基于 Stackelberg-Markov 非对等三方博弈模型的移动目标防御技术

陈子涵 程 光

(东南大学网络空间安全学院 南京 211189)

(教育部计算机网络和信息集成重点实验室(东南大学) 南京 211189)

摘 要 易攻难守是当前网络安全面临的核心问题之一.移动目标防御技术对被保护系统的攻击面实施动态持续性变换,从而降低攻击者成功的概率.随着移动目标防御领域的研究进展,如何优化动态防御策略,实现成本和收益均衡的智能防御已经成为关键的研究点.移动目标防御的现有优化模型往往存在对目的性较强的实际攻守场景描述不当、无法预测参与者后续策略、缺乏对系统用户的考虑等问题.针对这些问题,本文创新性的将用户作为移动目标防御博弈中的第三方参与者,结合 Stackelberg 博弈和 Markov 模型来构建非对等三方博弈,以确定移动目标防御的最优策略.数学分析和仿真实验表明,本文提出的模型能够兼顾防御者和用户的成本和收益,避免过度的防御和不适宜的防御,有效的实现防御策略智能决策.

关键词 移动目标防御;Stackelberg 博弈;Markov 模型;非对等三方博弈;智能决策;效用平衡

中图法分类号 TP393 **DOI 号** 10.11897/SP.J.1016.2020.00512

Moving Target Defense Technology using Stackelberg Markov Asymmetrical Trilateral Game Model

CHEN Zi-Han CHENG Guang

(*School of Cyber Science and Engineering, Southeast University, Nanjing* 211189)

(*Key Laboratory of Computer Network and Information Integration of Ministry of Education (Southeast University), Nanjing* 211189)

Abstract With the continuous development of Internet technology and the gradual expansion of the Internet scale, an increasing number of network attacks and security incidents bring more severe challenges to cybersecurity. The core problem of the situation is the imbalance between the offensive and defensive sides due to information asymmetry. Moving Target Defense (MTD) technology is a kind of technology that hopes to confuse attackers by implementing continuous and dynamic changes, and implement these changes on the attack surface of the protected system, so as to increase the attack costs and reduce the attack success rate. With the development of the research in this field, how to optimize the dynamic defense strategy and realize the balance of costs and benefits in intelligent defense has become the focus of the existing research on MTD technology. However, existing decision-making schemes or optimization models often have problems such as improper description of the actual attack and defense scenarios with strong purposes, insufficient prediction of the state of the attack and defense process, inability to deal with unknown types of attacks, lack of quantitative evaluation of utility, and lack of consideration for system users. In

收稿日期:2018-10-31;在线出版日期:2019-05-28.本课题得到国家重点研发计划项目课题(2018YFB1800602, 2017YFB1800602)、教育部-中国移动科研基金(MCM20180506)、赛尔网络下一代互联网技术创新项目(NGIICS20190101,NGII20170406)资助.**陈子涵**,博士研究生,主要研究方向为网络空间安全、网络测量与行为分析、MTD 博弈、信任体系架构.E-mail:zhchen@njnet.edu.cn.**程 光**(通信作者),博士,教授,博士生导师,中国计算机学会(CCF)会员,主要研究领域为 SDN 网络测量与管理、网络流量测量与大数据分析、僵尸网络、APT 攻击检测.E-mail:gcheng@njnet.edu.cn.

view of the above problems, we innovatively take users as the third participant in MTD game, and put forward a MTD technology based on the asymmetrical trilateral game model with the combination of Stackelberg game and Markov model. In this model, we first define the structure of the game model: asymmetrical trilateral game. The offensive and defensive sides participate in the game as equal players, while the user is at a disadvantage as a third player in the game and has an asymmetrical relationship with the other two participants. In addition, we also assume that the user, as an independent participant, does not interact with the attacker and the defender, that is, the user does not aware of the existence of the attack and defense game and does not care about the result of attack and defense, but only focuses on the functions and performances of the current system. Then, under the assumption of the trilateral game model, we adopt the traditional Stackelberg game model of sequential offense and defense, and put forward the trilateral sequential game model of "attacker first, defender later, and user stochastically". Specifically, we summarize the game in each timestep into a Markov process unit, so that the whole series of the trilateral game process is built into a Markov model in the timeline. According to this, we define the three participants' respective utility and the relationship between them, in which the cost and benefit of the defender and the user are both taken into account. Through the first-order game prediction and the k-order game prediction, the optimal strategy of MTD under the current time step is determined. Compared with the existing game between offensive and defensive sides, this model optimizes the user experience of using the system, improves the adaptability of the model, makes the optimization convergence process more consistent with the actual scenes and can be applied to the complex attack scenes, which has a better effect of process state prediction. Mathematical analysis and simulation experiments demonstrate that the model can give consideration to both the defender and the user by balancing their utilities, avoid excessive defense and inappropriate defense strategies, so as to effectively realize the intelligent defense strategy decision-making.

Keywords　moving target defense; Stackelberg game; Markov model; asymmetrical trilateral game; intelligent decision-making; utility balance

1　引　言

随着互联网技术的蓬勃发展，通讯信息网络基础设施深远地影响着人们的生产和生活行为，成为了与电力网络、交通网络、金融网络等并列的具备国家战略高度的关键基础设施. 网络空间已经成为陆海天空之后的第五疆域，与国家安全、社会稳定和民众利益息息相关. 近年来，层出不穷的网络攻击手段和网络安全事件使得网络安全面临着严峻挑战[1]. 因此，构建网络攻击武器库和建设网络安全防御体系成为各国竞相研究的重点. 然而，攻击者与防御者之间博弈的整体态势始终存在"易攻难守"的不对等关系. 攻击者具备时间优势对目标反复做攻击测试和渗透测试，只需要找到一个有效的攻击点就可以实施攻击并持续扩大战果，甚至扩展至其他相似的被攻击环境中；而防御方则需要对所有可能的攻击点和攻击方式加以防护.

为改变网络空间这种"易攻难守"的博弈态势，美国政府开始发展"改变游戏规则"的革命性技术，以实现积极主动的网络防御. 移动目标防御[2]（Moving Target Defense，MTD）是美国针对防御者当前劣势所提出的"改变游戏规则"的网络安全发展方向之一，期望通过实施持续、动态的变化迷惑攻击者，增加其攻击成本，降低其攻击成功率. 需要注意的是，MTD 是一种防御机制的设计指导思想，而非具体机制. 这一思想指导了诸多具体防御机制的构建，衍生出如 IP address mutation[3]、MT6D[4]、ChameleonSoft[5]、MAS[6] 等.

传统的 MTD 研究主要着眼于防御方如何在特定的博弈模型假设下寻求最优防御策略的决策. 然而，无论是顺序博弈模型[7]还是随机博弈模型[8]都具有各自明显的局限性. 顺序博弈模型主要考虑攻守决策和攻守行为具备确定的逻辑关系. 其需要大

量的先验知识,且对攻守过程状态的预测不足,无法应对未知类型的攻击.为防御未知类型的攻击,人们利用随机博弈模型,假设攻守方进行随机博弈,并常常基于 Markov 过程描述攻守关系.然而,随机博弈模型难以描述实际的系统攻守状况,缺乏实用价值.更重要的是,这两类博弈模型仅考虑攻守双方之间的博弈关系,没有考虑到系统中用户的行为,缺乏量化评估该防御策略对于用户的成本与收益的分析,难以应用于实际的复杂攻击场景.

针对以上问题,本文提出了一种基于 Stackelberg 博弈和 Markov 模型的非对等三方博弈模型的移动目标防御技术(Moving Target Defense Technology using Stackelberg Markov Asymmetrical Trilateral Game Model,SMATG-MTD).我们首先定义了博弈模型的结构:非对等三方博弈.攻击者和防御者作为对等双方参与博弈,而用户在博弈关系中作为第三方处于劣势,与其他两方居于不对等的关系.并且我们额外假设用户作为独立参与者,不和攻击者和防御者产生信息交互,即用户并不知道攻防博弈的存在,也不关心攻防的结果,只关注系统的功能和性能.然后,在三方博弈模型的假设下,我们采用了传统顺序攻防中的 Stackelberg 博弈模型,提出"攻击者先、防御者后、用户可先可后"的三方顺序博弈模型.具体来说,我们将每一个时间步(timestep)的一次博弈归纳为一次马尔科夫过程的行为,从而在时间轴上将整个一系列的三方博弈过程构建为一个马尔科夫模型.

与现有的攻守双方的博弈对比,本文所提出方法的贡献如下:(1)创新性的将系统用户作为以第三方参与者的身份加入到攻防博弈过程中,额外优化了用户的系统使用体验,更加符合实际的攻守场景,提高了模型适应度;(2)在攻击者、防御者、用户的三方博弈模型中,应用了 Stackelberg 博弈模型,使优化收敛过程更符合实际的攻守顺序博弈场景;(3)参与三方每个时间步的一次博弈行为都被建模为时间轴上一次 Markov 过程行为,体现了博弈的随机性,提高了三方博弈模型的过程状态预测效果.

本文第 1 节介绍移动目标防御的背景知识和研究意义;第 2 节阐述移动目标防御攻防双方博弈的相关工作,并介绍如何实现防御策略的有效性评估;第 3 节提出 Stackelberg-Markov 非对等三方博弈模型,引入用户作为攻守双方博弈的第三方,并且将用户定义为与攻守双方不对等的特殊参与者.在该非对等三方博弈模型中,应用了 Stackelberg 博弈和 Markov 博弈理论;第 4 节在三方博弈模型假设下,研究参与者的成本和收益计算以及互相之间的作用与影响.最终提出防御策略智能决策方案,兼顾防御者和用户两者的成本与收益;第 5 节通过数学分析与实验评估的方式来验证该模型和方案的有效性;最后一节总结全文工作,并展望将来的研究.

2 相关工作

2.1 移动目标防御技术

移动目标防御是由美国针对当前"易守难攻"的态势所提出的一个"改变游戏规则"的网络安全发展方向,期望通过实施持续、动态的变化迷惑攻击者,以增加其攻击的成本和复杂度.值得注意的是,移动目标防御不是某一种具体的防御方法,而是一种设计指导思想.这一思想可应用到被保护系统的某一属性上,衍生出多种具体的防御机制[1].

现在已经存在大量的 MTD 相关的防御机制和研究成果,当前主流的研究方向主要集中于 MTD 的攻防博弈模型的研究[7-10]和将 MTD 具体的防御机制应用于实际场景中的研究[11-12].

对于移动目标,当前也没有一个权威的统一定义,仅在美国白宫国防安全委员会的进展报告中提到移动目标是可在多个维度上移动以降低攻击者优势并增加弹性的系统.对于移动目标防御,当前也不存在统一的定义,其目标是通过持续变换系统呈现在攻击者面前的攻击面,从而有效增加攻击者想要探测目标脆弱性的代价.

移动目标防御技术作为新兴的一套防御技术,随着防御机制和博弈模型的研究与发展,将逐步系统化地应用于电网、数据存储网络、内部局域网等各种网络环境,以应对各式各样的攻击,改变"易守难攻"现状.

2.2 Stackelberg 攻守双方博弈

Stackelberg 模型是由德国经济学家斯塔克尔伯格(H. Von Stackelberg)于 20 世纪 30 年代提出的一套经济学博弈模型.它是一个产量领导模型,由领导性厂商先决定产量,然后其他厂商观察到该产量从而制定自己的产量的过程.其中领导性厂商充分地了解自己决策之后其他厂商会如何行动.

该模型在近年被很好的引入到攻守双方博弈之中.在现有的研究中,Stackelberg 攻守博弈中的领导者(leader)对应于防御者,而跟随者(follower)对应于攻击者[7,13-14].防御者优先决策,预先选择一套

系统环境从而攻击者可以了解到防御者的系统环境及其漏洞从而做出一系列的攻击决策.

钱震等人[14]将 Stackelberg 攻守双方博弈应用在 Web 应用场景中,防御者作为领导者首先选择一套系统环境,然后攻击者和防御者都知道该环境的漏洞、对应的攻击方式和修复方案.防御者根据可能出现的攻击方案和对应的修复方案进行博弈矩阵的效用计算,从而发现最大可能性出现的攻击从而做出最合理的修复方案.该方案需要足够的先验知识和预定义好的漏洞风险与攻守效用,在实际场景中很难实现,并且是一种被动防御的方案.

Vadlamudi 等人[7]同样将 Stackelberg 攻守双方博弈应用在 Web 应用场景中.防御者应用了 MTD 技术构建了多套环境并从中进行切换,而攻击者的策略相对有限(针对已侦查到的环境),但是可以进行攻击策略的混合.两者进行一次随机博弈过程,防御者将当前系统切换至某一环境,而攻击者采取一种攻击策略(包括混合策略).攻击策略与环境相匹配则成功攻击,反之则成功防御.该方案同样需要足够的先验知识,且无法应对未知类型攻击.

Feng 等人[11]提出了一种通过信息披露(information disclosure)的方式来进一步提高 MTD 的效果.通过使用以防御者为主导、攻击者为从者的贝叶斯劝说模型(Bayesian Persuasion model),防御者可以设计一种信号机制来利用 MTD 产生的系统不确定性来进一步影响攻击者的行为.通过此种方式构建一种 Bayesian-Stackelberg 博弈基础之上的具有联合迁移(migration)和信号(signaling)策略模型.理论分析计算结果表明,该模型可以提高 MTD 的效果,使之更加有效.但是该方案缺乏实验证明,且缺乏对博弈过程的预测能力.

2.3　Markov 攻守双方博弈

Markov 模型是一种状态间转移仅依赖于前 n 个状态的过程的模型.Markov 模型作为一种统计模型在近年被很好的引入到攻守双方博弈中.在现有的研究中,每一次攻守双方博弈构成一个决策单元,产生一个状态,根据攻守双方的策略向量可以得到攻守双方的状态转移矩阵,将概率引入则构成了整体的博弈状态转移矩阵[8-9].从而根据 Markov 模型计算博弈的效用从而模拟一次攻守双方的博弈.

Maleki 等人[8]将 MTD 中攻守双方博弈定义为一种基于 Markov 模型的博弈,在一个时间步(timestep)内攻守双方互相做出一个决策并且根据双方的决策产生一个攻防结果(0-攻击未成功、1-攻击成功).文献通过一个随机算法来输出攻击策略、防御策略和效用的一个三元组,并反复使用随机算法来模拟双方的决策,来计算出第一次攻击成功所需要的时间.文献以单主机 IP 跳变模型和多主机隐藏模型来证明了该 Markov 博弈模型的可行性,并且从复杂度分析的计算出发证明了该 MTD Markov 模型提高了攻击者成功攻击的平均成本.该 Markov 移动目标防御攻防博弈模型具有一定的理论使用价值,但是对于实际的复杂的系统环境,如何降低 Markov 模型的容量和计算成本是一个需要考虑的问题.

Lei 等人[9]将 Markov 模型应用于 MTD 的基础之上,并且通过引入攻击面(Attack Surface,AS)和探测面(Exploration Surface,ES)的概念来重新定义了针对网络攻防的动态 MG-MTD(Markov Game MTD)模型以提高 MTD 防御效果和防御精准程度(accuracy).通过变换 AS 和 ES 的方式来增加了移动目标防御决策的普适性(universality),最终兼顾防御者的成本和收益,从而实现最优防御策略决策.文献中通过 AS 与 ES 的变换提高了 MTD 的普适性,从而使之可以应对多种网络攻击,但是文献中对防御者成本和收益的定义相对抽象,如何将该模型与实际系统中的防御者成本与收益关联是一个需要考虑的问题.

后续 Lei 等人[15]针对完全信息假设和传统的矩阵博弈结构在应用未知策略的攻防对抗下难以适用的问题,提出了 MTD 下的不完全信息 Markov 博弈理论方法 IIMG-MTD(Incomplete Iformation Markov Game MTD)来实现最优防御策略生成.不完全信息理论相较于之前的方法,更好地体现了现有的攻击者利用未知漏洞而防御者可以使用自适应 MTD 的攻守环境.结合 MG-MTD,可以准确的描述在多状态和多阶段下的 MTD 不完全信息特征.作者通过将均衡策略选择转换为非线性规划问题来简化了计算复杂度.通过数学计算,理论上证明了该算法可以构建有效而实用的最佳防御策略.但是该研究仍未解决单纯的 Markov 博弈无法体现实际攻守状况的问题,且如何在不完全博弈中摆脱贝叶斯法则的限制也是需要研究问题.

2.4　移动目标防御机制有效性评估

移动目标防御机制的有效性评估是 MTD 中十分重要的一个部分,它用于评估和比较不同防御机制的有效性,从而为后续的 MTD 防御机制设计提供参考.

Zhuang 等人[16]首先提出了可以采用仿真模拟的方法来评估 MTD 防御机制的有效性，并通过 NeSSi2 创建模拟网络. 通过随机改变网络节点属性与周期性发起攻击来判断机制的有效性. 该方法具有良好的参考价值，现今仿真模拟工具较多，可以使用 NS-3 等工具进行仿真模拟. 在后续的工作中，Zhuang 等人[17]提出了通过构建分析模型来计算在同一网络中的 MTD 节点与常规节点之间被攻击的概率，从而可以直观的比较不同防御机制参数的作用.

Carroll 等人[18]和 Luo 等人[19]延续了使用模型进行分析的思路，分别提出在瓮模型下分析网络地址变换和端口跳变技术的方案. 研究发现地址变换技术仅能保护带有少量脆弱节点的网络，且在实际中网络地址变换成本较高；反之，端口跳变能够有效提高防御侦查的能力，如果系统中易被攻击服务较少且端口池较大，端口跳变可以较好地应用.

Evans 等人[20]和 Han 等人[21]以及 Lei 等人[15]通过理论分析的方式评估了防御机制的有效性. Evans 等人通过理论分析方法分析了动态多样化技术在多种常见主被动攻击下的效果；Han 等人通过网络传播动力学理论对能够改变攻防结构、能够改变攻防能力以及能够同时改变攻防结构和能力 3 类 MTD 技术进行了评估，并分析其有效性；Lei 等人则是通过理论分析和数学计算的方式来证明研究中提出的不完全信息 Markov 博弈理论方法的有效性. 因此，理论分析具有较好的效果，但是配合模型或模拟则更加具有说服力.

Okhravi 等人[22]和 Clark 等人[23]则采用了基于混合方法的评估. 虽然使用混合方法进行评估更为有效，但是现有的混合评估方案具有单一性和特殊性，且缺乏对效能的评估和不同机制之间的比较. 因此，后续的评估需要考虑到多种方法的混合，且需要综合考虑防御的成本和效果.

3 Stackelberg-Markov 非对等三方博弈模型

传统的移动目标防御中的博弈都是攻守双方的博弈，本文提出了一种基于 Stackelberg 博弈和 Markov 模型的非对等三方博弈模型. 本文有以下两条前提假设：

前提假设 1. 博弈的任何参与者都是合乎理性的，会综合考虑自身成本与收益而做出最有利于自身的决策.

前提假设 2. 防御者可以通过态势感知的手段获取部分或全部的关于攻击者的此次攻击的信息，至少获取到攻击行为的产生.

3.1 引入用户的非对等三方博弈模型

传统的攻守双方博弈并没有考虑到系统中的用户，只考虑防御者和攻击者的博弈可能会导致防御者做出效用很高的但是对系统功能和性能影响极大的决策(如：直接切断系统网络从而抵御一切网络攻击，但是会造成系统服务的离线). 系统的主要目的是为特定的用户群体提供服务，因此需要将用户考虑到攻守当中.

定义 1. 防御者(Defender). 防御者全称为系统维护与防御集合，是一个由防御人员、防御策略和防御系统与设施组成的集合. 在本模型中，防御者主要指的是防御策略的智能决策系统. 其目的是在受到外界攻击的时候可以在一定程度上抵御住攻击，并且尽力维持系统和服务的功能与性能.

定义 2. 攻击者(Attacker). 攻击者全称为攻击系统单元实体. 对于防御者而言，不需要考虑攻击者具体是什么，只需要通过态势感知的手段，得到关于某次攻击的信息，从而进行后续的决策，更加符合信息不对等的实际场景.

定义 3. 用户(User). 全称为系统使用用户. 是一个特殊的博弈参与者，用户无法获知当前攻击者和防御者的行为和决策情况，并且对两者的行为和决策并不关心. 用户只关心当前系统状态——即系统和服务的功能与性能.

将用户引入系统作为参与者并同时考虑防御者的成本和收益以及用户的成本和收益，可以使防御者的防御策略的选择更加的精准和科学.

定义 4. 攻击行为(Attack Behavior). 攻击行为指的是由攻击者发起的、目标为侵犯防御者或用户利益的、与其他攻击行为具有逻辑关联性的行为；攻击行为是可以被防御者通过态势感知手段获取的，可体现为发生攻击行为与未发生攻击行为的二分类结果.

定义 5. 攻击策略(Attack Strategy). 攻击策略是攻击行为的内在逻辑体现，攻击行为一定会映射至某攻击策略，但并非一一对应关系.

定义 6. 攻击类型(Attack Type). 攻击类型是攻击策略的外在表征形式，未知攻击策略表征为未知攻击类型.

3.2 Stackelberg 博弈在非对等三方博弈中的应用

在 SMATG-MTD 中，Stackelberg 博弈中的领

导者对应于攻击者，而跟随者对应于防御者. 由于攻击者的侦查行为对于防御者而言是不可知的，因此可以将侦查行为定义为必然发生的攻击行为. 通过态势感知的手段可以预先知道攻击者的攻击行为(即使不知道具体的攻击策略). 那么攻击者则预先进行决策，而攻击者也知道防御者会根据攻击者的决策进行一系列的决策，体现出防御者在博弈中的劣势.

定义 7.　参与者策略容器. 攻击者攻击策略向量 $\boldsymbol{A}=\{a_0,a_1,a_2,a_3,\cdots\}$(攻击策略不能为空，攻击策略为空视为未发生攻击，$a_0$ 用于归纳防御者的未知攻击类型). 攻击者在发动攻击之前需要进行侦查行为，从而获取系统的漏洞，并根据漏洞制定一系列的攻击策略. 防御者防御策略二维数组 $\boldsymbol{D}=\{\boldsymbol{D}_0,\boldsymbol{D}_1,\boldsymbol{D}_2,\boldsymbol{D}_3,\cdots\}$($\boldsymbol{D}_0$ 表示攻击行为已知但攻击类型未知的防御策略向量). 其中每一个元素对应于一种攻击所对应的防御策略向量 $\boldsymbol{D}_i=\{\varepsilon,d_{i1},d_{i2},d_{i3},\cdots\}$($\varepsilon$ 表示防御者并未选择一个防御策略，防御策略为空). 用户行为策略向量 $\boldsymbol{U}=\{\varepsilon,u_1,u_2,u_3,\cdots\}$($\varepsilon$ 表示用户并未选择一个行为策略，行为策略为空). 根据先验知识，攻击者攻击策略向量长度应等于防御者防御策略二维数组中防御策略向量的个数.

定义 8.　参与者策略概率容器. 对于每个参与者而言，对于自身的策略容器中策略的选择需要遵循一定的规则，每一个规则都有可能被选择，而每一个规则都对应于一个被选择的概率. 攻击者攻击策略向量在特定的攻击状态 S_i 下具有相同的策略向量但是具有不同的概率分布，从而构成矩阵形式 $\boldsymbol{P}(A^S)=\{\boldsymbol{P}(A^{S_0}),\boldsymbol{P}(A^{S_1}),\boldsymbol{P}(A^{S_2}),\cdots\}$. 其中每个向量 $\boldsymbol{P}(A^{S_i})=\{\boldsymbol{P}(\varepsilon^{S_i}),\boldsymbol{P}(a_1^{S_i}),\boldsymbol{P}(a_2^{S_i}),\cdots\}$. 每一个攻击策略对应的防御者防御策略的防御策略向量概率分布 $\boldsymbol{P}(\boldsymbol{D}_i)=\{\boldsymbol{P}(\varepsilon),\boldsymbol{P}(d_{i1}),\boldsymbol{P}(d_{i2}),\cdots\}$，若干向量构成二维数组 $\boldsymbol{P}(\boldsymbol{D})=\{\boldsymbol{P}(D_0),\boldsymbol{P}(D_1),\boldsymbol{P}(D_2),\cdots\}$. 对于防御者根据攻击者特定的攻击行为而选择特定的防御策略向量的过程则不由概率分布决定，而直接通过一一对应的匹配关系决定. 对于用户而言，用户行为策略向量概率分布 $\boldsymbol{P}(\boldsymbol{U})=\{\boldsymbol{P}(\varepsilon),\boldsymbol{P}(u_1),\boldsymbol{P}(u_2),\cdots\}$，只和系统状态有关.

定义 9.　时间步(timestep). 同一攻击者进行一次攻击和进行下一次攻击的时间间隔区间. 倘若防御者未能在攻击者的一次攻击与下一次攻击之间做出决策，则体现为防御者防御策略为空(等效于未知攻击类型下的未做出防御决策，$d_{00}=\varepsilon$).

如图 1 所示，在一个时间步内：

(1) 攻击者首先对系统进行侦查，获取了系统脆弱性与系统当前状态.

(2) 攻击者根据当前攻击状态 S_i 从其对应的攻击策略向量概率分布 $\boldsymbol{P}(A^{S_i})$ 中选取一个攻击策略向量中的策略 a_i 发动一次攻击.

(3) 用户发现系统功能失效或性能下降，从用户行为策略向量中选取一个策略 u_p.

(4) 防御者通过态势感知的手段发现攻击者使用了攻击策略 a_i，针对 a_i 从其对应的防御策略二维数组 $\boldsymbol{D}$ 中选取 a_i 对应的防御策略向量 $\boldsymbol{D}_i$，根据防御策略向量概率分布 $\boldsymbol{P}(\boldsymbol{D}_i)=\{\boldsymbol{P}(\varepsilon),\boldsymbol{P}(d_{i1}),\boldsymbol{P}(d_{i2}),\cdots\}$ 选取一个策略 d_{ij}；或者防御者感知到了攻击行为，但不知道攻击策略. 则根据未知攻击类型防御策略向量概率分布 $\boldsymbol{P}(\boldsymbol{D}_0)=\{\boldsymbol{P}(\varepsilon),\boldsymbol{P}(d_{01}),\boldsymbol{P}(d_{02}),\cdots\}$ 选取一个防御策略 d_{0a}(若 $d_{ij}=d_{i0}=\varepsilon$，则表示防御者未能做出决策或防御者选择了不做任何操作).

(5) 用户发现系统功能失效或性能下降，从用户行为策略向量中选取一个策略 u_q.

(6) 系统状态与攻防效用评估，得到此时间步的攻击结果状态.

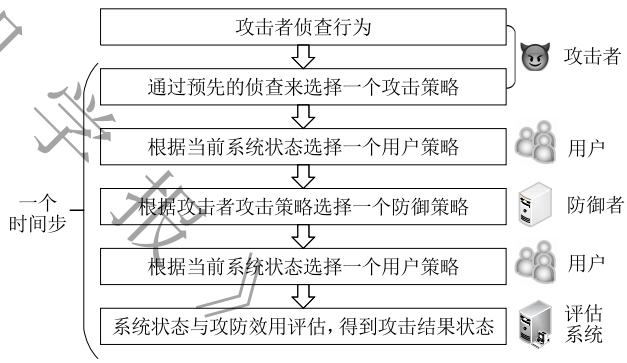

图 1　Stackelberg 三方博弈时间步示意图

3.3　Markov 博弈在非对等三方博弈中的应用

Markov 决策过程(Markov Decision Process，MDP)是指决策者根据每个时刻观察到的状态，从可用的行为集合中按照一定的概率选用行为的过程. 系统下一步的状态是随机的，状态转移概率具有 Markov 性，即下一时刻的状态只与当前时刻状态相关. Markov 博弈是由博弈论和 MDP 综合而来，综合考虑多个参加者的决策.

攻防双方的互动通常不会一蹴而就，很有可能涉及多个步骤，每个步骤双方采取的策略都以系统当时的状态为决策依据，下一个系统状态由双方的策略组合共同决定. 因此，在时间上有先后顺序的一连串博弈作为一个完整过程，符合 Markov 决策过

程特征.鉴于此,本文拟将连续时间步内的单元博弈构成的连续博弈过程,构建为一个攻、防、用户三方参与的 Markov 博弈模型.

定义 10. 攻防 Markov 链.整个攻防博弈过程,构成一个 Markov 决策过程,也就是一个 Markov 链.由于攻防博弈是以攻击作为导向的,因此,其中每一个状态节点都由一个 Stackelberg 博弈最终表现出来的攻击结果状态所构成.为了更加契合实际模型和降低计算时间复杂度,我们选取一阶 Markov 攻击链作为本攻防 Markov 链的模型.也就是下一攻击状态只能由上一攻击状态决定.在第一次态势感知获取到攻击类型的时候,便可以生成 Markov 攻击链.倘若第一次态势感知只获取到攻击行为而未知攻击类型,则将 Markov 攻击链的长度定为常见攻击类型 Markov 攻击链长度的均值.1 阶 Markov 攻击链如图 2 所示.

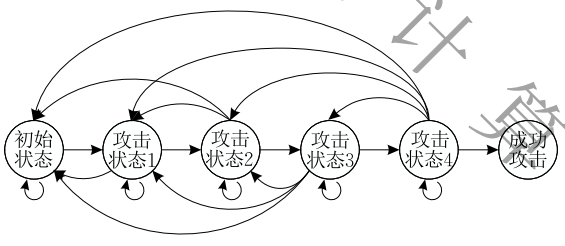

图 2 一阶 Markov 攻击链

定义 11. Markov 三方博弈模型.Markov 三方博弈模型是一个九元组$\langle \boldsymbol{S},\boldsymbol{A},\boldsymbol{D},\boldsymbol{U},T,\mu_{\boldsymbol{D}},\mu_{\boldsymbol{A}},\mu_{\boldsymbol{U}},\beta\rangle$.其中,$\boldsymbol{S}$ 是系统状态集,也就是攻击状态集;$\boldsymbol{A}$ 是攻击者动作集,体现为 Stackelberg 博弈中攻击者攻击策略向量;$\boldsymbol{D}$ 是防御者动作集,体现为 Stackelberg 博弈防御者防御策略向量之集合;$\boldsymbol{U}$ 是用户动作集,体现为 Stackelberg 博弈用户行为策略向量;$T:\boldsymbol{S}\times\boldsymbol{A}\times\boldsymbol{D}\times\boldsymbol{U}\rightarrow Prob(\boldsymbol{S})$ 是攻击状态转移函数,对每个攻击状态和博弈参与者动作组合,给出下一个攻击状态的概率分布,即 $P(s'|s,a,d,u)$,根据该攻击状态转移函数,可以构建状态转移矩阵 $\boldsymbol{M}_{\text{Attack}}$;$\mu_{\boldsymbol{D}}:\boldsymbol{S}\times\boldsymbol{A}\times\boldsymbol{D}\times\boldsymbol{U}\rightarrow R$ 是防御者回报函数(R 为实数集),类似地,$\mu_{\boldsymbol{A}}:\boldsymbol{S}\times\boldsymbol{A}\times\boldsymbol{D}\times\boldsymbol{U}\rightarrow R$、$\mu_{\boldsymbol{U}}:\boldsymbol{S}\times\boldsymbol{A}\times\boldsymbol{D}\times\boldsymbol{U}\rightarrow R$ 分别是攻击者和用户回报函数;$\beta(0<\beta\leq 1)$ 为折扣因子,表示预期回报的折扣.

定义 12. K 阶 Markov 博弈预测.对于 Markov 三方博弈而言,单纯的分析已经发生的事件并不能体现博弈的逻辑,需要在当前决策的时候对将来其他参与者的行为和策略进行预测,从而使决策最优.通过博弈决策树预测模型来预测当前步的之后 K 步的情况从而对各个决策的权重进行修正.

4 兼顾成本和收益的最优防御策略

4.1 非对等三方博弈中参与者各自的成本和收益

对于攻击者而言,攻击者只希望对目标系统进行拒绝服务、窃密或破坏等攻击行为,因此攻击者的成本对应于实施某项攻击策略所产生的开销 $\delta_{\boldsymbol{A}}=CostA(a_i)$,而攻击者的收益则为该项攻击策略在受到防御者的防御和用户行为的作用之后,对目标系统产生的攻击影响 $\theta_{\boldsymbol{A}}=GainA(a_i,d_{ij})$.在前文定义的标准 Stackelberg 博弈过程中,用户进行两次决策,在模型构建时均进行考虑.因此,综合考虑攻击者的效用 $\mu_{\boldsymbol{A}}=GainA(a_i,d_{ij})-CostA(a_i)$.

同样,对于防御者而言,本文防御者的目的是对已存在的攻击行为进行防御,不存在攻击行为情况下的预先防御不在本文的考虑范畴之内.因此,防御者的防御某项攻击时(可能只知攻击行为而未知攻击策略)的成本 $\delta_{\boldsymbol{D}}=CostD(a_i,d_{ij})$,而防御者的收益则是取决于能否防御下此次攻击并防御到何种程度 $\theta_{\boldsymbol{D}}=GainD(a_i,d_{ij})$.因此,综合考虑防御者的效用 $\mu_{\boldsymbol{D}}=GainD(a_i,d_{ij})-CostD(a_i,d_{ij})$.

对于用户而言,用户的成本和收益仅取决于系统当前的状态和用户所做出的决策.对于用户而言,本来就要使用该系统,因此用户不存在收益的概念,即 $\mu_{\boldsymbol{U}}=-\delta_{\boldsymbol{U}}$.所以用户只存在当系统状态发生改变之后产生的损失,称为用户的成本 $\delta_{\boldsymbol{U}}=CostU(u_p,u_q)$.用户的损失仅由用户的行为决定,对系统功能或性能产生的影响归为防御者的损失.

4.2 非对等三方博弈中参与者之间的效用关系

非对等三方博弈中,由于用户的地位和攻守双方并不对等,因此将三者放在一起考虑并不妥当,应从两两效用关系入手.

首先考虑攻击者和防御者两位博弈参与者之间的效用关系.攻守双方作为不存在合作的对立双方,应构成零和博弈,双方均选取当前对自身最有利的策略,从而构成纳什均衡,即 $\mu_{\boldsymbol{A}}+\mu_{\boldsymbol{D}}=0$.

用户的效用收到防御者和攻击者策略的影响,只要防御者或攻击者做出了不为空的策略,那么该策略一定会对系统的功能和性能产生影响.因此 $\mu_{\boldsymbol{U}}=-\delta_{\boldsymbol{U}}$ 与 $\mu_{\boldsymbol{A}}$ 或 $\mu_{\boldsymbol{D}}$ 无关,仅由系统状态的改变计算得到.

4.3 兼顾防御者和用户效用的防御策略智能决策

兼顾防御者和用户效用也就是在进行防御策略选取的时候使防御者的效用与用户的效用的代数和

最小. 两者代数和具有许多的计算方式，本文选取最简单的加权代数和计算方式作为防御者和用户效用的兼顾平衡方式.

假设在当前系统中，对防御者分配的权重为 β，$0<\beta\leq 1$，那么分配给用户的权重就为 $1-\beta$. 所以综合考虑的效用 $\mu_{Both}=\beta\times\mu_{\boldsymbol{D}}+(1-\beta)\times\mu_{\boldsymbol{U}}$.

假设对当前系统状态进行评估通过评估函数来进行，本文模型将评估测度抽象化为系统的功能与性能，因此在攻击者攻击策略产生的系统状态影响的差值为

$$\Delta E(a_i)=Eval(func,perf|a_i,s_k)-Eval(func,perf|s_k) \tag{1}$$

同理，在此后防御者防御策略产生的系统状态影响的差值为

$$\Delta E(d_{ij})=Eval(func,perf|d_{ij},s_k)-Eval(func,perf|s_k) \tag{2}$$

最终 Markov 攻击链攻击状态变换影响的差值为

$$\Delta E(s_k\rightarrow s_h)=EvalS(s_h)-EvalS(s_k) \tag{3}$$

因此，用户的效用

$$\begin{aligned}\mu_{\boldsymbol{U}}&=-CostU(u_p,u_q)\\&=-Cost(u_p)-Cost(u_q)\end{aligned} \tag{4}$$

攻击者的效用

$$\mu_{\boldsymbol{A}}=-\alpha\times(\Delta E(a_i)+\Delta E(d_{ij}))+(1-\alpha)\times(\Delta E(s_k\rightarrow s_h))-CostA(a_i) \tag{5}$$

其中 α 用于平衡对系统产生影响的收益和攻击状态改变的收益.

由于攻击者和防御者构成零和博弈，对于已知类型攻击而言，$CostA(a_i)$ 是可以预设为先验知识的，因此可以利用 $\mu_{\boldsymbol{D}}=-\mu_{\boldsymbol{A}}$ 来计算当前时间步的综合效用.

$$\mu_{Both}=-\beta\times[-\alpha\times(\Delta E(a_i)+\Delta E(d_{ij}))+(1-\alpha)\times(\Delta E(s_k\rightarrow s_h))-CostA(a_i)]-(1-\beta)\times[Cost(u_p)+Cost(u_q)] \tag{6}$$

而在攻击类型未知的情况下，$CostA(a_i)=CostA(a_0)$ 则取为所有已知攻击成本的平均值.

以上仅考虑到兼顾防御者和用户的效用的计算，接下来则是需要根据效用计算和 Markov 博弈预测来选取做出当前的防御策略的智能决策.

根据 Markov 状态转移方程：

$$T:\boldsymbol{S}\times\boldsymbol{A}\times\boldsymbol{D}\times\boldsymbol{U}\rightarrow Prob(\boldsymbol{S}) \tag{7}$$

和攻击者、防御者、用户三方各方选择策略的概率分布，可以进行 K 阶 Markov 博弈预测，计算出当前选择某一防御策略之后，在 K 个时间步后的综合收益的数学期望，从而选择在当前情况下的最优策略. 状态转移方程的结果向量应符合多项式分布.

假设当前攻击状态为 S_h，攻击者已经选取了攻击策略 a_i，用户选取了用户策略 u_p，防御者在向量 $\boldsymbol{D}_i$ 中有若干策略可选. 倘若选择了防御策略 d_{ij}，则可以计算出状态转移矩阵中当前状态 S_h 转移到其他状态的概率分布 $P(h\rightarrow i)$，$i=0,1,\cdots$. 此时基于防御策略 d_{ij} 对下一个时间步的综合效用的预测值（1 阶 Markov 博弈预测）为

$$\hat{\mu}_{pre1}(d_{ij}|a_i,u_p)=\Big[\sum_q\boldsymbol{P}(u_q)\times\sum_i^{h+1}\sum_e\sum_f\sum_x\sum_y\boldsymbol{P}(h\rightarrow i)\times\boldsymbol{P}(a_e)\times\boldsymbol{P}(d_{ef})\times\boldsymbol{P}(u_x)\times\boldsymbol{P}(u_y)\times\hat{\mu}_{Both}\Big]\Big/\boldsymbol{P}(d_{ij}) \tag{8}$$

对于不同的防御策略 d_{ij}，$d=0,1,\cdots$，分别计算 $\hat{\mu}_{pre1}(d_{ij}|a_i,u_p)$，选择其中效用值最大的策略（均小于等于 0），则为当前防御者的最优策略.

对于 K 阶 Markov 博弈预测，只需要在计算 1 阶 Markov 博弈预测的基础之上，继续进行预测计算即可. 最终选择其中效用值最高的策略作为当前智能决策的策略（效用值为负）.

$$\hat{\mu}_{preK}(d_{ij}|a_i,u_p)=\hat{\mu}_{pre1}(d_{ij}|a_i,u_p)\times\prod_{f=2}^{K}\Big|\sum_{i=0}^{t}\sum_e\sum_f\sum_x\sum_y[P(s_{cur}\rightarrow i)\times P(a_e)\times P(d_{ef})\times P(u_x)\times P(u_y)\times\hat{\mu}_{Both}]\Big| \tag{9}$$

4.4　非独立性用户情况下的 SMATG

在用户为非独立性用户的情况下（即用户与防御者并非独立双方，用户与防御者存在信息共享和统筹规划，包括用户与防御者为一体），用户与防御者之间可以通过交流来决定双方共同的效用最大化. 在忽略交流成本的情况下，防御者可以通过与用户交流来确定以系统防御效果为主还是用户损失最小化，也就是可以根据实际遭遇到的攻击情况对防御者分配权重 β 进行调整.

且该情况下防御者可以更加智能的进行决策或安排用户进行决策、更加清楚地知晓用户行为对系统状态带来的影响、使用户损失计算度更加准确且可以降低由于用户行为与概率分布的不确定性所带来的预测计算开销.

由于第 3 节理性假设的存在，不考虑攻击者与用户存在信息共享与统筹规划的情况.

5　数学分析

5.1　算法复杂度分析

按照标准的大 O 算法复杂度分析方法，取每一

个和“加”操作为单元操作.从计算分布来看,状态转移矩阵、各方策略影响度及其概率分布以及系统状态变更所带来的系统网络属性变化均可以先验统计计算,因此不用将其统计到算法的复杂度当中.

因此本模型的算法复杂度分析直接考虑 K 阶 Markov 博弈预测公式出发.

首先从 1 阶 Markov 博弈预测公式来看,由于多阶取值求和的存在,时间复杂度 $T(n)$与求和阶的个数和每一个被求和的单元值的个数相关,即

$$T(n)=O(u_q)\times O(h\rightarrow i)\times O(a)\times O(d)\times O(u_x)\times O(u_y)\times T(\mu_{Both}) \tag{10}$$

因为对系统攻击状态转移的各种参数值均可以预先计算,所以所有 $T(eval)$均可视为在 $O(1)$时间复杂度范围内完成,$T(n)=O(n^6)$.

而对于 K 阶预测,则是在前一阶的基础之上进行运算,所以 $T(n)=O(n^{6K})$,但 K 的值非常有限.在简化计算模型的情况下,时间复杂度可以降低为 $O(n^{5K})$.

虽然看上去时间复杂度非常高,但是攻击状态、各方策略等数量都是很有限的.根据业界统计,以单一攻击种类为例,各参数离散可取值规模局限于两位数内,且大部分为个位数.因此,从工程角度来说,本文模型运算时间非常有限,工程上可行.

5.2 与其他模型与算法的比较

将本文所提出的模型及其算法与文献[7-9,11,14-15]提出的博弈模型从信息需求、博弈类型、过程状态预测性、场景描述性、场景适应性等方面进行分析比较,结果如表 1 所示由于各个博弈模型应用的场景不同,因此没有进行算法复杂度的比较,且各博弈模型所求解的可操作性都较好,直接给出防御者应该选择的防御策略,因此不予比较.过程状态预测性代表对后续参与者行为的预测能力,该能力能够对当前决策进行反馈修正,以提高正确决策的概率.场景描述性体现了对实际攻防场景的描述能力,良好的场景描述性可以提高工程化应用价值.场景适应性则代表模型可以应用的场景范围,场景适应性越好,可以解决的攻防场景就越复杂.

表 1 与现有 MTD 模型的比较

MTD 模型	信息需求	博弈类型	博弈参与者	过程状态预测性	场景描述性	场景适应性
文献[7]	完全信息	Bayesian Stackelberg 博弈	攻击者与防御者	缺乏	好	简单 MTD 场景
文献[8]	完全信息	单 Markov 链模型	攻击者与防御者	一般	一般	IP 跳变场景
文献[9]	完全信息	完全信息 Markov 博弈	攻击者与防御者	好	一般	简单 MTD 场景
文献[11]	完全信息	信号博弈策略与 Stackelberg 博弈混合	攻击者与防御者	一般	好	简单 MTD 场景
文献[14]	完全信息	Bayesian Stackelberg 博弈	攻击者与防御者	缺乏	一般	非 MTD 场景
文献[15]	不完全信息	不完全信息 Markov 博弈	攻击者与防御者	好	一般	不完全信息 MTD 场景
SMATG-MTD	未知攻击类型的不完全信息	Stackelberg 博弈与 Markov 博弈混合	攻击者、防御者与用户	好	好	复杂 MTD 场景

5.3 数学实例分析

以数据窃取攻击和高度保密性和大量用户使用的信息系统为例,传统被动静态的防御手段对此类攻击基本没有抵抗能力,甚至难以发现此类攻击.

常规的通过随机变换等方式进行的 MTD 防御方案,没有考虑到攻守双方之间的逻辑关系,定时地或随机地进行环境变换导致成本十分高昂,虽然能一定程度上防御数据窃取攻击,但是防御方开销过大,在需要保证可用性和性能的信息系统上并不合适.

在 SMATG-MTD 中,防御者系统通过态势感知的方式获取到攻击方进行数据窃取攻击的行为和具体攻击类型,从而在每一步通过运算和预测选出最优的策略,从而应对数据窃取攻击.

如图 3 所示,首先攻击者在完成侦查行为之后,对系统进行扫描从而发现漏洞,但是对系统的扫描造成了系统参数的异常,被态势感知模块所感知到并通知防御系统,防御系统开始对攻击者的策略进行分析;接着防御者进行防御策略的决策,通过已有的先验知识构建博弈决策决策树预测模型从而获取每一种防御者策略可能在 K 步 Stackelberg 决策之后所造成的状态,并且从中选择一个最优的策略.图中选择的则是端口跳变策略,而在选择端口跳变策略之后,用户由于出现短暂的访问则选择了减少访问的策略.若该次扫描攻击者成功找到了可以利用的提权漏洞,那么防御者会通过态势感知获知到攻击者的提权行为,通过再一次进行计算与预测,做出了对攻击者有关 IP 网段进行封禁的操作,长此以往,防御攻击者进

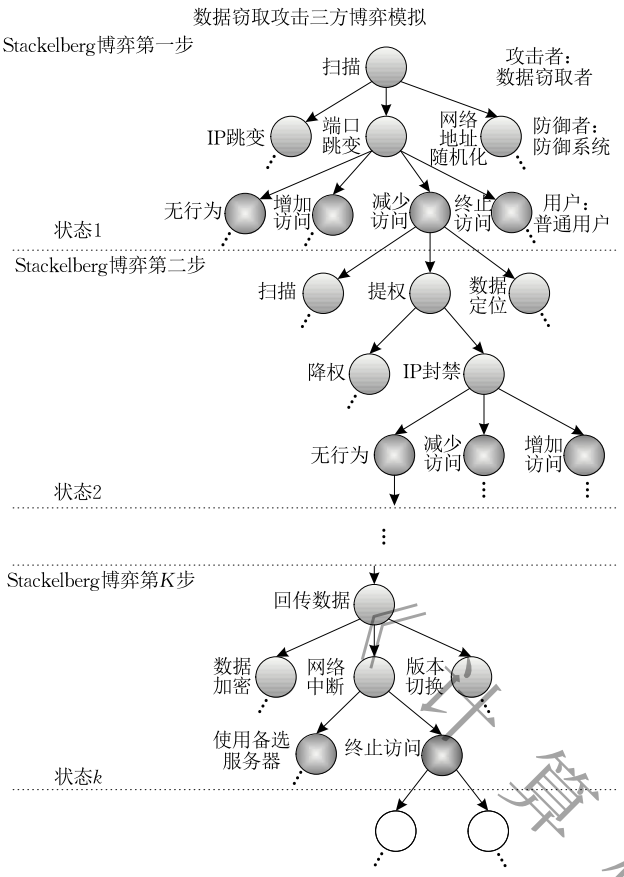

图 3　数据窃取攻击三方博弈模拟图

行的每一次攻击，且防御策略主要为 MTD 动态攻击面转移技术.

接下来将通过数学分析实例来证明 SMATG-MTD 的效果.

在本实例中，系统状态评估函数：

$$Eval(func,perf|s_k),k=0,1,\cdots \tag{11}$$

中 $func$ 功能参数由当前系统功能可用数量 λ 决定；$perf$ 性能参数由当前使用该系统的用户数量 N_{usr}、当前可用服务网络平均时延 $\bar{T}_{delay}$ 和系统平均吞吐量 $\bar{\theta}_{thr}$ 决定，系统每一次遭遇策略（任何一方的决策）都会记录并修正该策略对系统功能与性能的影响，以作为后续计算的先验知识.

攻击状态评估函数对于不同种攻击行为具有不同的表现形式，攻击状态评估的结果由进入下一攻击状态对系统运行产生的影响、对系统被攻破概率分布的影响和两者所造成的实际损失决定. 根据 Verizon 等机构的 2018 年数据泄露与窃取调查报告，数据窃取攻击对被攻击方造成的损失主要来自于数据本身之价值. 因而，数据窃取攻击状态评估函数应着重考虑对系统被攻破概率分布的影响与数据被窃取所造成的实际损失. 为保证简洁性，本实例中将数据价值等额化（每种数据被窃取的损失额度是相同的），仅考虑攻击状态对系统被攻破概率分布的影响. 以统计报告中攻击策略、攻击状态与最终攻击成功的统计数量来看，随着攻击的深入，系统被攻破的概率逐渐提高. 将攻击状态评估函数定义为线性函数 $EvalS(S_k)=Para_S\times k$，通过上述数据进行拟合后 $Para_S$ 为 45.17，因而攻击状态评估函数 $EvalS(S_k)=45\times k$.

系统状态如何评估与度量并不在本文的研究内容之内，因此现仅考虑最简单的线性无关模型.

$$\begin{aligned}&Eval(func,perf|s_k)\\&=\rho\times Eval(func|s_k)+\\&\quad(1-\rho)\times Eval(perf|s_k)\\&=\rho\times W_\lambda\times\lambda+(1-\rho)\times\\&\quad(W_N\times N_{usr}-W_T\times\bar{T}_{delay}+W_{\bar{\theta}}\times\bar{\theta}_{thr})\end{aligned} \tag{12}$$

信息系统中较为看重数据的保密性，因此，各个参数对应的权重控制变量如下式所示，用于控制各个变量的权重和调整不同参数之间单位差距所带来的影响：

$$\begin{aligned}&\alpha=0.2,\ \beta=0.7,\ \rho=0.5,\\&W_\lambda=20,\ W_N=5,\ W_T=0.3,\ W_{\bar{\theta}}=0.01\end{aligned} \tag{13}$$

为了简化计算，默认系统用户仅在攻守双方完成各自的决策之后进行一次决策，即 $U_p=\varepsilon$.

如表 2 所示，通过先验知识可以得到该环境下的攻击状态基础转移矩阵.

表 2　攻击状态基础转移矩阵

状态	S_0	S_1	S_2	S_3	S_4	S_F
S_0	0.2	0.8	0	0	0	0
S_1	0.5	0.1	0.4	0	0	0
S_2	0.2	0.05	0.3	0.45	0	0
S_3	0.1	0.05	0.05	0.2	0.6	0
S_4	0.01	0.01	0.02	0.06	0.1	0.8
S_F	0	0	0	0	0	1

对于状态转移函数 $T:\boldsymbol{S}\times\boldsymbol{A}\times\boldsymbol{D}\times\boldsymbol{U}\to Prob(\boldsymbol{S})$ 而言，在一阶 Markov 攻击链中，后以攻击状态只能由前一攻击状态得到，攻击者攻击策略在当前的攻击状态下具有针对性，且难以直接准确构建多项式分布攻击状态分布. 因此，在本实例中，采用正态分布对每一特定攻击状态下的攻击状态转移概率进行拟合.

$$Prob(\boldsymbol{S})\sim N(m,n^2),$$

$$f(x)=\frac{1}{\sqrt{2\pi}n}\mathrm{e}^{\left(-\frac{(x-m)^2}{2n^2}\right)} \tag{14}$$

其中，不同的参与者策略之间在时间步内博弈的结果，将会影响到正态分布的均值 m；可转移到的

攻击状态个数(基础转移矩阵中当前行转移概率不为 0 的状态个数)则会影响到正态分布的方差 n^2. 而当前状态下最终攻击状态转移的概率值,则由根据 x 算出的分布概率值向量和基础转移矩阵决定. 体现出不同防御策略之间的防御效果差异.

不妨令 h 等于当前可选状态的个数,x 为目标转移状态的编号.

$$\boldsymbol{PVal}_{\text{State}}=Prob(S_k\rightarrow S_h)=f(x)\times M_{\text{Base}}(S_k\rightarrow S_h),0\leq h\leq k+1 \quad (15)$$

然后对状态转移概率向量 $\boldsymbol{PVal}_{\text{State}}$ 进行归一化处理,使其所有概率和为 1,则得到了当前博弈结果下攻击状态所对应的状态转移概率.

对于攻击者而言,攻击者每一时间步的攻击目标都是希望能够使当前攻击状态转移到下一攻击状态,因此攻击者对正态分布均值 m 的影响度 $I_a=x+1$. 防御者对 m 的影响度则由先验知识决定,根据在相同攻击策略情况下的防御仿真实验结果,进行归一化处理并约减为至多一位的小数之后,得到如表 3 所示的防御策略影响度表. 其中策略代号所对应的具体防御策略在表 5 中提供.

表 3 防御者策略影响度

策略代号	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}
影响度	0	−0.6	−0.5	−0.8	−2
策略代号	d_{20}	d_{21}	d_{22}	d_{23}	d_{24}
影响度	0	−1.4	−1	−1.5	−1
策略代号	d_{30}	d_{31}	d_{32}	d_{33}	
影响度	0	−1	−3	−2	
策略代号	d_{40}	d_{41}	d_{42}	d_{43}	
影响度	0	−1.5	−3	−2.5	
策略代号	d_{50}	d_{51}	d_{52}	d_{53}	
影响度	0	−1	−7	−3	

在特殊情况下,当 $m<0$ 时,表明防御者该策略可以完全防御整个攻击行为,攻击状态应转移到初始状态.

在已知攻击状态和攻击类型时,通过先验知识来计算特定攻击状态下的某一攻击策略的发生概率,为了便于计算,仅选取常见的攻击类型与攻击策略作为先验知识. 如表 4 所示,现有数据窃取攻击需要经过最少以下几步操作:扫描获取可以进行渗透的端口及漏洞,通过漏洞进行提权操作以获取 root 权限或组权限,通过数据定位操作寻找到数据所在目录,通过代码注入对数据进行隐秘化处理,最后将处理过的数据回传至攻击主机以实现数据窃取. 根据 CERNET 边界采集并确认的 217 份网络攻击样本数据结合 Verizon 等机构的 2018 年数据泄露与窃取调查报告可以发现,在攻击初期攻击者存在扫描漏洞或以合法用户身份定位数据的操作,其分布比例如下表中策略概率所示. 其他攻击状态下的攻击者策略概率分布则是由该样本数据的后续攻击流量进行推演产生的,该数据经过规约并约减至两位小数.

表 4 攻击者策略与概率分布

策略代号		a_1	a_2	a_3	a_4	a_5
策略描述		扫描	提权	数据定位	代码注入	数据回传
攻击状态	S_0	0.95	0	0.05	0	0
	S_1	0.26	0.68	0.05	0.01	0
	S_2	0.1	0.1	0.7	0.1	0
	S_3	0.07	0.05	0.15	0.73	0
	S_4	0.05	0	0.1	0.05	0.8

现有 MTD 攻击面变换策略覆盖网络、数据、软件、操作系统等层面,在遭遇网络攻击时使用网络层面的攻击面变换策略(如 IP 跳变、端口跳变等),可以有效的在攻击早期对网络攻击进行防范;而数据层面、软件层面、操作系统层面等层面的攻击面变换虽然也可以有效的防御网络攻击,不过在数据窃取前期攻击的前提下,效果均不如网络层面的攻击面变换(在以数据为目标的攻击下,前期隐藏被攻击主机比隐藏数据或漏洞更加有效);反之,在攻击中后期使用其他层面的 MTD 策略则也更优,因而产生了如表 5 所示的防御策略编排,其中策略选择概率则是由先验知识归纳而成,随着系统运行会逐渐修正,初始值影响较小.

表 5 防御者策略与概率分布

代号	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}
描述	无行为	IP 跳变	端口跳变	网络地址随机化	网络架构更改
概率	0.05	0.4	0.4	0.05	0.1
代号	d_{20}	d_{21}	d_{22}	d_{23}	d_{24}
描述	无行为	IP 封禁	降权	IP 跳变	用户重认证
概率	0.05	0.3	0.3	0.1	0.25
代号	d_{30}	d_{31}	d_{32}	d_{33}	
描述	无行为	数据加密	数据迁移	数据随机化	
概率	0.05	0.5	0.2	0.25	
代号	d_{40}	d_{41}	d_{42}	d_{43}	
描述	无行为	IP 跳变	进程锁定	可疑进程终止	
概率	0.05	0.15	0.4	0.4	
代号	d_{50}	d_{51}	d_{52}	d_{53}	
描述	无行为	数据加密	网络中断	软件版本切换	
概率	0.05	0.3	0.6	0.1	

同理,用户策略与概率分布如表 6 所示.

表 6　用户策略与概率分布

	策略代号	u_0	u_1	u_2	u_3
	策略描述	无行为	增加访问	减少访问	终止访问
系统状态	性能略微下降	0.8	0.05	0.14	0.01
	性能部分下降	0.3	0.01	0.66	0.05
	性能大幅下降	0.1	0	0.8	0.1
	部分服务离线	0.01	0	0.64	0.35
	系统离线	0	0	0	1

假设当前态势感知系统感知到攻击者在对系统进行数据窃取攻击，且进行扫描操作，防御系统则根据 SMATG-MTD 模型进行一阶博弈预测计算，计算结果如表 7 所示.

表 7　一阶博弈预测计算结果

策略代号	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}
策略描述	无行为	IP 跳变	端口跳变	网络地址随机化	网络架构更改
用户无行为	−246.56	−146.05	−237.51	−2631.88	−1642.03
用户增加访问	−246.56	−5.63	−8.69	−93.60	−56.80
用户减少访问	−1627.27	−421.50	−624.40	−6564.95	−3885.04
用户终止访问	−164.37	−35.73	−51.16	−526.69	−304.65
一阶预测效用和	−2284.75	−608.90	−921.77	−9817.12	−5888.52

而后在一阶博弈预测的基础之上，进行 K 阶博弈预测计算($K=3$，结果保留至整数).

表 8　3 阶博弈预测计算结果

策略代号	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}
策略描述	无行为	IP 跳变	端口跳变	网络地址随机化	网络架构更改
3 阶预测效用和	−26914	−3716	−4029	−34676	−18318

从表 8 中结果可以看出，通过 SMATG-MTD 模型计算之后，应选择 IP 跳变的防御策略. 该防御策略是当前先验知识下的最优结果，在此次时间步中 Stackelberg 博弈结束后，可以对该策略的有效性进行评估，从而对攻击者、防御者和用户的策略分布进行修正，以逐渐完善先验知识. 从该数学实例可以看出，对实际的高度保密性的、大量用户使用的信息系统而言，在合理配置各项先验参数的情况下，可以很好的抵御数据窃取攻击. 只需要将该数学模型通过编程构建为部署在控制服务器上的主控模块和部署在数据服务器上的从控模块，主控模块在控制服务器上进行防御策略的博弈计算，将结果发送给从控模块以实现防御策略的实施.

5.4　仿真实验

本文通过 MATLAB 仿真平台实现如图 4 所示的以上数学实例的仿真，该实验环境包括攻击者、防御者与用户三个单元，每个单元之间通过管道(pipe)进行数据传输. 假设态势感知可以完全实现，体现为攻击者将在每次进行策略选择之后将攻击策略通过管道告知防御者；攻击者与防御者也完全知道当前系统所处的攻击状态，状态转移概率由基础状态转移矩阵、攻击者对正态分布均值 m 的影响度 I_a 与防御者对 m 的影响度 I_d 联合运算所决定(由防御者进行运算，作为攻击状态是否转移的判定基础)；且其中所有概率选择的实际结果均由轮盘选择算法决定.

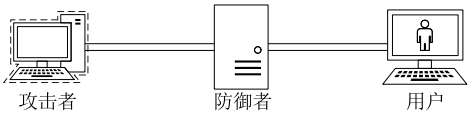

图 4　仿真实验架构示意图

仿真实验结果表明，从初始状态 S_0 到达 S_1 攻击状态进行了 3 次攻击，表 9 为在 S_1 攻击状态下的 1 阶博弈预测结果.

表 9　S_1 攻击状态下的 1 阶博弈预测计算结果

策略代号	d_{20}	d_{21}	d_{22}	d_{23}	d_{24}
策略描述	无行为	IP 封禁	降权	IP 跳变	用户重认证
1 阶预测效用和	−3454.68	−1279.62	−1856.3	−7258.66	−3591.40

后续通过攻击收敛运算，数学期望上攻击者完成攻击需要进行 2234 次攻击行为，实际使用 1629 次，具有良好的防御效果.

6　总结与展望

本文在传统双方攻守博弈的基础上创新性地引入了用户作为第三参与者，并且融合了 Stackelberg 博弈和 Markov 模型，以构建 Stackelberg-Markov 非对等三方博弈模型. 然后根据该模型定义了三方各自的效用以及效用之间的关系. 该模型相比于现有的攻守双方的博弈而言，更符合实际的场景并能应用于复杂攻击场景，且具有较好的过程状态预测效果. 数学分析与仿真实验表明，SMATG-MTD 能够兼顾防御者和用户的成本和收益，避免过度的防御和不适宜的防御，从而有效地实现防御策略智能决策.

但是该模型在模型预测方面的计算复杂度较大.虽然各种策略影响下的状态转移矩阵均可预先计算,不过对于较庞大的博弈策略空间而言,仍会有较高的预测成本;且该模型需要较多的先验知识,且对系统状态的评估也需要通过其他模型来支撑.后续研究重点应着眼于如何降低计算复杂度、如何度量和评估系统的状态并引入自适应策略概率分布修正机制.

参 考 文 献

[1] Cai Gui-Lin, Wang Bao-Sheng, Wang Tian-Zuo, et al. Research and development of moving target defense technology. Journal of Computer Research & Development, 2016, 53(5): 968-987(in Chinese)
(蔡桂林, 王宝生, 王天佐等. 移动目标防御技术研究进展. 计算机研究与发展, 2016, 53(5): 968-987)

[2] Zhou Yu-Yang, Cheng Guang, Guo Chun-Sheng, et al. A survey on attack surface dynamic transfer technology based on moving target defense. Journal of Software, 2018, 29(9): 2799-2820(in Chinese)
(周余阳, 程光, 郭春生等. 移动目标防御的攻击面动态转移技术研究综述. 软件学报, 2018, 29(9): 2799-2820)

[3] Al-Shaer E, Duan Q, Jafarian J H. Random host mutation for moving target defense//Proceedings of the International Conference on Security and Privacy in Communication Systems. Padua, Italy, 2012: 310-327

[4] Dunlop M, Groat S, Urbanski W, et al. MT6D: A moving target IPv6 defense//Proceedings of the Military Communications Conference. Baltimore, USA, 2012: 1321-1326

[5] Azab M, Hassan R, Eltoweissy M. ChameleonSoft: A moving target defense system//Proceedings of the International Conference on Collaborative Computing: Networking, Applications and Worksharing. Orlando, USA, 2011: 241-250

[6] Jajodia S, Ghosh A K, Swarup V, et al. Moving Target Defense: Creating Asymmetric Uncertainty for Cyber Threats. Berlin, Germany: Springer Science & Business Media, 2011

[7] Vadlamudi S G, Sengupta S, Taguinod M, et al. Moving target defense for web applications using bayesian Stackelberg games//Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems. Singapore, 2016: 1377-1378

[8] Maleki H, Valizadeh S, Koch W, et al. Markov modeling of moving target defense games//Proceedings of the ACM Workshop on Moving Target Defense. Vienna, Austria, 2016: 81-92

[9] Lei Cheng, Ma Duo-He, Zhang Hong-Qi. Optimal strategy selection for moving target defense based on Markov game. IEEE Access, 2017, 5(99): 156-169

[10] Zhu Q. Game-theoretic approach to feedback-driven multi-stage moving target defense//Proceedings of the International Conference on Decision and Game Theory for Security. New York, USA, 2013: 246-263

[11] Feng X, Zheng Z, Cansever D, et al. A signaling game model for moving target defense//Proceedings of the IEEE INFOCOM 2017-IEEE Conference on Computer Communications. Atlanta, USA, 2017: 1-9

[12] Anderson N, Mitchell R, Chen I R. Parameterizing moving target defenses//Proceedings of the IFIP International Conference on New Technologies, Mobility and Security. Larnaca, Cyprus, 2016: 1-6

[13] Vorobeychik Y, Singh S. Computing Stackelberg equilibria in discounted stochastic games//Proceedings of the 26th AAAI Conference on Artificial Intelligence. Toronto, Canada, 2012: 1-7

[14] Qian Zhen, Wei Cheng-Jian, Wang Kai, et al. Web security research base on Bayesian Stackelberg game. Application of Electronic Technique, 2015, 41(12): 124-128(in Chinese)
(钱震, 蔚承建, 王开等. 基于贝叶斯 Stackelberg 博弈的 Web 安全问题研究. 电子技术应用, 2015, 41(12): 124-128)

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

[16] Zhuang R, Zhang S, DeLoach S A, et al. Simulation-based approaches to studying effectiveness of moving-target network defense//Proceedings of the National Symposium on Moving Target Research. Annapolis, USA, 2012: 246

[17] Zhuang R, Deloach S A, Ou X. A model for analyzing the effect of moving target defenses on enterprise networks//Proceedings of the Cyber and Information Security Research Conference. Oak Ridge, USA, 2014: 73-76

[18] Carroll T E, Crouse M, Fulp E W, et al. Analysis of network address shuffling as a moving target defense//Proceedings of the IEEE International Conference on Communications. Sydney, Australia, 2014: 701-706

[19] Luo Y B, Wang B S, Cai G L. Effectiveness of port hopping as a moving target defense//Proceedings of the International Conference on Security Technology. Haikou, China, 2014: 7-10

[20] Evans D, Nguyen-Tuong A, Knight J. Moving Target Defense: Effectiveness of Moving Target Defenses. New York, USA: Springer, 2011

[21] Han Y, Lu W, Xu S. Characterizing the power of moving target defense via cyber epidemic dynamics//Proceedings of the 2014 Symposium and Bootcamp on the Science of Security. Raleigh, USA, 2014: 10

[22] Okhravi H, Riordan J, Carter K. Research in Attacks, Intrusions and Defenses: Quantitative evaluation of dynamic platform techniques as a defensive mechanism. Berlin, Germany: Springer International Publishing, 2014: 405-425

[23] Clark A, Sun K, Poovendran R. Effectiveness of IP address randomization in decoy-based moving target defense//Proceedings of the 52nd IEEE Conference on Decision and Control. Palazzo dei Congressi, Italy, 2013:678-685

CHEN Zi-Han, Ph. D. candidate. His research interests include cyber security, network measurement & behavior analysis, MTD game, trust architecture.

CHENG Guang, Ph. D., professor, Ph. D. supervisior. His research interests include SDN network measurement and management, network traffic measurement and big data analysis, botnet, APT attack detection.

Background

Moving Target Defense technology is proposed in recent years as a kind of new proactive defense technology that hopes to confuse attackers by implementing continuous and dynamic changes which are used to defend the attacks by switching the attack surface.

However, existing MTD games have simple game models while the participators are limited to the attacker and the defender. And existing MTD models cannot fitting the description of the actual system well, lacking prediction of the state of offense-defense process and short of ability to apply to complex attack scenarios.

In this paper, we proposed an asymmetrical trilateral game model based on Stackelberg game and Markov model used in MTD (SMATG-MTD). First, we innovatively introduce to take the user as the third participant to the MTD game. The offensive and defensive sides participate in the game as equal players, while the user is at a disadvantage as a third player in the game and has an asymmetrical relationship with the other two participants. As a result, the asymmetrical trilateral game is constructed. And the model is able to take both user and defender into consideration. Second, we combine Stackelberg game and Markov model to better describe the actual offensive-defensive scenarios and enhance the predictability of game model. Third, to raise the ability to apply to complex attack scenarios and realize intelligently MTD defense strategy making, this paper presents a *K*-orders Markov game prediction method with a utility calculation method using multi-measure evaluation. Compared with the existing game between offensive and defensive sides, SMATG-MTD optimizes the user experience of using the system, improves the adaptability of the model, makes the optimization convergence process more consistent with the actual scenes and can be applied to the complex attack scenes, which has a better effect of process state prediction. Mathematical analysis and simulation experiments demonstrate that the model can give consideration to both the defender and the user by balancing their utilities, avoid excessive defense and inappropriate defense strategies, so as to effectively realize the intelligent defense strategy decision-making.

However, there is still a large computational complexity in model prediction, which would cost much in larger game strategy space. Therefore, the subsequent research should focus on how to reduce the computational complexity, how to measure and evaluate the state of the system, and introduce the adaptive strategy probability distribution correction mechanism.

This paper is supported by National Key R&D Program of China (2018YFB1800602, 2017YFB1800602), Ministry of Education-China Mobile Research Fund Project (MCM20180506), the CERNET Innovation Project (NGICS20190101, NGII20170406). This paper is part of the topic for intelligently MTD defense strategy making.

The research team has focused on MTD mechanism for years, and some papers in this field have published in highly-ranked journals.