在 $G(K)$ 的某段运行过程 j_λ，子博弈 $G(j_\lambda)$ 的限制策略组合 $(P^k_{\mathrm{MTA}^*},P^k_{\mathrm{MTD}^*})|j_\lambda$ 是 $G(j_\lambda)$ 的纳什均衡。

不同移动目标攻防策略的选取会影响每阶段博弈情况，根据 Markov 决策准则，局中人必有一个 Markov 最优响应策略[18]。因此，如果 $\{(U^k_{\mathrm{MTA}^*},U^k_{\mathrm{MTD}^*})\,|\,1\leqslant k\leqslant T\}$ 为 Markov 最优响应策略，那么 $(U^k_{\mathrm{MTA}^*},U^k_{\mathrm{MTD}^*})$ 使目标准则函数 $R_S(P^k_{\mathrm{MTA}},P^k_{\mathrm{MTD}})$ 对任意阶段 k 均满足式(2)所示条件。

$$(U^k_{\mathrm{MTA}^*},U^k_{\mathrm{MTD}^*})\in\arg\max R_S(P^k_{\mathrm{MTA}},P^k_{\mathrm{MTD}})=$$
$$\arg\max[(U^k_{\mathrm{MTA}^*},U^k_{\mathrm{MTD}^*})+\beta\sum_{S'}f(S,P^k_{\mathrm{MTA}},P^k_{\mathrm{MTD}},S')R_{S'}]\tag{2}$$

定理 1　多阶段 Markov 攻防时间博弈 MTG-MTD 存在混合策略下的纳什均衡。

证明　MTG-MTD 博弈由多个独立且相似的单阶段不完全信息动态博弈构成。一方面，由于每个独立的单阶段不完全信息动态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[19]。另一方面，由多阶段 Markov 时间博弈模型的定义，依据转移概率和收益函数可知，存在与 MTG-MTD 等价的有限 Markov 博弈，且收益函数为凸函数。依据有限 Markov 博弈的均衡策略存在性定理[20]，存在混合策略下的纳什均衡。证毕。

4.2　博弈均衡求解

4.2.1　单阶段时间博弈均衡求解

首先，给出单阶段时间博弈均衡的求解过程和步骤，参照完全信息动态博弈的相关理论知识，移动目标攻防双方对攻击面的控制权争夺具有先后顺序，先行动的一方的各种信息会被另一方完全掌握，因而后行动的一方可以根据对方的信息进行相应调整以最大化自身利益。

针对本文完全信息动态移动目标攻防场景，引入泽尔腾的子博弈精炼纳什均衡思想方法[21]，去除均衡中的不可置信威胁策略的纳什均衡，得出合理的预测结果。不失一般性，子博弈精炼纳什均衡的每个信息集上的均衡结果均为最优策略。

移动目标攻防双方在不同策略组合下的收益矩阵可以用图 3 的博弈树直观展示。假设博弈开始时刻由移动目标攻击者控制攻击面，随后移动目标防御者实施策略，争夺攻击面的控制权，单阶段博弈总时间为 T。

4.2.2　多阶段 Markov 时间博弈均衡求解

引入折现因子，将未来收益折算成基于初始阶段的折现收益，在此基础上，将博弈均衡策略的求解问题转化为非线性规划（NLP2, nonlinear programming second）最优值问题，求解多阶段均衡策略 B^* 及其收益 U^*。

对于 $K=\{1,\cdots,n\}$，$n\in\mathbb{N}$，$B=\{P^k_{\mathrm{MTA}},P^k_{\mathrm{MTD}}\}$，

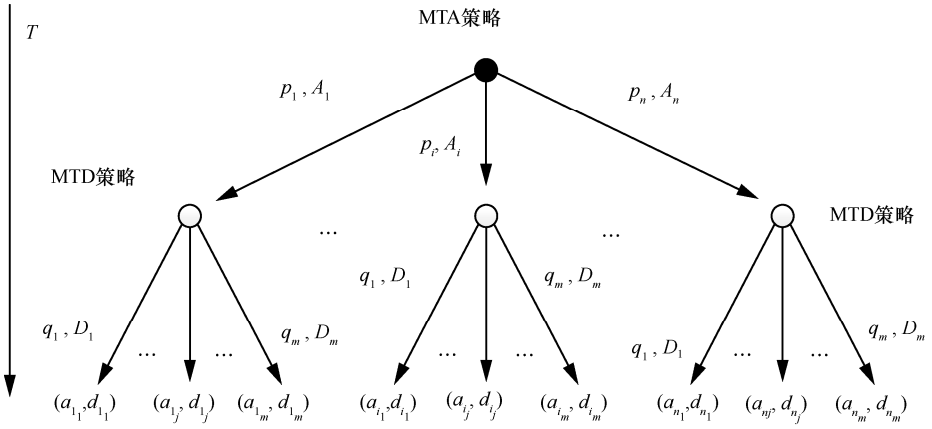

图 3　网络攻防时间博弈树

有目标函数为

$$\min \sum_{k\in K}\sum_{S_i\in S} R_i^k - U_i^k(B) - \beta\sum_{S'} f(S,B,S')R_{S'}$$

约束条件为

$$\forall k\in K, \forall S_i\in S, \forall B_i\in B$$

$$\text{s.t.}\quad R_i^k \geqslant U_i^k(B) + \beta\sum_{S'} f(S,B,S')R_{S'}$$

$$\forall k\in K, \forall S_i\in S, \sum_{B_i\in B} B^*(B) = 1$$

$$\forall k\in K, \forall S_i\in S, B_i\in B, B^*(B)\geqslant 0$$

求解上述目标函数可以得到最优解集合 $B^* = (P^k_{\text{MTA}^*}, P^k_{\text{MTD}^*})$。依据博弈理论，混合策略 $B^* = (P^k_{\text{MTA}^*}, P^k_{\text{MTD}^*})$ 是第 k 阶段攻防双方的最优选择，因此，$P^k_{\text{MTD}^*}$ 即为最优移动目标防御策略。

4.3 最优策略选取算法

基于移动目标攻防场景下多阶段 Markov 时间博弈模型及其子博弈精炼纳什均衡的研究，给出多阶段 Markov 时间博弈的最优主动防御策略选取算法。

算法 1 多阶段 Markov 时间博弈的最优防御策略选取算法

输入 多阶段 Markov 时间博弈模型 MTG- MTD

输出 多阶段最优移动目标防御策略

begin

初始化 MTG-MTD $=(N,K,S,f,B,\eta,U,\beta)$；

初始化状态转移概率 $f_{ij} = f(S_j \mid S_i)$；

构建攻防动作空间 $B = \{P^k_{\text{MTA}}, P^k_{\text{MTD}}\}$；

构建安全状态集合 $S = \{S_1, S_2, \cdots, S_K\}$；

构建信念集合 $\eta = \{\eta^k_{\text{MTA}_i}, \eta^k_{\text{MTD}_i}\}$，$\sum_{i=1}^{m}\eta^k_{\text{MTA}_i} = 1$，$\sum_{j=1}^{l}\eta^k_{\text{MTD}_j} = 1$；

for ($k = 1$; $k \leqslant n$; k++)

{

构建收益函数 $U_{\text{MTA}}(S^k, P^k_{\text{MTA}_i}, P^k_{\text{MTD}_j})$ = AE+DC − AC，$U_{\text{MTD}}(S^k, P^k_{\text{MTA}_i}, P^k_{\text{MTD}_j}) = \text{DE} + \text{AC} - \text{DC}$；

构建平均收益函数 $\bar{U}^k_{\text{MTA}} = \sum_{j=1}^{n}\eta^k_{\text{MTA}_j}U^k_{\text{MTA}}$，$\bar{U}^k_{\text{MTD}} = \sum_{i=1}^{n}\eta^k_{\text{MTD}_i}U^k_{\text{MTD}}$；

}

构建折现收益函数 $\beta\sum_{S'} f(S, P^k_{\text{MTA}}, P^k_{\text{MTD}}, S')R_{S'}$；

构建目标函数 $\min \sum_{k\in K}\sum_{S_i\in S} R_i^k - U_i^k(B) - \beta\sum_{S'} f(S, B, S')R_{S'}$；

return $P^k_{\text{MTD}^*}$；

end

算法的时间复杂度为 $O(k(m+n)^2)$，空间复杂度为 $O(knm)$，表 3 展示了本文提出的最优策略选取方法与其他最优策略选取方法的比较结果。在移动目标攻防对抗中，Manadhata 等[22]仅讨论了单阶段博弈。Clark 等[23]虽然将博弈模型扩展到多阶段，但仍不能揭示移动目标攻防对抗的多状态和多阶段过程。Lei 等[8]结合 Markov 决策过程理论和动态博弈描述了多状态和多阶段特征。上述研究成果均采用历史数据与专家经验量化收益计算，本文针对 MTD 攻防过程的动态连续特性，将时间因素加入收益度量能够提高收益计算的准确性。与上述方法相比，MTG-MTD 是基于 Markov 时间博弈建立的，完美地展示了移动目标攻防过程的对立性、动态性及自适应性的特征。在最优策略选取方面，本文分析了时间因素对攻防成本和收益的影响，并将最优策略选取问题转化为非线性规划问题求解，在降低复杂度的同时大大增加了不同的应用场景下的通用性。

5 应用实例分析

5.1 应用实例

本节通过应用实例验证 MTG-MTD 最优防御策略选取算法的有效性，利用软件定义网络（SDN, software defined network）的部分节点拓扑搭建了实验网络环境，系统结构如图 4 所示。其中，LDAP 服务器、FTP 服务器、Linux 数据库等控制服务器作为移动目标防御策略的应用目标，同时移动目标攻击者可以通过网络等途径访问控制服务器，它们的连通性通过表 4 中的访问控制策略来确定，应用服务器作为控制服务器的应用提供者。移动目标攻击者具有对应用服务器的用户级访问权限，其目标是窃取存储在 Linux 数据库服务器中的敏感信息。

移动目标攻击者的可能的攻击路径如下。

路径 1：应用服务器→LDAP 服务器→Linux 数据库。

路径 2：应用服务器→LDAP 服务器→FTP 服务器→Linux 数据库。

表 3　不同策略选取方法对比分析

方法	收益量化	动态性	博弈类型	均衡求解	最优选取算法
文献[22]方法	历史数据	单阶段	静态博弈	简单	未给出
文献[23]方法	历史数据	多阶段	动态博弈	详细	给出
文献[8]方法	历史数据	多阶段	Markov 矩阵博弈	详细	给出
本文方法	历史数据+时间因素	多阶段	Markov 时间博弈	详细	给出

表 4　访问控制策略

端点名称	移动目标攻击者	应用服务器	LDAP 服务器	FTP 服务器	Linux 数据库
移动目标攻击者	local	IIS	—	—	—
应用服务器	—	local	all	all	Squid LICQ
LDAP 服务器	—	—	local	all	Squid LICQ
FTP 服务器	—	IIS	all	local	—
Linux 数据库	—	—	all	all	local

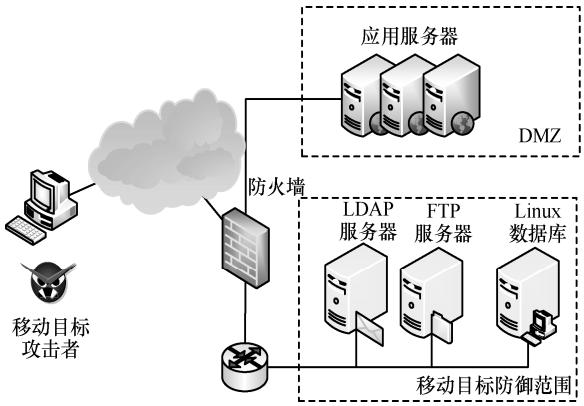

图 4　实验系统结构示意

1) 初始化参数

令 $S=\{S_1, S_2, S_3, S_4\}$表示网络阶段状态。其中，S_1 是移动目标攻击者利用应用服务器的漏洞，并获得其 root 权限的阶段状态；S_2 和 S_3 分别是移动目标攻击者利用 LDAP 服务器和 FTP 服务器的漏洞获得 Linux 数据库访问权限的阶段状态；S_4 是攻击者通过利用 Linux 数据库的漏洞获得 root 权限的阶段状态。本实验中 MTG-MTD 的折扣率为$\beta=0.7$。

2) 构建策略空间，状态转移概率和收益矩阵

表5显示了每个网络状态下的移动目标攻防策略。$P_{\mathrm{MTA}}=\{P_{\mathrm{MTA}_1}, P_{\mathrm{MTA}_2}, P_{\mathrm{MTA}_3}, P_{\mathrm{MTA}_4}, P_{\mathrm{MTA}_5}, P_{\mathrm{MTA}_6}, P_{\mathrm{MTA}_7}, P_{\mathrm{MTA}_8}\}$表示移动目标攻击者控制攻击面，相关 MTA 策略集合如表 1 所示。$P_{\mathrm{MTD}}=\{P_{\mathrm{MTD}_1}, P_{\mathrm{MTD}_2}, P_{\mathrm{MTD}_3}\}$表示移动目标防御者控制着攻击面，其中，$P_{\mathrm{MTD}_1}=$ {IP(C类),Port(64512),Timing(fixed)}表示 MTD 在固定周期中变换 IP 地址和端口号，括号中的内容表示相应变换元素的取值范围，IP（C 类）表示 IP 的变换取值为 C 类 IP 地址空间，Port(64512)表示端口变换取值为 64512，Timing(fixed)和 Timing(random)分别表示 MTD 固定变换时机和随机变换时机，$P_{\mathrm{MTD}_2}=$ {IP(C类),Port(64512),Timing(random)} 表示 MTD 在随机周期中变换 IP 地址和端口号，$P_{\mathrm{MTD}_3}=$ {Forwarding Path,Timing(fixed)} 表示 MTD 在固定周期内变换转发路径，括号中的内容表示相应变换元素的取值范围。同时，网络状态转移概率具体如表 6 所示。依据 3.2 节移动目标攻防收益的计算方法，表 7 给出了移动目标攻防收益矩阵。

表 5　不同网络状态下的移动目标攻防策略

网络状态	攻防策略
S_1	$P^1_{\mathrm{MTA}}=\{P_{\mathrm{MTA}_1}, P_{\mathrm{MTA}_2}, P_{\mathrm{MTA}_3}\}$ $P^1_{\mathrm{MTD}}=\{\mathrm{IDS}, P_{\mathrm{MTD}_1}, P_{\mathrm{MTD}_3}\}$
S_2	$P^2_{\mathrm{MTA}}=\{P_{\mathrm{MTA}_2}, P_{\mathrm{MTA}_4}, P_{\mathrm{MTA}_5}\}$ $P^2_{\mathrm{MTD}}=\{\mathrm{patch\ upgrade},$ $P_{\mathrm{MTD}_3}, P_{\mathrm{MTD}_1}+P_{\mathrm{MTD}_3}\}$
S_3	$P^3_{\mathrm{MTA}}=\{P_{\mathrm{MTA}_1}, P_{\mathrm{MTA}_6}, P_{\mathrm{MTA}_7}\}$ $P^3_{\mathrm{MTD}}=\{P_{\mathrm{MTD}_1}, P_{\mathrm{MTD}_5}, P_{\mathrm{MTD}_2}+P_{\mathrm{MTD}_3}\}$
S_4	$P^4_{\mathrm{MTA}}=\{P_{\mathrm{MTA}_3}, P_{\mathrm{MTA}_7}, P_{\mathrm{MTA}_8}\}$ $P^4_{\mathrm{MTD}}=\{P_{\mathrm{MTD}_2}, \mathrm{IDS, close\ service}\}$

3) 选取 MTG-MTD 模型的最优策略

在选取最优策略之前，将最优策略选取问题等价转化为非线性规划问题。在此基础上，利用所提

表 6 网络系统状态转移概率

网络状态	转移概率
S_1	$f_{12}=0.25(P_{\mathrm{MTA}_1},P_{\mathrm{MTD}_1})$，$f_{13}=0.36(P_{\mathrm{MTA}_2},P_{\mathrm{MTD}_1})$，$f_{14}=0.85(P_{\mathrm{MTA}_3},P_{\mathrm{MTD}_1})$
S_2	$f_{21}=0.9(P_{\mathrm{MTA}_1},P_{\mathrm{MTD}_2})$，$f_{23}=0.16(P_{\mathrm{MTA}_2},P_{\mathrm{MTD}_2})$，$f_{24}=0.38(P_{\mathrm{MTA}_3},P_{\mathrm{MTD}_2})$
S_3	$f_{32}=0.9(P_{\mathrm{MTA}_1},P_{\mathrm{MTD}_3})$
S_4	$f_{43}=0.9(P_{\mathrm{MTA}_1},P_{\mathrm{MTD}_4})$，$f_{42}=0.8(P_{\mathrm{MTA}_2},P_{\mathrm{MTD}_4})$

表 7 移动目标攻防策略收益矩阵

网络状态	MTA 收益	MTD 收益
S_1	$\begin{bmatrix}21&42&22\\19&15&18\\0&0&0\end{bmatrix}$	$\begin{bmatrix}-21&-44&-22\\-19&-15&-18\\-36&-11&-6\end{bmatrix}$
S_2	$\begin{bmatrix}53&62&34\\16&27&21\\15&17&9\end{bmatrix}$	$\begin{bmatrix}-53&-62&-34\\-16&-27&-21\\-16&-17&-9\end{bmatrix}$
S_3	$\begin{bmatrix}36&21&18\\23&26&21\\12&7&13\end{bmatrix}$	$\begin{bmatrix}-36&-21&-18\\-23&-26&-21\\-12&-7&-13\end{bmatrix}$
S_4	$\begin{bmatrix}35&18&0\\27&13&0\\6&20&0\end{bmatrix}$	$\begin{bmatrix}-35&-18&0\\-27&-13&0\\-6&-20&0\end{bmatrix}$

算法及交互式的线性和通用优化求解器（LINGO, linear interactive and general optimizer）求解最优策略。表 8 给出了攻防双方及其相应收益的最优策略。

约束条件为

$$R_1-(21P^1_{\mathrm{MTD}_1}+42P^1_{\mathrm{MTD}_2}+22P^1_{\mathrm{MTD}_3})-0.7(0.25R_2)\geqslant 0$$

$$R_1-(19P^1_{\mathrm{MTD}_1}+15P^1_{\mathrm{MTD}_2}+18P^1_{\mathrm{MTD}_3})-0.7(0.35R_3)\geqslant 0$$

$$R_2-(53P^2_{\mathrm{MTD}_1}+62P^2_{\mathrm{MTD}_2}+34P^2_{\mathrm{MTD}_3})-0.7(0.9R_1)\geqslant 0$$

$$R_2-(16P^2_{\mathrm{MTD}_1}+27P^2_{\mathrm{MTD}_2}+21P^2_{\mathrm{MTD}_3})-0.7(0.16R_3)\geqslant 0$$

$$R_2-(15P^2_{\mathrm{MTD}_1}+17P^2_{\mathrm{MTD}_2}+19P^2_{\mathrm{MTD}_3})-0.7(0.38R_4)\geqslant 0$$

$$R_3-(36P^3_{\mathrm{MTD}_1}+21P^3_{\mathrm{MTD}_2}+18P^3_{\mathrm{MTD}_3})-0.7(0.63R_2)\geqslant 0$$

$$R_4-(35P^4_{\mathrm{MTD}_1}+18P^4_{\mathrm{MTD}_2}+27P^4_{\mathrm{MTD}_3})-0.7(0.9R_3)\geqslant 0$$

$$R_4-(13P^4_{\mathrm{MTD}_1}+6P^4_{\mathrm{MTD}_2}+20P^4_{\mathrm{MTD}_3})-0.7(0.8R_2)\geqslant 0$$

$$\sum_{i=1}^{3}P^k_{\mathrm{MTA}_i}=1,\sum_{j=1}^{3}P^k_{\mathrm{MTD}_j}=1,P^k_{\mathrm{MTA}_i}\geqslant 0,P^k_{\mathrm{MTD}_j}\geqslant 0,i,j=\{1,2,3\}$$

目标函数为

$$\min\{[R_1-(21P^1_{\mathrm{MTA}_1}P^1_{\mathrm{MTD}_1}+42P^1_{\mathrm{MTA}_1}P^1_{\mathrm{MTD}_2}+22P^1_{\mathrm{MTA}_1}P^1_{\mathrm{MTD}_3}+19P^1_{\mathrm{MTA}_2}P^1_{\mathrm{MTD}_1}+15P^1_{\mathrm{MTA}_2}P^1_{\mathrm{MTD}_2}+18P^1_{\mathrm{MTA}_2}P^1_{\mathrm{MTD}_3})-0.7(0.25R_2+0.36R_3+0.85R_4)]+[R_2-(53P^2_{\mathrm{MTA}_1}P^2_{\mathrm{MTD}_1}+62P^2_{\mathrm{MTA}_1}P^2_{\mathrm{MTD}_2}+34P^2_{\mathrm{MTA}_1}P^2_{\mathrm{MTD}_3}+16P^2_{\mathrm{MTA}_2}P^2_{\mathrm{MTD}_1}+27P^2_{\mathrm{MTA}_2}P^2_{\mathrm{MTD}_2}+21P^2_{\mathrm{MTA}_2}P^2_{\mathrm{MTD}_3}+15P^2_{\mathrm{MTA}_3}P^2_{\mathrm{MTD}_1}+17P^2_{\mathrm{MTA}_3}P^2_{\mathrm{MTD}_2}+19P^2_{\mathrm{MTA}_3}P^2_{\mathrm{MTD}_3})-0.7(0.9R_1+0.16R_3+0.38R_4)]+[R_3-(36P^3_{\mathrm{MTA}_1}P^3_{\mathrm{MTD}_1}+21P^3_{\mathrm{MTA}_1}P^3_{\mathrm{MTD}_2}+18P^3_{\mathrm{MTA}_1}P^3_{\mathrm{MTD}_3}+23P^3_{\mathrm{MTA}_2}P^3_{\mathrm{MTD}_1}+26P^3_{\mathrm{MTA}_2}P^3_{\mathrm{MTD}_2}+21P^3_{\mathrm{MTA}_2}P^3_{\mathrm{MTD}_3}+12P^3_{\mathrm{MTA}_3}P^3_{\mathrm{MTD}_1}+7P^3_{\mathrm{MTA}_3}P^3_{\mathrm{MTD}_2}+13P^3_{\mathrm{MTA}_3}P^3_{\mathrm{MTD}_3})-0.63R_2]+[R_4-(35P^4_{\mathrm{MTA}_1}P^4_{\mathrm{MTD}_1}+18P^4_{\mathrm{MTA}_1}P^4_{\mathrm{MTD}_2}+27P^4_{\mathrm{MTA}_2}P^4_{\mathrm{MTD}_1}+13P^4_{\mathrm{MTA}_2}P^4_{\mathrm{MTD}_2}+6P^4_{\mathrm{MTA}_3}P^4_{\mathrm{MTD}_1}+20P^4_{\mathrm{MTA}_3}P^4_{\mathrm{MTD}_2})-0.7(0.9R_3+0.8R_2)]\}$$

表 8 移动目标攻防策略和收益

网络阶段状态	MTA 策略	MTD 策略	MTA 收益	MTD 收益
S_1	[0.31,0.29,0.4]	[0.07,0.48,0.45]	187.9	–187.9
S_2	[0.32, 0.32, 0.36]	[0.17,0.39,0.44]	107.3	–107.3
S_3	[0.35, 0.38, 0.27]	[0.12,0.36,0.52]	473.5	–473.5
S_4	[0.12, 0.42, 0.46]	[0.33, 0.09, 0.58]	601.3	–601.3

5.2 结果分析

通过对移动目标防御模型均衡和收益分析，可以得出以下移动目标攻防过程的一般规律。

1) 由于防御实施效果的针对性，应该尽可能实施成本低且防御效果佳的 MTD 策略，针对特定的移动目标攻击，应实施适当的移动目标防御。例如在状态 S_1，攻击者的主要攻击手段是利用自身的动态变换规避常规的入侵检测系统，因而 IDS 对于上述攻击无效；相反地，实施移动目标防御可以有效抵御此类攻击。

2) 由于攻击的持续性，要尽可能避免攻击者与目标系统建立通信控制连接，否则很难采取有效防御策略。例如在状态 S_4，当攻击者已经入侵目标系统，并且进行后续攻击开发时，IDS 等传统防御手段对于攻击防御无效，并且移动目标防御的效果也不理想，此时最佳策略为关闭服务。

由于单阶段博弈过程由时间博弈所刻画，使博弈场景更贴近有实际网络攻防过程，相较于矩阵博弈，本文所采用的时间博弈可以更好地刻画博弈动态性，同时利用 Markov 决策过程刻画多阶段性，从而帮助网络安全管理人员更好地决策。

6 结束语

本文基于多阶段 Markov 时间博弈模型研究了移动目标攻防策略选取问题，主要工作如下。在分析移动目标攻防过程的基础上，构建了 Markov 时间博弈模型，具备分析多阶段–多状态攻防行为的能力；基于折扣总收益设计了移动目标防御博弈的目标准则函数，实现了对多阶段攻防博弈的量化分析；提出了基于非线性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。研究成果对于在多阶段移动目标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理论模型支持。

当前网络攻防策略集合均与时间无关，需要将时间作为策略因素考虑，因此对于攻防策略行动问题时机的研究是下一步开展的主要研究方向。

参考文献：

[1] MITROPOULOS D , LOURIDAS P , POLYCHRONAKIS M, et al. Defending against web application attacks: approaches, challenges and implications[J]. IEEE Transactions on Dependable and Secure Computing, 2017:1.

[2] ZHENG J, NAMIN A S. A survey on the moving target defense strategies: an architectural perspective[J]. Journal of Computer Science and Technology, 2019, 34(1):207-233.

[3] CAI G L , WANG B S , XING Q Q . Game theoretic analysis for the mechanism of moving target defense[J]. Frontiers of Information Technology & Electronic Engineering, 2017, 18(12):2017-2034.

[4] 姜伟, 方滨兴, 田志宏. 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, 2013,32(4):818-827.
JIANG W, FANG B X, TIAN Z H. Defense strategies selection based on attack-defense game model[J]. Chinese Journal of Computers, 2013, 47(12):818-827.

[5] 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 W Q, WANG H, LIU J H. 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.

[6] MANADHATA P K. Game theoretic approaches to attack surface shifting[M]. New York: Springer, 2013: 1-13.

[7] VADLAMUDI S G, SENGUPTA S, TAGUINOD M, et al. Moving target defense for web applications using Bayesian Stackelberg games[C]//The 2016 International Conference on Autonomous Agents & Multiagent Systems. International Foundation for Autonomous Agents and Multiagent Systems, 2016: 1377-1378.

[8] 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.

[9] MALEKI H, VALIZADEH M H, KOCH W, et al. Markov modeling of moving target defense games[J]. Journal of Cryptology, 2016: 47-83.

[10] JAJODIA S, GHOSH A K, SWARUP V, et al. Moving target defense: creating asymmetric uncertainty for cyber threats[J]. Springer Ebooks, 2011, 54.

[11] LEI C, ZHANG H Q, WANG L M, et al. Incomplete information Markov game theoretic approach to strategy generation for moving target defense[J]. 2018, 116:184-199.

[12] ZHENG J J, NAMIN A S. A survey on the moving target defense strategies: an architectural perspective[J]. Journal of Computer Science and Technology, 2019, 34(1): 207-233.

[13] 谭晶磊,张红旗,雷程,等.面向 SDN 的移动目标防御技术研究进展[J]. 网络与信息安全学报,2018,4(7):1-12.
TAN J L, ZHANG H Q, LEI C, et al. Research progress on moving target defense for SDN[J]. Chinese Journal of Network and Information Security, 2018, 4(7): 1-12.

[14] DIJK M V, ARI JUELS, ALINA OPREA, et al. FlipIt: the game of “stealthy takeover”[J]. Journal of Cryptology, 2013, 26(4):655-713.

[15] ZHENG J, SIAMI NAMIN A. A Markov decision process to determine optimal policies in moving target[C]//The 2018 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2018: 2321-2323.

[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.

[17] LEI C, MA D H, ZHANG H Q. Optimal strategy selection for moving target defense based on Markov game[J]. IEEE Access, 2017, PP(99):1.

[18] 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.

[19] CHEN M, SAAD W, YIN C. Virtual reality over wireless networks: quality-of-service model and learning-based resource management[J]. IEEE Transactions on Communications, 2018, 66(11):5621-5635.

[20] NILIM A, GHAOUI L E. Robust control of Markov decision processes with uncertain transition matrices[J]. Operations Research, 2016, 53(5): 780-798.

[21] SULEIMAN R. On gamesmen and fair men: explaining fairness in non-cooperative bargaining games[J]. Royal Society Open Science, 2018, 5(2):171709.

[22] MANADHATA P K. Game theoretic approaches to attack surface shifting[M]. New York: Springer, 2013: 1-13.

[23] CLARK A, SUN K, BUSHNELL L, et al. A game-theoretic approach to IP address randomization in decoy-based cyber defense[C]// International Conference on Decision and Game Theory for Security. Springer, 2015: 3-21.

[作者简介]

谭晶磊（1994– ），男，山东章丘人，信息工程大学博士生，主要研究方向为网络信息安全、移动目标防御、攻防博弈对抗等。

张恒巍（1978– ），男，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。

张红旗（1962– ），男，河北遵化人，博士，信息工程大学教授、博士生导师，主要研究方向为网络安全、移动目标防御、等级保护和信息安全管理等。

金辉（1988– ），男，北京人，信息工程大学硕士生，主要研究方向为网络信息安全等。

雷程（1989– ），男，北京人，信息工程大学博士生，主要研究方向为网络信息安全、移动目标防御、数据安全交换和网络流指纹等。