中国科学: 信息科学 2024 年 第 54 卷 第 1 期: 210–226
SCIENTIA SINICA Informationis
论文

基于前景理论的行为安全博弈

路石, 杨浩*, 姜斌

南京航空航天大学自动化学院, 南京 211106
* 通信作者. E-mail: haoyang@nuaa.edu.cn

收稿日期: 2023–01–02; 修回日期: 2023–02–12; 接受日期: 2023–04–04; 网络出版日期: 2024–01–15

国家自然科学基金 (批准号: 62073165, 62233009) 资助项目

摘要 本文基于前景理论在博弈的框架下研究了行为感知概率对网络化系统中攻防资源配置的影响. 首先, 在理性决策情形下, 通过融合物理系统、执行器攻击和资源分配模型, 构建了一个新的安全博弈框架. 其次, 针对决策者依靠感知概率而非真实概率进行主观决策的情形, 基于前景理论构造了行为安全博弈模型. 然后, 建立了安全博弈和行为安全博弈下期望收益函数关于攻防资源凹凸性的充要条件, 并深入分析和比较了安全博弈和行为安全博弈下攻防资源最优配置的存在性和唯一性以及行为概率对安全博弈的影响. 最后以无人机群为例进行了仿真, 验证了所提方法的有效性.

关键词 执行器攻击, 前景理论, 资源分配, 安全博弈, 行为安全博弈

1 引言

网络化系统由于其在实际生活中的广泛应用, 如智能电网、网络通信、无人机群等, 引起了广泛的关注. 由于网络系统信息交互的开放性和复杂性, 如何保证网络的安全运行是一个巨大的挑战 [1~3].

现代网络系统的特点是物理层、网络层和决策 (人) 层组成的层次结构 [4]. 不同类型的安全问题同样显现在这 3 个层面: 在物理层, 设备内部的故障会破坏系统的稳定性和性能 [5~9]; 在网络层, 各种网络化攻击, 如虚假数据注入 [10~12]、拒绝服务 [13,14] 和隐身攻击 [15], 可以通过操纵测量数据和控制输入命令, 严重损害系统性能和完整性; 在决策层, 社会偏好 [16]、前景理论 [17] 等工具表征了人的主观非理性, 展示人们往往主观地或者恶意地不选择使系统性能最优的决策.

攻防对抗是研究网络系统安全的一个关键问题. 安全博弈用于建模恶意破坏网络系统安全性的攻击者和防御它的个体之间的交互关系, 为研究攻防对抗提供一个分析框架. 不同类型的安全博弈模型被构建用来解决网络化系统的安全问题. 通过建模检测器与被检测的虚假数据注入攻击之间的对抗式交互关系为两人零和博弈设计检测器 [18], 这与经典的恶意行为检测和识别方法不同. 针对物理系统的

引用格式: 路石, 杨浩, 姜斌. 基于前景理论的行为安全博弈. 中国科学: 信息科学, 2024, 54: 210–226, doi: 10.1360/SSI-2023-0002
Lu S, Yang H, Jiang B. Prospect theory-based behavioral security game (in Chinese). Sci Sin Inform, 2024, 54: 210–226, doi: 10.1360/SSI-2023-0002

执行器被损害的情形, 零和博弈模型还可以用于设计执行器攻击下攻击方和防御方的最优策略[19]. 文献 [20] 利用斯塔克尔伯格博弈还解决了攻防对抗中的资源分配问题.

建立基于人的行为的安全博弈模型是寻求更加切合现实的均衡解的必要过程. 攻防过程均是由人操作和参与构成的一种对抗行为, 因此研究人的行为对攻防对抗的影响是十分必要的. 针对多个防御者协同对抗攻击的情形, 具有主观恶意行为的防御者可能会在某一时刻将最优防御策略切换到次优甚至最坏的防御策略从而损害整个系统的性能. 此外, 人的主观非理性是心理学中一个重要的研究分支, 其表征了人在不同场景下由于自身心理活动引起的不合理行为. 主观非理性行为在经济博弈论中已经被充分研究. 例如, 前景理论[17] 从实证研究出发, 从人的心理特质、行为特征揭示了影响选择行为的非理性心理因素, 表征了人具有损失规避和风险偏好的特征. 随机动态学习也是基于人的行为建立的一类学习行为, 揭示了噪声等因素的影响下, 人在学习的过程中具有选择非最优策略的可能性. 文献 [21,22] 基于前景理论初步建立了行为安全博弈模型, 探讨了主观非理性对决策过程的影响.

如何求解纳什 (Nash) 均衡是研究安全博弈和行为安全博弈中的一个关键问题. 直观的方法是对博弈的收益函数提供凸假设[22,23]. 凸性不仅保证了 Nash 均衡的存在, 而且提高了搜索 Nash 均衡的速度. 在凸假设下, 求解连续策略集中由 Karush-Kuhn-Tucker (KKT)[24] 条件构成的优化问题, 可以得到 Nash 均衡. 此外, 还可以使用梯度下降方法快速搜索 Nash 均衡, 适当的迭代参数设置可以加速趋向 Nash 均衡[25]. 在不提供凸假设的情况下, 一些复杂的搜索算法, 如神经网络学习[26]、粒子群优化[27]、强化学习[28], 可以通过遍历大量策略来搜索 Nash 均衡.

上述研究存在 3 个局限性: (a) 在资源配置问题中, 很少考虑三层网络系统下决策者的非理性行为; (b) 攻击者需要了解物理系统的实时状态信息, 而这些信息在资源有限的情况下往往很难获得[29]; (c) 资源的最优配置是基于凸假设得到的, 但这个假设在实际问题中是否成立需要验证.

基于上述观察结果, 本文旨在研究系统实时状态未知的情况下, 行为概率对网络化系统中安全资源配置问题的影响. 首先针对理性的攻击者和防御者构建安全博弈模型, 当攻击者和防御者具有主观非理性行为, 即高估低概率和低估高概率时, 将安全博弈扩展为行为安全博弈模型. 本文的主要贡献总结如下:

(1) 本文利用随机切换律的线性时不变切换系统建模了系统状态未知情形下的执行器攻击模型, 并结合感知决策层的预算分配模型, 构造了具有一般性期望收益函数的两方安全博弈框架, 展示了物理层以及网络层对决策层资源分配的影响机理, 为研究任意攻击下资源分配问题提供了有效的建模方案. 此外, 基于前景理论, 将安全博弈中建立的期望收益函数扩展为非线性加权函数建模攻防双方的主观非理性行为, 进而建立行为安全博弈下攻击方和防御方感知期望收益函数;

(2) 本文基于建立的安全博弈以及行为安全博弈模型建立了期望收益函数关于资源配置凹凸性的判别准则, 为研究其他类型攻击下期望收益函数关于资源分配凹凸性的准则提供了实用的研究方案. 此外, 本文基于凹凸性准则, 还深入分析了安全博弈和行为安全博弈下资源最优配置的存在唯一性以及行为概率对最优资源配置的影响.

本文的其余部分组织如下. 第 2 节描述了攻防框架. 第 3 节建立了安全博弈和行为安全博弈模型. 第 4 节分别讨论了安全博弈和行为安全博弈下资源最优配置的存在唯一性, 并分析了行为概率的影响. 第 5 节给出了仿真验证结果. 第 6 节展示了结论.

符号. $\mathbb{R}$ 表示实数的集合; $\mathbb{R}^+$ 表示正实数的集合; $\mathbb{R}^m$ 表示 m- 维实向量的集合; $\mathbb{R}^m_{\geqslant 0}$ 表示 m- 维非负实向量的集合; $\mathbb{R}^{m\times n}$ 表示 $m\times n$ 实矩阵的集合; $(\cdot)'$ 表示矩阵的转置; $\nabla_x(\cdot)$ 和 $\nabla^2_x(\cdot)$ 分别表示对变量 x 的一阶偏导数和二阶偏导数; $\dot{x}(t)$ 表示 x 对时间 t 的导数.

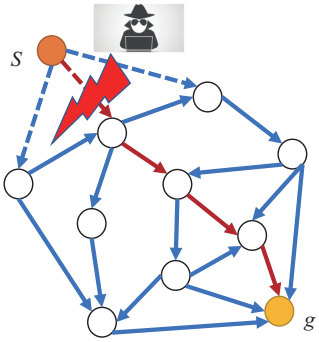

图 1 (网络版彩图) 攻击图概述

Figure 1 (Color online) Overview of attack graph

2 攻防框架

本节基于图论介绍了网络化系统中从源节点到目标的完整攻击过程, 以及执行器攻击对每个节点的破坏细节. 进一步给出了攻击和防御预算的分配模型.

2.1 攻击图

网络化系统的安全问题可以通过攻击图表征, 其描述了攻击者从源节点到目标对一系列脆弱节点进行攻击的过程 [22,30]. 考虑一个有向攻击图 $G=\{V,E\}$, 其由一组节点 V 和一组有向边 E 组成 (参见图 1). V 中的每个节点 i 代表一个资产, 其中它可以从节点 j 沿着 E 中的有向边 (j,i) 被攻击.

在攻击图 G 中, 构建了一个虚拟节点 s 作为攻击的起始, 它不属于网络的一部分, 即 $s\notin V$. 与 s 接触的节点 k 是脆弱的, 即沿着可行边 (s,k) 易被破坏. 节点 i $(i\neq s)$ 可以被攻击, 前提是与 i 接触的节点已被成功攻击, 即跳板攻击 [31]. 令 $\mathcal{P}_i$ 表示从源节点 s 到 i 的有向路径集, 其中路径 $\mathcal{P}\in\mathcal{P}_i$ 是边 $\{(s,k),\ldots,(j,i)\}$ 的集合. 如果攻击者沿着一条可行路径 $\mathcal{P}\in\mathcal{P}_g$ 成功地妥协了一系列节点, 目标 g 最终可以被攻击. 如果攻击在路径 $\mathcal{P}$ 上的任意中间节点上失败, 则目标 g 不能被成功攻击.

2.2 执行器攻击

采用具有随机切换信号的线性时不变切换系统描述了执行器攻击破坏节点 i 的过程:

$$\dot{x}_i(t)=A_ix_i(t)+B_{i,\delta(t)}u_i(t),\ x_i(t_{i,0})=x_{i0}, \tag{1}$$

其中 $x_i\in\mathbb{R}^{n_i}$ 表示系统状态, $u_i\in\mathbb{R}^{n_i}$ 是系统控制输入, $A_i\in\mathbb{R}^{n_i\times n_i}$ 和 $B_{i,\delta(t)}\in\mathbb{R}^{n_i\times m_i}$ 是定常矩阵. 切换律 $\delta(t):=[0,\infty)\to\mathcal{N}:=\{1,2\}$ 是分段常数且右连续的. 模式 $\delta(t)=1$ 表示在 t 时刻无执行器攻击被激活的子系统 1. 模式 $\delta(t)=2$ 表示在 t 时刻执行器攻击下被激活的子系统 2. 令 $\{t_{i,k}\}$ 表示 $\mathcal{N}$ 中 k 的切换时间序列. 切换时刻的状态跳满足

$$x_i(t_{i,k})=\alpha_{i,\delta(t_{i,k}^-)}x_i(t_{i,k}^-),\ x_i(t_{i,k}^-)=\lim_{h\to 0}x_i(t_{i,k}-h), \tag{2}$$

其中 $\alpha_{i,\delta(t_{i,k}^-)}\in\mathbb{R}^{n_i\times n_i}$ 是定常矩阵.

执行器攻击会在 $[t_{i,1},t_{i,2}]$ 的时间间隔内导致执行器部分失效, 此时控制输入 $u_i(t)$ 被破坏为 $\rho_iu_i(t)$, 其中定常矩阵 $\rho_i\in\mathbb{R}^{m_i\times m_i}$ 表示由攻击者控制的执行器攻击参数. 因此 $B_{i,2}=\rho_iB_{i,1}$. 状态跳 (2) 表示执行器攻击在 $t_{i,1}$ 处出现, 在 $t_{i,2}$ 处消失时导致的状态重置, 其中突变程度由矩阵 $\alpha_{i,\delta(t_{i,k}^-)}$ 表征.

假设 $(A_i, B_{i,1})$ 是可控的, 控制输入 $u_i(t)$ 被限制为由可测函数组成的一类可容许控制. 因此, 存在状态反馈增益矩阵 K_i, 使得矩阵 $\mathcal{A}_i := A_i + B_{i,1}K_i$ 为 Hurwitz. 当攻击在有限时间 $t_{i,2}$ 处消失时, 系统切换到子系统 1, 因此系统 (1) 在原点处渐近稳定.

2.3 攻击者和防御者预算的分配

系统 (1) 的代价函数由一般线性二次函数给出,

$$J_i(x(t)) := \frac{1}{2}\int_{t_{i,0}}^{\infty}[x_i'(t)Q_{\delta(t)}x_i(t) + u_i'(t)R_{\delta(t)}u_i(t)]\mathrm{d}t, \tag{3}$$

其中 $Q_{\delta(t)} \in \mathbb{R}^{n_{\delta(t)}\times n_{\delta(t)}}$ 是非负定对称矩阵, $R_{\delta(t)} \in \mathbb{R}^{m_{\delta(t)}\times m_{\delta(t)}}$ 是正定对称矩阵. 代价函数 (3) 用二次函数的积分形式来度量能耗 [32]. 在无执行器攻击的情况下, 每个节点 i 的代价满足

$$\hat{J}_i(x(t)) := \frac{1}{2}\int_{t_{i,0}}^{\infty}x_i'(t)\bar{Q}_1x_i(t)\mathrm{d}t,$$

其中 $\bar{Q}_1 := Q_1 + K_i'R_1K_i$. 在执行器攻击下, 系统 (1) 的运行时间可分为 $[t_{i,0}, t_{i,1})$, $[t_{i,1}, t_{i,2})$ 和 $[t_{i,2}, \infty)$ 三个阶段. 那么每个节点 i 的代价可以重写为

$$\tilde{J}_i(x(t), t_{i,1}, t_{i,2}) := \frac{1}{2}\int_{t_{i,0}}^{t_{i,1}}x_i'(t)\bar{Q}_1x_i(t)\mathrm{d}t + \frac{1}{2}\int_{t_{i,1}}^{t_{i,2}}x_i'(t)\bar{Q}_2x_i(t)\mathrm{d}t + \frac{1}{2}\int_{t_{i,2}}^{\infty}x_i'(t)\bar{Q}_1x_i(t)\mathrm{d}t, \tag{4}$$

其中 $\bar{Q}_2 := Q_2 + K_i'\rho_i'R_2\rho_iK_i$. 为了简单起见, 简写符号 $\tilde{J}_i(x(t), t_{i,1}, t_{i,2})$ 为 $\tilde{J}_i$. 假设防御者有一个有限的安全预算 $\bar{\Gamma} \in \mathbb{R}^+$, 其保证图 G 中除了虚拟节点 s 之外的所有节点的消耗. 令 $|V|$ 表示防御者负责的节点数. 所有节点的总代价不大于预算, 即 $\sum_{i=1}^{|V|}\tilde{J}_i \leqslant \bar{\Gamma}$. 定义集合 V 上的防御预算分配 $\psi := [\psi_1, \ldots, \psi_{|V|}]'$, 对角形矩阵 $\delta := \mathrm{diag}\{\sqrt{\hat{J}_1}, \ldots, \sqrt{\hat{J}_{|V|}}\}$. 令 $\bar{\Theta} := \{\psi \in \mathbb{R}_{\geqslant 0}^{|V|} | \mathbf{1}'\psi = \bar{\Gamma}, \psi' \psi I \geqslant \delta\}$ 表示可行资源分配集, 其包含集合 V 上所有可能的非负分配. 对任意 $i \in V$, 如果 $\varphi_i = \psi_i - \hat{J}_i \geqslant 0$, 则系统 (1) 的状态在无攻击时可以到达原点.

攻击者通过攻击一条可行路径 $\mathcal{P} \in \mathcal{P}_g$ 上的节点序列来妥协目标. 每次攻击的代价定义为

$$\bar{J}_i(x(t), t_{i,1}, t_{i,2}) := \frac{1}{2}\int_{t_{i,1}}^{t_{i,2}}x_i'(t)\bar{R}_ix_i(t)\mathrm{d}t, \tag{5}$$

其中 $\bar{R}_i \in \mathbb{R}^{m_i\times m_i}$ 是正定对称矩阵. 类似地, 简写符号 $\bar{J}_i(x(t), t_{i,1}, t_{i,2})$ 为 $\bar{J}_i$. 假设攻击者有限的预算为 $\Gamma \in \mathbb{R}^+$, 用来保证所有攻击的消耗. 定义包含在 $\mathcal{P}$ (虚拟节点 s 除外) 上的节点的攻击预算分配为 $\omega := [\omega_1, \ldots, \omega_{|\mathcal{P}|}]$, 其中 $|\mathcal{P}|$ 是路径 $\mathcal{P}$ (s 除外) 上的节点数. 则代价 $\bar{J}_i = \omega_i$. 令 $\Theta := \{\omega \in \mathbb{R}_{\geqslant 0}^{|\mathcal{P}|} | \mathbf{1}^{\mathrm{T}}\omega = \Gamma\}$ 表示攻击者的可行资源分配集, 它包含路径 $\mathcal{P}$ 上所有可能的非负分配. 注意, 攻击预算 Γ 只分配给攻击路径上包含的节点, 其他节点的攻击消耗为零.

3 安全博弈模型和行为安全博弈模型

本节建立了理性攻击者和理性防御者的两人安全博弈模型, 在此基础上进一步建立了行为概率加权下的行为安全博弈模型.

3.1 安全博弈模型

攻击各个节点的具体过程如下所示. 首先通过两个场景说明执行器攻击具有可行时间域的合理性. 第一, 目标只有进入数据可传输或武器发射范围后才会受到攻击. 第二, 任务在有限时间内完成. 因此, 执行器攻击在有限的间隔 $[t_0, T]$ 内发起, 才可以破坏任务, 其中常数 t_0 代表可行时间域的初始时刻. 根据攻击图, 所有攻击均遵循时间序列. 因此, 与源节点 s 接触的节点 1 将在 $[t_0, T]$ 的时间间隔内受到攻击. 假设攻击在 $t_{1,1}$ 时发起, 并在 $t_{1,2}$ 时结束, 节点 1 被成功攻破. 沿着一条可行路径, 接续的节点可以在 $[t_{1,2}, T]$ 时间间隔内被攻击. 通过重复上述过程, 可以在 $[t_{i-1,2}, T]$ 内攻击节点 i, 其中 $t_0 = t_{0,2} \leqslant t_{1,1} \leqslant t_{1,2} \leqslant \cdots \leqslant t_{i,1} \leqslant \cdots \leqslant T$.

如果无攻击时, 系统 (1) 是稳定的, 执行器最终将系统状态驱动到原点. 在执行器攻击下, 系统的状态 (1) 可能偏离原点或缓慢收敛到原点. 我们定义, 如果系统 (1) 的状态不能在分配的防御预算 ψ_i 下最终到达原点, 则对节点 i 的攻击是成功的. 否则, 节点 i 是安全的. 因此, 在节点 i 中, 当且仅当实际代价大于分配的预算, 即 $\tilde{J}_i > \psi_i$ 时, 攻击是成功的. 令 $\Phi_i := \{t_{i,1} | \tilde{J}_i(x(t), t_{i,1}, t_{i,2}) > \psi_i, t_{i-1,2} \leqslant t_{i,1} \leqslant \bar{t}_{i,1}, \bar{J}_i(x(t), \bar{t}_{i,1}, T) = \omega_i\}$ 表示节点 i 可被成功攻击的可行时域. 定义可行时域的长度为 $|\Phi_i|$. 为了计算此范围, 我们重新定义 $\Phi_i := \{t_{i,1} | \tilde{J}_i(x(t), t_{i,1}, t_{i,2}) \geqslant \psi_i, t_{i-1,2} \leqslant t_{i,1} \leqslant \bar{t}_{i,1}, \bar{J}_i(x(t), \bar{t}_{i,1}, T) = \omega_i\}$, 其中满足 $\tilde{J}_i = \psi_i$ 的有限点对区间长度没有影响. 由于在有限预算 ω_i 下, 攻击者很难检测到每个节点 i 的实时状态 $x_i(t)$, 因此其只在时间区间 $t_{i,1} \in [t_{i-1,2}, T]$ 内随机攻击节点 i 的执行器. 那么, 节点 i 以概率

$$p_i = (T - t_0 - t_{i-1,2})^{-1} |\Phi_i| \tag{6}$$

被成功地攻击.

整个攻击过程通过取图 G 中的一条完整路径 $\mathcal{P}$ 来说明. 在 $\mathcal{P}$ 中首先通过攻击与源节点 s 接触的节点 i 的执行器而使其受到损害. 只有在攻击成功的情况下, 与节点 i 接触的 $\mathcal{P}$ 中的下一个节点 j 才会被攻击. 通过重复上述过程, 直到 $\mathcal{P}$ 上除 g 外的所有节点都被成功入侵, 目标 g 才可以被攻击. 因此, 对于 $\mathcal{P}_g$ 中的每一条攻击路径 $\mathcal{P}$, 目标被成功攻击的概率为

$$h_{\mathcal{P}} := \prod_{i \in \mathcal{P}} p_i. \tag{7}$$

上式代表节点 i 沿边 (k, i) 被破坏的条件概率, 前提是节点 k 已经被成功破坏.

目标被成功攻破后, 攻击者获得收益为

$$\pi(\mathcal{P}) := \sum_{i \in \mathcal{P}} L_i, \tag{8}$$

其中已知常数 $L_i \in \mathbb{R}$ 表示成功妥协节点 i 后获得的收益. 例如, 在通信覆盖问题中, 攻击者可以通过成功地破坏一个节点来减少覆盖面积 L_i [33]. 最终, 攻击者或防御者的期望收益可以被定义为

$$R(\omega_{\mathcal{P}}, \psi) := h_{\mathcal{P}} \cdot \pi(\mathcal{P}), \tag{9}$$

其中 $\omega_{\mathcal{P}}$ 表示在路径 $\mathcal{P} \in \mathcal{P}_g$ 上的攻击预算分配.

攻击者通过在集合 $\mathcal{P}_g$ 中选择最优路径并合理分配预算来最大化期望收益 (9), 而防御者通过合理分配安全预算到每个节点来最小化损失 (9). 因此, 攻击和防御预算的最优分配问题可以描述为

$$\min_{\psi \in \bar{\Theta}} \max_{\mathcal{P} \in \mathcal{P}_g} \max_{\omega \in \Theta} R(\omega_{\mathcal{P}}, \psi). \tag{10}$$

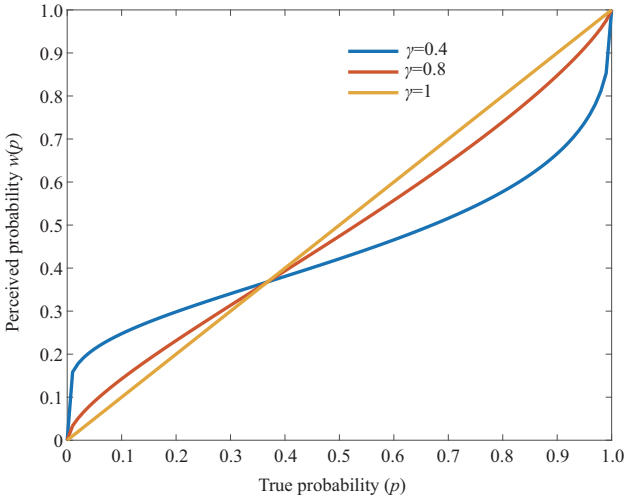

图 2 (网络版彩图) 非线性权重函数

Figure 2 (Color online) Nonlinear weighting function

这样的资源分配问题可以被建模为 (非行为) 安全博弈, 其中所有可行配置 Θ 和 $\bar{\Theta}$ 构成攻击者和防御者的策略集, 其纳什均衡 (NE) 定义如下 [34].

定义1 一个二元策略 $(\omega^*_{\mathcal{P}^*}, \psi^*)$ 构成安全博弈的 Nash 均衡如果这个二元策略组对任意 $\omega_{\mathcal{P}} \in \Theta, \psi \in \bar{\Theta}, \mathcal{P} \in \mathcal{P}_g$ 满足 $R(\omega_{\mathcal{P}}, \psi^*) \leqslant R(\omega^*_{\mathcal{P}^*}, \psi^*) \leqslant R(\omega^*_{\mathcal{P}^*}, \psi)$.

实际上, 攻击方可能会选择非最优路径, 防御方也可能由于非理性行为而不遵循最优分配. 接下来, 我们引入这种行为, 并通过行为安全博弈将其建模到资源分配问题中.

3.2 行为安全博弈模型

3.2.1 前景理论

前景理论揭示了人具有损失规避和风险偏好的特征, 而且通过高估低概率和低估高概率始终错误地感知真实概率 [17]. 更具体地说, 人类将真实的概率 $p \in [0,1]$ 感知为 $\theta(p) \in [0,1]$, 其中 $\theta(\cdot)$ 是一个概率加权函数. 一个常用的概率加权函数由 Prelec 在文献 [35] 中提出,

$$\theta(p) = \exp\left[-(-\ln(p))^{\gamma}\right],\ p \in [0,1], \tag{11}$$

其中参数 $\gamma \in (0,1]$ 表示概率敏感的程度. 非理性程度随着参数 γ 的减小而增大, 如图 2 所示. 如果 $\gamma = 1$, 那么对于任意 $p \in [0,1]$, 函数 $\theta(p) = p$, 其表征了理性下的策略更新.

接下来, 通过将概率加权函数 (11) 融合到 3.1 小节建立的安全博弈中来对行为安全博弈建模.

3.2.2 感知期望收益函数

攻击者和防御者主观上通过高估低概率和低估高概率来偏离真实的概率. 结合期望收益函数 (6) 和概率加权函数 (11), 可知每个节点 i 下攻击者和防御者的感知概率函数满足

$$\text{攻击者}:\ \bar{\theta}(p_i) = \exp\left[-(-\ln(p_i))^{\bar{\gamma}_i}\right],\ \text{防御者}:\ \tilde{\theta}(p_i) = \exp\left[-(-\ln(p_i))^{\tilde{\gamma}_i}\right], \tag{12}$$

其中参数 $\bar{\gamma}_i$ 和 $\tilde{\gamma}_i$ 分别表示攻击者和防御者的概率敏感程度. 注意, 攻击者和防御者对于成功攻破每个节点的真实概率可能有不同程度的主观心理.

攻击者感知到沿着路径 $\mathcal{P}\in\mathcal{P}_g$ 目标 g 以概率 $\bar{h}_{\mathcal{P}}:=\prod_{i\in\mathcal{P}}\bar{\theta}(p_i)$ 被成功攻击. 则攻击者和防御者的感知期望收益函数可以被定义为

$$\bar{R}(\omega_{\mathcal{P}},\psi):=\bar{h}_{\mathcal{P}}\cdot\pi(\mathcal{P})=\exp\left[-\sum_{i=1}^{|\mathcal{P}|}(-\ln(p_i))^{\bar{\gamma}_i}\right]\cdot\pi(\mathcal{P}),\tag{13}$$

$$\tilde{R}(\omega_{\mathcal{P}},\psi):=\tilde{h}_{\mathcal{P}}\cdot\pi(\mathcal{P})=\exp\left[-\sum_{i=1}^{|\mathcal{P}|}(-\ln(p_i))^{\tilde{\gamma}_i}\right]\cdot\pi(\mathcal{P}),\tag{14}$$

其中收益 $\pi(\mathcal{P})$ 仍然满足 (8). 特别地, 如果每个节点参数 $\bar{\gamma}_i$ 和 $\tilde{\gamma}_i$ 均等于 1, 则感知的期望收益函数 (13) 和 (14) 与安全博弈中的期望收益函数 (9) 相同.

在行为概率加权下, 攻击者根据感知期望收益函数 (13) 选择感知到的最优路径并分配预算, 而防御者则根据感知期望收益函数 (14) 分配预算. 此外, 攻击者和防御者可能感知到不同的最优路径. 行为概率加权下, 预算的最优配置问题可以描述为

$$\begin{aligned}&\text{(i)}\ \max_{\omega_{\mathcal{P}}\in\Theta,\mathcal{P}\in\mathcal{P}_g}\bar{R}(\omega_{\mathcal{P}},\bar{\psi}),\ \bar{\psi}\in\arg\min_{\psi\in\bar{\Theta}}\bar{R}(\omega_{\mathcal{P}},\psi),\\&\text{(ii)}\ \min_{\psi\in\bar{\Theta}}\tilde{R}(\bar{\omega}_{\bar{\mathcal{P}}},\psi),\ \bar{\omega}_{\bar{\mathcal{P}}}\in\arg\max_{\omega_{\mathcal{P}}\in\Theta,\mathcal{P}\in\mathcal{P}_g}\tilde{R}(\omega_{\mathcal{P}},\psi).\end{aligned}\tag{15}$$

这样的资源分配问题构成行为安全博弈. 由于攻击者和防御者不同的主观非理性行为, 安全博弈中的单目标优化问题 (10) 变成了行为安全博弈中的多目标优化问题 (15). 基于定义 1, 定义行为安全博弈中的 Nash 均衡 $(\bar{\omega}^*_{\bar{\mathcal{P}}^*},\bar{\psi}^*)$ 如下.

定义2 一个二元策略 $(\bar{\omega}^*_{\bar{\mathcal{P}}^*},\bar{\psi}^*)$ 构成行为安全博弈的 Nash 均衡如果这个二元策略组对任意 $\bar{\omega}_{\bar{\mathcal{P}}}\in\Theta,\psi\in\bar{\Theta},\bar{\mathcal{P}}\in\mathcal{P}_g$ 满足 $\bar{R}(\bar{\omega}_{\bar{\mathcal{P}}},\psi^*)\leqslant\bar{R}(\bar{\omega}^*_{\bar{\mathcal{P}}^*},\psi^*)\leqslant\bar{R}(\bar{\omega}^*_{\bar{\mathcal{P}}^*},\psi)$, 对任意 $\omega_{\mathcal{P}}\in\Theta,\bar{\psi}\in\bar{\Theta},\mathcal{P}\in\mathcal{P}_g$ 满足 $\tilde{R}(\omega_{\mathcal{P}},\bar{\psi}^*)\leqslant\tilde{R}(\omega^*_{\mathcal{P}^*},\bar{\psi}^*)\leqslant\tilde{R}(\omega^*_{\mathcal{P}^*},\bar{\psi})$.

与定义 1 不同, 行为安全博弈中攻击和防御资源的分配分别基于各自的感知期望收益函数. 这相当于分别基于函数 (13) 和 (14) 寻求 Nash 均衡 $(\bar{\omega}^*_{\bar{\mathcal{P}}^*},\psi^*)$ 和 $(\omega^*_{\mathcal{P}^*},\bar{\psi}^*)$, 最后把它们组合成行为安全博弈下的 Nash 均衡 $(\bar{\omega}^*_{\bar{\mathcal{P}}^*},\bar{\psi}^*)$.

基于期望收益函数 (9) 和感知的期望收益函数 (13) 和 (14), 本文将深入分析安全博弈和行为安全博弈下最优预算分配的存在性和唯一性, 其中将特别强调行为概率的影响.

4 安全博弈和行为安全博弈中最优配置的存在唯一性

本节分别分析了安全博弈和行为安全博弈中最优配置的存在性, 并进一步讨论了它们的唯一性.

4.1 安全博弈中最优配置的存在唯一性

首先说明了安全博弈中最优配置的存在性. 对于路径 $\mathcal{P}\in\mathcal{P}_g$, 固定分配 $\omega_{\mathcal{P}}$, 在闭区间 $\bar{\Theta}$ 内必然存在分配 $\bar{\psi}$ 使得收益函数 (9) 满足 $R(\omega_{\mathcal{P}},\bar{\psi})\leqslant R(\omega_{\mathcal{P}},\psi)$. 同样地, 固定分配 ψ, 在闭区间 Θ 内必然存在分配 $\bar{\omega}_{\mathcal{P}}$ 使得收益函数 (9) 满足 $R(\omega_{\mathcal{P}},\psi)\leqslant R(\bar{\omega}_{\mathcal{P}},\psi)$. 综合考虑, 必然存在 $\bar{\omega}_{\mathcal{P}}$ 和 $\bar{\psi}$, 使得收益函数 (9) 满足 $R(\omega_{\mathcal{P}},\bar{\psi})\leqslant R(\bar{\omega}_{\mathcal{P}},\bar{\psi})\leqslant R(\bar{\omega}_{\mathcal{P}},\psi)$. 通过遍历所有路径, 我们可以获得 Nash 均衡 $(\omega^*_{\mathcal{P}^*},\psi^*)$.

此外, 可能存在多个使得期望收益最大化的最优路径. 如果攻击者在 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 两个不完全重叠的路径上获得的不同收益满足 $R(\omega^*_{\mathcal{P}_1},\psi^*)<R(\omega^*_{\mathcal{P}_2},\psi^*)$, 防御者则从 $\mathcal{P}_1$ 上的节点转移出一部分预算,

并将其分散到 $\mathcal{P}_2$ 的节点上, 直到攻击者在这两条路径上获得相同的最大收益. 这降低了攻击成功概率, 从而减少了损失. 因此, Nash 均衡在安全博弈中可能不是唯一的.

接下来, 我们关注最优分配 $\omega^*_{\mathcal{P}^*}$ 和 ψ^* 在每条最优路径 $\mathcal{P}^*$ 上的唯一性. 在证明它之前, 首先需要提出保证期望收益函数 (9) 关于攻防预算分配凹凸性的充要条件.

基于代价函数 (4) 和 (5), 通过求解耦合方程 $\tilde{J}_i = \psi_i$ 和 $\bar{J}_i = \omega_i$, 可得到 Φ_i 的边界. 对于给定分配 ω_i 和 ψ_i, 可能存在多个可行时间间隔, 使得代价函数 (4) 满足 $\tilde{J}_i \geqslant \psi_i$. 令 $[t^j_{i,1}, t^{j+1}_{i,1}]$ 表示节点 i 可以被成功攻击的时间间隔, 其中 $j \in \Delta := \{0, 2, \ldots, 2h-2\}$, $h \in \mathbb{R}^+$ 表示可行的时间间隔个数, 时间 $t_{i-1,2} \leqslant t^0_{i,1}$, $t^{2h-1}_{i,1} \leqslant T - \bar{t}_{i,1}$. 如果代价函数在初始时刻 $t_{i,0}$ 满足 $\tilde{J}_i \geqslant \psi_i$, 则 $t^0_{i,1} = t_{i,0}$. 因此, 节点 i 被成功攻击的概率满足

$$p_i = (T - t_0 - t_{i-1,2})^{-1} \sum_{j\in\Delta} (t^{j+1}_{i,1} - t^j_{i,1}). \tag{16}$$

一些符号的定义如下. 令 $\alpha_{i,1}\alpha^{-1}_{i,2} = I$, 其中 $\alpha^{-1}_{i,2}$ 表示矩阵 $\alpha_{i,2}$ 的逆. 令 (1) $f_1(t) := x'_i(t)\bar{R}_i x_i(t)$; (2) $g_1(t) := x'_i(t)(\alpha^{-1'}_{i,2}\bar{Q}_1\alpha_{i,2} - \bar{Q}_2)x_i(t)$; (3) $H_1(t) := (f_1(t))^{-1}g_1(t)$; (4) $H_2(t_1, t_2) := (f_1(t_2))^{-2}g_1(t_1)$ $\cdot g_1(t_2)[(g_1(t_2))^{-1}g_2(t_2) - f_2(t_2)(f_1(t_2))^{-1}]$. 对于固定攻击预算 ω_i, 代价函数 (4) 关于 $t_{i,1}$ 的一阶偏微分满足 $\nabla_{t_{i,1}}\tilde{J}_i = g_1(t_{i,1}) - g_1(t_{i,2})f_1(t_{i,1})(f_1(t_{i,2}))^{-1}$, 代价函数 (4) 关于 $t_{i,1}$ 的二阶偏微分满足 $\nabla^2_{t_{i,1}}\tilde{J}_i = g_2(t_{i,1}) - g_2(t_{i,2})\left(f_1(t_{i,1})\right)^2\left(f_1(t_{i,2})\right)^{-2} - H_1(t_{i,2})\left[f_2(t_{i,1}) - \left(f_1(t_{i,1})\right)^2 f_2(t_{i,2})\left(f_1(t_{i,2})\right)^{-2}\right]$. 为了避免冲突, 如果攻击时刻在初始时间 $t_{i,0}$ 或结束时间 T, 代价函数 (4) 满足 $\tilde{J}_i > \psi_i$, 我们定义 $\nabla_{t_{i,0}}\tilde{J}_i = \nabla_T\tilde{J}_i = 0$, $\nabla^2_{t_{i,0}}\tilde{J}_i = \nabla^2_T\tilde{J}_i = 0$. 令示性函数 I_j 为

$$I_j = \begin{cases} 1, & j \text{ 是偶数}, \\ -1, & j \text{ 是奇数}. \end{cases}$$

定义 $f_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) := \sum_{j\in\Upsilon} I_j(\nabla^2_{t^j_{i,1}}\tilde{J}_i)(\nabla_{t^j_{i,1}}\tilde{J}_i)^{-3}$, $g_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) := \sum_{j\in\Upsilon} I_j(H_2(t^j_{i,1}, t^j_{i,2})(\nabla_{t^j_{i,1}}J_i)^{-2} - (H_1(t^j_{i,2}))^2(\nabla^2_{t^j_{i,1}}\tilde{J}_i)(\nabla_{t^j_{i,1}}\tilde{J}_i)^{-3})$, 其中 $\Upsilon := \{0, 1, \ldots, 2h-1\}$.

引理1 在安全博弈中, 期望收益函数 (9) 关于分配 ω 是严格凹的当且仅当系统 (1) 满足

$$\max_{t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}} g_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) < 0, \ \forall i \in V. \tag{17}$$

期望收益函数 (9) 关于分配 ψ 是严格凸的当且仅当系统 (1) 满足

$$\min_{t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}} f_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) > 0, \ \forall i \in V, \tag{18}$$

其中 $t^j_{i,1}, t^j_{i,2}$ 对于任意 $j, k \in \Upsilon$ 满足 $\tilde{J}(x(t), t^j_{i,1}, t^j_{i,2}) = \tilde{J}(x(t), t^k_{i,1}, t^k_{i,2})$.

证明 通过 Hessian 矩阵的正定或负定来寻求保证期望收益函数 (9) 关于攻击和防御预算分配凹凸性的充要条件. 首先证明期望收益函数 (9) 关于分配 ψ 的凸性. $R(\omega_{\mathcal{P}}, \psi)$ 关于分配 ψ 的二阶偏微分满足

$$\nabla^2_\psi R(\omega_{\mathcal{P}}, \psi) = \pi(\mathcal{P}) \cdot \begin{pmatrix} p_{-1}\nabla^2_{\psi_1}p_1 & 0 & \cdots & 0 \\ 0 & p_{-2}\nabla^2_{\psi_2}p_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{-|\mathcal{P}|}\nabla^2_{\psi_{\mathcal{P}}}p_{|\mathcal{P}|} \end{pmatrix},$$

其中 p_{-i} 表示除节点 i 外沿路径 $\mathcal{P}$ 的成功攻击概率. 因此, 期望收益函数 (9) 关于分配 ψ 是严格凸的当且仅当概率函数 p_i 对于任意 $i \in V$ 下分配 ψ_i 是严格凸的, 即

$$\nabla^2_{\psi_i} p_i > 0. \tag{19}$$

因此, 我们寻求保证 (19) 的充要条件. 根据式 (16), 可得

$$\nabla^2_{\psi_i} p_i = \left(\nabla^2_{\psi_i}\left(\sum_j I_j t^j_{i,1}\right)\right)(T - t_{i-1,2})^{-1} = \left(\sum_j I_j(\nabla^2_{\psi_i} t^j_{i,1})\right)(T - t_{i-1,2})^{-1}. \tag{20}$$

对于分配 ψ_i, 代价函数 (4) 在边界 $t^j_{i,1}$ 或 $t^{j+1}_{i,1}$ 满足 $\tilde{J}_i = \psi_i$. 由此得出 $\tilde{J}_i$ 对于分配 ψ_i 的一阶偏微分满足

$$\frac{1}{2} g_1(t^j_{i,1}) \nabla_{\psi_i} t^j_{i,1} - \frac{1}{2} g_1(t^j_{i,2}) \nabla_{t^j_{i,1}} t^j_{i,2} \nabla_{\psi_i} t^j_{i,1} = 1. \tag{21}$$

而 $\tilde{J}_i$ 关于时间 $t_{i,1}$ 的一阶微分满足

$$\nabla_{t_{i,1}} \tilde{J}_i = \frac{1}{2} g_1(t_{i,1}) - \frac{1}{2} g_1(t_{i,2}) \nabla_{t^j_{i,1}} t^j_{i,2}. \tag{22}$$

此外, 对于固定分配 ω_i, 代价函数 (5) 满足 $\bar{J}_i(x(t)) = \omega_i$. 可得 $t^j_{i,2}$ 关于 $t^j_{i,1}$ 的一阶偏微分为

$$\nabla_{t^j_{i,1}} t^j_{i,2} = f_1(t^j_{i,1})(f_1(t^j_{i,2}))^{-1}. \tag{23}$$

基于式 (21), (22) 和 (23), 可得

$$\nabla_{\psi_i} t^j_{i,1} = \frac{1}{2}(\nabla_{t^j_{i,1}} \tilde{J}_i)^{-1}. \tag{24}$$

进而, $t^j_{i,1}$ 对 ψ_i 的二阶偏微分满足

$$\nabla^2_{\psi_i} t^j_{i,1} = -\frac{1}{4}(\nabla_{t^j_{i,1}} \tilde{J}_i)^{-3} \nabla^2_{t^j_{i,1}} \tilde{J}_i. \tag{25}$$

通过将式 (25) 代入式 (20), p_i 关于分配 ψ_i 的二阶偏微分为

$$\nabla^2_{\psi_i} p_i = \frac{1}{4} f(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2})(T - t_{i-1,2})^{-1}. \tag{26}$$

因此, 函数 p_i 对于分配 ψ_i 是严格凸的, 即 $\nabla^2_{\psi_i} p_i > 0$, 当且仅当 $f_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) > 0$. 而期望收益函数 (9) 对于分配 ψ 是严格凸的当且仅当 $f_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) > 0$ 对于任意 $i \in V$. 通过相同的方法, 同样可知期望收益函数 (9) 关于分配 ω 是严格凹的当且仅当 $g_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) < 0$ 对于任意 $i \in V$.

引理 1 在不假设概率函数 $h_{\mathcal{P}}$ 关于预算分配凹凸性的情况下, 给出了验证它们的充要条件. 条件 (17) 和 (18) 可以通过求解以下优化问题来验证:

$$\begin{aligned}
&\min_{t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}} f_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) \\
&\max_{t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}} g_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) \\
\text{s.t.}\ \ &\tilde{J}(x(t), t^j_{i,1}, t^j_{i,2}) = \tilde{J}(x(t), t^k_{i,1}, t^k_{i,2}),\ \forall j, k \in \Upsilon, \\
&0 \leqslant t^0_{i,1} \leqslant t^1_{i,1} \leqslant \cdots \leqslant t^{2h-1}_{i,1} \leqslant T,\ t^j_{i,1} \leqslant t^j_{i,2}.
\end{aligned} \tag{27}$$

该条件是变量受限下的最优值求解问题, 可通过拉格朗日乘子法转换为非受限下的最优值求解问题. 特别地, 因为条件 (17) 和 (18) 是关于资源分配凹凸性的条件, 系统参数的变化仅仅影响攻击时间域以及相应的攻击成功概率, 对上述结果无影响. 接下来, 基于引理 1, 证明了最优配置的唯一性.

定理1 在安全博弈中, 对任意节点 $i \in V$, 如果系统 (1) 满足条件 (17) 和 (18), 那么最优分配 $\omega^*_{\mathcal{P}^*}$ 和 ψ^* 是唯一的.

证明 根据 1.3 小节的描述, 每个节点的预算 ψ_i 是由正常消耗 $\hat{J}_i(x(t))$ 和防御消耗 φ_i 构成的. 消耗 $\hat{J}_i(x(t))$ 由方程 (3) 决定, 并且是固定的. 下面仅证明防御代价 φ_i 的唯一性. 首先证明每个具有非零分配 φ_i 的节点必属于最优路径. 假设存在一个非零分配 φ_i 的节点不属于最优路径, 那么防御者将从 φ_i 中转移出一部分预算分配到最优攻击路线上的节点, 从而降低攻击成功的概率. 因此每个具有非零分配 φ_i 的节点必须属于最优路径. 此外, 在每条最优路径上至少存在一个非零分配 φ_i 的节点. 接下来, 利用期望收益函数 (9) 的凹凸性证明最优解的唯一性. 假设安全博弈沿最优路径 $\mathcal{P}^*$ 有两个不同的纳什均衡, 分别定义为 $(\omega^1_{\mathcal{P}^*}, \psi^1)$ 和 $(\omega^2_{\mathcal{P}^*}, \psi^2)$. 根据上述声明, 如果期望收益函数 (9) 关于 ψ 是严格凸的, 则存在分配 ψ^3 使得

$$\begin{aligned}\ln R(\omega^3_{\mathcal{P}^*}, \psi^3) &\leqslant \alpha \ln R(\omega^3_{\mathcal{P}^*}, \psi^1) + (1-\alpha)\ln R(\omega^3_{\mathcal{P}^*}, \psi^2)\\ &< \alpha \ln R(\omega^1_{\mathcal{P}^*}, \psi^1) + (1-\alpha)\ln R(\omega^2_{\mathcal{P}^*}, \psi^2),\end{aligned}$$

其中 $\alpha \in (0,1)$. 这意味着分配 ψ^3 在最优路径 $\mathcal{P}^*$ 上优于 ψ^1 和 ψ^2, 这与上面的假设相矛盾. 另外, 这个结果在每个最优路径上都是有效的. 因此, 最优分配 ψ^* 对于防御者来说是唯一的. 此外, 根据 (9) 对于攻击预算的严格凹性, 我们也可以用同样的方法证明最优分配 $\omega^*_{\mathcal{P}^*}$ 是唯一的.

注释1 保证资源分配的凹凸性对求解 Nash 均衡具有重要的意义. 首先, 凹凸性可以保证最优资源分配的唯一性, 如果最优分配是不唯一的, 存在攻击者和防御者的策略在多个最优解之间不断切换的可能性, 从而影响系统的稳定性, 因此凹凸性具有保证稳定性的作用; 其次, 利用凹凸性对于求解 Nash 均衡具有巨大的助力. 如果条件 (17) 和 (18) 被满足, 期望收益函数 (9) 关于资源分配的凹凸性得到保证, 最优资源分配唯一, 此时可以通过先假定最优路径, 然后通过 KKT 条件求解最优分配方案, 或者通过梯度下降的方法去搜索 Nash 均衡. 此外, 凹凸性还影响寻找 Nash 均衡的速度. 如果凹凸性无法得到保证, 则全局最优解可能不唯一, 且可能存在局部最优解, 此时不能利用 KKT 条件直接求得全局最优解, 而且梯度下降算法极有可能陷入局部最优解.

4.2 行为安全博弈中最优配置的存在唯一性

类似于安全博弈, 分配 $(\bar{\omega}^*_{\mathcal{P}^*}, \psi^*)$ 和 $(\omega^*_{\mathcal{P}^*}, \bar{\psi}^*)$ 可以通过遍历所有可行路径单独寻找. 因此, 行为安全博弈的 Nash 均衡 $(\bar{\omega}^*_{\mathcal{P}^*}, \bar{\psi}^*)$ 也是存在的. 为了在行为安全博弈中寻求最优分配, 首先建立了收益函数 $\bar{R}(\omega_{\mathcal{P}}, \psi)$ 关于分配 ω 严格凹和收益函数 $\tilde{R}(\omega_{\mathcal{P}}, \psi)$ 关于分配 ψ 严格凸的充要条件. 定义 $\bar{f}_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}, t) := f_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) + \bar{e}(p_i)(\sum_{j\in\Upsilon} I_j(\nabla_{t^j_{i,1}}\tilde{J}_i)^{-1})^2(T-t)^{-1}$, $\bar{g}_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}, t) := g_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}) + \bar{e}(p_i)(\sum_{j\in\Upsilon} I_j H_1(t^j_{i,2})(\nabla_{t^j_{i,1}}\tilde{J}_i)^{-1})^2(T-t)^{-1}$, 其中 $\bar{e}(p_i) := (1-\bar{\gamma}_i)(-\ln(p_i))^{-1} + \bar{\gamma}_i(-\ln(p_i)^{\bar{\gamma}_i - 1})$, $\tilde{e}(p_i) := (1-\tilde{\gamma}_i)(-\ln(p_i))^{-1} + \tilde{\gamma}_i(-\ln(p_i)^{\tilde{\gamma}_i - 1})$.

引理2 在行为安全博弈中, 期望收益函数 (13) 关于分配 ω 是严格凹的当且仅当系统 (1) 满足

$$\max_{t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}, t} \bar{g}_i(t^0_{i,1}, t^0_{i,2}, t^1_{i,1}, \ldots, t^{2h-1}_{i,2}, t) < 0,\ \forall i \in V. \tag{28}$$

期望收益函数 (14) 关于分配 ψ 是严格凸的当且仅当系统 (1) 满足

$$\min_{t^0_{i,1},t^0_{i,2},t^1_{i,1},\ldots,t^{2h-1}_{i,2},t} \bar{f}_i(t^0_{i,1},t^0_{i,2},t^1_{i,1},\ldots,t^{2h-1}_{i,2},t)>0,\ \forall i\in V, \tag{29}$$

其中 $t^j_{i,1},t^j_{i,2}$ 对于任意 $j,k\in\Upsilon$ 满足 $\tilde{J}(x(t),t^j_{i,1},t^j_{i,2})=\tilde{J}(x(t),t^k_{i,1},t^k_{i,2})$.

证明 通过与引理 1 相同的分析, 可知概率函数 $\bar{h}_{\mathcal{P}}$ 对于分配 ω 是严格凹的当且仅当概率函数 $\bar{\theta}(p_i)$ 对于分配 ω_i 是严格凹的. 令 $\bar{h}(p_i):=-(-\ln(p_i))^{\bar{\gamma}}$. 对于概率函数 $\bar{\theta}(p_i)$, 可知

$$\nabla^2_{\omega_i}\bar{\theta}(p_i)=(\nabla^2_{\omega_i}\bar{h}(p_i)+(\nabla_{\omega_i}\bar{h}(p_i))^2)\exp(\bar{h}(p_i)).$$

因此, 概率函数 $\bar{\theta}(p_i)$ 对于分配 ω_i 是严格凹的当且仅当 $\nabla^2_{\omega_i}\bar{h}(p_i)+(\nabla_{\omega_i}\bar{h}(p_i))^2<0$. 对于概率函数 $\bar{h}(p_i))$, 可知

$$\nabla_{\omega_i}\bar{h}(p_i)=\bar{\gamma}(-\ln(p_i))^{\bar{\gamma}-1}\nabla_{\omega_i}p_i,$$

$$\nabla^2_{\omega_i}\bar{h}(p_i)=\bar{\gamma}(-\ln(p_i))^{\bar{\gamma}-1}[(1-\bar{\gamma})(-\ln(p_i))^{-1}(\nabla_{\omega_i}p_i)^2+\nabla^2_{\omega_i}p_i].$$

由此可得 $\nabla^2_{\omega_i}\bar{h}(p_i)+(\nabla_{\omega_i}\bar{h}(p_i))^2=\bar{\gamma}(-\ln(p_i))^{\bar{\gamma}-1}[\bar{e}(p_i)(\nabla_{\omega_i}p_i)^2+\nabla^2_{\omega_i}p_i]$. 根据引理 1 的证明, 进一步得到 $\nabla^2_{\omega_i}\bar{h}(p_i)+(\nabla_{\omega_i}\bar{h}(p_i))^2=\bar{\gamma}(-\ln(p_i))^{\bar{\gamma}-1}(T-t_{i-1,2})^{-1}\bar{g}_i(t^0_{i,1},t^0_{i,2},t^1_{i,1},\ldots,t^{2h-1}_{i,2},t)$. 因此, 函数 $\bar{\theta}(p_i)$ 对于分配 ω_i 是严格凹的当且仅当 $\bar{g}_i(t^0_{i,1},t^0_{i,2},t^1_{i,1},\ldots,t^{2h-1}_{i,2},t)<0$ 对任意 $t^j_{i,1},t^j_{i,2},\ j\in\Upsilon$. 进一步, 概率函数 $\bar{h}_{\mathcal{P}}$ 对于分配 ω 是严格凹的当且仅当条件 (28) 被满足. 通过同样的方法, 可知概率函数 $\tilde{h}(p_i)$ 关于分配 ψ 是严格凸的当且仅当条件 (29) 被满足.

条件 (28) 和 (29) 可以通过解决类似于 (27) 的优化问题来验证.

定理2 在行为安全博弈中, 对于任意 $i\in V$, 如果系统 (1) 满足条件 (28) 和 (29), 则最优分配 $\bar{\omega}^*_{\mathcal{P}^*}$ 和 $\bar{\psi}^*$ 是唯一的.

定理 2 可以利用与定理 1 相同的方法证明, 因此在此省略. 不管感知到的最优路径是否相同, 条件 (28) 和 (29) 可以保证最优分配 $(\bar{\omega}^*_{\mathcal{P}^*},\bar{\psi}^*)$ 的唯一性. 通过与安全博弈比较, 概率敏感性的影响主要体现为以下几个方面:

- 行为感知概率可能导致博弈框架由零和博弈转变为非零和博弈.
- 行为感知概率可能导致攻击者和防御者不选择真实的 Nash 均衡.
- 行为感知概率导致凹凸条件依赖于前一次攻击的结束时间.
- 如果最优分配 ψ^* 在安全博弈中是唯一的, 那么最优分配 $\bar{\psi}^*$ 在行为安全博弈中是唯一的. 如果最优分配 $\bar{\omega}^*_{\mathcal{P}^*}$ 在行为安全博弈中是唯一的, 那么最优分配 $\omega^*_{\mathcal{P}^*}$ 在安全博弈中是唯一的.

当安全博弈和行为安全博弈下的凹凸性条件被满足时, 可通过 KKT 条件 [24] 或者一些搜索算法寻找 Nash 均衡 [26~28], 由于篇幅限制, 不在本文一一赘述.

注释2 在实际工程应用中, 大多数决策问题需要及时迅速得到 Nash 均衡解, 但是本文研究的资源分配问题不需要实时决策, 在攻防开始前, 资源已经按照预设的模型合适地分配到各个节点, 在攻防的过程中不涉及资源分配策略的演化, 因此对求解 Nash 的实时性要求很低. 另外, 本文研究的结果主要涉及最优分配的唯一性结果. 如果最优分配是唯一的, 求解 Nash 均衡的复杂度会极大地被降低. 例如, 在利用 KKT 条件求解 Nash 均衡的过程中, 先假设路径 $\mathcal{P}_1$ 为最优路径, 由上述模型可知, 与其不相交的路径均为最优路径, 因此不需要遍历所有的可行路径寻找 Nash 均衡解. 令 Γ 表示与任意

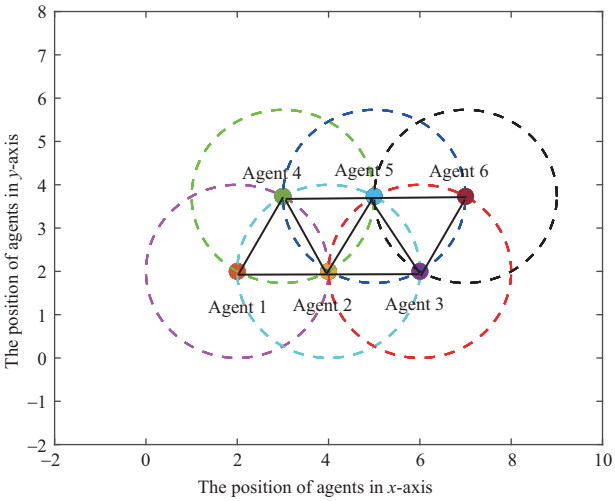

图 3 (网络版彩图) 无人机编队

Figure 3 (Color online) Formation of unmanned aerial vehicles

路径具有交集 (即除源节点和目标外具有共同的节点) 的路径集合, 则安全博弈下利用 KKT 算法求解 Nash 均衡的复杂度最大为 $\mathcal{O}(|\Gamma|)+1$, 而行为安全博弈下算法的复杂度最大为 $2\mathcal{O}(|\Gamma|)+2$. 当凹凸性条件不被满足时, 则需要搜索的时间较长且难以确定. 此外, 搜索 Nash 均衡的时间还与攻击图中的所有可行的路径数量有关, 随着节点数量的增加, 可行路径的数量越大, 搜索到可行解的时间越长.

注释3 在信息不完备或者噪声等因素的干扰下, 敌对方具有惯性决策的可能性, 其沿袭上一时刻的策略不改变. 但是防御方是在攻击开始前根据自己感知的攻击成功概率选择资源最优配置方案, 因此攻击方的惯性决策并不会影响防御方资源的最优配置方案. 理性情形下, 当防御者按照最优方案布置安全资源后, 防御者更加乐于见到攻击者由于各种因素作出惯性决策, 这会使得防御者的损失更低. 但是主观非理性情形下攻击者作出惯性策略可能会导致防御者更大的损失. 由于资源的配置是一次性的, 所以在攻击过程中发生惯性决策等情形时, 防御者无法利用资源重置的方法应对这个现象, 但是通过网络的快速拓扑重构等手段在损失一定利益的基础上可以有效地阻止这类情形.

注释4 本文的结论可拓展到类似情形下构造凹凸性判据的问题, 基于本文的研究过程拓展的主要流程如下: 首先建模攻击模型, 结合系统模型以及决策模型给出安全博弈或者行为安全博弈的收益函数模型, 进而利用本文的分析方法研究收益函数关于资源分配凹凸性的条件以及最优资源配置的存在唯一性.

5 案例研究: 无人机编队问题

本节利用无人机蜂群验证了所提出的结果. 如图 3 所示, 蜂群由 6 个标记为 $1,\ldots,6$ 的无人机组成. 这些无人机构成了一个分布式网络通信系统, 其中每架无人机只与其相邻的无人机通信. 假设每架无人机的最大通信距离 $d_{\max}=2$. 为了实现最大的通信覆盖, 相邻无人机的距离满足 $d_{ij}=d_{\max}=2$, $j\in\mathcal{N}_i$. 这个蜂群的任务是移动到指定的位置, 并再次组成相同的编队.

考虑每架无人机的二阶模型如下:

$$\begin{cases} \dot{x}_i = x_i, \\ \dot{v}_{x_i} = a_{x_i}, \end{cases} \quad \begin{cases} \dot{y}_i = y_i, \\ \dot{v}_{y_i} = a_{y_i}, \end{cases} \tag{30}$$

其中 (x_i, y_i), (v_{x_i}, v_{y_i}) 和 (a_{x_i}, a_{y_i}) 表示 $\mathbb{R}^2$ 平面上 x 和 y 方向的位置、速度和加速度. 令 $z_i = [x_i, y_i, x_i, y_i]$, $u_i = [a_{x_i}, a_{y_i}]$. 系统 (30) 可以重写为

$$\dot{z}_i = A_i z_i + B_i u_i, \tag{31}$$

其中

$$A_i = \begin{Bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{Bmatrix}, \ B_i = \begin{Bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{Bmatrix}.$$

线性状态反馈控制器设计为 $u_i = K_i z_i$, 其中增益矩阵 K_i 为

$$K_i = -\begin{Bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{Bmatrix}. \tag{32}$$

由于矩阵 $A_i + B_i K_i$ 为 Hurwitz, 在控制器 u_i 的作用下, 系统状态收敛到原点. 我们假定每个无人机基于其目标为原点的独立的 $\mathbb{R}^2$- 平面测量自己的状态 z_i, 因此所有无人机具有相同的初始状态 $z_i(t_{i,0}) = (2, 0, 2, 0)$.

攻击者的目的是破坏蜂群的通信覆盖. 在飞行过程中, 通过攻击无人机的执行器使其无法达到指定位置, 而与相邻无人机断开链路, 则认定攻击是成功的. 假定攻击的起始位置为无人机 1, 终端位置是无人机 6. 攻击者可以沿不同的路径依次攻击无人机 1~6, 使得通信覆盖范围尽可能减少. 此外, 假设攻击者的能量预算为 4, 防御方的能量总预算为 42, 其中分配给每架无人机的初始预算为 6.5, 以保证其正常移动的能量消耗, 剩余防御预算额外分配给部分无人机, 使其在受到攻击时能够到达目的地. 根据代价函数 (3), 可知 $\hat{J}_i(x(t)) < 6.5$. 因此在没有攻击的情况下, 每架无人机都可以在该分配下到达原点, 并重新组成相同的编队.

假设 Q_i, R_i 和 $\bar{R}_i$ 是具有适当维数的单位矩阵. 在执行器攻击下, 将虚假数据注入到每个无人机中, 使无人机突然接收到虚假的位置和速度信息, 其中状态 z_i 的突变程度由参数 α_i 来测量

$$\alpha_i = \begin{Bmatrix} 0.9 & 0 & 0 & 0 \\ 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0.9 & 0 \\ 0 & 0 & 0 & 0.9 \end{Bmatrix}. \tag{33}$$

此外, 攻击矩阵满足

$$\rho_1 = \rho_4 = \rho_5 = \rho_6 = \begin{Bmatrix} 0.4 & 0.5 \\ -0.1 & 0.1 \end{Bmatrix}, \ \rho_2 = \rho_3 = \begin{Bmatrix} 0.1 & -0.1 \\ 0.5 & 0.4 \end{Bmatrix}. \tag{34}$$

表 1 不同路径下的最优分配

Table 1 Optimal allocations under different paths

Path	Payoff	Probability	Expected payoff	The optimal allocation of attack budgets	The optimal allocation of defense budgets
$1\to2\to3\to6$	10^3	0.26^4	4.57	$\omega_1,\omega_2,\omega_3,\omega_6=1$	$\psi_1,\ldots,\psi_6=7$
$1\to4\to5\to6$	10^3	0.26^4	4.57	$\omega_1,\omega_4,\omega_5,\omega_6=1$	$\psi_1,\ldots,\psi_6=7$
$1\to2\to4\to3\to5\to6$	10^3	0.18^6	0.034	$\omega_1,\ldots,\omega_6=2/3$	$\psi_1,\ldots,\psi_6=7$
$1\to2\to4\to3\to5\to6$	10^6	0.18^6	34	$\omega_1,\ldots,\omega_6=2/3$	$\psi_1,\ldots,\psi_6=7$

攻击需要在有限的时间内进行, 以避免无人机到达原点后再次构成编队. 为便于计算, 假设每个无人机具有相同的可被攻击时域 T. 令 $T=5$, 因此攻击者在 $[0,5]$ 的时间间隔内持续攻击各个无人机执行器. 接下来, 两种情况被分别讨论: (i) 当 6 号无人机无法接收到其他无人机的通信时, 蜂群固定损失为 L; (ii) 随着被孤立的无人机数量的增加, 蜂群的损失越大. 基于这两种情况, 分别考虑静态收益和动态收益.

情形一 (静态收益): 从表 1 中可以看出, 如果攻击者在任意路径上只获得一个常数收益 $L=10^3$, 则攻击者选择最短路径从无人机 1 攻击到无人机 6, 即 $1\to2\to3\to6$, $1\to4\to5\to6$ 或者 $1\to2\to5\to6$. 与其他路径相比, 攻击者可以在这些路径上以最大的概率攻击无人机 6, 并获得最大的期望收益. 通过检验, 在上述参数下满足条件 (17) 和 (18), 因此最优分配是唯一的. 其分配方案是将攻击预算平均分配给所选攻击路径上的无人机, 并将防御预算平均分配给每架无人机. 图 4 展示了在攻击路径 $1\to2\to3\to6$ 上无人机的飞行轨迹和无人机 6 被成功破坏后的编队构型, 其中一些相邻无人机之间的距离超过上限 2.

情形二 (动态收益): 如果攻击者随着成功入侵无人机数量的增加而获得更高的收益, 则攻击者可能会选择最长的路径以获得最大的预期收益, 即使第 6 架无人机在这条路径上以最低概率被成功入侵. 与收益不变的情况相比, 表 1 显示, 当在 $1\to2\to4\to3\to5\to6$ 路径上收益 $\pi(\mathcal{P})=10^6$ 时, 攻击者在该路径上取得的期望收益最大.

其次, 考虑攻击方具有非理性, 而防御方是理性的. 由于防御方不具有行为概率, 因此防御预算的分配始终保持不变. 对于常数收益 $L=10^3$, 如果行为概率加权参数满足 $\bar{\gamma}_1=\cdots=\bar{\gamma}_6=0.3$, 攻击者仍然在三条最短路径中选择一条路径攻击 6 号无人机. 通过检验, 在上述参数下满足条件 (28) 和 (29), 则最优分配存在且唯一. 安全博弈和行为安全博弈在攻防预算的最优分配上没有差异. 如果行为概率加权参数满足 $\bar{\gamma}_1=\bar{\gamma}_2=\bar{\gamma}_3=\bar{\gamma}_6=0.3$ 和 $\bar{\gamma}_4=\bar{\gamma}_5=0.6$, 分别讨论如下两种情况: (i) 时间 $T=5$. 攻击者只选择感知到的最优路径 $1\to2\to3\to6$. 由于攻击者在最优路径上的每个节点上的概率敏感性相同, 因此攻击预算的分配也保持不变. 最优分配方案见表 2; (ii) 时间 $T=2$. 由于 $p\geqslant\exp(-1)$, 函数 $\theta(p)$ 是关于 γ 的增函数, 因此攻击者会选择感知到的最优路径 $1\to4\to5\to6$, 在该路径下 Nash 均衡是存在的, 如表 2 所示. 通过比较这两种情况, 说明时间边界 T 对资源分配有影响.

上述结果展示了行为概率对最优资源分配的影响. 当概率敏感度满足上述给定参数时, 期望收益函数满足条件 (28) 和 (29), 因此关于任意防御和攻击资源的分配具有凹凸性, 此时利用 KKT 条件或者梯度下降方法可快速得到最优解. 此外, 当概率参数满足 $\bar{\gamma}_1=0.1$, $\bar{\gamma}_2=\bar{\gamma}_3=0.2$, $\bar{\gamma}_4=\bar{\gamma}_5=0.4$ 和 $\bar{\gamma}_6=0.6$ 时, 根据定理 2, 可知条件 (28) 和 (29) 不满足, 此时全局最优解不能通过 KKT 条件直接

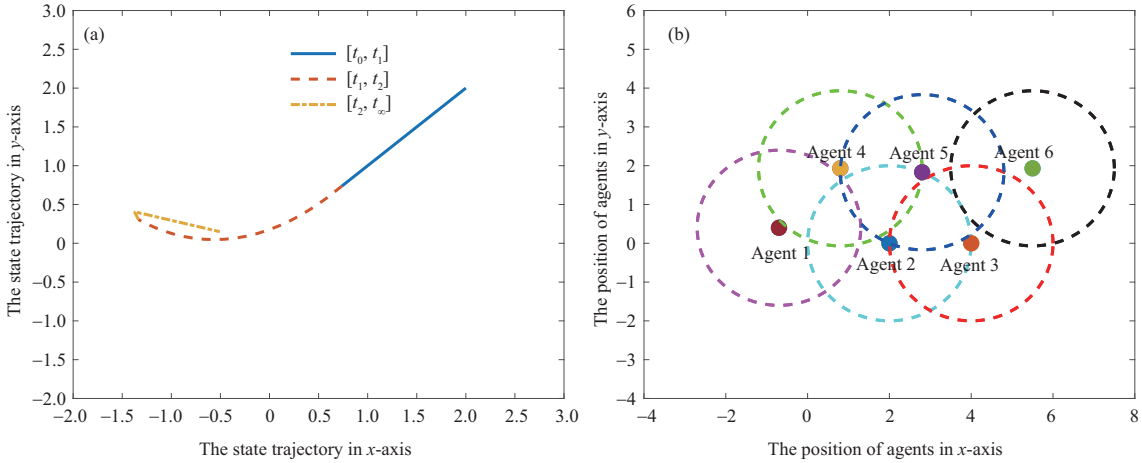

图 4　(网络版彩图) 蜂群沿路径 $1 \to 2 \to 3 \to 6$ 被攻击后无人机 1 的轨迹 (a) 和编队构型 (b)

Figure 4　(Color online) Agent 1's trajectory (a) and the destroyed formation (b) when the drones are attacked along the path $1 \to 2 \to 3 \to 6$

表 2　不同时间界下的最优分配

Table 2　Optimal allocations under different time bounds

Time T	Probability sensitivity	The optimal path	Probability	Perceived probability	Expected payoff	Perceived expected payoff	The optimal allocation of attack budgets	The optimal allocation of defense budgets
5	$\gamma_1, \ldots, \gamma_6 = 0.3$	$1 \to 2 \to 3 \to 6$, $1 \to 4 \to 5 \to 6$, $1 \to 2 \to 5 \to 6$	0.26^4	0.335^4	4.57	12.59	$\omega_1, \omega_2, \omega_3, \omega_6 = 1$, $\omega_1, \omega_4, \omega_5, \omega_6 = 1$, $\omega_1, \omega_2, \omega_5, \omega_6 = 1$,	$\psi_1, \ldots, \psi_6 = 7$
	$\gamma_1, \gamma_2, \gamma_3, \gamma_6 = 0.3$, $\gamma_4, \gamma_5 = 0.6$	$1 \to 2 \to 3 \to 6$	0.26^4	0.335^4	4.57	12.59	$\omega_1, \omega_2, \omega_3, \omega_6 = 1$	$\psi_1, \ldots, \psi_6 = 7$
2	$\gamma_1, \ldots, \gamma_6 = 0.3$	$1 \to 2 \to 3 \to 6$, $1 \to 4 \to 5 \to 6$, $1 \to 2 \to 5 \to 6$	0.65^4	0.46^4	178.5	44.77	$\omega_1, \omega_2, \omega_3, \omega_6 = 1$, $\omega_1, \omega_4, \omega_5, \omega_6 = 1$, $\omega_1, \omega_2, \omega_5, \omega_6 = 1$,	$\psi_1, \ldots, \psi_6 = 7$
	$\gamma_1, \gamma_2, \gamma_3, \gamma_6 = 0.3$, $\gamma_4, \gamma_5 = 0.6$	$1 \to 4 \to 5 \to 6$	0.61^4	0.44^4	127.18	39.2	$\omega_1, \omega_6 = 1.04$, $\omega_4, \omega_5 = 0.96$	$\psi_1, \ldots, \psi_6 = 7$

得到, 需要通过梯度下降算法分别从不同初值出发搜索局部最优资源分配方案, 进而得到全局最优资源分配方案, 此时则需要耗费较长的搜索时间.

6　结论

本文通过构造安全博弈研究了网络化系统中的攻防资源分配问题, 并利用前景理论建立行为安全博弈进一步讨论了行为感知概率对最优分配的影响, 其中行为感知概率不仅可能会降低攻击者的最大期望收益, 而且还可能改变最优配置的唯一性. 一个有趣的未来研究是设计调控机制, 使行为决策者的真实收益恢复到理性情形下的最优水平.

参考文献

1　Chen X, Makki K, Yen K, et al. Sensor network security: a survey. IEEE Commun Surv Tutorials, 2009, 11: 52–73
2　Humayed A, Lin J, Li F, et al. Cyber-physical systems security—a survey. IEEE Internet Things J, 2017, 4: 1802–1831
3　Abdelkader M, Güler S, Jaleel H, et al. Aerial swarms: recent applications and challenges. Curr Robot Rep, 2021, 2: 309–320

4 Zhu Q Y, Başar T. Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: games-in-games principle for optimal cross-layer resilient control systems. IEEE Control Syst Mag, 2015, 35: 46–65
5 Blanke M, Kinnaert M, Staroswiecki M, et al. Diagnosis and Fault-Tolerant Control. Berlin: Springer, 2006
6 Zhang Y M, Jiang J. Bibliographical review on reconfigurable fault-tolerant control systems. Annu Rev Control, 2008, 32: 229–252
7 Panagi P, Polycarpou M M. Distributed fault accommodation for a class of interconnected nonlinear systems with partial communication. IEEE Trans Automat Contr, 2011, 56: 2962–2967
8 Yang H, Jiang B, Staroswiecki M, et al. Fault recoverability and fault tolerant control for a class of interconnected nonlinear systems. Automatica, 2015, 54: 49–55
9 Yang H, Han Q L, Ge X, et al. Fault-tolerant cooperative control of multiagent systems: a survey of trends and methodologies. IEEE Trans Ind Inf, 2020, 16: 4–17
10 Jin X, Haddad W M, Yucelen T. An adaptive control architecture for mitigating sensor and actuator attacks in cyber-physical systems. IEEE Trans Automat Contr, 2017, 62: 6058–6064
11 Wu G Y, Wang G, Sun J, et al. Optimal partial feedback attacks in cyber-physical power systems. IEEE Trans Automat Contr, 2020, 65: 3919–3926
12 Ye D, Zhang T Y. Summation detector for false data-injection attack in cyber-physical systems. IEEE Trans Cybern, 2020, 50: 2338–2345
13 Zhang X M, Han Q L, Ge X, et al. Resilient control design based on a sampled-data model for a class of networked control systems under denial-of-service attacks. IEEE Trans Cybern, 2020, 50: 3616–3626
14 Xu W Y, Hu G Q, Ho D W C, et al. Distributed secure cooperative control under denial-of-service attacks from multiple adversaries. IEEE Trans Cybern, 2020, 50: 3458–3467
15 Zhang Q R, Liu K, Xia Y Q, et al. Optimal stealthy deception attack against cyber-physical systems. IEEE Trans Cybern, 2020, 50: 3963–3972
16 Zhao Z, Lu H, Cai D, et al. User preference learning for online social recommendation. IEEE Trans Knowl Data Eng, 2016, 28: 2522–2534
17 Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. In: Handbook of the Fundamentals of Financial Decision Making: Part I. Singapore: World Scientific, 2013. 99–127
18 Zhang R C, Venkitasubramaniam P. False data injection and detection in LQG systems: a game theoretic approach. IEEE Trans Control Netw Syst, 2019, 7: 338–348
19 Wu C W, Li X L, Pan W, et al. Zero-sum game-based optimal secure control under actuator attacks. IEEE Trans Automat Contr, 2021, 66: 3773–3780
20 Li Y Z, Shi D W, Chen T W. False data injection attacks on networked control systems: a Stackelberg game analysis. IEEE Trans Automat Contr, 2018, 63: 3503–3509
21 Sanjab A, Saad W, Basar T. A game of drones: cyber-physical security of time-critical UAV applications with cumulative prospect theory perceptions and valuations. IEEE Trans Commun, 2020, 68: 6990–7006
22 Abdallah M, Naghizadeh P, Hota A R, et al. Behavioral and game-theoretic security investments in interdependent systems modeled by attack graphs. IEEE Trans Control Netw Syst, 2020, 7: 1585–1596
23 Zhu M, Frazzoli E. Distributed robust adaptive equilibrium computation for generalized convex games. Automatica, 2016, 63: 82–91
24 Avriel M. Nonlinear Programming: Analysis and Methods. New York: Dover Publications, 2003
25 Ye M J, Hu G Q. Distributed Nash equilibrium seeking by a consensus based approach. IEEE Trans Automat Contr, 2017, 62: 4811–4818
26 Li S, He J B, Li Y M, et al. Distributed recurrent neural networks for cooperative control of manipulators: a game-theoretic perspective. IEEE Trans Neural Netw Learn Syst, 2017, 28: 415–426
27 Zhao C H, Guo D H. Particle swarm optimization algorithm with self-organizing mapping for Nash equilibrium strategy in application of multiobjective optimization. IEEE Trans Neural Netw Learn Syst, 2021, 32: 5179–5193
28 Nguyen T T, Nguyen N D, Nahavandi S. Deep reinforcement learning for multiagent systems: a review of challenges, solutions, and applications. IEEE Trans Cybern, 2020, 50: 3826–3839
29 Kim J, Tong L, Thomas R J. Subspace methods for data attack on state estimation: a data driven approach. IEEE Trans Signal Process, 2014, 63: 1102–1114

30 Homer J, Zhang S, Ou X M, et al. Aggregating vulnerability metrics in enterprise networks using attack graphs. J Comput Security, 2013, 21: 561–597
31 Zonouz S, Rogers K M, Berthier R, et al. SCPSE: security-oriented cyber-physical state estimation for power grid critical infrastructures. IEEE Trans Smart Grid, 2012, 3: 1790–1799
32 Lewis F L, Vrabie D, Syrmos V L. Optimal Control. Hoboken: John Wiley & Sons, 2012
33 Lalropuia K C, Gupta V. A Bayesian game model and network availability model for small cells under denial of service (DoS) attack in 5G wireless communication network. Wireless Netw, 2020, 26: 557–572
34 Başar T, Olsder G J. Dynamic Noncooperative Game Theory. Philadelphia: SIAM, 1998
35 Prelec D. The probability weighting function. Econometrica, 1998, 66: 497–527

Prospect theory-based behavioral security game

Shi LU, Hao YANG* & Bin JIANG

College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106*, China*
* Corresponding author. E-mail: haoyang@nuaa.edu.cn

Abstract This paper studies the influence of behavioral perceived probability on the allocation of attack and defense resources in networked systems based on prospect theory under a game framework. For rational decision-makers, a new security game framework is constructed by integrating the models of the physical system, network attack, and resource allocation. Then, for irrational decision-makers whose decisions depend on perceived rather than true probability, a behavioral security game model is established using prospect theory. Furthermore, in security and behavioral security games, a sufficient and necessary condition is provided to guarantee the convexity and concavity of the expected payoff function with respect to attack and defense resources. Based on this condition, the existence and uniqueness of the optimal allocation of attack and defense resources are analyzed in depth and compared for security and behavioral security games. Furthermore, the effect of behavioral probability on security game is presented. Finally, a simulation example of an unmanned aerial vehicle swarm is taken to demonstrate the effectiveness of the proposed results.

Keywords actuator attacks, prospect theory, resource allocation, security game, behavioral security game