随机斯塔克尔伯格非对称安全博弈中的零行列式策略

Zhaoyang Cheng*Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Beijing, 100190, ChinaSchool of Mathematical Sciences, University of Chinese Academy of Sciences,Beijing, 100049, China

Guanpu Chen†JD Explore Academy, JD.com Inc, Beijing, 100176, China

Yiguang Hong‡
Department of Control Science and Engineering, Tongji University, Shanghai, 201804, China
Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, 210201, China

arXiv:2301.02543v1 [cs.GT] 5 Jan 2023

在随机斯塔克尔伯格非对称安全博弈中，强斯塔克尔伯格均衡（SSE）策略是防御者面对采用最佳响应（BR）策略的攻击者时，获取最高效用的一种常用选择。然而，攻击者可能是有限理性玩家，会结合采用最佳响应策略和一种固定的顽固策略。在这种情况下，由于顽固策略的影响，SSE策略可能无法维持防御性能。本文重点研究防御者如何采用单边控制的零行列式（ZD）策略来应对有限理性攻击者。首先，我们验证了防御者的零行列式策略存在。随后，我们研究防御者的零行列式策略对抗有限理性攻击者时的表现，并与SSE策略进行比较。具体而言，当攻击者的策略接近最佳响应策略时，与SSE策略相比，零行列式策略使防御者承受的损失有界。相反，当攻击者的策略接近顽固策略时，零行列式策略能为防御者带来比SSE策略更高的防御性能。

一、引言

随机安全博弈在信息物理系统（CPS）、无人机（UAV）和移动目标防御（MTD）等众多领域中越来越受到关注 [1–4]。随机斯塔克尔伯格非对称安全博弈是刻画防御者面临攻击者持续性威胁时参与者行为的重要类别之一。作为 [5, 6], 中讨论的一种基础模型，攻击者倾向于在观察防御者的策略后选择最佳响应（BR）策略，而防御者则旨在考虑攻击者的因素以最大化自身效用。实际上，防御者作为领导者，在引导攻击者决策方面具有优势，并根据对攻击者最佳响应策略的预测选择最优策略。相应的均衡被定义为强斯塔克尔伯格均衡（SSE）[5, 7, 8]。

然而，受主观或客观因素影响，参与者在实际中未必总是完全理性 [9, 10] 。典型情况是，有限理性攻击者在实际中不会严格采用最佳响应策略。这可能是由于攻击者的观测受限、环境干扰，或攻击者的模仿行为所致。例如，在移动目标防御问题中，攻击者可能因管理员（防御者）的干扰而无法直接观测到特定的防御策略 [2, 11]。在无人机系统中，恶意无人机可能因野外障碍物的阻挡而无法观测到合法无人机（防御者）的位置或飞行姿态 [3]。在信息物理系统中，攻击者可能会设计隐蔽攻击方案，而非采用最佳响应策略，以躲避防御者的故障检测 [1, 12]。

当有限理性攻击者失去实现最佳响应策略的能力或兴趣时，在大多数情况下，它很可能会转而采用固定顽固策略。例如，在信息物理系统安全[13] 中，玩家倾向于采用顽固策略，以避免被诱导至不满意的结果；在移动目标防御问题 [14]中，若玩家无法及时计算最佳响应策略，则可能选择固定的可信策略。面对顽固攻击者，SSE策略不再被视为防御者的最优解。事实上，有限理性攻击者可能采用由最佳响应策略和顽固策略构成的混合策略，这类有限理性玩家在安全问题中十分常见。例如，在信息物理系统中，攻击者需要考虑自身数据采集系统的失效概率，以避免潜在损失，因此会在采用顽固策略和作为跟随者行动之间犹豫不决 [15]。在无人机安全问题中，无人机在不同阶段也面临不同选择：穿越森林等复杂地形时，它可能会失去防御者的位置，而在平原上则能完全观察到防御者[3]。因此，潜在偏好、内在认知和可用资源等各种因素，使有限理性攻击者成为防御者不可忽视的对象。

显然，由于有限理性攻击者具有固执的一面，原有的 SSE 策略可能不再适用于防御者。因此，

* chengzhaoyang@amss.ac.cn
† chengp@amss.ac.cn
‡ yghong@iss.ac.cn

考虑其他策略来帮助防御者维持其防御性能至关重要。幸运的是，零行列式（ZD）策略提供了一种强有力的方法，使一方能够单方面强制参与者的期望效用满足某种有利关系，而不论对手选择何种策略。 [16] 在迭代囚徒困境（IPD）中提出了这一策略，零行列式（ZD）策略意味着一方能够单方面强制两位参与者的期望效用遵循线性关系。此后，各种 ZD 策略已得到广泛研究，旨在促进合作或在公共物品博弈（PGG）、人机交互（HCI）、演化博弈等领域单方面实施勒索 [17–23]。

此外，非对称矩阵博弈比对称博弈更贴近现实，而由于偏好不同，解决非对称博弈也面临一些挑战 [24, 25]。目前，在对称博弈中应用零行列式策略的突破还不多。例如，一些研究采用零行列式策略，在迭代数据交易困境博弈中说服服务提供商进行合作 [26]，并将其部署为物联网设备中的特殊主动防御策略 [27]。鉴于非对称博弈在安全领域的普遍性和重要性，有必要探索安全场景下零行列式策略在非对称博弈中的表现，因为针对 IPD 中零行列式策略的原始分析无法直接应用于非对称安全博弈。

受此启发，本文旨在揭示在随机斯塔克尔伯格非对称安全博弈中，防御者能否采用零行列式策略对抗有限理性攻击者，以弥补 SSE策略的不足。为此，我们表明，零行列式策略的表现优于 SSE策略。本文的主要贡献总结如下。

- 我们在非对称安全博弈中应用ZD策略。我们验证了ZD策略的存在性，以确保防御者能够采用ZD策略。此外，针对两种特殊攻击者，我们考察了ZD策略相较于SSE策略的防御表现。具体而言，面对采用最佳响应策略的攻击者，ZD策略相较于SSE策略会带来有界的效用损失；而面对顽固攻击者时，ZD策略表现良好，为防御者带来的效用高于SSE策略。

- 我们进一步分析有界理性的攻击者采用混合策略的一般情形。这一扩展也得到了类似的可接受结果。当攻击者的策略接近最佳响应策略时，我们为防御者提供适当的零行列式策略，使其防御性能相较于SSE策略的损失保持有界，同时节省计算资源。此外，当攻击者接近顽固攻击者时，我们展示了适用于防御者的零行列式策略，使其防御性能高于SSE策略。

- 我们通过两项实验验证了研究结果：为防御者提供适当的零行列式策略，并将其与SSE策略 [5]进行比较。首先，我们展示了其在移动目标防御问题中的表现。在这类问题中，有界理性的攻击者可以直接观察防御者的策略，并推导出明确的最佳响应策略 [2, 11]。随后，我们展示了其在信息物理系统问题中的表现。与所考虑的移动目标防御问题相比，这类问题的设定更复杂，但也更贴近实际：攻击者只能观察玩家行动历史，并根据虚拟博弈和Q-learning等特定机制计算最佳响应策略[7, 28]。

II. 随机斯塔克尔伯格非对称安全博弈

众所周知，在具有上一阶段记忆的随机非对称安全博弈中，攻击者旨在入侵两个目标，而防御者则在每个阶段阻止攻击 [2, 29]。考虑随机斯塔克尔伯格非对称安全博弈 $\mathcal{G} = \{\mathcal{S}, N, \mathcal{D}, \mathcal{A}, r, P\}$。$\mathcal{S} = \{11, 12, 21, 22\}$ 是状态集合，由先前的攻击目标和防御目标组成。 $N = \{d, a\}$是参与者集合。 $\mathcal{D} = \{1, 2\}$ 和 $\mathcal{A} = \{1, 2\}$ 分别是防御者的行动集合和攻击者的行动集合。$\mathbf{r} = \{r_d, r_a\}$ 是参与者的奖励集合，其中 $r_i : \mathcal{D} \times \mathcal{A} \to \mathbb{R}, i \in N$。此外， P: $\mathcal{S} \times \mathcal{S} \times \mathcal{D} \times \mathcal{A} \to [0, 1]$是转移函数，其中 $P(s'|s, d, a)$表示参与者采取 d, a时，从当前状态 s 转移到下一状态 $s' \in \mathcal{S}$ 的概率，而$\sum P(s'|s, d, a) = 1_{s' \in \mathcal{S}}$

对于 $s \in \mathcal{S}, d \in \mathcal{D}$ 和 $a \in \mathcal{A}$。

在这个安全博弈中，由于状态表示前一轮参与者的行动，因此，对于任意 $s \in \mathcal{S}$， $P(s'|s, d, a) = 1$ 当且仅当$s' = (da)$。因此，下一状态取决于参与者的策略和当前状态。为方便起见，将 $P(s'|s)$ 记为从状态 s转移到状态s' 的状态转移概率，其中 $s', s \in \mathcal{S}$。此外，在博弈$\mathcal{G}$中，每个玩家的策略取决于当前状态，这也是一种记忆一策略。防御者的策略是一个概率分布 π_d，其中 $\pi_d(d|s) \in \Delta\mathcal{D}$， $\Delta\mathcal{D}$ 表示定义在空间 $\mathcal{D}$上的概率单纯形。类似地，攻击者的策略为π_a，其中 $\pi_a(a|s) \in \Delta\mathcal{A}$。因此， $P(s, s') = \pi_d(d|s)\pi_a(a|s)$，其中 $s' = (da)$。令 $M = \{P(s|s')\}_{s,s' \in \mathcal{S}}$ 为该安全博弈的状态转移矩阵。如[16, 30], 中所述，我们继续研究正则矩阵 M。

表 I. 效用矩阵

		攻击者	
		1	2
防御者	1	(U_{11}^d, U_{11}^a)	(U_{12}^d, U_{12}^a)
	2	(U_{21}^d, U_{21}^a)	(U_{22}^d, U_{22}^a)

在 $\mathcal{G}$的阶段 t，每个玩家观察当前状态s_t，并根据其策略采取行动。防御者选择行动 $d_t \in \mathcal{D}$，而攻击者选择行动 $a_t \in \mathcal{A}$。防御者在阶段 t的奖励记为 $r_d(d_t, a_t) = U^d_{d_t a_t}$，其中 $U^d_{d_t a_t}$表示防御者保护目标 d_t而攻击者入侵目标 a_t时防御者的效用。类似地，攻击者在阶段 t 的奖励记为 $r_a(d_t, a_t) = U^a_{d_t a_t}$。每个阶段的效用矩阵见表 I。

重复安全博弈中的预期长期效用记为

$$U_d(\pi_d, \pi_a) = \mathbb{E}\left(\lim_{T\to\infty} \sum_{t=0}^{T} \frac{r_d(d_t, a_t)}{T}\right),$$
$$U_a(\pi_d, \pi_a) = \mathbb{E}\left(\lim_{T\to\infty} \sum_{t=0}^{T} \frac{r_a(d_t, a_t)}{T}\right),$$

其中，$\{d_t \sim \pi_d(\cdot|s_t)\}_{t\geqslant 0}$、$\{a_t \sim \pi_a(\cdot|s_t)\}_{t\geqslant 0}$和$\{s_t \sim P(\cdot|s_{t-1}, d_{t-1}, a_{t-1})\}_{t>0}$描述了每个阶段状态与行动的演化。此外，$s_0$ 是从 $\mathcal{S}$中随机采样得到的初始状态；由于 M 收敛，因此对于任意 s_0，每位玩家的预期效用都相同。

假设1 效用满足$\min\{U^d_{11}, U^d_{22}\} > \max\{U^d_{12}, U^d_{21}\}$。此外，$U^a_{11} < U^a_{12}$，且 $U^a_{22} < U^a_{21}$。

与 IPD [16, 31], 不同，这种安全博弈中的非对称性源于实际的安全机制。具体而言，防御者倾向于抵御攻击，即保护脆弱目标；攻击者则倾向于入侵未受保护的目标。上述情况代表了安全场景中的一大类非对称博弈，可概括为假设1。文献中已对各种安全博弈开展了大量相关研究 [6, 32, 33]。

III. 有限理性的攻击者

在随机斯塔克尔伯格非对称安全博弈中，防御者是领导者，会预先宣布策略；攻击者则是跟随者，在观测到防御者的策略后选择自己的策略。在大多数情况下，攻击者获得防御者的策略后，可能会选择最佳响应策略。

观测到防御者的策略 π_d后，攻击者可能会按如下方式选择最佳响应策略 [7]:

$$\pi^{BR}_a(\pi_d) \in \mathbf{BR}(\pi_d) = \operatorname*{argmax}_{\pi_a \in \Delta\mathcal{B}} U_a(\pi_d, \pi_a).$$

不失一般性，如果存在多个选项，跟随者可以选择对领导者最有利的方式打破平局。在这种情况下，防御者以攻击者为考量，力求最大化自身效用；此时的均衡被定义为强斯塔克尔伯格均衡（SSE） [5, 7, 8]。

定义 1 策略组合 $(\pi^{SSE}_d, \pi^{SSE}_a)$ 称为 $\mathcal{G}$ 的 *SSE*，当且仅当

$$(\pi^{SSE}_d, \pi^{SSE}_a) \in \operatorname*{argmax}_{\pi_d, \pi_a \in \mathbf{BR}(\pi_d)} U_d(\pi_d, \pi_a).$$

当攻击者在观测到防御者的策略后选择最佳响应策略时，SSE策略 π^{SSE}_d对防御者而言是最优的，而且防御者在引导攻击者的策略决策方面占有优势。此外，如果攻击者观测的不是防御者的策略本身，而是玩家行动历史，那么攻击者也可以通过虚拟博弈和 Q 学习方法 [7, 28]等方法选择最佳响应策略。

然而，由于主观或客观因素，例如攻击者观测能力受限、环境干扰以及攻击者的模仿行为 [1, 2, 11]，攻击者在安全问题中未必总会选择最佳响应策略。实际上，攻击者可能会转而采用其他策略。由于其潜在偏好、内在认知和可用资源 [13, 14]，一种不受防御者影响的固定顽固策略，是攻击者最有可能选择的策略之一。

在该安全博弈中，将顽固策略记为π^*_a，而相应的攻击者实际上是一个顽固型玩家。事实上，攻击者一旦发现最具吸引力的目标，就打算保持其行动不变。例如，在移动目标防御中，始终存在对黑客而言最脆弱的目标，并且黑客一旦发现目标 [34]，就无意改变其攻击目标。此外，当无人机的视野有限且缺乏探测其他目标的资源时，它往往会持续攻击当前的最优目标 [35]。不失一般性，我们认为存在一个比其他目标更吸引攻击者的目标，并将上述内容总结为以下假设，该假设也在 [13, 34, 35] 中得到了广泛考虑。

假设2 目标 1 对攻击者而言比目标2更具吸引力，即 $\pi^*_a(1|11) = \pi^*_a(1|21) = 1$。

事实上，无论是最佳响应策略还是顽固策略，都未必是攻击者唯一的最优选择。攻击者可能会采用由这两种策略组成的混合策略。在这种情况下，攻击者实际上被称为有限理性玩家，这在现实中并不少见。例如，由于移动目标防御中的数据采集系统出现错误，以及无人机 [3, 15, 36] 中的数据缺失，攻击者可能会在最佳响应策略和顽固策略之间犹豫。因此，我们将混合策略表述如下。

定义2 如果攻击者的策略 $\pi^\lambda_a(\pi_d, \pi^*_a)$ 满足以下条件，则称其为有限理性策略

$$\pi^\lambda_a(\pi_d, \pi^*_a) = \lambda\pi^{BR}_a(\pi_d) + (1-\lambda)\pi^*_a, \lambda \in [0, 1].$$

对于防御者策略 π_d，我们认为有限理性攻击者采用最佳响应策略

$$\Gamma_1 = \{\lambda \in [0,1] | (U_d^{SSE} - U_{12}^d) D(\mathbf{1}) \lambda^4 - A(1-\lambda)^4 - B\lambda(1-\lambda) \geqslant 0\},$$
$$\Gamma_2 = \{\lambda \in [0,1] | (U_{12}^d - U_d^{SSE}) D(\mathbf{1}) \lambda^4 + A(1-\lambda)^4 - B\lambda(1-\lambda) \geqslant 0\}, \tag{1}$$

$$H(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda) = \frac{\left(U_d^{SSE} - U_{12}^d\right) C(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, 1)\lambda^4 - A(1-\lambda)^4 + B\lambda(1-\lambda)}{C(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda)}. \tag{2}$$

$\pi_a^{BR}(\pi_d) \in$**最佳响应**(π_d) 的概率为 λ，采用顽固策略 π_a^* 的概率为 $1-\lambda$ [10, 37]。因此，当攻击者选择顽固策略时，防御者便失去了引导攻击者做出策略决策的优势；由于 SSE 策略包含顽固因素，它可能无法维持防御性能。因此，面对有限理性攻击者时，SSE 策略不再适用于防御者。研究其他策略以帮助防御者维持防御性能十分重要。

IV. 零行列式策略的性能

本节介绍随机斯塔克尔伯格非对称安全博弈中防御者的零行列式策略。首先，我们给出防御者零行列式策略的定义，并分析该策略存在的条件。此外，我们还将探讨零行列式策略与SSE策略相比的表现。

A. 防御者的零行列式策略

由 [16], 提出的零行列式策略，指一名玩家能够单方面约束两名玩家的期望效用满足线性关系；这种策略已被广泛研究，用于促进公共物品博弈（PGG）、人机交互（HCI）和演化博弈中的合作，或实施单方面敲诈 [17–20]。对于这一随机斯塔尔克尔伯格非对称安全博弈 G, 防御者策略 [2, 16, 31] 的零行列式策略定义如下:

定义 3 如果防御者策略 π_d^{ZD}满足

$$\pi_d^{ZD}(1) = \eta \mathbf{S}^d + \beta \mathbf{S}^a + \gamma \mathbf{1}_4 + \hat{\pi},$$
$$\pi_d^{ZD}(2) = 1 - \pi_d^{ZD}(1), \tag{3}$$

其中 $\eta, \beta, \gamma \in \mathbb{R}$ 和 $\hat{\pi} = [1,1,0,0]^T$。

令 $\pi_d^{ZD}(k) = [\pi_d(k|11) \ \pi_d(k|12) \ \pi_d(k|21) \ \pi_d(k|22)]^T$, $k \in (1), (2)$, 以及 $\mathbf{S}^l = [U_{11}^l, U_{12}^l, U_{21}^l, U_{22}^l]^T$, $l \in \{d, a\}$。防御者所有可行的零行列式策略表示为以下集合

$$\Xi = \{\pi_d^{ZD} \in \Delta\mathcal{A} | \pi_d^{ZD}(1) = \eta \mathbf{S}^d + \beta \mathbf{S}^a + \gamma \mathbf{1}_4 + \hat{\pi}, \pi_d^{ZD}(2) = 1 - \pi_d^{ZD}(1), \eta, \beta, \gamma \in \mathbb{R}\}.$$

如果防御者采用式 (3) 中的零行列式（ZD）策略，那么参与者的期望效用将受到线性关系的约束，这称为零行列式（ZD）：

$$\eta U_d(\pi_d^{ZD}, \pi_a) + \beta U_a(\pi_d^{ZD}, \pi_a) + \gamma = 0, \ \forall \pi_a \in \Delta\mathcal{B}.$$

将 $\pi^{ZD}(\eta, \beta, \gamma)$ 作为相应的零行列式策略。借助零行列式策略对参与者效用的单边强制作用，我们旨在探究，面对有限理性攻击者时，防御者能否采用零行列式策略，比原有的 SSE策略更好地维持其防御性能。

接下来，我们将探究零行列式策略是否存在，以确保防御者能够采用该策略。

B. 零行列式策略的存在性

事实上，零行列式策略无法强制实现两个参与者效用之间任意的线性关系，因为该策略必须属于实施者的策略集。因此，由零行列式策略强制实现的可行线性关系对于进一步分析至关重要。以下引理给出了参与者效用之间的线性关系具有可行性的充要条件，其证明见附录 B。

引理1 在假设*1*下，存在一种能够强制实现 $\eta U_d + \beta U_a + \gamma = 0$ 的零行列式策略，当且仅当满足以下两个不等式之一。

$$\max_{s \in \{11,12\}} \eta U_s^d + \beta U_s^a \leqslant -\gamma \leqslant \min_{s \in \{21,22\}} \eta U_s^d + \beta U_s^a. \tag{4}$$

$$\max_{s \in \{21,22\}} \eta U_s^d + \beta U_s^a \leqslant -\gamma \leqslant \min_{s \in \{11,12\}} \eta U_s^d + \beta U_s^a. \tag{5}$$

引理1意味着，防御者可以采用零行列式策略 $\pi^{ZD}(\eta, \beta, \gamma)$，其中 η、β和 γ满足(4)或(5)，从而强制参与者的期望效用之间呈现理想的线性关系。实际上，引理1拓展了零行列式策略的应用范围，因为安全博弈中的收益矩阵不像迭代囚徒困境博弈中的收益矩阵那样对称 [38]。此外，引理1还涵盖以下情形。

- 若 $U_{11}^a > U_{12}^a$，防御者便可单方面限制攻击者的效用。若防御者采用$\pi^{ZD}(\eta, \beta, \gamma)$，且满足 $\eta = 0$, $\beta \neq 0$ 和 $U_{12}^a \leqslant -\frac{\gamma}{\beta} \leqslant U_{11}^a$，则防御者的零行列式策略可单方面将攻击者的效用限制为 $-\frac{\gamma}{\beta} \in [U_{11}^a, U_{12}^a]$，这与迭代囚徒困境博弈中的均衡器策略相同[16, 39]。

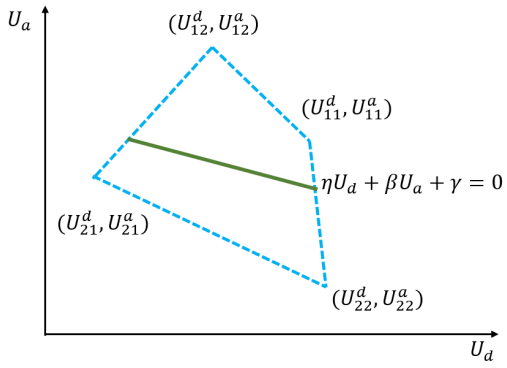


图1. $(U_{11}^d,\ U_{11}^a)$ 和 $(U_{12}^d,\ U_{12}^a)$ 位于直线 $\eta U_a + \beta U_b + \gamma = 0$ 的一侧，而 $(U_{21}^d,\ U_{21}^a)$ 和 $(U_{22}^d,\ U_{22}^a)$ 位于另一侧。

- 若 $U_{12}^d > U_{22}^d$，防御者便可单方面限制自身效用。若防御者采用 $\pi^{ZD}(\eta,\beta,\gamma)$，且满足 $\eta \neq 0, \beta = 0$ 和 $U_{22}^d \leqslant -\frac{\gamma}{\eta} \leqslant U_{12}^d$，则防御者可将防御者的效用单方面限制在 U_{22}^d 至 U_{12}^d 之间，这与移动目标防御问题中的结果一致 [27]。

基于 ZD策略所施加的可行线性关系的存在条件，我们可以进一步分析安全博弈 $\mathcal{G}$ 中是否至少存在一个ZD策略。为简化起见，对于任意 $x_1,\ x_2,\ y_1,\ y_2 \in \mathbb{R}$ 且满足 $x_1 \neq x_2$ 的情形，记

$$\Gamma^-(x_1,y_1,x_2,y_2) = \{(x,y)|y - y_1 \leqslant \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)\},$$
$$\Gamma^+(x_1,y_1,x_2,y_2) = \{(x,y)|y - y_1 \geqslant \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)\}.$$

实际上，Γ^- (Γ^+) 是经过点 $(x_1,\ y_1)$ 和 $(x_2,\ y_2)$ 的直线的上方（下方）区域。下面的定理给出了 $\mathcal{G}$ 中存在ZD策略的一个充分条件，其证明见附录 C。

定理 1 在假设*1*下，若以下两个关系之一成立，则在 $\mathcal{G}$ 中至少存在一个防御者的零行列式策略。

$$(U_{21}^d, U_{21}^a), (U_{22}^d, U_{22}^a) \in \Gamma^+(U_{11}^d, U_{11}^a, U_{12}^d, U_{12}^a).$$
$$(U_{21}^d, U_{21}^a), (U_{22}^d, U_{22}^a) \in \Gamma^-(U_{11}^d, U_{11}^a, U_{12}^d, U_{12}^a).$$

定理 1 表明，在安全博弈 $\mathcal{G}$ 中，只要点 (U_{21}^d, U_{21}^a), (U_{22}^d, U_{22}^a) 位于经过点 (U_{11}^d, U_{11}^a) 和 (U_{12}^d, U_{12}^a) 的直线同一侧，防御者就能选择一种零行列式策略，如图 1 所示。因此，本文重点研究至少存在一种零行列式策略的情形，因为为防御者选择零行列式策略的前提是这种策略存在。事实上，这两个条件可以互换。例如，如果 $(U_{21}^d, U_{21}^a), (U_{22}^d, U_{22}^a) \in \Gamma^-(U_{11}^d, U_{11}^a, U_{12}^d, U_{12}^a)$，那么可以通过交换 (U_{11}^d, U_{11}^a) 与 (U_{22}^d, U_{22}^a) 的值，并交换 (U_{12}^d, U_{12}^a) 与 (U_{21}^d, U_{21}^a) 的值，得到 $(U_{21}^d, U_{21}^a), (U_{22}^d, U_{22}^a) \in \Gamma^-(U_{11}^d, U_{11}^a, U_{12}^d, U_{12}^a)$。因此，不失一般性，本文其余结果均在条件 $(U_{21}^d, U_{21}^a), (U_{22}^d, U_{22}^a) \in \Gamma^-(U_{11}^d, U_{11}^a, U_{12}^d, U_{12}^a)$ 下成立。

C. 两种特殊情况下的零行列式策略

为便于理解，我们从两种特殊情况开始：$\lambda = 1$，其中攻击者采用最佳响应策略[7], ；以及 $\lambda = 0$，其中攻击者是固执型玩家，采用固执型策略 [13, 14]。

当 $\lambda = 1$时，攻击者在观察到防御者策略 π_d后选择 **最佳响应**(π_d)。回忆一下，$(\pi_d^{SSE},\ \pi_a^{SSE})$ 是一个 SSE，而 U_d^{SSE}是防御者采用 SSE策略时获得的效用。

引理2 在假设*1*下，对于任意 $\pi_d^{ZD} \in \Xi$ 和 $\pi_a \in BR(\pi_d^{ZD})$, $U_d(\pi_d^{ZD}, \pi_a) \leqslant U_d(\pi_d^{SSE}, \pi_a^{SSE})$。

引理2表明，面对采用最佳响应策略的攻击者时，SSE策略总能为防御者带来最高效用。在这种情况下，零行列式策略的性能上限无法超过防御者采用SSE策略时所获得的效用。尽管如此，下面的定理指出，与SSE策略相比，零行列式策略的损失有界，其证明见附录D。

定理2 在假设*1*下，

$$\min_{\pi_d^{ZD} \in \Xi} U_d\left(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE})\right) - U_d\left(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD})\right)$$
$$= \begin{cases} 0, & \text{if } U_{11}^d \geqslant U_{21}^d, \\ U_d^{SSE} - U_{12}^d, & \text{if } U_{11}^d < U_{21}^d. \end{cases}$$

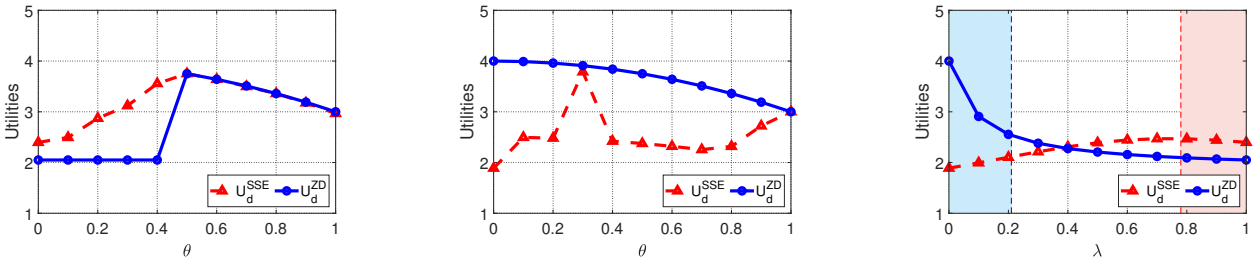
对应的零行列式策略为 π_d^{ZD}

$$= \begin{cases} \pi^{ZD}(-k_1, 1, k_1 U_{11}^d - U_{11}^a), & \text{if } U_{11}^d \geqslant U_{22}^d,\ U_{11}^a \geqslant U_{21}^a, \\ \pi^{ZD}(0, 1, -U_{21}^a), & \text{if } U_{11}^d < U_{22}^d,\ U_{11}^a \geqslant U_{21}^a, \\ \pi^{ZD}(-k_2, 1, k_2 U_{12}^d - U_{12}^a), & \text{otherwise,} \end{cases}$$

$$\text{where } 0 \leqslant k_1 \leqslant \frac{U_{11}^a - U_{21}^a}{U_{11}^d - U_{21}^d}, \text{ and } \frac{U_{22}^a - U_{12}^a}{U_{22}^d - U_{12}^d} \leqslant k_2 \leqslant \frac{U_{11}^a - U_{12}^a}{U_{11}^d - U_{12}^d}.$$

定理2表明，与SSE策略相比，防御者可以采用零行列式策略，在效用上承受可接受的损失。一方面，当 $U_{11}^a \geqslant U_{21}^a$ 时，定理2中的零行列式策略就是SSE策略，为防御者带来与SSE策略相同的效用。另一方面，如果防御者能够承受有界损失，那么采用相应的零行列式策略也是不错的选择，因为这样可以避免复杂的SSE策略计算；推导SSE策略需要求解一个双层优化问题。

当 $\lambda = 0$时，攻击者选择顽固策略π_a^*。由于顽固攻击者不



(a) 采用最佳响应策略的攻击者 (b) 采用顽固策略的攻击者 (c) 有限理性攻击者

图2. 移动目标防御问题中零行列式策略与SSE策略的性能比较。红色虚线表示采用SSE策略时防御者的期望效用，蓝色实线表示采用相应零行列式策略时防御者的期望效用。在(c)中，红色（蓝色）区域表示根据定理4（定理5）得出的 Γ_1 （Γ_2）界限。

会按照防御者的预期选择最佳响应策略，SSE策略可能无法为防御者带来可接受的效用。在这种情况下，由于零行列式策略能够单方面约束参与者效用，零行列式策略可以恰好发挥关键作用，为防御者强制实现期望的效用，甚至带来高于原SSE策略的效用；以下定理的证明见附录E

定理3 在假设*1*和假设*2*下，存在一个零行列式策略 $\pi_d^{ZD} = \pi^{ZD}(-k, 1, kU_{12}^d - U_{12}^a)$，满足 $k = \frac{U_{11}^a - U_{12}^a}{U_{11}^d - U_{12}^d}$ 使得

$$U_d(\pi_d^{ZD}, \pi_a^*) \geqslant U_d(\pi_d^{SSE}, \pi_a^*).$$

事实上，零行列式策略 $\pi^{ZD}(-k,\ 1,\ kU_{12}^d - U_{12}^a)$ 强制参与者的期望效用之间形成一条经过 $(U_{12}^a,\ U_{12}^b)$ 和 $(U_{11}^d,\ U_{11}^a)$ 的线性关系。零行列式策略 $\pi^{ZD}(-k,\ 1,\ kU_{12}^d - U_{12}^a)$ 的下确界不低于任何其他零行列式策略的下确界，包括那些能够单方面设定防御者效用 [27] 的策略。此外，根据其证明，面对顽固攻击者时，与SSE策略相比，该零行列式策略 π_d^{ZD}会提高防御者的效用 $U_{11}^d - \frac{U_{11}^d \pi_d^{SSE}(1|21) + U_{21}^d \pi_d^{SSE}(2|11)}{\pi_d^{SSE}(2|11) + \pi_d^{SSE}(1|21)}$。

D. 一般情形下的零行列式策略

现在考虑一般情形，**其中** $\lambda \in [0, 1]$。在此情形下，有限理性攻击者以概率 λ 选择最佳响应策略，以概率 $1 - \lambda$选择顽固策略π_a^*，即定义2中的 $\pi_a^{\lambda}(\pi_d,\ \pi_a^*)$。直观而言，当 λ 接近 1 时，ZD策略可能使防御者获得与定理2中类似的表现；当 λ 接近 0 时，则可能获得与定理3中类似的表现。事实上，本小节的一个主要结果如下述定理所示，其证明见附录 F。

定理4 在假设*1*和假设*2*下，对于 *(1)* 中的 $\lambda \in \Gamma_1$，有

$$\min_{\pi_d^{ZD} \in \Xi} U_d\left(\pi_d^{SSE}, \pi_a^{\lambda}(\pi_d^{SSE}, \pi_a^*)\right) - U_d\left(\pi_d^{ZD}, \pi_a^{\lambda}(\pi_d^{ZD}, \pi_a^*)\right)$$
$$\leqslant \begin{cases} 0, & \text{if } U_{11}^d \geqslant U_{22}^d,\ U_{11}^a \geqslant U_{21}^a, \\ H(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda), & \text{otherwise}, \end{cases}$$

其中，$H(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda)$ 在式 (*2*) 中定义，其参数见附录*A*。相应的零行列式策略为 π_d^{ZD}

$$= \begin{cases} \pi^{ZD}(-k_1, 1, k_1 U_{11}^d - U_{11}^a), & \text{if } U_{11}^d \geqslant U_{22}^d,\ U_{11}^a \geqslant U_{21}^a, \\ \pi^{ZD}(-k_2, 1, k_2 U_{12}^d - U_{12}^a), & \text{otherwise}, \end{cases}$$

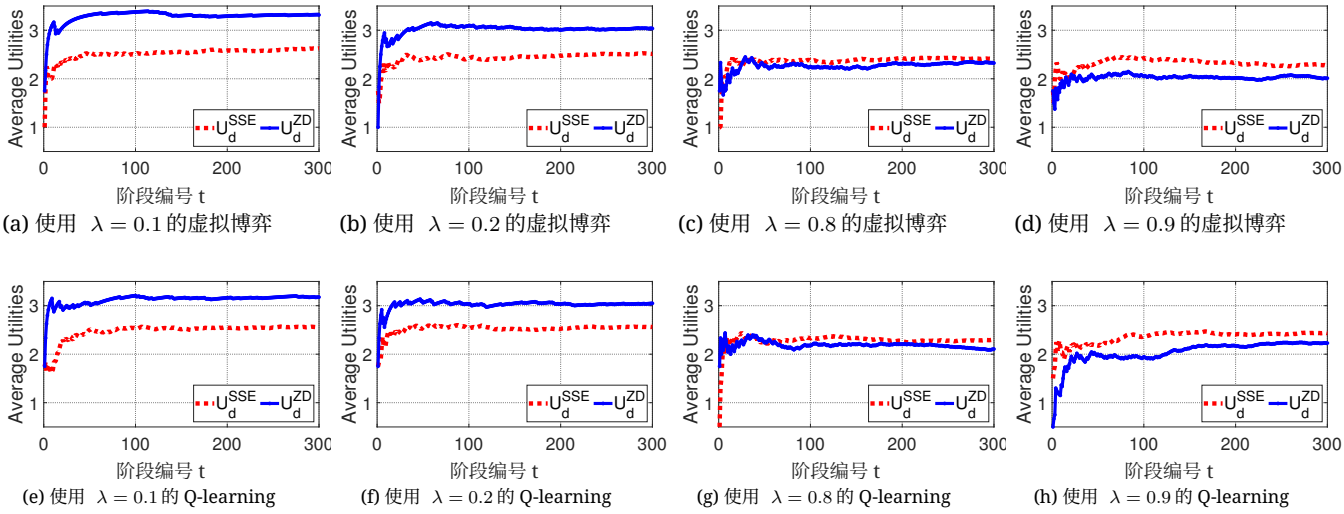
where $0 \leqslant k_1 \leqslant \frac{U_{11}^a - U_{21}^a}{U_{11}^d - U_{21}^d}$, *and* $k_2 = \frac{U_{11}^a - U_{12}^a}{U_{11}^d - U_{12}^d}$.

定理4为防御者提供了集合 Γ_1，在该集合中，零行列式策略带来的防御性能损失是有界且可容忍的，尽管零行列式策略无法优于SSE策略。因此，对于具有 $\lambda \in \Gamma_1$的有限理性攻击者，如果防御者不太在意效用略有损失，那么防御者可以采用相应的零行列式策略，因为采用定理4中的零行列式策略，可以避免为求解SSE策略而耗费大量资源去处理双层优化问题。此外，如果 $U_{11}^a \geqslant U_{21}^a$ 和$U_{11}^d \geqslant U_{22}^d$，零行列式策略能够为防御者带来与SSE策略相同的效用，这意味着防御者仍然可以采用零行列式策略。

尽管 Γ_1 看起来很复杂，难以验证，但可以轻松确认 λ 的一些典型取值是否属于 Γ_1。例如，$\lambda = 1$始终属于 Γ_1。在这种情况下，攻击者倾向于采取最佳响应策略，这与定理2一致。实际上，当 λ 接近1时，λ 属于 Γ_1，这意味着攻击者倾向于选择最佳响应策略。此外，我们提供了 Γ_1的一个子集，防御者可以轻松验证该子集。

推论1 在假设*1*和假设*2*下，如果 $\lambda \in [\frac{1}{2}, 1]$且

$$4(U_{11}^d - U_{21}^d)\pi_d^{SSE}(2|11)(1-\lambda)^4$$
$$\leqslant \left(\frac{1}{4}(U_d^{SSE} - U_{12}^d)D(\boldsymbol{1}) - B\right)\left(\pi_d^{SSE}(2|11) + \pi_d^{SSE}(1|21)\right),$$



(a) 使用 $\lambda = 0.1$ 的虚拟博弈 (b) 使用 $\lambda = 0.2$ 的虚拟博弈 (c) 使用 $\lambda = 0.8$ 的虚拟博弈 (d) 使用 $\lambda = 0.9$ 的虚拟博弈

(e) 使用 $\lambda = 0.1$ 的 Q-learning (f) 使用 $\lambda = 0.2$ 的 Q-learning (g) 使用 $\lambda = 0.8$ 的 Q-learning (h) 使用 $\lambda = 0.9$ 的 Q-learning

图3. 在采用不同机制的信息物理系统中，零行列式策略与SSE策略的性能比较。红色虚线表示采用SSE策略时防御者的平均效用，蓝色实线表示采用定理4或定理5中相应零行列式策略时防御者的平均效用。

then

$$\min_{\pi_d^{ZD}\in\Xi} U_d\left(\pi_d^{SSE},\pi_a^{\lambda}(\pi_d^{SSE},\pi_a^*)\right) - U_d\left(\pi_d^{ZD},\pi_a^{\lambda}(\pi_d^{ZD},\pi_a^*)\right)$$
$$\leqslant \begin{cases} 0, & \text{if } U_{11}^a \geqslant U_{21}^a,\ U_{11}^d \geqslant U_{22}^d, \\ H(\pi_d^{ZD},\pi_d^{SSE},\pi_a^*,\lambda), & \text{otherwise.} \end{cases}$$

最后，我们类似地考虑 λ接近0的情况，即攻击者倾向于在以下定理中采取顽固策略，其证明见附录G。

定理5 在假设*1*和假设*2*下，若 $\lambda \in \Gamma_2$ $\pi_d^{ZD} = in()$，则存在一个零行列式策略 $(,,kU_{21}^a - U_{21}^b)$，使得

$$U_d(\pi_d^{ZD},\pi_a^{\lambda}(\pi_d^{ZD},\pi_a^*)) \geqslant U_d(\pi_d^{SSE},\pi_a^{\lambda}(\pi_d^{SSE},\pi_a^*)),$$

where $k = \frac{U_{11}^b - U_{12}^b}{U_{11}^a - U_{12}^a}$.

定理5给出了 Γ_2，供防御者采用零行列式策略，以获得高于原始SSE策略的效用。值得注意的是，相应的零行列式策略表现出色。如果 $\lambda \in \Gamma_2$，防御者可以放心地选择相应的零行列式策略，因为与SSE策略相比，该策略能带来更高的防御性能。

显然，$\lambda = 0$ 始终属于 Γ_2，这与定理3的结果一致。实际上，当 λ 接近0时，λ属于 Γ_2，这意味着攻击者趋于成为顽固攻击者。此外，我们还为防御者提供了 Γ_2 的一个子集，防御者可以轻松验证这一子集。

推论2 在假设*1*和假设*2*下，如果 $\lambda \in [0, \frac{1}{2}]$且

$$4(U_d^{SSE} - U_{12}^a)D(\mathbf{1})\lambda^4 \leqslant \frac{1}{4}A - B,$$

那么存在 $\pi_d^{ZD} \in \Xi$ 使得

$$U_d(\pi_d^{ZD},\pi_a^{\lambda}(\pi_d^{ZD},\pi_a^*)) \geqslant U_d(\pi_d^{SSE},\pi_a^{\lambda}(\pi_d^{SSE},\pi_a^*)).$$

V. 应用

为作说明，我们通过实验验证零行列式策略能够帮助防御者在面对有限理性攻击者时维持其防御性能，并以原始SSE策略作为基准。

A. 移动目标防御问题

考虑一个移动目标防御问题，其中攻击者可以直接观察防御者策略，并采取其明确的最佳响应策略 [2, 11]。设 Y_i 为防御者将防御资源从目标 i 转移到另一个目标所需的成本，设 C_i 为攻击者入侵目标 i的成本。与 [27]，类似，我们也使用平均值来近似转移成本。此外，设 R_i^d 和 R_i^a分别为状态 $i \in \mathcal{S}$下两位玩家的奖励和损失。因此，移动目标防御问题中的效用矩阵如表 II 所示。对于任意 $s \in \mathcal{S}$，设 $R_s^d = d_s^1\theta + d_s^0$和 $R_s^a = a_s^1\theta + a_s^0$，其中 $\theta \in [0,1]$、d_s^k 和 a_s^k 分别是任意 $s \in \mathcal{S}$和 $k \in \{1,2\}$下参与者奖励与损失中的参数。

如图2(a)所示，对于采用最佳响应策略的攻击者，采用SSE策略的防御者的期望效用始终高于采用零行列式策略的防御者。此外，如图2(b)所示，对于采用顽固策略的攻击者，采用定理3中的零行列式策略的防御者的期望效用始终高于采用SSE策略的防御者。此外，在图

表 II. 移动目标防御问题中的效用矩阵

		攻击者	
		1	2
防御者	1	$(R_{11}^d - \frac{Y_2}{2},\ R_{11}^a - C_1)$	$(R_{12}^d - \frac{Y_2}{2},\ R_{12}^a - C_2)$
	2	$(R_{21}^d - \frac{Y_1}{2},\ R_{21}^a - C_1)$	$(R_{22}^d - \frac{Y_1}{2},\ R_{22}^a - C_2)$

2(c)中，当 $\lambda < 0.21$时， U_d^{ZD}始终高于 U_d^{SSE}，这与定理5一致。此外，当 $\lambda > 0.78$时， U_d^{SSE}始终高于 U_d^{ZD}，这与定理4一致，且两种效用之间的差异有界且可以容忍。

B. 信息物理系统问题

这里考虑一个信息物理系统问题，其中防御者是系统管理员，攻击者是干扰者或窃听者。与之前的实验不同，攻击者只能观察玩家行动历史，而无法直接获取防御者的策略。攻击者采用基于虚拟博弈 [28]或 Q 学习方法 [7], 的最佳响应策略，具体细节见附录 H。

如图 3(a)-(f) 所示，当 λ 接近 0 时，例如图 3(a)、(e) 中的 0.1，以及图 3(b)、(f) 中的 0.2，采用零行列式策略的防御者平均效用高于采用 SSE 策略的防御者。此外，当 λ 接近 1 时，例如图 3(c)、(g) 中的 0.8，以及图 3(d)、(h) 中的 0.9，尽管采用零行列式策略的平均效用低于采用 SSE 策略的平均效用，但损失足够小，可以接受。因此，防御者也可以采用零行列式策略，以有限且可容忍的损失维持其防御性能，并避免 SSE 策略所需的复杂计算。

VI. 讨论

本文聚焦于随机斯塔克尔伯格非对称安全博弈。由于有限理性攻击者内部存在固执性因素，我们研究了零行列式策略的防御性能，并分析了在这种情况下，零行列式策略能否弥补SSE策略的不足。此外，我们还通过为防御者采用合适的零行列式策略进行实验，以支持我们的方法。

事实上，我们的结果可以扩展到一些涉及多个目标的安全问题。例如，考虑安全博弈中的目标可分为两类的情况。对于属于同一类别的任意两个目标，每个玩家的效用都相同；而对于属于不同类别的任意两个目标，玩家的效用则不同。在这种情况下，防御者仍可采用本文中类似的 ZD策略来提升防御效果。

事实上，在一般的多目标非对称随机安全博弈中，针对有限理性攻击者应用 ZD策略仍面临挑战。我们从一个简单的角度说明这一障碍。在多目标设定下，防御者的期望效用由 λ 中的复杂多项式构成，而每个多项式都是一个 $(n^2 \times n^2)$ 矩阵的行列式，且其次数很高。在一般的多目标安全博弈中应用 ZD策略仍是一个开放问题，值得进一步深入探索。

致谢

本研究得到中国国家重点研发计划（编号：2022YFA1004700）、国家自然科学基金（编号：62173250）以及上海市重大科技项目（编号：2021SHZDZX0100）的资助。

附录A：符号说明

对于任意 $\pi_d \in \Delta\mathcal{D}$， π_a^1， $\pi_a^2 \in \Delta\mathcal{A}$ 和 **函数** $= [f_1, f_2, f_3, f_4]^T \in \mathbb{R}^4$，记

$$D(\pi_d, \pi_a^1, \pi_a^2, \mathbf{f}) = \begin{bmatrix} \pi_d(1|11)\pi_a^2(1|11) - 1 & \pi_d(1|11) - 1 & \pi_a^1(1|11) - 1 & f_1 \\ \pi_d(1|12)\pi_a^2(1|12) & \pi_d(1|12) - 1 & \pi_a^1(1|12) & f_2 \\ \pi_d(1|21)\pi_a^2(1|21) & \pi_d(1|21) & \pi_a^1(1|21) - 1 & f_3 \\ \pi_d(1|22)\pi_a^2(1|22) & \pi_d(1|22) & \pi_a^1(1|22) & f_4 \end{bmatrix}. \tag{A1}$$

为方便起见 因此，我们取 $D(\pi_d, \pi_a, \mathbf{f}) = D(\pi_d, \pi_a, \pi_a, \mathbf{f})$ 和 $D(\mathbf{f}) = \max_{\pi_d, \pi_a^1, \pi_a^2} D(\pi_d, \pi_a^1, \pi_a^2, \mathbf{f})$。记

$$J(\pi_d, \pi_a, \mathbf{f}) = D(\pi_d, \pi_a^{BR}(\pi_d), \pi_a, \mathbf{f}) + D(\pi_d, \pi_a, \pi_a^{BR}(\pi_d), \mathbf{f}),$$

and

$$C(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda) = D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1}) \cdot D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1}),$$

为简化表达，此外，取

$$\begin{aligned}
A =& U_{11}^d - \frac{U_{11}^d \pi_d^{SSE}(1|21) + U_{21}^d \pi_d^{SSE}(2|11)}{\pi_d^{SSE}(2|11) + \pi_d^{SSE}(1|21)}, \\
B_1 =& \max_{\pi_d^{SSE}, \pi_d^{ZD}} \max_{\mathbf{f} \in \{\mathbf{1}, S^d\}} \frac{1}{2} \left| D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{f}) - D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{f}) \right|, \\
B_2 =& \max_{\pi_d^{SSE}, \pi_d^{ZD}} \left| D(\pi_d^{ZD}, \pi_a^*, S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) + D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) J(\pi_d^{ZD}, \pi_a^*, S^d) - D(\pi_d^{SSE}, \pi_a^*, S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) \right. \\
& \left. - D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) J(\pi_d^{SSE}, \pi_a^*, S^d) \right|, \\
B_3 =& \max_{\pi_d^{SSE}, \pi_d^{ZD}} \left| J(\pi_d^{ZD}, \pi_a^*, S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) - J(\pi_d^{ZD}, \pi_a^*, S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) \right|, \\
B =& \max \left\{ B_1, B_2, \frac{1}{2} B_3 \right\}.
\end{aligned}$$

附录 B：引理1的证明

充分性：考虑零行列式策略 π_d 强制实现 $\eta U_d + \beta U_a + \gamma = 0$。因此，$\pi_d(\mathbf{1}) = \phi(\eta \mathbf{S}^d + \beta \mathbf{S}^a + \gamma) + \hat{}\pi$，其中 $\phi \neq 0$。由于 $\pi_d \in \Xi$，以下不等式成立:

$$-1 \leqslant \phi(\eta U_i^d + \beta U_i^a + \gamma) \leqslant 0, i \in \{11, 12\}, \tag{B1a}$$

$$0 \leqslant \phi(\eta U_j^d + \beta U_j^a + \gamma) \leqslant 1, j \in \{21, 22\}. \tag{B1b}$$

若 $\phi > 0$，则由 (B1a) 中右侧的不等式和 (B1b) 中左侧的不等式可得

$$\begin{aligned}
\eta U_i^d + \beta U_i^a + \gamma \leqslant 0, i \in \{11, 12\}, \\
\eta U_j^d + \beta U_j^a + \gamma \geqslant 0, j \in \{21, 22\},
\end{aligned} \tag{B2}$$

这意味着 $\max(\eta U_{11}^d + \beta U_{11}^a, \eta U_{12}^d + \beta U_{12}^a) \leqslant -\gamma \leqslant \min(\eta U_{21}^d + \beta U_{21}^a, \eta U_{22}^d + \beta U_{22}^a)$。类似地，若 $\phi < 0$，可得 $\max(\eta U_{21}^d + \beta U_{21}^a, \eta U_{22}^d + \beta U_{22}^a) \leqslant -\gamma \leqslant \min(\eta U_{11}^d + \beta U_{11}^a, \eta U_{12}^d + \beta U_{12}^a)$。

必要性：考虑 $\max(\eta U_{11}^d + \beta U_{11}^a, \eta U_{12}^d + \beta U_{12}^a) \leqslant -\gamma \leqslant \min(\eta U_{21}^d + \beta U_{21}^d, \eta U_{22}^d + \beta U_{22}^a)$。由此得到 (B2)。如果 (B2) 中的所有不等式都不严格成立，那么零行列式策略便能强制实现 $\eta U_d + \beta U_a + \gamma = 0$。否则，取$\phi = \max\{|\eta U_i^d + \beta U_i^a + \gamma|\}_{i \in \{11,12,21,22\}}$，可得

$$\begin{aligned}
-1 \leqslant \frac{1}{\phi}(\eta U_i^d + \beta U_i^a + \gamma) \leqslant 0, i \in \{11, 12\}, \\
0 \leqslant \frac{1}{\phi}(\eta U_j^d + \beta U_j^a + \gamma) \leqslant 1, j \in \{21, 22\}.
\end{aligned}$$

因此，零行列式策略 $\pi_d^{ZD}(\frac{\eta}{\phi}, \frac{\beta}{\phi}, \frac{\gamma}{\phi})$ 是可行的，并且能强制实现 $\eta U_d + \beta U_a + \gamma = 0$。类似地，对于 $\max_{s \in \{21,22\}} \eta U_s^d + \beta U_s^a \leqslant -\gamma \leqslant \min_{s \in \{11,12\}} \eta U_s^d + \beta U_s^a$，结论也成立

附录 C：定理 1 的证明

考虑 $(U_{21}^d, U_{21}^a), (U_{22}^d, U_{22}^a) \in \Gamma^+(U_{11}^d, U_{11}^a, U_{12}^d, U_{12}^a)$。取 $\eta = -\frac{U_{21}^a - U_{11}^a}{U_{21}^d - U_{11}^d}$, $\beta = 1$, $\gamma = U_{11}^d \frac{U_{21}^a - U_{11}^a}{U_{21}^d - U_{11}^d} - U_{11}^a$，可得

$$\max(\eta U_{11}^a + \beta U_{11}^b, \eta U_{12}^a + \beta U_{12}^b) \leqslant -\gamma \leqslant \min(\eta U_{21}^a + \beta U_{21}^b, \eta U_{22}^a + \beta U_{22}^b).$$

同理，结论对于 $(U_{21}^d,\ U_{21}^a),\ (U_{22}^d,\ U_{22}^a) \in \Gamma^-(U_{11}^d,\ U_{11}^a,\ U_{12}^d,\ U_{12}^a)$ 也成立。因此，防御者至少存在一种零行列式策略。

附录D：定理2的证明

1）当 $U_{11}^d \geqslant U_{21}^d$ 且 $U_{11}^a \geqslant U_{21}^a$ 时，根据引理1，ZD策略 $\pi^{ZD}(-k_1,\ 1,\ k_1U_{11}^d - U_{11}^a)$ 对防御者可行，其中 $0 \leqslant k_1 \leqslant \frac{U_{11}^a - U_{21}^a}{U_{11}^d - U_{21}^d}$。该ZD策略强制实现 $U_a - U_{11}^a = k_1(U_d - U_{11}^d)$。由于 $k_1 \geqslant 0$，当攻击者观察到防御者策略时，攻击者的最优效用为 U_{11}^a 。在这种情况下，防御者的效用为 U_{11}^d。注意，由于 $U_{11}^d \leqslant \max\{U_{12}^d, U_{21}^d, U_{22}^d\}$，$U_{11}^d$ 对防御者而言也同样是最优的。于是 $U_d^{SSE} \leqslant U_{11}^d$。根据引理2，$U_d^{SSE} = U_d(\pi_d^{ZD},\ \pi_a^{BR}(\pi_d^{ZD}))$。

2）当 $U_{11}^d < U_{21}^d$ 和 $U_{11}^a \geqslant U_{21}^a$ 时，根据引理1，零行列式策略 $\pi^{ZD}(0, 1, -U_{21}^a)$ 对防御者而言也是可行的。该零行列式策略可确保 $U_a - U_{21}^a = 0$。由于存在多个选项时，攻击者总会在平局时选择对防御者最有利的方案，因此攻击者会选择可确保 $U_a = U_{21}^a$ 和 $U_d = (U_{21}^a - U_{11}^a)(U_{22}^d - U_{11}^d)\frac{}{U_{22}^2 - U_{11}^2} + U_{11}^d$ 的策略。因此，$U_d^{SSE} \geqslant (U_{21}^a - U_{11}^a)(U_{22}^d - U_{11}^d)\frac{}{U_{22}^2 - U_{11}^2} + U_{11}^d$。此外，假设防御者选择SSE策略时，其效用高于 $(U_{21}^a - U_{11}^a)(U_{22}^d - U_{11}^d)\frac{}{U_{22}^2 - U_{11}^2} + U_{11}^d$，即 $U_d^{SSE} > (U_{21}^a - U_{11}^a)(U_{22}^d - U_{11}^d)\frac{}{U_{22}^2 - U_{11}^2} + U_{11}^d$。由于 $U_a^{SSE} - U_{11}^a \leqslant (U_{22}^a - U_{11}^a)(U_d^{SSE} - U_{11}^d)\frac{}{U_{22}^d - U_{11}^d}$ 和 $\frac{U_{22}^a - U_{11}^a}{U_{22}^d - U_{11}^d} < 0$，我们有 $U_a^{SSE} < U_{12}^a$。实际上，对于任意防御者策略 π_d，当攻击者选择满足 $\pi_a(1|11) = \pi_a(1|21) = \pi_a(2|12) = \pi_a(2|22) = 1$，$U_a(\pi_d, \pi_a) = U_{11}^a \frac{\pi_d(1|21)}{\pi_d(2|11)+\pi_d(1|21)} + U_{21}^a \frac{\pi_d(2|11)}{\pi_d(2|11)+\pi_d(1|21)} \geqslant U_{21}^a$ 的策略 π_a 时，攻击者总能找到一种策略，使其效用不低于 U_{12}^a，这与 $U_a^{SSE} < U_{12}^a$ 矛盾。因此，$U_d^{SSE} = (U_{21}^a - U_{11}^a)(U_{22}^d - U_{11}^d)\frac{}{U_{22}^2 - U_{11}^2} + U_{11}^d = U_d(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}))$。

3）当 $U_{11}^a < U_{21}^a$ 时，根据引理1，ZD策略 $\pi^{ZD}(-k_2,\ 1,\ k_2U_{12}^d - U_{12}^a)$ 对防御者而言是可行的，其中 $\frac{U_{22}^a - U_{12}^a}{U_{22}^d - U_{12}^d} \leqslant k_2 \leqslant \frac{U_{11}^a - U_{12}^a}{U_{11}^d - U_{12}^d}$。因此，观察到防御者策略后，攻击者的最优效用为 U_{12}^a。在这种情况下，防御者相应的效用为 U_{12}^d，且 $U_d(\pi_d^{ZD},\ \pi_a^{BR}(\pi_d^{ZD})) - U_d(\pi_d^{SSE},\ \pi_a^{BR}(\pi_d^{SSE})) = U_d^{SSE} - U_{12}^d$。此外，根据引理1，对于任何能够强制实现 $\eta U_d + \beta U_a + \gamma = 0$ 的ZD策略 π_d，当 $U_{11}^a < U_{21}^a$ 时，$\eta \cdot \beta > 0$ 始终成立。因此，攻击者的最佳响应策略也会使防御者的效用最小化，且 $\max_{\pi_d^{ZD} \in \Xi} U_d(\pi_d^{ZD},\ \pi_a^{BR}(\pi_d^{ZD})) = U_{12}^d$。那么 $\min_{\pi_d^{ZD} \in \Xi} U_d(\pi_d^{SSE},\ \pi_a^{BR}(\pi_d^{SSE})) - U_d(\pi_d^{ZD},\ \pi_a^{BR}(\pi_d^{ZD})) = U_d^{SSE} - U_{12}^d$.

Therefore, $\min_{\pi_d^{ZD} \in \Xi} U_d\left(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE})\right) - U_d\left(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD})\right) = \begin{cases} 0, & \text{if } U_{11}^d \geqslant U_{21}^d, \\ U_d^{SSE} - U_{12}^d, & \text{if } U_{11}^d < U_{21}^d. \end{cases}$

附录E：定理3的证明

根据 [16]，$U_d(\pi_d, \pi_a) = D(\pi_d, \pi_a, S^d)\, D(\pi_d, \pi_a, 1)$ 和 $U_a(\pi_d, \pi_a) = D(\pi_d, \pi_a, S^a)\, D(\pi_d, \pi_a, 1)$。对于具有 $\pi_a^*(1|11) = \pi_a^*(1|21) = 1$ 的顽固攻击者，我们有

$$U_d(\pi_d, \pi_a^*) = U_{11}^d \frac{\pi_d(1|21)}{\pi_d(2|11) + \pi_d(1|21)} + U_{21}^d \frac{\pi_d(2|11)}{\pi_d(2|11) + \pi_d(1|21)},$$

and

$$U_a(\pi_d, \pi_a^*) = U_{11}^a \frac{\pi_d(1|21)}{\pi_d(2|11) + \pi_d(1|21)} + U_{21}^a \frac{\pi_d(2|11)}{\pi_d(2|11) + \pi_d(1|21)},$$

对于任意 $\pi_a^*(1|s) = 1$ 或 0，其中 $s \in \{12, 22\}$。注意，$U_d(\pi_d, \pi_a^*)$ 和 $U_a(\pi_d, \pi_a^*)$ 关于 $\pi_a^*(1|s) \in [0, 1]$ 是单调的。因此，$U_d(\pi_d^{SSE}, \pi_a^*) = \frac{U_{11}^d \pi_d^{SSE}(1|21) + U_{21}^d \pi_d^{SSE}(2|11)}{\pi_d^{SSE}(2|11) + \pi_d^{SSE}(1|21)}$。取 $\pi_d^{ZD} = \pi^{ZD}(-k, 1, kU_{12}^d - U_{12}^a)$，其中 $k = \frac{U_{11}^a - U_{12}^a}{U_{11}^d - U_{12}^d}$，则有 $U_d(\pi_d^{ZD}, \pi_a^*) = U_{11}^d$，这意味着 $U_d(\pi_d^{ZD}, \pi_a^*) - U_d(\pi_d^{SSE}, \pi_a^*) = U_{11}^d - \frac{U_{11}^d \pi_d^{SSE}(1|21) + U_{21}^d \pi_d^{SSE}(2|11)}{\pi_d^{SSE}(2|11) + \pi_d^{SSE}(1|21)} \geqslant 0$.

附录 F：定理4的证明

如果 $U_{11}^d \geqslant U_{22}^d$ 和 $U_{11}^a \geqslant U_{21}^a$，根据引理1，零行列式策略 $\pi^{ZD}(-k_1,\ 1,\ k_1U_{11}^d - U_{11}^a)$ 对防御者而言是可行的。此外，根据定理2，该零行列式策略也是一种SSE策略。因此，面对有限理性攻击者时，该零行列式策略为防御者带来的效用与SSE策略相同。因此，

$$\min_{\pi_d^{ZD} \in \Xi} U_d\left(\pi_d^{SSE}, \pi_a^{\lambda}(\pi_d^{SSE}, \pi_a^*)\right) - U_d\left(\pi_d^{ZD}, \pi_a^{\lambda}(\pi_d^{ZD}, \pi_a^*)\right) \leqslant 0, \quad \text{if } U_{11}^d \geqslant U_{22}^d,\ U_{11}^a \geqslant U_{21}^a.$$

否则，根据引理1，零行列式策略 $\pi^{ZD}(-k_2,\ 1,\ k_2U_{12}^d-U_{12}^a)$ 对防御者而言是可行的。根据定理2，$U_d(\pi_d^{ZD},\ \pi_a^{BR}(\pi_d^{ZD}))-U_d(\pi_d^{SSE},\ \pi_a^{BR}(\pi_d^{SSE}))=-(U_d^{SSE}-U_{12}^d)$。那么

$$\frac{D(\pi_d^{ZD},\pi_a^{BR}(\pi_d^{ZD}),S^d)}{D(\pi_d^{ZD},\pi_a^{BR}(\pi_d^{ZD}),\mathbf{1})}-\frac{D(\pi_d^{SSE},\pi_a^{BR}(\pi_d^{SSE}),S^d)}{D(\pi_d^{SSE},\pi_a^{BR}(\pi_d^{SSE}),\mathbf{1})}=-(U_d^{SSE}-U_{12}^d).$$

因此，

$$\begin{aligned}&D(\pi_d^{ZD},\pi_a^{BR}(\pi_d^{ZD}),S^d)D(\pi_d^{SSE},\pi_a^{BR}(\pi_d^{SSE}),\mathbf{1})-D(\pi_d^{SSE},\pi_a^{BR}(\pi_d^{SSE}),S^d)D(\pi_d^{ZD},\pi_a^{BR}(\pi_d^{ZD}),\mathbf{1})\\=&-(U_d^{SSE}-U_{12}^d)D(\pi_d^{ZD},\pi_a^{BR}(\pi_d^{ZD}),\mathbf{1})D(\pi_d^{SSE},\pi_a^{BR}(\pi_d^{SSE}),\mathbf{1})\end{aligned}\tag{F1}$$

根据定理3，$U_d(\pi_d^{ZD},\ \pi_a^*)-U_d(\pi_d^{SSE},\ \pi_a^*)=U_{11}^a-\frac{U_{11}^a\pi_d^{SSE}(1|21)+U_{21}^a\pi_d^{SSE}(2|11)\pi_d^{SSE}(2|11)+\pi_d^{SSE}(1|21)}{}$。类似地，

$$\frac{D(\pi_d^{ZD},\pi_a^*,S^d)}{D(\pi_d^{ZD},\pi_a^*,\mathbf{1})}-\frac{D(\pi_d^{SSE},\pi_a^*,S^d)}{D(\pi_d^{SSE},\pi_a^*,\mathbf{1})}=U_{11}^a-\frac{U_{11}^a\pi_d^{SSE}(1|21)+U_{21}^a\pi_d^{SSE}(2|11)}{\pi_d^{SSE}(2|11)+\pi_d^{SSE}(1|21)}.\tag{F2}$$

回忆

$$U_d(\pi_d^{ZD},\lambda\pi_a^{BR}(\pi_d^{ZD})+(1-\lambda)\pi_a^*)=\frac{D(\pi_d^{ZD},\lambda\pi_a^{BR}(\pi_d^{ZD})+(1-\lambda)\pi_a^*,S^d)}{D(\pi_d^{ZD},\lambda\pi_a^{BR}(\pi_d^{ZD})+(1-\lambda)\pi_a^*,\mathbf{1})},$$

$$U_d(\pi_d^{SSE},\lambda\pi_a^{BR}(\pi_d^{SSE})+(1-\lambda)\pi_a^*)=\frac{D(\pi_d^{SSE},\lambda\pi_a^{BR}(\pi_d^{SSE})+(1-\lambda)\pi_a^*,S^d)}{D(\pi_d^{SSE},\lambda\pi_a^{BR}(\pi_d^{SSE})+(1-\lambda)\pi_a^*,\mathbf{1})}.$$

根据 [16]，$C(\pi_d^{ZD},\ \pi_d^{SSE},\ \pi_a^*,\ \lambda)\neq 0$，其定义见附录A。不失一般性，我们考虑 $C(\pi_d^{ZD},\ \pi_d^{SSE},\ \pi_a^*,\ \lambda)>0$。因此，

$$\begin{aligned}&U_d(\pi_d^{SSE},\lambda\pi_a^{BR}(\pi_d^{SSE})+(1-\lambda)\pi_a^*)-U_d(\pi_d^{ZD},\lambda\pi_a^{BR}(\pi_d^{ZD})+(1-\lambda)\pi_a^*)\\=&\frac{D(\pi_d^{SSE},\lambda\pi_a^{BR}(\pi_d^{SSE})+(1-\lambda)\pi_a^*,S^d)}{D(\pi_d^{SSE},\lambda\pi_a^{BR}(\pi_d^{SSE})+(1-\lambda)\pi_a^*,\mathbf{1})}-\frac{D(\pi_d^{ZD},\lambda\pi_a^{BR}(\pi_d^{ZD})+(1-\lambda)\pi_a^*,S^d)}{D(\pi_d^{ZD},\lambda\pi_a^{BR}(\pi_d^{ZD})+(1-\lambda)\pi_a^*,\mathbf{1})}\\=&\frac{1}{C(\pi_d^{ZD},\pi_d^{SSE},\pi_a^*,\lambda)}\left(D(\pi_d^{SSE},\lambda\pi_a^{BR}(\pi_d^{SSE})+(1-\lambda)\pi_a^*,S^d)D(\pi_d^{ZD},\lambda\pi_a^{BR}(\pi_d^{ZD})+(1-\lambda)\pi_a^*,\mathbf{1})\right.\\&\left.-D(\pi_d^{ZD},\lambda\pi_a^{BR}(\pi_d^{ZD})+(1-\lambda)\pi_a^*,S^d)D(\pi_d^{SSE},\lambda\pi_a^{BR}(\pi_d^{SSE})+(1-\lambda)\pi_a^*,\mathbf{1})\right).\end{aligned}$$

实际上，对于任意 $\pi_d\in\Delta\mathcal{D},\ \pi_a^1,\ \pi_a^2\in\Delta\mathcal{A},\ \lambda\in[0,1]$，以及$\mathbf{f}=[f_1,f_2,f_3,f_4]^T\in\mathbb{R}^4$，

$$\begin{aligned}&D(\pi_d,\lambda\pi_a^1+(1-\lambda)\pi_a^2,\mathbf{f})\\=&det\begin{bmatrix}\pi_d(1|11)(\lambda\pi_a^1(1|11)+(1-\lambda)\pi_a^2(1|11))-1 & \pi_d(1|11)-1 & \lambda\pi_a^1(1|11)+(1-\lambda)\pi_a^2(1|11)-1 & f_1\\ \pi_d(1|12)(\lambda\pi_a^1(1|12)+(1-\lambda)\pi_a^2(1|12)) & \pi_d(1|12)-1 & \lambda\pi_a^1(1|12)+(1-\lambda)\pi_a^2(1|12) & f_2\\ \pi_d(1|21)(\lambda\pi_a^1(1|21)+(1-\lambda)\pi_a^2(1|21)) & \pi_d(1|21) & \lambda\pi_a^1(1|21)+(1-\lambda)\pi_a^2(1|21)-1 & f_3\\ \pi_d(1|22)(\lambda\pi_a^1(1|22)+(1-\lambda)\pi_a^2(1|22)) & \pi_d(1|22) & \lambda\pi_a^1(1|22)+(1-\lambda)\pi_a^2(1|22) & f_4\end{bmatrix}\\=&det\begin{bmatrix}\pi_d(1|11)(\lambda\pi_a^1(1|11)+(1-\lambda)\pi_a^2(1|11))-1 & \pi_d(1|11)-1 & \lambda\pi_a^1(1|11)-\lambda & f_1\\ \pi_d(1|12)(\lambda\pi_a^1(1|12)+(1-\lambda)\pi_a^2(1|12)) & \pi_d(1|12)-1 & \lambda\pi_a^1(1|12) & f_2\\ \pi_d(1|21)(\lambda\pi_a^1(1|21)+(1-\lambda)\pi_a^2(1|21)) & \pi_d(1|21) & \lambda\pi_a^1(1|21)-\lambda & f_3\\ \pi_d(1|22)(\lambda\pi_a^1(1|22)+(1-\lambda)\pi_a^2(1|22)) & \pi_d(1|22) & \lambda\pi_a^1(1|22) & f_4\end{bmatrix}\\&+det\begin{bmatrix}\pi_d(1|11)(\lambda\pi_a^1(1|11)+(1-\lambda)\pi_a^2(1|11))-1 & \pi_d(1|11)-1 & (1-\lambda)\pi_a^2(1|11)-(1-\lambda) & f_1\\ \pi_d(1|12)(\lambda\pi_a^1(1|12)+(1-\lambda)\pi_a^2(1|12)) & \pi_d(1|12)-1 & (1-\lambda)\pi_a^2(1|12) & f_2\\ \pi_d(1|21)(\lambda\pi_a^1(1|21)+(1-\lambda)\pi_a^2(1|21)) & \pi_d(1|21) & (1-\lambda)\pi_a^2(1|21)-(1-\lambda) & f_3\\ \pi_d(1|22)(\lambda\pi_a^1(1|22)+(1-\lambda)\pi_a^2(1|22)) & \pi_d(1|22) & (1-\lambda)\pi_a^2(1|22) & f_4\end{bmatrix}\\=&\lambda^2D(\pi_d,\pi_a^1,\mathbf{f})+(1-\lambda)^2D(\pi_d,\pi_a^2,\mathbf{f})+\lambda(1-\lambda)(D(\pi_d,\pi_a^1,\pi_a^2,\mathbf{f})+D(\pi_d,\pi_a^2,\pi_a^1,\mathbf{f})),\end{aligned}$$

其中，$D(\pi_d,\pi_a^1,\pi_a^2,\mathbf{f})$ 如式（A1）所示。

根据上述方程，我们得到

$$
\begin{aligned}
&D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d) D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
=&\big(\lambda^2 D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d) + (1-\lambda)^2 D(\pi_d^{ZD}, \pi_a^*, S^d) + \lambda(1-\lambda)(D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \pi_a^*, S^d) \\
&+D(\pi_d^{ZD}, \pi_a^*, \pi_a^{BR}(\pi_d^{ZD}), S^d)\big) \times \big(\lambda^2 D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1}) + (1-\lambda)^2 D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) \\
&+\lambda(1-\lambda)(D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \pi_a^*, \mathbf{1}) + D(\pi_d^{SSE}, \pi_a^*, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1})\big) \\
=&\lambda^4 D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d) D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1}) + (1-\lambda)^4 D(\pi_d^{ZD}, \pi_a^*, S^d) D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) \\
&+\lambda^2(1-\lambda)^2 D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d) D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) + \lambda^2(1-\lambda)^2 D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d) D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) \\
&+\lambda^3(1-\lambda)\left(D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) + D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1}) J(\pi_d^{ZD}, \pi_a^*, S^d)\right) \\
&+\lambda(1-\lambda)^3\left(D(\pi_d^{ZD}, \pi_a^*, S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) + D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) J(\pi_d^{ZD}, \pi_a^*, S^d)\right) \\
&+\lambda^2(1-\lambda)^2 J(\pi_d^{ZD}, \pi_a^*, S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{1}).
\end{aligned} \tag{F3}
$$

同理,

$$
\begin{aligned}
&D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d) D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
=&\big(\lambda^2 D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d) + (1-\lambda)^2 D(\pi_d^{SSE}, \pi_a^*, S^d) + \lambda(1-\lambda)(D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \pi_a^*, S^d) \\
&+D(\pi_d^{SSE}, \pi_a^*, \pi_a^{BR}(\pi_d^{SSE}), S^d)\big) \times \big(\lambda^2 D(\pi_d^{ZD}, \pi_a^{BR}(\pi_a^{ZD}), \mathbf{1}) + (1-\lambda)^2 D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) \\
&+\lambda(1-\lambda)(D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \pi_a^*, \mathbf{1}) + D(\pi_d^{ZD}, \pi_a^*, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1})\big) \\
=&\lambda^4 D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d) D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1}) + (1-\lambda)^4 D(\pi_d^{SSE}, \pi_a^*, S^d) D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) \\
&+\lambda^2(1-\lambda)^2 D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d) D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) + \lambda^2(1-\lambda)^2 D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d) D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) \\
&+\lambda^3(1-\lambda)\left(D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) + D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1}) J(\pi_d^{SSE}, \pi_a^*, S^d)\right) \\
&+\lambda(1-\lambda)^3\left(D(\pi_d^{SSE}, \pi_a^*, S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) + D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) J(\pi_d^{SSE}, \pi_a^*, S^d)\right) \\
&+\lambda^2(1-\lambda)^2 J(\pi_d^{SSE}, \pi_a^*, S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{1}).
\end{aligned} \tag{F4}
$$

将上述两个方程相减,

$$
\begin{aligned}
&D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d) D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
&- D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d) D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
=&\lambda^4\left(D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d) D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1}) - D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d) D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1})\right) \\
&+(1-\lambda)^4\left(D(\pi_d^{SSE}, \pi_a^*, S^d) D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) - D(\pi_d^{ZD}, \pi_a^*, S^d) D(\pi_d^{SSE}, \pi_a^*, \mathbf{1})\right) - g,
\end{aligned}
$$

其中

$$
\begin{aligned}
g =&\lambda^3(1-\lambda)\left(D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) + D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1}) J(\pi_d^{ZD}, \pi_a^*, S^d)\right) \\
&-\lambda^3(1-\lambda)\left(D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) + D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1}) J(\pi_d^{SSE}, \pi_a^*, S^d)\right) \\
&+\lambda(1-\lambda)^3\left(D(\pi_d^{ZD}, \pi_a^*, S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) + D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) J(\pi_d^{ZD}, \pi_a^*, S^d)\right) \\
&-\lambda(1-\lambda)^3\left(D(\pi_d^{SSE}, \pi_a^*, S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) + D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) J(\pi_d^{SSE}, \pi_a^*, S^d)\right) \\
&+\lambda^2(1-\lambda)^2 J(\pi_d^{ZD}, \pi_a^*, S^d) J(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) \\
&-\lambda^2(1-\lambda)^2 J(\pi_d^{SSE}, \pi_a^*, S^d) J(\pi_d^{ZD}, \pi_a^*, \mathbf{1}).
\end{aligned} \tag{F5}
$$

回忆附录A中 B_1, B_2, B_3 和 B 的定义，我们有

$$
\begin{aligned}
|g| \leqslant& B\lambda^3(1-\lambda) + B\lambda(1-\lambda)^3 + 2B\lambda(1-\lambda)^2 \\
=&B\lambda(1-\lambda)(\lambda^2 + (1-\lambda)^2 + 2\lambda(1-\lambda)) \\
=&B\lambda(1-\lambda)(\lambda + (1-\lambda))^2 \\
=&B\lambda(1-\lambda).
\end{aligned} \tag{F6}
$$

回忆 (F1) 和 (F2)，可得

$$\begin{aligned}
&D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d)D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
&- D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d)D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
\geqslant &\lambda^4 \left(U_d^{SSE} - U_{12}^d\right) D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1}) D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1}) \\
&- (1-\lambda)^4 \left(U_{11}^d - \frac{U_{11}^d \pi_d^{SSE}(1|21) + U_{21}^d \pi_d^{SSE}(2|11)}{\pi_d^{SSE}(2|11) + \pi_d^{SSE}(1|21)}\right) D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) \\
&- B\lambda(1-\lambda)).
\end{aligned}$$

实际上，$D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) = \frac{1\cdot\pi_d^{ZD}(1|21)+1\cdot\pi_d^{ZD}(2|11)}{\pi_d^{ZD}(2|11)+\pi_d^{ZD}(1|21)} = 1$，且 $D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) = \frac{1\cdot\pi_d^{SSE}(1|21)+1\cdot\pi_d^{SSE}(2|11)}{\pi_d^{SSE}(2|11)+\pi_d^{SSE}(1|21)} = 1$。那么$D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1})D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1}) \leqslant D(\mathbf{1})$。因此，对于 $\lambda \in \Gamma_1$,

$$\begin{aligned}
&D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d)D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
&- D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d)D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
\geqslant &\lambda^4 \left(U_d^{SSE} - U_{12}^d\right) D(\mathbf{1}) - A(1-\lambda)^4 - B\lambda(1-\lambda).
\end{aligned}$$

对于任意 $\lambda \in \Gamma_1 = \{\lambda \in [0,1] | (U_d^{SSE} - U_{12}^d)D(\mathbf{1})\lambda^4 - A(1-\lambda)^4 - B\lambda(1-\lambda) \geqslant 0\}$，有 $U_d\left(\pi_d^{SSE}, \pi_d^\lambda(\pi_d^{SSE}, \pi_d^*)\right) \geqslant U_d\left(\pi_d^{ZD}, \pi_d^\lambda(\pi_d^{ZD}, \pi_d^*)\right)$. 此外,

$$\begin{aligned}
&U_d(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*) - U_d(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*) \\
=&\frac{D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d)}{D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1})} - \frac{D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d)}{D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1})} \\
=&\frac{1}{C(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda)} \big(D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d)D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
&- D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d)D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1})\big) \\
\leqslant&\frac{\left(U_d^{SSE} - U_{12}^d\right) C(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, 1)\lambda^4 - A(1-\lambda)^4 + B\lambda(1-\lambda)}{C(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda)} \\
=&H(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda),
\end{aligned}$$

其中，$H(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda)$ 在 (2) 中定义。因此,

$$\min_{\pi_d^{ZD} \in \Xi} U_d\left(\pi_d^{SSE}, \pi_a^\lambda(\pi_d^{SSE}, \pi_a^*)\right) - U_d\left(\pi_d^{ZD}, \pi_a^\lambda(\pi_d^{ZD}, \pi_a^*)\right) \leqslant \begin{cases} 0, & \text{if } U_{11}^d \geqslant U_{22}^d,\ U_{11}^a \geqslant U_{21}^a, \\ H(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda), & \text{otherwise.} \end{cases}$$

附录G：定理5的证明

根据引理1，防御者可以采用 ZD策略 $\pi_d^{ZD} = \pi^{ZD}(-k, \mathbf{1}, kU_{21}^d - U_{21}^a)$。与定理4证明中的分析类似，我们也可以得到 (F1) 和 (F2)。于是

$$\begin{aligned}
&\big(D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d)D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
&- D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d)D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1})\big) \\
=&\lambda^4 \left(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), S^d)D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1}) - D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), S^d)D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1})\right) \\
&+ (1-\lambda)^4 \left(D(\pi_d^{ZD}, \pi_a^*, S^d)D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) - D(\pi_d^{SSE}, \pi_a^*, S^d)D(\pi_d^{ZD}, \pi_a^*, \mathbf{1})\right) + g,
\end{aligned}$$

其中， g 如 (F5) 所示。根据 (F6)，由于 $|g| \leqslant B\lambda(1-\lambda)$，我们有

$$
\begin{aligned}
& D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d) D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
& - D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d) D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1}) \\
\geqslant & \lambda^4 \left(U_{12}^d - U_d^{SSE}\right) D(\pi_d^{ZD}, \pi_a^{BR}(\pi_d^{ZD}), \mathbf{1}) D(\pi_d^{SSE}, \pi_a^{BR}(\pi_d^{SSE}), \mathbf{1}) \\
& + (1-\lambda)^4 \left(U_{11}^d - \frac{U_{11}^d \pi_d^{SSE}(1|21) + U_{21}^d \pi_d^{SSE}(2|11)}{\pi_d^{SSE}(2|11) + \pi_d^{SSE}(1|21)}\right) D(\pi_d^{ZD}, \pi_a^*, \mathbf{1}) D(\pi_d^{SSE}, \pi_a^*, \mathbf{1}) \\
& - B\lambda(1-\lambda)) \\
\geqslant & \lambda^4 \left(U_{12}^d - U_d^{SSE}\right) D(\mathbf{1}) + A(1-\lambda)^4 - B\lambda(1-\lambda).
\end{aligned}
$$

因此，

$$
\begin{aligned}
& U_d(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*) - U_d(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*) \\
= & \frac{D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d)}{D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1})} - \frac{D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d)}{D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1})} \\
= & \frac{1}{C(\pi_d^{ZD}, \pi_d^{SSE}, \pi_a^*, \lambda)} \left(D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, S^d) D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, \mathbf{1})\right. \\
& \left. - D(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*, S^d) D(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*, \mathbf{1})\right).
\end{aligned}
$$

因此，

$$
U_d(\pi_d^{ZD}, \lambda\pi_a^{BR}(\pi_d^{ZD}) + (1-\lambda)\pi_a^*) \geqslant U_d(\pi_d^{SSE}, \lambda\pi_a^{BR}(\pi_d^{SSE}) + (1-\lambda)\pi_a^*),
$$

由此得出结论。

附录 H：算法

我们在应用部分展示了上述算法的详细内容。这里，我们根据玩家行动历史，分别采用基于 [28] 的虚拟博弈方法和基于 [7] 的 Q 学习方法，为攻击者制定最佳响应策略。

算法 1 有限理性攻击者的虚拟博弈 **输入：** 理性因子： λ，顽固策略： π_a^*。 **初始化：** 对于所有 $d \in \mathcal{D}$、 $s \in \mathcal{S}$，防御者的策略频率 $\hat{\pi}_d(d|s) = 0$，其平均收益为 $U_d = 0$。 **1：** 对于 $t = 1$、2、$\cdots$，执行 **2：** 防御者采取 $d_t \sim \pi_{d\circ}$ **3：** 攻击者采取 $a_t \sim \pi_{a\circ}$ **4：** 到达状态 $s(t)$，玩家获得收益 $r_d(d_t, a_t)$ 和 $r_a(d_t, a_t)$。 **5：** $\hat{\pi}_d(d_t|s_{t-1}) = (_{t-1})\hat{\pi}_d(_{d_t|s_{t-1}+1})$。 **6：** $\pi_a = \lambda BR(\hat{\pi}_d) + (1-\lambda)\pi_{a\circ}^*$ **7：** $U_d = \sum_{t}^{} r_d(_{d_i,a_i})_{t\circ i=1}$ **8：** 结束循环

算法 2 有限理性攻击者的 **Q-learning** **输入：** 理性因子： λ，顽固策略： π_a^*、 ϵ_1和 $\epsilon_{2\circ}$ **初始化：** $Q(s, a) = 0$，其中 $s \in \mathcal{S}$、$a \in \mathcal{A}$、 $\bar{r}_a = 0$和 $U_d = 0$。 **1：** 对于 $t = 1$、**2**、$\cdots$，执行 **2：** 防御者采取 $d_t \sim \pi_{d\circ}$ **3：** 攻击者根据 $Q(s(t-1)\ b)$，采用（$1-\lambda$）-贪心策略采取 $a_{t\circ}$ **4：** 到达状态 $s(t)$。获取 $r_d(d_t, a_t)$和 $r_a(d_t, a_t)$。 **5：** $\delta = r_a(d_t, a_t) - \bar{r}_a + \max_{a'} Q(s(t)\ a') - Q(s(t-1)\ a_t)$。 **6：** $Q(s(t-1)\ a_t) = Q(s(t-1)\ a_t) + \epsilon_1\delta_\circ$ **7：** 如果 $Q(s(t-1)\ a_t) = \max_{b'} Q(s(t-1)\ a)$，则 **8：** $\bar{r}_a = (1-\epsilon_2)\bar{r}_a + \epsilon_2\,{}^{(t-1)\bar{r}_a + r_a}(_{d_t,a_t})_{t\circ}$ **9：** 结束条件判断 **10：** $U_d = \sum_{t}^{} r_d(_{d_i,a_i})_{t\ \circ i=1}$ **11：** 结束循环

[1] K. Xiao、C. Zhu、J. Xie、Y. Zhou、X. Zhu 和 W.Zhang，《信息物理系统中针对隐蔽恶意软件传播的动态防御策略》，载于 *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications* (2018) 第 1790–1798 页。[2] X. Feng、 Z. Zheng、 D. Cansever、 A. Swami 和 P. Mo-

hapatra，移动目标防御的信号博弈模型，载于 *IEEE INFOCOM 2017-IEEE Conference on Computer Communications* 2017 1 9 (IEEE，) pp. – 。[3] E. Bondi、H. Oh、H. Xu、F. Fang、B. Dilkina和M. Tambe，是否发出信号：在安全与可持续性信号博弈中利用不确定的实时信息，载于 *Proceedings of the AAAI Conference on Artificial Intelligence* 34 2020 1369 ，第( ) 卷，pp. – 1377。[4] D. Mutzari、J. Gan和S. Kraus，多防御者安全博弈中的联盟形成，载于 *Proceedings of the AAAI Conference on Artificial Intelligence* 2021 ( ) pp. 5603–5610。[5]Y. Vorobeychik和S. Singh，折扣随机博弈中斯塔克尔伯格均衡的计算，载于 *Proceedings of the AAAI Conference on Artificial Intelligence* 26 ，第 (2012)卷，pp. 1478–1484。[6] D. Korzhyk、Z. Yin、C. Kiekintveld、V. Conitzer和M. Tambe，安全博弈中的斯塔克尔伯格与纳什：对可互换性、等价性和唯一性的扩展研究，《人工智能研究杂志》**41**，297（2011）。[7] K. Li和D. Hao，重复公共物品博弈中的合作执行与抗合谋性，载于 *Proceedings of the AAAI Conference on Artificial Intelligence* 33 ，第 (2019)卷，pp. 2085–2092。[8] Z. Cheng、G. Chen和Y. Hong，基于超博弈框架的单领导者—多跟随者斯塔克尔伯格安全博弈，《IEEE信息取证与安全汇刊》**17**，954（2022）。[9] H. A. Simon，有限理性，载于 *Utility and Probability* 1990 15 18 (施普林格，) pp. – 。[10] A. X. Jiang、T. H. Nguyen、M. Tambe和A. D. Procaccia，单调极大极小：针对有限理性跟随者的稳健斯塔克尔伯格解，载于 *International Conference on Decision and Game Theory for Security* 2013 119 139 (施普林格，) pp. – 。[11] M. Carvalho和R. Ford，计算机网络的移动目标防御，《IEEE安全与隐私》**12**，73（2014）。[12] G. Chen、Y. Ming、Y. Hong和P. Yi，具有不确定耦合约束的 ε-广义纳什均衡分布式算法，《自动化》**123**，109313（2021）。[13] Q. D. La、T. Q. Quek、J. Lee、S. Jin和H. Zhu，物联网蜜罐网络中的欺骗攻击与防御博弈，《IEEE物联网期刊》**3**，1025（2016）。[14] K. Nayak、S. Kumar、A. Miller和E. Shi，顽固挖矿：自私挖矿的推广及其与日蚀攻击的结合，载于 *2016 IEEE European Symposium on Security and Privacy* 2016 305 320 (IEEE，) pp. – 。[15] A. Sanjab 和 W. Saad，信息物理系统安全中的有限理性：应用于智能电网保护的博弈论分析，载于 *2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids* 2016 1 6 (IEEE，) pp. – 。[16] W. H. Press 和 F. J. Dyson，迭代囚徒困境包含能够击败任何进化对手的策略，《美国国家科学院院刊》**109**，10409（2012）。[17] Z. Wang、Y. Zhou、J. W. Lien、J. Zheng和B. Xu，在迭代囚徒困境中，敲诈胜过慷慨，《自然·通讯》**7**，1（2016）。[18] C. Hilbe

，M. A. Nowak ，和 K. Sigmund ，进化

关于迭代囚徒困境博弈中的敲诈，《美国国家科学院院刊》**110**，6913（2013）。[19] A. Govaert 和 M. Cao，具有折扣收益的重复多人社会困境中的零行列式策略，IEEE 自动控制汇刊 **66**，4575（2021）。[20] D. Hao、Z. Rong 和 T. Zhou，不确定性下的敲诈：噪声博弈中的零行列式策略，《物理评论E》**91**，052803（2015）。[21] A. Engel 和 A. Feigel，用于冲突升级且不涉及信息传递的单一均衡策略，《物理评论E》**98**，012415（2018）。[22] A. Szolnoki 和 M. Perc，结构化种群中敲诈行为的演化，《物理评论E》**89**，022804（2014）。[23] X. Chen、L. Wang 和 F. Fu，零行列式策略的几何结构，arXiv 预印本 arXiv:2208.02486（2022）。[24] M. A. Taha 和 A. Ghoneim，重复非对称博弈中的零行列式策略，《应用数学与计算》**369**，124862（2020）。[25] A. McAvoy 和 C. Hauert，非对称演化博弈，《公共科学图书馆·计算生物学》**11**，e1004349（2015）。[26] K. Sooksatra、W. Li、B. Mei、A. Alrawais、S. Wang 和 J. Yu，利用零行列式策略解决非对称不完全信息下的数据交易困境，载于 *International Conference on Wireless Algorithms Systems and Applications* 2018 425 437 （施普林格，）第 – 页。[27] S. Wang、H. Shi、Q. Hu、B. Lin 和 X. Cheng，基于零行列式理论的物联网移动目标防御，IEEE《物联网期刊》**7**，661（2019）。[28] S. Qiu、X. Wei、J. Ye、Z. Wang 和 Z. Yang，结构化转移零和马尔可夫博弈中可证明高效的虚拟博弈策略优化，载于 *International Conference on Machine Learning*（机器学习研究论文集，2021）第 8715–8725 页。[29] Q. Guo、J. Gan、F. Fang、L. Tran-Thanh、M. Tambe 和 B. An，论安全博弈中斯塔克尔伯格均衡的可诱导性，载于 *Proceedings of the AAAI Conference on Artificial Intelligence* 33 2019 2020 2028 ，卷 （ ）第 – 页。[30] E. Akin，迭代囚徒困境：良好策略及其动力学，《遍历理论》，《动力系统进展》，77（2016）。[31] A. Mamiya 和 G. Ichinose，重复博弈中存在观测误差时的零行列式策略，《物理评论E》**102**，032115（2020）。[32] T. Nguyen 和 H. Xu，斯塔克尔伯格安全博弈中的模仿型攻击者欺骗，载于 *Proceedings of the 28th International Joint Conference on Artificial Intelligence*（2019）第 528–534 页。[33] Q. Guo、B. An、B. Bo˘sansk´y 和 C. Kiekintveld，比较安全博弈中的策略性保密与斯塔克尔伯格承诺，载于 *Proceedings of the 26th International Joint Conference on Artificial Intelligence* 2017 （ ）第 3691–3699 页。[34] R. Zhuang、S. A. DeLoach 和 X. Ou，迈向移动目标防御理论，载于 *Proceedings of the first ACM Workshop on Moving Target Defense* 2014 31 40 （ ）第 – 页。[35] X. Shan 和 J. Zhuang，在多阶段、多目标序贯博弈中对抗策略型攻击者的累积防御资源分配建模，《可靠性工程与系统安全》**179**，12（2018）。[36] Y. Gal 和 Z. Ghahramani，将丢弃法作为贝叶斯近似：在深度学习中表示模型不确定性，载于（机器学习研究论文集，）第 – 页。

[37] A. ˙Zychowski 和 J. Ma´ndziuk，在安全博弈中学习攻击者的有限理性模型，载于 *International Conference on Neural Information Processing*（施普林格，2021）第 530–539 页。

[38]C. Hilbe、A. Traulsen 和 K. Sigmund，《伙伴还是对手？迭代囚徒困境的策略》，《博弈与经济行为》**92**，41（2015）。

[39] Z. Cheng、G. Chen 和 Y. Hong，误判对迭代囚徒困境中零行列式策略的影响，《科学报告》**12**，1（2022）。