基于信号博弈的移动目标防御最优策略选取方法

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. 信息保障技术重点实验室，北京 100093）

摘 要：针对移动目标防御最优策略选取问题，从攻击面转换（ASS）和探测面扩展（ESE）的角度形式化来定义防御策略，阐释了防御原理；采用动态对抗和有限信息的视角对网络攻防行为进行研究，在分析攻防博弈类型和攻防过程的基础上，构建了基于信号博弈的移动目标防御模型；改进了攻防策略量化计算方法，提出了精炼贝叶斯均衡求解算法，并通过对博弈均衡的分析设计了最优防御策略选取算法。仿真实验验证了所提模型和方法的有效性。

关键词：网络安全；移动目标防御；信号博弈；精炼贝叶斯均衡；防御策略选取

中图分类号：TP309

文献标识码：A

doi: 10.11959/j.issn.1000−436x.2019125

Optimal strategy selection method for moving target defense based on signaling game

JIANG Lyu[1], ZHANG Hengwei[1,2], WANG Jindong[1]

1. The Third Institute, Strategic Support Force Information Engineering University, Zhengzhou 450001, China

2. Science and Technology on Information Assurance Laboratory, Beijing 100093, China

Abstract: To solve the problem of the optimal strategy selection for moving target defense, the defense strategy was defined formally, the defense principle from the perspective of attack surface shifting and exploration surface enlarging was taken into account. Then, network attack-defense behaviors were analyzed from the sight of dynamic confrontation and bounded information. According to the analysis of attack-defense game types and confrontation process, the moving target defense model based on signaling game was constructed. Meanwhile, the method to quantify strategies was improved and the solution of perfect Bayesian equilibrium was proposed. Furthermore, the optimal defense strategy selection algorithm was designed by the equilibrium analysis. Finally, the simulation demonstrates the effectiveness and feasibility of the proposed optimal strategy and selection method.

Key words: network security, moving target defense, signaling game, perfect Bayesian equilibrium, defense strategy selection

1 引言

网络系统作为国家关键基础设施，对电力、交通、金融、能源、航运等重要领域的有效运转具有关键支撑作用[1]。然而频繁爆发的网络安全事件表明，网络安全形势十分严峻。网络攻防对抗中“易攻难守”的特点突出，攻击者具有时间优势、信息不对称优势和成本优势[2]。移动目标防御（MTD, moving target defense）作为主动防御技术，能够有效地提升防御效能，其核心思想是利用多样化、随时间持续变化的机制和策略，来增加网络攻击的复杂度和成本，降低网络系统脆弱性暴露和被攻击的概率，提升防御能力[3]。防御策略选取方法是提升防御效能的关键和最佳途径，也是移动目标防御领域的研究重点。

收稿日期：2018−12−12；**修回日期**：2019−04−28

通信作者：张恒巍，zhw11qd@163.com

基金项目：国家自然科学基金资助项目（No.61521003, No.61572517）；河南省科技攻关计划基金资助项目（No.182102210144）

Foundation Items: The National Natural Science Foundation of China (No.61521003, No.61572517), The Science and Technology Research Project of Henan Province (No.182102210144)

网络安全的本质是攻防对抗，因此从攻防对抗的角度出发，研究探索网络攻防分析和防御决策方法体系具有重要的现实意义[4]。博弈理论与网络攻防所具有的目标对立性、关系非合作性和策略依存性十分吻合[5]。目前，运用博弈理论开展移动目标防御决策研究已经取得部分成果。文献[6]提出了一种基于完全信息静态博弈的 MTD 模型，但是由于采用完全信息条件，在描述实际攻防过程时准确性不足。文献[7]对此加以改进，采用不完全信息静态博弈研究 MTD 的防御机理。文献[8]进一步针对 MTD 与 Web 平台的适配性，提出了基于不完全信息静态博弈的最优防御策略选取方法，增强了网络防御效能。但是，以上成果均以静态博弈模型为基础，由于网络攻防具有动态性，攻防双方同时行动的限制条件很难满足。因此，动态不完全信息博弈更加符合实际，研究成果的实用性和指导意义更大。

从网络攻防实际出发，信号博弈模型（SGM, signaling game model）因为可以准确地描述情报信息对攻防双方策略选择的关键作用而受到研究者的特别关注。文献[9]针对 DDoS 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，设计了防御决策算法。文献[10]通过分析精炼贝叶斯均衡，对信息安全威胁定量评估进行了研究。文献[11]采用信号博弈来建模攻防场景，设计了一种信息安全防御机制。但是，上述理论研究成果未与具体防御机制结合，在网络防御实践应用方面存在不足。

本文以信号博弈理论和移动目标防御原理为基础，探索结合 2 种方法优势的主动防御机制，首先，从攻击面转换（ASS, attack surface shifting）和探测面扩展（ESE, exploration surface enlarging）的角度形式化来定义移动目标防御策略，通过主动改变目标系统的资源属性，提升系统的防御主动性。其次，针对攻防信息的不完全性和攻防过程的动态性，基于对防御者释放信号机制的一般性分析，采用防御者为信号发送者、攻击者为信号接收者的结构对网络攻防过程进行建模和博弈分析，提出精炼贝叶斯均衡求解方法，设计最优主动防御策略选取算法。最后，利用仿真实验验证了本文模型和方法的有效性，通过分析实验结果，总结结合信号博弈和移动目标防御方法实施综合性主动防御的特点规律。

2　移动目标防御信号博弈模型

2.1　移动目标防御机制

在网络攻防对抗中，如果防御者能够通过主动行为对攻击者的决策与行动实施干扰或影响，则体现了主动防御思想，具有更好的防御效果[12]。移动目标防御是一种典型的主动防御技术，通过增加网络系统的动态性和随机性，从时空维度提高网络系统结构的不可预测性，削弱和降低攻击者在网络对抗中的优势[13]。

定义 1　攻击面（AS, attack surface）是指防御者为了防止攻击者利用某些系统资源脆弱性成功发起攻击所需转移或变换的资源集合，由攻击面维度及其取值构成，记为 AS={ASD,ASV}。其中，系统攻击面的维度为 $\mathrm{ASD}=\{\mathrm{asd}_1,\mathrm{asd}_2,\cdots,\mathrm{asd}_k\}$，表示可能被利用的资源集合，如网络服务端口、服务协议等；$\mathrm{ASV}=\{\mathrm{asv}_1,\mathrm{asv}_2,\cdots,\mathrm{asv}_k\}$ 表示攻击面维度的取值。

定义 2　探测面（ES, exploration surface）是指攻击者为了能够进入目标系统并实现攻击目的所需探索的系统资源集合，由探测面维度及其取值范围构成，即 ES={ESD,ESV}。其中，探测面维度为 $\mathrm{ESD}=\{\mathrm{esd}_1,\mathrm{esd}_2,\cdots,\mathrm{esd}_l\}$，表示攻击者所探测到的系统资源集合，即目标系统资源配置属性，如系统指纹、数据存储位置等；$\mathrm{ESV}=\{\mathrm{esv}_1,\mathrm{esv}_2,\cdots,\ \mathrm{esv}_l\}$ 表示攻击者所探测到的系统资源维度的取值范围。

参考文献[13-14]，给出 2 种主要防御手段，即攻击面转换和探测面扩展的定义。

定义 3　攻击面转换是指在 t 时刻，目标系统满足以下 2 个条件之一，则说明目标系统攻击面发生了转换。

1) $\exists \Delta t>0, \mathrm{ASD}^t-\mathrm{ASD}^{t+\Delta t}\neq\varnothing$，即攻击面维度发生改变，简称攻击面转移。

2) $\exists \Delta t>0, (\mathrm{ASD}^t=\mathrm{ASD}^{t+\Delta t})\wedge(\mathrm{ASV}_i^t\neq\mathrm{ASV}_i^{t+\Delta t})$，即攻击面维度取值发生变化，简称攻击面变换。

定义 4　探测面扩展是指 t 时刻，目标系统满足以下 2 个条件之一，则说明探测面发生了扩展。

1) $\exists \Delta t>0, \mathrm{ESD}^t\subset\mathrm{ESD}^{t+\Delta t}$，即通过增加探测面维度扩展探测面。

2) $\exists \Delta t>0, \mathrm{ESD}^t=\mathrm{ESD}^{t+\Delta t}, \mathrm{ESV}^t\subset\mathrm{ESV}^{t+\Delta t}$，即通过增加探测面维度的取值范围扩展探测面。

MTD 通过灵活组合使用探测面扩展和攻击面转移、攻击面变换等手段，能够有效增强目标系统的动态性、多样性和不确定性，提高目标系统的攻击难度。

2.2 基于信号博弈的移动目标防御分析

分析网络攻防对抗实际场景可知，一方面，因为网络信息系统的开放互联需求、服务监管的要求等约束，网络系统的防御能力、防御策略，甚至防御设备都是公开信息；另一方面，攻击者往往采用嗅探、扫描等技术手段主动收集网络系统的防御情报。上述由攻击者抓取或防御者释放的有关防御信息是攻击行动决策和规划的重要依据，本文将其定义为防御信号。从主动防御思想出发，防御者通过有选择地主动释放真实描述网络系统的信息（真实信号）或与网络系统真实情况不一致的信息（虚假信号），影响或制约攻击者的情报判断和行动规划，增强攻击者对目标的认知难度，提升防御效果。

由于网络对抗中防御者一般会事先部署和实施防御策略，本文将防御者定义为信号博弈的 leader 和信号发送者，攻击者定义为信号博弈的 follower 和信号接收者。结合 MTD 机制在实际攻防对抗过程中的特点，该类型博弈具有如下特征。

1) 主动性

攻防双方不会将己方关键的博弈策略信息告知对方，攻击者可以通过网络探测手段获取目标系统的脆弱性信息；防御者通过转换攻击面中网络资源的脆弱性维度和取值或者扩展探测面的维度空间和取值范围，可以降低或避免系统脆弱性暴露的可能。因此，作为信号发送方的防御者通过主动释放虚假攻击面或者探测面信息，欺骗、迷惑作为信号接收方的攻击者。相比单纯的主动防御手段，利用信号博弈机制可以增强防御者在攻防过程中的主动性，提升 MTD 的防御效果。

2) 不完全信息性

由于攻防双方都希望在对抗过程中占据信息优势，攻击者与防御者会尽可能地减少自身博弈信息的暴露程度。因此，攻防双方的对抗博弈具有不完全信息性。

3) 动态性

在攻防过程中，攻防双方的行动有先有后，攻击者在接收到防御者释放的防御信号后，基于自身对防御类型的先验知识和后验判断，采取相应的攻击策略，防御者根据攻击策略采取针对性防御策略。因此，攻防博弈具有动态性。

由于防御者无法预知遭受攻击的时间，MTD 实际应用中一般采取固定周期或动态周期机制变换防御策略来抵御攻击。MTD 防御策略的内容主要包括改变攻击面和探测面的维度空间、不同维度取值范围和变化频率。其中，AS/ES 的维度空间代表各种网络系统资源；AS/ES 的维度取值范围代表不同系统资源属性的取值范围；变化频率代表单位时间内 AS/ES 的维度和取值范围改变的次数，包括固定频率和动态频率 2 种方式。维度空间和取值范围越大，变化频率越高，表明系统结构的动态性和随机性越强，攻击者越难发现并有效利用系统脆弱性。综上，将 MTD 防御策略形式化描述为一个五元组 (asd,asv,esd,esv,sf)，其中，asd、asv、esd、esv 的定义见 2.1 节，sf 代表变化频率。

2.3 博弈模型构建

本文采用信号博弈刻画在攻防博弈过程中防御者采取 MTD 防御策略，同时主动释放防御信号来欺骗、迷惑和诱导攻击者，进而增强其对防御类型的不确定性，提升防御效能。本文定义防御者是信号发送方，攻击者是信号接收方。

定义 5 移动目标防御信号博弈模型（MTDSG, moving target defense signaling game model）可以表示为七元组 $(N,T,M,B,p,\tilde{p},U)$，各参数定义如下。

1) $N=\{N_a,N_d\}$ 是博弈模型局中人集合，N_a 为攻击者，N_d 为防御者。

2) $T=(T_a,T_d)$ 为局中人 N_a 和 N_d 的类型空间。其中，$T_d=\{t_d^1,t_d^2,\cdots,t_d^n\}$ 表示防御者的类型集合，$n\geqslant 1$，$n\in \mathrm{N}^+$，防御类型 T_d 是私有信息，攻击者对于 T_d 的概率分布仅有先验知识，$T_a=\{t_a\}$ 表示攻击者类型。

3) M 为防御信号空间，$M=\{m_1,m_2,\cdots,m_n\}$ 由防御者释放，信号名称与防御者类型对应，防御者可以自主选择发送真实信号或虚假信号。

4) $B=(D,A)$ 是博弈双方的策略空间。$D=(d_1,d_2,\cdots,d_g)$ 为防御者的 MTD 策略集合，$g\geqslant 1$，$g\in \mathrm{N}^+$，其中 $d_i=(\mathrm{asd}_i,\mathrm{asv}_i,\mathrm{esd}_i,\mathrm{esv}_i,\mathrm{sf})$；$A=\{a_1,a_2,\cdots,a_h\}$ 为攻击策略，$h\geqslant 1$，$h\in \mathrm{N}^+$。

5) p 是攻击者的先验信念集合，表示攻击者对防御类型的先验知识，记为 $p(T_d)=(p_1,p_2,\cdots,p_n)$，$p_i=p(t_d^i)\geqslant 0$，满足 $\sum_{i=1}^{n}p_i=1$。

6) $\tilde{p}$ 是攻击者的后验信念集合，$\tilde{p}_{ij}=\tilde{p}(t_d^i\mid m_j)$ 表示攻击者接收到防御信号 m_j，使用贝叶斯法则修正后对防御类型 t_d^i 的后验推断，i、$j\in[1,n]$。

7) $U=\{U_d,U_a\}$ 是博弈双方收益函数集合。

2.4 博弈策略收益量化

攻防策略收益量化直接影响博弈分析和均衡计算，也是最优防御策略选取的基础。结合文献[11,15]，本文对攻防策略收益进行量化计算。

定义 6 系统损失代价（SDC, system damage cost）是指攻击者发动攻击后给系统带来的损失，一般由攻防策略共同决定，可认为是攻防策略的回报，通常用系统资源重要程度（C, criticality）、攻击致命度（AL, attack lethality）、安全属性损害（SAD, security attribute damage）进行描述，一般将防御策略实施后降低的系统损失代价作为防御回报。攻击成本（AC, attack cost）是指攻击者发现并利用系统资源脆弱性发动攻击所付出的成本，通常包括对探测面的信息侦测和情报收集，以及发动攻击行为所需的时间和系统软硬件资源等。防御成本（DC, defense cost）是指防御者为隐藏自身防御类型信息，以及实施 MTD 策略所需的时间和系统软硬件资源。具体定义及计算方法参考文献[16-17]。

定义 7 信号成本（SC, signal cost）是指防御者主动释放虚假信号，用以欺骗、迷惑和诱导攻击者所付出的代价。

依据不同等级防御策略的差异度量信号成本。参考文献[11,18]，根据防御策略能够应对的攻击行动的权限不同，将 SC 分为 3 个等级，取值范围分别为 $\mathrm{SC}_1 \in [1,50)$、$\mathrm{SC}_2 \in [50,100)$、$\mathrm{SC}_3 \in [100,200)$。基于上述定义，参考本文改进的收益量化方法，可以得到攻防双方的期望收益分别为

$$U_a(t_d^i, m_j, a_h) = \mathrm{SDC}(a_h) - \mathrm{AC}(a_h) \tag{1}$$

$$U_d(t_d^i, m_j, a_h) = \mathrm{SDC}(a_h) - \mathrm{DC}(t_d^i) - \mathrm{SC}(m_j) \tag{2}$$

由于同一防御等级的防御策略成本大致相同，可以认为它们的防御收益也基本一致，若某个防御等级有 m 个防御策略，则可以认为防御者采用等概率 $\beta_k = \dfrac{1}{m}$ 选择该等级的第 k 个防御策略，得到防御者在该防御等级下的期望收益为

$$U_d(t_d^i) = \sum_{k=1}^{m} \beta_k U_d(t_d^i, m_j, a_h) \tag{3}$$

3 博弈均衡计算与防御策略选取

3.1 博弈均衡计算

定义 8 MTDSG 的精炼贝叶斯均衡由策略组合 $(m^*(t_d), a^*(m))$ 与后验概率 $\tilde{p}(t_d \mid m)$ 组成，并且满足以下条件。

1) $a^*(m) \in \arg\max_{a \in A} \sum \tilde{p}(t_d \mid m)\, U_a(t_d, m, a)$。

2) $m^*(t_d) \in \arg\max_{m \in M} U_d(t_d, a^*(m), m)$。

3) $\tilde{p}^*(t_d \mid m)$ 是攻击者根据先验概率 p、接收到的防御信号 m 和攻击者的最优策略 $a^*(m)$ 得到的。

上述定义中，条件 1)表示给定后验概率 $\tilde{p}(t_d \mid m)$，攻击者针对防御者释放的防御信号做出的最优攻击策略；条件 2)表示在攻防双方完全理性的条件下，防御者预测到攻击者最优策略 $a^*(m)$，防御者选择最优防御类型；条件 3)是攻击者运用贝叶斯法则修正先验概率得到后验概率的过程。

精炼贝叶斯均衡的具体计算过程可参考 4.2 节收益计算与 4.3 节均衡求解与防御策略选取。

3.2 最优防御策略选取算法及对比分析

基于上述研究，给出基于信号博弈的移动目标防御最优策略选取算法，如算法 1 所示。

算法 1 基于信号博弈的移动目标防御最优策略选取算法

输入 MTDSG

输出 最优防御策略 $m^*(t_d)$

1) 初始化 $\mathrm{MTDSG} = (N, T, M, B, p, \tilde{p}, U)$；

2) 构建防御类型空间 $T_d = \{t_d^1, t_d^2, \cdots, t_d^n \mid n \geqslant 1\}$；

3) 构建防御信号空间 $M = \{m_j, 1 \leqslant j \leqslant n\}$；

4) 构建防御策略集合 $D = (d_1, d_2, \cdots, d_g)$，$g \geqslant 1$；

5) 构建攻击策略集合 $A = (a_1, a_2, \cdots, a_h)$，$h \geqslant 1$；

6) 对防御类型 t_d^i 及防御信号 m_j，$U_d(t_d^i, m_j, a_k) = \mathrm{SDC}(a_k) - \mathrm{DC}(t_d^i) - \mathrm{SC}(m_j)$；

7) 对攻击策略 $a_k \in A$，$U_a(t_d^i, m_j, a_k) = \mathrm{SDC}(a_k) - \mathrm{AC}(a_k)$；

8) 攻击者建立后验概率推断 $P(t_d \mid m)$；

9) 最优攻击策略 $a^*(m) \in \arg\max_{a \in A} \sum p(t_d \mid m) \cdot U_a(t_d, m, a)$；

10) 最优防御策略 $m^*(t_d) \in \arg\max_{m \in M} U_d(t_d, a^*(m), m)$；

11) 根据最优攻防策略求出满足贝叶斯法则的攻击者对防御类型的后验概率推断 $\tilde{P}^*(t_d \mid m)$；

12) if $P(t_d \mid m)$ 与 $\tilde{P}^*(t_d \mid m)$ 不冲突

13) then 求得该博弈模型下的精炼贝叶斯均衡 $(m^*(t_d), a^*(m), \tilde{p}^*(t_d \mid m))$；

14) return $m^*(t_d)$；

15) end if

分析均衡求解算法可知，计算时间复杂度主要取决于精炼贝叶斯均衡解的计算过程，根据 3.1 节的分析，设防御类型空间和防御信号空间的大小为 n，令 $u = \max(g, h)$，由动态博弈理论[19]可知，计算均衡的平均时间复杂度为 $O(u^3 + n^2 + 2n)$。存储空间消耗主要集中在策略收益和均衡求解中间值的存储上，为 $O(un)$。

将本文提出的模型及方法和其他文献进行对比，具体如表 1 所示，其中，博弈类型是指攻防双方对博弈信息的掌握情况及博弈行为顺序；局中人类型是指在博弈模型中博弈参与者是否区分不同类型和类型的多少，实际攻防对抗过程中攻防双方在策略收益、行为成本等方面存在差异，局中人类型细化可以更好地提高攻防策略选取的准确性与针对性；信号机制是指攻防分析是否采用信号博弈以及设定的信号发送方；模型通用性是指博弈类型和攻防策略的数量是否能够扩展；均衡求解是指文献中是否给出了求解博弈均衡的具体方法，如果没有将严重削弱实用性。

通过分析已有方法可知，文献[6]是基于局中人具有完全信息的假设基础之上的，但实际攻防过程中，由于攻防双方侦察能力有限，同时攻防双方均会减少或隐藏自身攻防信息暴露，因此网络攻防对抗一般应作为不完全信息博弈考虑。文献[7]是将攻防对抗作为静态博弈进行研究，不符合实际攻防场景的动态性特征。文献[9,17,20]以攻击者作为信号发送方进行建模研究，基于 2.2 节的分析，防御者释放信号的情况具有更强的一般性，同时防御者可以利用信号机制主动欺骗、迷惑、诱导攻击者，实施主动防御，具有更好的防御效果。文献[21]将基于 MTD 的攻防对抗抽象为马尔可夫博弈，无法分析攻防信息对博弈过程和结果的影响。

通过分析可知，对比文献[6-7]中攻防模型的静态性和信息完全性假设，本文方法基于动态攻防过程和不完全攻防信息开展研究。对比文献[21]采用马尔可夫博弈描述攻防过程，本文方法采用信号博弈进行建模研究，可以准确描述情报信息对攻防双方策略选择的关键作用。对比文献[9,17,20]采用攻击者作为信号发送方，本文方法使用防御者作为信号发送方，增强了方法的一般性，并能够刻画主动释放防御信号来欺骗、迷惑和诱导攻击者的主动防御机制。

本文工作针对攻防实际，基于移动目标防御原理设计防御策略，通过主动改变目标系统的资源属性，增强了系统结构的动态性和随机性，增大了攻击者对目标系统及其防御措施的认知难度，提升了防御效能。同时，基于主动防御思想，结合信号博弈和移动目标防御的优点实施综合防御，利用主动释放针对性防御信号实现了防御欺骗，不但增加了防御样式和手段，进一步提升了防御主动性，而且虚假防御信号的欺骗、迷惑作用可以扰乱和延迟攻击行动，为移动目标防御策略的实施提供反应时间。

4 仿真实验与分析

4.1 仿真环境描述

为了验证 MTDSG 模型和最优防御策略选取算法的可行性与有效性，构建如图 1 所示的典型网络系统[16]开展仿真实验。实验网络系统主要由业务网、接入网和外部互联网构成，其中业务网主要包

表 1 模型与方法对比分析

方法	博弈类型	局中人类型	信号机制	模型通用性	均衡求解	具体应用
文献[6]	完全信息静态博弈	1	—	一般	详细	攻防分析
文献[7]	不完全信息静态博弈	2	—	差	详细	策略选取
文献[9,17]	信号博弈	*n*	攻击者	一般	简单	机制设计
文献[20]	信号博弈	2	攻击者	差	无	—
文献[21]	马尔科夫博弈	*n*	—	较差	简单	攻防分析
本文模型及方法	信号博弈	*n*	防御者	较好	详细	策略选取

括业务服务器、数据服务器和客户端，接入网主要包括 IIS 网络服务器和网络防御设备。

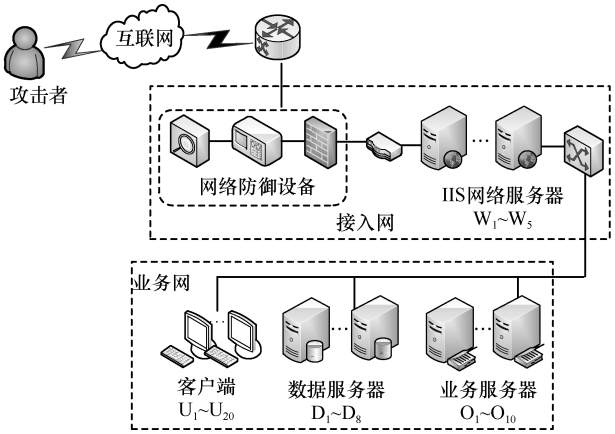


图 1　实验网络系统拓扑

实验网络系统配置相应的访问控制策略规定网络节点之间的访问权限[12]。利用 Nessus 漏洞扫描工具获得网络系统中各节点的资源脆弱性，如表 2 所示。

表 2　　网络节点脆弱性

序号	节点	服务	脆弱性
1	W_1	IIS	LSASS process
2	W_2	FTP	FTP rhost overwrite
3	O_1	RPC	Code injection
4	O_7	RSH	Rsh login
5	U_9	NETBOIS-SSN	Nullsession
6	U_{13}	FTP	Wu-FTP Sockprintf
7	D_5	SQUID PROXY	Squid port scan
8	D_7	SQL DB	Oracle TNS

参考文献[22-24]和美国 MIT 林肯实验室攻防行为数据库[18]，给出攻击和防御动作信息，分别如表 3 和表 4 所示。为方便攻防博弈分析，设防御者类型 $T_d=\{$高能力防御等级 t_d^1，低能力防御等级 $t_d^2\}$，对应的防御信号为 $M=\{m_1,m_2\}$。

表 3　　原子攻击描述

原子攻击名称	分类	攻击成本 AC	攻击致命度 AL
e_1:remote buffer overflow	root	100	9
e_2:install Trojan	probe	80	3
e_3:steal account and crack it	user	140	5
e_4: send abnormal data to GIOP	root	50	8
e_5: LPC to LSASS process	probe	40	2
e_6: FTP rhost attack	root	120	10
e_7: Oracle TNS Listener	root	90	8
e_8: shutdown Database server	user	150	6
e_9:SR-hard blood	root	120	8

4.2　收益计算

借鉴文献[4,22]的方法，综合考虑投入成本和专家意见，设定攻击策略为 $a_1\{e_1,e_4,e_9\}$、$a_2\{e_3,e_5,e_7\}$、$a_3\{e_2,e_6,e_8\}$；自然以概率 (p_h,p_l) 选择防御类型 (t_d^1,t_d^2)，攻击者对防御类型的先验概率 $(p_h,p_l)=(0.3,0.7)$；攻击者收到信号 m_1 后对防御类型 (t_d^1,t_d^2) 的后验修正概率为 $(\alpha,1-\alpha)$，攻击者收到信号 m_2 对防御类型 (t_d^1,t_d^2) 的后验修正概率为 $(\beta,1-\beta)$；防御信号成本 $(SC_1,SC_2,SC_3)=(40,100,180)$，网络系统的安全属性代价 SAD(高, 中, 低)$=(50,30,20)$，其中 IIS 网络服务器、

表 4　　防御策略描述

防御策略		策略描述	MTD 类型	防御成本
t_d^1	d_1	IP switch, data storage enlarge，动态频率	攻击面变换+探测面扩展	220
	d_2	protocol switch, port counterchange, 固定频率	攻击面转移+攻击面变换	190
	d_3	fingerprint switch, Renew root data，动态频率	攻击面转移	170
	d_4	port enlarge, protocol switch, 固定频率	攻击面转移+探测面扩展	160
t_d^2	d_5	route enlarge, install oracle patches，动态频率	探测面扩展	120
	d_6	port switch, patch SSH on FTP，固定频率	攻击面转移	115
	d_7	add physical resources, add address blacklist	探测面扩展	125
	d_8	limit packet to ports, limit ICMP/SYN packets	无	110

业务服务器、数据服务器的 SAD 分别为 20、30、50，资源重要度 C 分别为 3、3、5。

参考文献[11,16]和式(1)~式(3)，分别计算攻击者与防御者的博弈收益。当防御者类型为高等级防御 t_d^1，防御策略为 d_1，释放防御信号 m_1，攻击者采取攻击策略 a_1 时，攻防双方收益为

$$U_a(t_d^1,m_1,a_1,d_1)=\mathrm{SDC}(a_1,d_1)-\mathrm{AC}(a_1)=3\,170-270=2\,900$$

$$U_d(t_d^1,m_1,a_1,d_1)=\mathrm{SDC}(a_1,d_1)-\mathrm{DC}(d_1)-\mathrm{SC}(m_1)=3\,170-220-40=2\,910$$

同理，当防御者采取策略 d_2、d_3、d_4 时，攻防双方收益分别为 $U_a(t_d^1,m_1,a_1,d_2)=2\,910$，$U_d(t_d^1,m_1,a_1,d_2)=3\,240$；$U_a(t_d^1,m_1,a_1,d_3)=2\,870$，$U_d(t_d^1,m_1,a_1,d_3)=3\,200$；$U_a(t_d^1,m_1,a_1,d_4)=2\,930$，$U_d(t_d^1,m_1,a_1,d_4)=3\,120$。综上，该场景下攻击者和防御者的平均收益分别为 $\bar{U}_a(t_d^1,m_1,a_1)=2\,902.5$，$\bar{U}_d(t_d^1,m_1,a_1)=3\,117.5$。同理求得其余攻防策略对应的收益，攻防博弈如图 2 所示。

为了验证仿真实验中 MTD 策略的有效性与针对性，考虑攻防对抗过程中目标系统性能受到的影响程度，借鉴文献[25]中的量化计算方法对目标系统的服务质量进行分析，具体分析在不同攻防策略下，系统中 Web 服务和在线视频服务的平均时延（ADT, average delay time），以此反映系统的性能损失。实验分别对上述 2 种服务进行了 15 次测试，并与系统正常工作时的时延进行对比，获得在攻防对抗情况下的平均时延。具体数据如图 2 所示。

4.3　均衡求解与防御策略选取

按照 3.1 节给出的均衡计算步骤，求解 MTDSG 模型精炼贝叶斯均衡，并选取最优防御策略。

1)攻击者推断的最优攻击策略

$$a^*(m)\in\arg\max_a\sum\tilde{p}(t_d\mid m)\ U_a(t_d,m,a)$$

当 $m=m_1$ 时，有

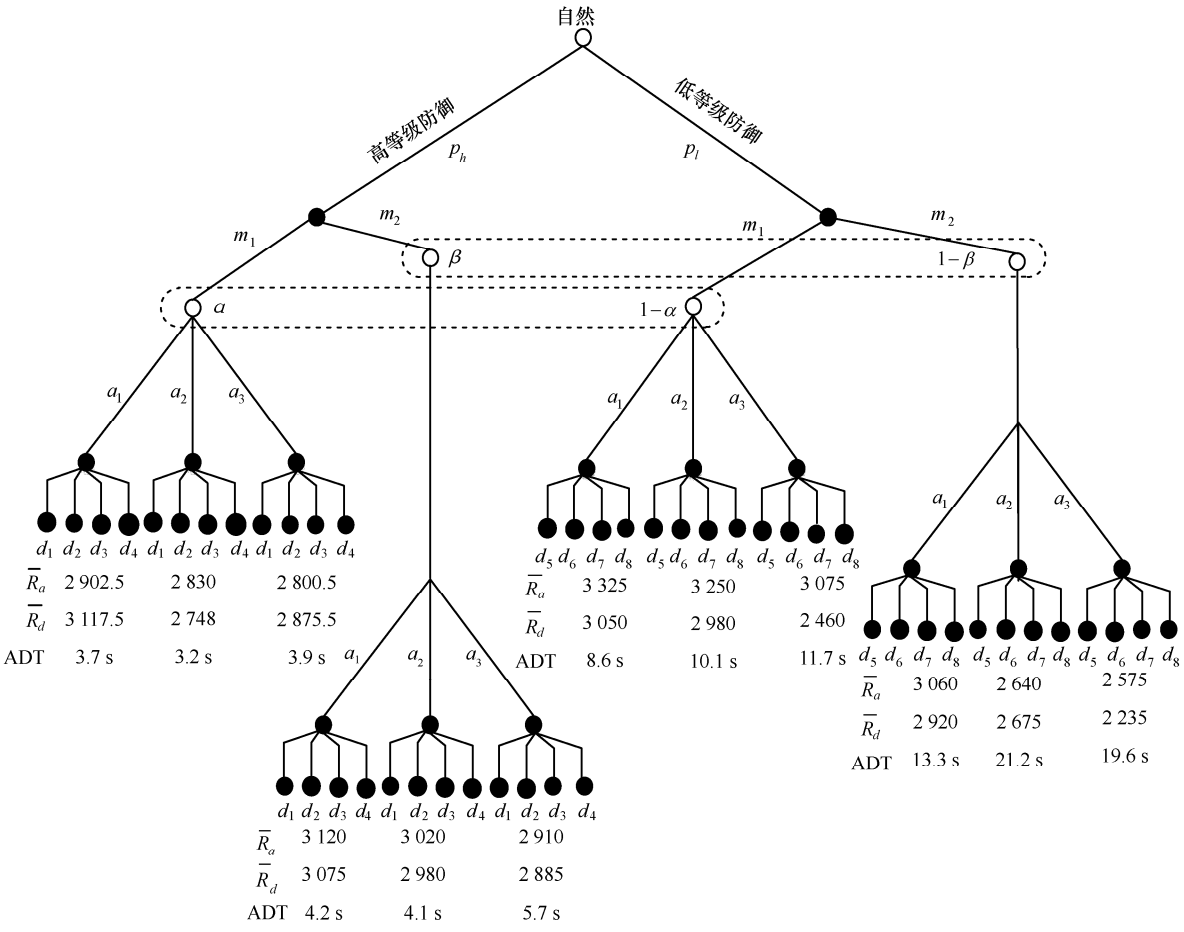


图 2　攻防博弈树

$$\arg\max_{a\in A}\sum_{t}\tilde{p}(t_d\mid m_1)R_a(t_d,m_1,a)=$$

$$\max\{\tilde{p}(t_d^1\mid m_1)U_a(t_1,m_1,a_1)+\tilde{p}(t_d^2\mid m_1)U_a(t_d^2,m_1,a_1),$$

$$\tilde{p}(t_d^1\mid m_1)U_a(t_d^1,m_1,a_2)+\tilde{p}(t_d^2\mid m_1)U_a(t_d^2,m_1,a_2),$$

$$\tilde{p}(t_d^1\mid m_1)U_a(t_d^1,m_1,a_3)+\tilde{p}(t_d^2\mid m_1)R_a(t_d^2,m_1,a_3)\}$$

根据上式，当 $\alpha\in[0,1]$， $a^*(m_1)=a_1$。

当 $m=m_2$ 时，同理有，当 $\beta\in[0,1]$，$a^*(m_2)=a_1$。

2)防御者推断的最优防御策略

$$m^*(t_d)\in\arg\max_{m\in M}U_d(t_d,a^*(m),m)$$

当 $t_d=t_d^1$ 时，有

$$m^*(t_d)\in\arg\max_{m\in M}U_d(t_d,a^*(m),m)=$$

$$\max\{U_d(t_d^1,a^*(m_1),m_1),U_d(t_d^1,a^*(m_2),m_2)\}$$

上式 $=\max\{U_d(t_d^1,a_1,m_1),U_d(t_d^1,a_1,m_2)\}$，即 $m^*(t_d^1)=m_1$。

当 $t_d=t_d^2$ 时，有

$$m^*(t_d)\in\arg\max_{m\in M}U_d(t_d,a^*(m),m)=$$

$$\max\{U_d(t_d^2,a^*(m_1),m_1),U_d(t_d^2,a^*(m_2),m_2)\}$$

上式 $=\max\{U_d(t_d^2,a_1,m_1),U_d(t_d^2,a_1,m_2)\}$，即 $m^*(t_d^2)=m_2$。

3) 防御类型的后验概率修正

从 1)和 2)可以得到此时的博弈均衡为 (t_d^1,m_1,a_1) 与 (t_d^2,m_2,a_1)，利用贝叶斯法则修正得到的后验概率为

$$\alpha=\tilde{p}(t_d^1\mid m_1)=\frac{p(m_1\mid t_d^1)p(t_d^1)}{p(m_1\mid t_d^1)p(t_d^1)+p(m_1\mid t_d^2)p(t_d^2)}=1$$

$$\beta=\tilde{p}(t_d^1\mid m_2)=\frac{p(m_2\mid t_d^1)p(t_d^1)}{p(m_2\mid t_d^1)p(t_d^1)+p(m_2\mid t_d^2)p(t_d^2)}=0$$

由精炼贝叶斯均衡的定义可知，(t_d^1,m_1,a_1) 与 (t_d^2,m_2,a_1) 为分离均衡。后验概率是由攻击者使用贝叶斯法则从先验概率、接收到的信号和防御者的最优防御策略得到的，该概率作为当前阶段的后验概率，同时作为下一阶段攻击者的先验概率。

将上述 2 个分离均衡表示为统一形式，即 PE：$[(t_d^1,t_d^2)\to(m_1,m_2)\to(a_1,a_1),\alpha=1,\beta=0]$，在此均衡下，当防御者采取高等级防御类型 t_d^1 时，释放高等级防御信号 m_1，攻击者采取攻击策略 a_1，防御者平均收益为 3 117.5，系统服务时延为 3.7 s；当防御者采取低等级防御类型 t_d^2 时，防御者释放低等级防御信号 m_2，攻击者采取攻击策略 a_1，防御者平均收益为 3 050，目标系统服务时延为 8.6 s。最优防御策略是防御者采用高等级防御行动，同时释放高防御信号，表明防御者利用信号机制可以主动向外界传递防御信息，对攻击者的攻击意图起到了威慑作用，从而减小或消除攻击带来的损失，一定程度上达到了主动防御的效果。

4.4 实验分析

仿真实验采用 Matlab2013a 实现了防御策略选取算法，根据所得到的实验数据，通过对攻防博弈过程、攻防收益和博弈均衡的一般性分析，可以发现结合 MTD 和信号博弈进行综合性主动防御的规律。

1) MTD 策略的成本普遍较高，由表 4 可知，策略成本普遍大于传统防御策略，但防御者采取 MTD 策略时防御收益普遍大于其他传统防御策略。在仿真实验中，选择高等级防御类型 t_d^1，释放高等级防御信号 m_2，防御收益分别为(3 075, 2 980, 2 885)；选择低等级防御类型 t_d^2，释放高等级防御信号 m_2，防御收益分别为(2 920,2 675,2 235)。这说明相较于被动防御方式，MTD 策略通过主动改变目标系统的不确定性，能够增加攻击者的攻击难度，更加有效地抵御攻击行为。

2) 由于攻击者对防御者类型的后验推断直接影响攻防博弈过程和均衡求解，防御者可利用信号机制直接影响博弈过程和均衡策略求解。攻击者根据先验概率、防御者释放的防御信号和防御者的最优策略使用贝叶斯法则对防御类型进行修正。因此利用信号机制防御者能够影响攻击者的后验推断的形成，提高防御者在攻防对抗过程中的主动性。

3) 防御者采取 MTD 策略同时结合信号博弈可以更加有效地增强主动防御效果。从防御策略特点和防御信号作用的角度分析，MTD 策略是防御者通过目标系统脆弱性的随机化、动态化和不确定化，使防御者在攻防过程中获得基于目标系统自身的防御主动性；信号博弈可以使防御者通过主动选择及释放针对性信号，在攻防信息获取和认知领域实现对攻击者的欺骗、迷惑，削弱攻击者的信息优势，提升主动防御能力，并能为移动目标防御策略的实施提供准备和反应时间。

4) 低等级防御者通过信号机制发送诱导信号可以增强防御效果，提高防御收益。由图 2 可知，当目

标系统处于低等级防御时，使用高等级防御信号，防御者获得的收益为(3 050,2 980,2 460)；而使用低等级防御信号时，防御收益为(2 920,2 675,2 235)。

综上，低等级防御者伪装成高等级防御者可以对攻击者起到威慑、迷惑的效果，攻击者无法清晰地认识防御者的虚实情况，导致攻击者出于自身利益，一般采取保守的试探攻击，在一定程度上能够起到主动防御的效果。

5 结束语

移动目标防御是一种改变攻防不对称、网络防御“被动挨打”格局的前沿性主动防御技术，在实际应用中具有良好效果和巨大潜力。为进一步提升主动防御能力，本文将移动目标防御与信号博弈相结合，从攻击面转换和探测面扩展的角度定义防御策略，在分析攻防对抗过程的基础上，构建了基于信号博弈的移动目标防御模型，设计了博弈均衡求解方法和最优防御策略选取算法，通过仿真实验和数据分析，验证了所提模型和方法的有效性，总结了结合信号博弈和移动目标防御方法实施综合性主动防御的特点规律。本文的研究成果为信息受限的动态攻防过程中增强移动目标防御的效能提供了有效的模型方法，并能够对防御策略的选取提供指导。

下一步工作主要包括从参数的灵敏度和多属性分析角度改进移动目标防御策略的量化计算方法，提升博弈收益计算准确性；针对多阶段连续性网络攻防过程，结合随机博弈和微分博弈开展研究，提高模型与方法的适用范围。

参考文献：

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

[2] JAJODIA S, GHOSH A K, SWARUP V, et al. Moving target defense: creating asymmetric uncertainty for cyber threats[M]. Berlin: Springer Science Business Media, 2011.

[3] 蔡桂林, 王宝生, 王天佐, 等. 移动目标防御技术研究进展[J]. 计算机研究与发展, 2016, 53(5): 968-987.
CAI G L, WANG B S, WANG T Z, et al. Research on development of moving target defense technology[J]. Journal of Computer Research and Development, 2016, 53(5): 968-987.

[4] 刘效武, 王慧强, 吕宏武, 等. 网络安全态势认知融合感控模型[J]. 软件学报, 2016, 27(8): 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(8): 2099-2114.

[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(1): 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(1): 43-49.

[6] MANADHATA P K. 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[C]//ACM Workshop on Moving Target Defense. ACM, 2017: 21-30.

[8] VADLAMUDI S G, SENGUPTA S, KAMBHAMPATI S, et al. Moving target defense for Web applications using Bayesian Stackelberg games[J]. Adaptive Agents and Multi-Agents Systems, 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] 张恒巍, 余定坤, 韩继红, 等. 信号博弈网络安全威胁评估方法[J]. 西安电子科技大学学报, 2016, 43(3): 137-143.
ZHANG H W, YU D K, HAN J H, 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, 等. 基于攻防信号博弈模型的防御策略选取方法[J]. 通信学报, 2016, 37(5): 51-61.
ZHANG H W, YU D K, HAN J H, et al. Defense policies selection method based on attack-defense signaling game model[J]. Journal on Communications, 2016, 37(5): 51-61.

[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] BENZEL T. A strategic plan for cyber security research and development [J]. IEEE Security & Privacy, 2015, 13(4): 3-5.

[14] FENG X, ZHENG Z, CANSEVER D. A signaling game model for moving target defense[C]//2017 IEEE Conference on Computer Communications. IEEE, 2017: 1-9.

[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] HUANG S R, ZHANG H W, WANG J, et al. Markov differential game

for network defense decision-making method[J]. IEEE Access, 2018: 39621-39634.

[17] 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.

[18] GORDON L, LOEB M, LUCYSHYN W, et al. Computer crime and security survey[C]//2014 Computer Security Institute. 2014: 11-34.

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

[20] LIN J Q, LIU P, JING J W. Using signaling games to model the multi-step attack-defense scenarios on confidentiality[J]. Security Lecture Notes in Computer Science, 2017, 39(6): 118-137.

[21] MALEKI H, VALIZADEH S, KOCH W, et al. Markov modeling of moving target defense games[C]//ACM Workshop on Moving Target Defense. ACM, 2018: 104-110.

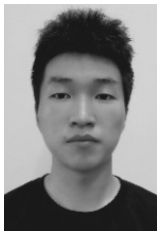
[22] ZHUANG R, BARDAS A G, DELOACH S A, et al. A theory of cyber attacks: a step towards analyzing MTD systems[C]//ACM Workshop on Moving Target Defense. ACM, 2017: 211-220.

[23] GAO X, ZHU Y F. Defense mechanism analysis based on signaling game model[C]//International Conference on Intelligent Human-Machine Systems and Cybernetics. IEEE, 2016: 414-417.

[24] FUDENBERG D, TIROLE J. Game theory [M]. Boston: Massachusetts Institute of Technology Press, 2012.

[25] ZHU Q, BAŞAR T. Game-theoretic approach to feedback-driven multi-stage moving target defense[C]//Decision and Game Theory for Security. Springer International Publishing, 2013: 246-263.

[作者简介]



蒋侣（1995− ），男，四川广安人，战略支援部队信息工程大学博士生，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（1978− ），男，河南洛阳人，博士，战略支援部队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。

王晋东（1966− ），男，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。