Imitative Follower Deception in Stackelberg Games*

Jiarui Gan[1], Haifeng Xu[2], Qingyu Guo[3], Long Tran-Thanh[4],
Zinovi Rabinovich[3], and Michael Wooldridge[1]

[1] University of Oxford
jiarui.gan@cs.ox.ac.uk, michael.wooldridge@cs.ox.ac.uk
[2] Harvard University
hxu@seas.harvard.edu
[3] Nanyang Technological University
QGUO005@e.ntu.edu.sg, zinovi@ntu.edu.sg
[4] University of Southampton
ltt08r@ecs.soton.ac.uk

Abstract. Uncertainty is one of the major challenges facing applications of game theory. In the context of Stackelberg games, various approaches have been proposed to deal with the leader's incomplete knowledge about the follower's payoffs, typically by gathering information from the leader's interaction with the follower. Unfortunately, these approaches rely crucially on the assumption that the follower will *not* strategically exploit this information asymmetry, i.e. the follower behaves truthfully during the interaction with respect to their actual payoffs. As we show in this paper, the follower may have strong incentives to deceitfully imitating the behavior of a different follower type and, in doing this, benefit significantly from subverting the leader into choosing a highly suboptimal strategy. This raises a fundamental question: how to design a leader strategy in the presence of such strategic followers? To answer this question, we put forward a basic model of Stackelberg games with *(imitative) follower deception* and show that the leader is indeed able to reduce the loss due to follower deception with carefully designed strategies. We then provide a systematic study of the problem of computing the optimal leader strategy and draw a relatively complete picture of the complexity landscape; essentially matching positive and negative complexity results are provided for natural variants of the model. Our intractability results are in sharp contrast to the situation with no deception, where the leader's optimal strategy can be computed in polynomial time, and thus illustrate the intrinsic difficulty of handling follower deception. Through simulations we also demonstrate empirically the benefit of considering follower deception.

1 Introduction

Recently, there is a growth of interest in Stackelberg games in the AI community. This trend is driven in part by a number of high-impact real-world applications in security domains [23]. Beyond that, Stackelberg game models also find many noteworthy applications in other problems, such as principal-agent contract design [15,24] and exam

* Jiarui Gan was supported by the EPSRC International Doctoral Scholars Grant EP/N509711/1. Zinovi Rabinovich's contribution was supported by the NTU SUG grant "Choice Manipulation and Security Games".

design for large-scale tests [11]. In a Stackelberg game, a leader commits to a mixed strategy and a follower best-responds after observing the leader's strategy. The Stackelberg equilibrium yields the optimal strategy the leader can commit to in this framework. As in many other game-theoretic applications, a key real-world challenge facing applications of Stackelberg games is that the leader may not have full information about the follower's payoffs for computing the equilibrium. To address this issue, various approaches have been proposed: When the leader can estimate the follower's payoffs to within certain intervals, leader strategies that are robust against small interval uncertainties are studied [10,9,12]; When the leader knows a distribution about the follower's payoffs/types, Bayesian Stackelberg equilibria are computed to maximize the leader's average utility [4,16,8]; When the leader can interact with the follower, a line of work examines how to learn the optimal leader strategy to commit to from the follower's best responses from the interaction [10,1,7,20,17].

Despite their differences, the above approaches share a common and crucial step — obtaining payoff-relevant information about the follower. However, knowing the leader's attempt at information gathering, a strategic follower would gain incentives to intentionally distort the information learned by the leader, in particular by deceitfully imitating the behavior of a different follower type — a phenomenon we term as *imitative follower deception*. Unfortunately, all aforementioned approaches ignore the possibility of such strategic behavior of the follower and adopt a simplistic assumption that any information the algorithm gathers about the follower is truthful.

As we will show in this paper, algorithms designed under the above assumption can be easily manipulated by the follower and may produce highly suboptimal leader strategy as a result. Consider perhaps the most basic learning setting where the follower has an uncertain type (i.e., a set of payoffs) $\theta \in \Theta$; knowing the set Θ of all possible types but not the actual θ, the leader wants to learn the optimal strategy against this follower. If the follower is truthful, the optimal strategy that the leader can learn would be the strong Stackelberg equilibrium against θ. However, when the follower may strategically deceive by behaving according to another type θ' (he may indeed have such an incentive, as we will show), the leader can only learn the optimal strategy against the fake type θ', which can be highly suboptimal. A fundamental question then is: what is the optimal strategy that the leader can learn or design, when facing a deceitful follower?

In this paper, we put forward a basic Stackelberg model in an attempt to formalize this question and aim to understand how imitative follower deception affects the leader's choice of strategies. We note that, even though there are other types of follower deception as well, imitative deception is arguably the simplest and most basic deception approach; we expect it to happen the most often in practice. In our model, the leader faces a follower with a type that is drawn from a finite set Θ. The follower knows his true type, but the leader only has a probabilistic belief about it. The leader commits to a *policy*, which is a "menu" that specifies the mixed strategy to play for each follower type. Knowing the leader's commitment, a follower of true type $\theta \in \Theta$ can imitate another type $\theta' \in \Theta$ — behaving consistently according to the payoffs of θ' — if this improves his utility. The leader learns the fake type θ' and plays the strategy specified by the policy for θ'. The optimal leader policy maximizes the leader's expected utility, taking into account the follower's deception in the loop, which is precisely the optimal leader strategy that can be learned in the deceptive setting.

Our Contribution. We provide a systematic study of computational aspects of the above model. We prove that either with or without *incentive compatibility* (IC) constraints, the optimal leader policy cannot be approximated within a meaningful ratio, unless P = NP. These results illustrate that deception is the fundamental reason of the hardness: when there is no deception, the problem can be solved in polynomial time simply by computing the optimal commitment to each follower type separately. Nevertheless, the leader's optimal policy can be formulated as an MILP (mixed integer linear program). Next, we observe that the leader can further improve her utility by randomizing her commitment (to mixed strategies) for each follower type. We call such a randomized commitment a *mixed policy*. We prove that it remains NP-hard to approximate the optimal mixed policy, but, interestingly, the problem becomes polynomial-time computable when IC constraints are posed. Finally, we demonstrate the benefit of considering follower deception via empirical evaluations.

Applications and Related Work. Our model can be applied directly to a number of applications. For example, in principal-agent contract design (a widely studied Stackelberg model in economics [15,24,3]), a principal may present a menu of contracts to an uncertain agent and ask the agent to choose the one that matches his true type. Naturally, the principal's menu must take into account the agent's misreport of his type. Similar ideas apply to exam design for large-scale tests, e.g., for MOOCs (massive open online courses), which have also been modeled as Stackelberg games [11]. In these situations, the tester may need to use different exams for different types of test takers who usually come from various education backgrounds and have different learning objectives. Again, exam designers will also need to take into account exam takers' misreport about their type to strategically conceal their strength in the hope of preferred tests.

Deception has been extensively studied in Stackelberg games, particularly its application in security domains [2,18,28,27,26,6,21]. However, all these works consider designing deceptive strategies for the *leader* (usually, a defender). In these models, it is the leader who has the informational advantage, whereas in our model the follower has the greater decision-pertinent knowledge and discloses it strategically. To the best of our knowledge, there has been very limited work studying *follower* deception in Stackelberg games. The most relevant to ours is perhaps [25], which studies signaling in Bayesian Stackelberg games (BSGs) and also considered the possibility that a follower may strategically misreport his type. However, the question they study is completely different from ours — they design signaling schemes for different follower types whereas we design the leader's mixed strategy for each follower type. The contrasting complexity between their model (polynomial-time solvable in normal form BSGs) and ours (NP-hard) also highlights the intrinsic difference. Even more importantly, though, in our research type reporting takes the conceptual form of a behaviour inducing message. In other words, we reverse the common information asymmetry (see e.g. [19,26,5]) between the leader and the follower.

2 A Motivating Example

We illustrate how the follower's deceptive behavior may result in highly suboptimal leader strategy, and how we can overcome this issue using carefully-designed leader

strategies. For illustrative purpose, we use a real-world security game example, though the underlying phenomenon also generalizes to other applications.

A security game is a Stackelberg game played between a *defender* (the leader) and an *attacker* (the follower). Concretely, we consider a security agency who wants to protect two conservation areas, i.e., areas 1 and 2, from a *poacher*'s attack. The defender can only patrol one of these two areas, while the poacher chooses one area to attack. At different periods of the year, the poacher's payoffs change due to price fluctuation of wildlife products on black markets. We capture this uncertainty using two possible types of poacher payoffs, i.e., *type* A and B. Assume that the defender's payoff is *independent* of the poacher's type. The poacher knows his true type, but the defender only has a prior belief that each type shows up with probability 0.5 (This can be obtained via, e.g., surveying past prices on black markets). The following table shows the payoff matrices of the defender (row player) and the two poacher types (column player).

	$1^{\times}$	$2^{\times}$
1	1	-1
2	-1	0.99

defender

	$1^{\times}$	$2^{\times}$
1	-1	$1/3$
2	3	-1

poacher type A

	$1^{\times}$	$2^{\times}$
1	-1	1
2	1	-1

poacher type B

The only difference between these poacher types is their value for successful attacks, which is affected by wildlife product prices at that time period. Standard calculation shows that if the poacher has type A, the optimal defender strategy is $(3/4, 1/4)$, i.e., patrolling areas 1 and 2 with probability $3/4$ and $1/4$, respectively, resulting in poacher utility 0 at both areas. By the standard assumption, the poacher thus breaks tie *in favor* of the defender and attacks area 1,[5] yielding defender utility $1/2$. Similarly, the optimal defender strategy is $(1/2, 1/2)$ against a type-B poacher, which induces the poacher to also attack area 1 and yields defender utility 0.

Ideally, the defender would like to play the optimal patrolling strategy against the poacher's true type at any time which, however, is unknown to the defender. Nevertheless, if the poacher were to behave truthfully according to his type, the defender can easily learn the poacher's type, e.g., by observing their (different) best responses to strategy $(3/5, 2/5)$. Indeed, previous approaches for learning the optimal leader strategy rely crucially on the assumption that the follower will truthfully respond (e.g., [10,1,7,20]). However, in this example, a type-A poacher has strong incentives to be untruthful: by imitating a type-B poacher and reacting exactly according to the type-B payoff structure, a type-A poacher is able to induce the defender to play $(1/2, 1/2)$ — the optimal strategy against a type-B poacher — which results in true utility $\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot (3) = 1$

[5] This assumption is without loss of generality. The corresponding solution concept, the *strong Stackelberg equilibrium*, remains the most widely adopted solution concept in the literature of Stackelberg games (see e.g., [14,4,23]). Normally, via infinitesimal strategy variations, the leader can induce the follower to play *any* best response action to the leader's benefit. For example, the defender can play $(\frac{3}{4} - \epsilon, \frac{1}{4} + \epsilon)$ with $\epsilon \to 0$, so that the poacher will strictly prefer to attack area 1 while the change to the defender's utility can be arbitrarily small. Our empirical evaluation in Section 6 will further confirm the robustness of our solutions against the tie-breaking issue.

for the type-A poacher. This strictly improves upon his previous utility 0 of reacting truthfully. Simple calculation shows that a type-B poacher does not benefit from deception and will behave truthfully. As a result, the defender will not be able to distinguish type A from B under the poacher's strategic deception.

The issue raised above is due to the fact that the defender focuses solely on optimizing her utility without considering the poacher's strategic behavior. We now show how the defender can overcome this issue by taking into account the poacher's deception. It happens that in this example the optimal policy for the defender is to still play $(3/4, 1/4)$ for type A and $(1/2, 1/2)$ for type B, but induce a type-B poacher to *break tie "against" the defender* by attacking area 2.[6] This slightly decreases the defender's utility against a type-B poacher (from 0 to -0.005), but it erases the incentive of deception of a type-A poacher. This is because if the type-A poacher attacks area 2 under defender strategy $(1/2, 1/2)$ now, his expected utility becomes -1 which is worse than his utility 0 under truthful behavior.

Remarks. (i) This example illustrates an intriguing phenomenon: when there is follower deception, it may be undesirable to always induce tie breaking in favor of the leader, since other tie breaking strategies may enforce more desirable follower behaviors. If the defender adopts this nuanced strategy, the poacher will have no incentive to deceive. Thus, from the perspective of machine learning, what we are designing can be viewed as the best strategy the leader can learn in the presence of follower deception.

(ii) One might wonder why the defender does not ignore the poacher's deception by simply playing a single strategy against all types of poachers, which is precisely the problem of optimal commitment in Bayesian Stackelberg games. This is because our more sophisticated approach can achieve better defender utility. In this example, the defender would play $(1/2, 1/2)$ in the Bayesian Stackelberg equilibrium, obtaining utility 0. However, the strategy we designed above yields defender utility $\frac{1}{4} - \frac{1}{400}$. A more formal result regarding the utility improvement will be presented in Proposition 1.

(iii) Perhaps counter-intuitively, follower deception is not always bad for the leader. Consider, e.g., a Stackelberg game version of the prisoner's dilemma shown on the right, where the row player is the leader and the column player the follower. If the follower deceives the leader into believing that he has utility 1, instead of -1, for the action profile (C, C) (annotated in red), the leader would feel reassured to commit to C, resulting in utility -1 for both players.

	C	D
C	$-1, -1$ (1)	$-3, \ 0$
D	$0, -3$	$-2, -2$

3 The Model

3.1 Stackelberg Game Basics

A Stackelberg game (SG) is played between a *leader* and a *follower*. A normal-form SG is given by two matrices $u^L, u^F \in \mathbb{R}^{m \times n}$, which are the payoff matrices of the leader (row player) and follower (column player), respectively. We use $u^L(i, j)$ to denote a generic entry of u^L, and use $[m] = \{1, ..., m\}$ to denote the set of the leader's pure strategies (also called *actions*). A leader *mixed* strategy is a probabilistic distribution

[6] This tie breaking can be induced by playing $(\frac{1}{2} + \epsilon, \frac{1}{2} - \epsilon)$ for an arbitrarily small ϵ.

over $[m]$, denoted by a vector $\mathbf{x} \in \Delta_m = \left\{\mathbf{p} \geq 0 : \sum_{i\in[m]} p_i = 1\right\}$. With slight abuse of notation, we denote by $u^L(\mathbf{x}, j)$ the expected utility $\sum_{i\in[m]} x_i \cdot u^L(i, j)$. Notation for the follower is the same by changing labels. We will sometimes write a pure strategy i in positions where a mixed strategy is expected, when we mean the pure strategy in its mixed strategy form, i.e., a basis vector $\mathbf{e}_i$.

In a Stackelberg game, the leader moves first by committing to a mixed strategy $\mathbf{x}$, and the follower then best responds to $\mathbf{x}$. W.l.o.g., it is assumed that the follower's best response is a pure strategy with ties, if any, broken in favor of the leader; such responses can be induced via an infinitesimal deviation from $\mathbf{x}$ [22]. Under this assumption, the leader mixed strategy that maximizes her expected utility leads to a *strong Stackelberg equilibrium* (SSE), which is the standard solution concept of Stackelberg games. Formally, denote by $BR(\mathbf{x}) := \arg\max_{j\in[n]} u^F(\mathbf{x}, j)$ the set of follower best responses to the leader strategy $\mathbf{x}$; the pair of strategies $\mathbf{x}^*$ and j^* forms an SSE iff

$$\langle \mathbf{x}^*, j^* \rangle \in \arg\max_{\mathbf{x}\in\Delta_m, j\in BR(\mathbf{x})} u^L(\mathbf{x}, j).$$

3.2 Stackelberg Game with Imitative Follower Deception

We now describe a basic model that captures (imitative) follower deception; the reader may refer back to our motivating example as an instantiation of this model. We consider a leader who faces a follower with an uncertain type, which falls in a discrete set Θ. Each type $\theta \in \Theta$ corresponds to a different follower payoff matrix $u^F_\theta \in \mathbb{R}^{m\times n}$. For ease of presentation, we will assume that the leader's payoff does *not* depend on θ, though we remark that all our results easily generalize to the case with type-dependent leader payoffs. To model the leader's prior knowledge regarding the follower's type, we adopt the classic Bayesian perspective and assume that the leader has a prior distribution π, i.e., each type θ appears with probability π_θ. Note that π, Θ and the payoff matrices are common knowledge.

We assume that the leader can commit to a *policy* — a "menu" that specifies a mixed strategy $\mathbf{x}^\theta$ to be played against follower type θ. In addition, we assume a follower best response $j^\theta \in BR_\theta(\mathbf{x}^\theta)$ is also specified in the policy, where $BR_\theta(\mathbf{x}) = \arg\max_j u^F_\theta(\mathbf{x}, j)$ denotes the *best response set* of a type-θ follower against leader strategy $\mathbf{x}$; hence, in the case of a tie, the follower will be induced to play this specific response. Note that, as in the motivating example, j^θ is not necessarily the one that maximizes the leader's utility with *this particular type* (unlike in the standard Stackelberg game model); rather, the leader may prefer to induce the follower to a carefully chosen action that discourages follower deception and improves the leader's overall utility.

After observing the leader's policy, a follower of type θ deceptively reports a type $\hat{\theta}$ to the leader, and the leader then plays the mixed strategy $\mathbf{x}^{\hat{\theta}}$ as specified by her policy. The follower then "best-responds" with $j^{\hat{\theta}}$ as if his type is $\hat{\theta}$. Naturally, the follower will report the $\hat{\theta}$ that maximizes her true expected utility, i.e., $\hat{\theta} = \arg\max_{\theta'\in\Theta} u^F_\theta(\mathbf{x}^{\theta'}, j^{\theta'})$, resulting in the leader to obtain utility $u^L(\mathbf{x}^{\hat{\theta}}, j^{\hat{\theta}})$. The reporting step in our model is straightforward in various applications (see Section 1), whereas in some other applications, reporting abstracts the process that the leader learns the follower's type by interacting with the follower (as in the motivating example).

1. The leader commits to a policy $\sigma = \{o^\theta\}_{\theta\in\Theta}$ that prescribes, for each reported follower type $\theta \in \Theta$, an *outcome* $o^\theta = \langle \mathbf{x}^\theta, j^\theta\rangle$ such that $j^\theta \in BR_\theta(\mathbf{x}) := \arg\max_j u_\theta^F(\mathbf{x}, j)$.
2. After observing the leader policy σ, a follower of type θ, who appears with probability π_θ, reports a best type $\hat{\theta}(\sigma) = \arg\max_{\theta'\in\Theta} u_\theta^F(\mathbf{x}^{\theta'}, j^{\theta'})$; same as the standard SSE assumption, we assume the follower breaks ties by reporting a type in favor of the leader. This results in expected leader utility:

$$U^L(\sigma) := \textstyle\sum_{\theta\in\Theta} \pi_\theta \cdot u^L(\mathbf{x}^{\hat{\theta}(\sigma)}, j^{\hat{\theta}(\sigma)}).$$

The computational task we examine in this paper is to compute the σ that maximizes U^L. For convenience, we will sometimes write $u_\theta^F(o) = u_\theta^F(\mathbf{x}, j)$ and $u^L(o) = u^L(\mathbf{x}, j)$ for an outcome $o = \langle \mathbf{x}, j\rangle$.

Incentive Compatibility (IC). One can impose in addition the IC constraints on σ so that reporting truthfully is a (weakly) dominant strategy for every follower type, i.e., $u_\theta^F(\mathbf{x}^\theta, j^\theta) \geq u_\theta^F(\mathbf{x}^\beta, j^\beta)$ for any $\beta \in \Theta$. We will consider model variants both with and without IC constraints. The following result is due to the fact that playing the same mixed strategy against all follower types is trivially IC, and it offers the same leader utility as the BSE does.

Proposition 1 *There always exists an IC policy, and the optimal IC policy achieves at least as much leader utility as that achieved in the Bayesian Stackelberg equilibrium.*

4 Computing the Optimal Policy

In this section, we study the complexity and algorithms for computing the optimal leader policy. We will refer to the problem of computing the optimal policy OPT and, when IC constraints are imposed, OPT-IC.

4.1 Hardness of Approximation

We show in Theorems 2 and 3 that it is NP-hard even just to approximate the optimal policy to within a meaningful ratio, with or without IC constraints, and even when the follower has only a small number of actions. These inapproximability results are essentially tight by Theorem 4. In contrast, an efficient algorithm can indeed be found for a small number of follower types, which we will show in Theorem 5 after we present an algorithm for the general case. All our approximations are *multiplicative*. We shift the leader's payoffs to be non-negative in order to analyze multiplicative ratios.

Theorem 2 *For any constant $\epsilon > 0$,* OPT *does not admit any polynomial-time $\frac{1}{(|\Theta|-1)^{1-\epsilon}}$-approximation algorithm unless P = NP, even when the follower has only three actions.*

Proof. We show a reduction from the MAX-INDEPENDENT-SET problem, which asks for the size of the maximum independent set in a graph $G = (V, E)$. A set of nodes $V' \subseteq V$ is an *independent set* of G if no edge in E connects any pair of nodes in V'; an independent set is maximum if there exists no other independent set with a larger

size. It is known that no polynomial-time $\frac{1}{|V|^{1-\epsilon}}$-approximation algorithm exists for MAX-INDEPENDENT-SET, unless P = NP [29].

We construct a Stackelberg game with $|V|+1$ follower types $\Theta = \{\theta_*\} \cup \{\theta_v : v \in V\}$, where each θ_v corresponds to node v. Let $\pi_{\theta_*} = 0$ and $\pi_\theta = \frac{1}{|\Theta|-1}$ for all other θ. The leader has $2|V|+1$ actions $\{a_v : v \in V\} \cup \{b_v : v \in V\} \cup \{a_0\}$; and the follower has three actions $\{1^\times, 2^\times, 3^\times\}$. The payoffs are given below, where $\mathcal{N}(v)$ denotes the set of neighbouring nodes of v, and all the empty entries are 0.

	$1^\times$	$2^\times$	$3^\times$
a_0	0.5	1	1
a_v		0.5	0.5
b_v		1	1
$a_{v'} : v' \in \mathcal{N}(v)$		1	1
otherwise		0.5	1

follower type θ_v

	$1^\times$	$2^\times$	$3^\times$
any i	1		

follower type θ_*

	$1^\times$	$2^\times$	$3^\times$
any i	1		

leader

The following are a few observations about the game:

- For follower type θ_*, $j = 1^\times$ strictly dominates all the other actions. A follower can only be induced to play $j = 1^\times$ when θ_* is reported, and the leader gets utility 1.
- For each θ_v, when they are reported, $j = 1^\times$ cannot be induced by any leader strategy because it is strictly dominated by $j = 2^\times$. As a result, the leader gets 0 once θ_v is reported.
- Given the above, the leader's overall utility is proportional to the number of follower types θ_v that are motivated to report θ_*.

Now for any independent set V' of size k in G, we can construct the following leader policy σ: $o^{\theta_*} = \langle a_0, 1^\times \rangle$, $o^{\theta_v} = \langle a_v, 2^\times \rangle$ if $v \in V'$, and $o^{\theta_v} = \langle b_v, 2^\times \rangle$ if $v \notin V'$. In this policy, all types θ_v with $v \in V'$ find reporting θ_* optimal, while all the others are better-off reporting truthfully. As a result, the leader receives overall utility $\frac{k}{|\Theta|-1}$.

Conversely, we claim that for any leader policy σ, nodes v's corresponding to types θ_v's that are incentivized to report θ_* must form an independent set. To see this, suppose for a contradiction that θ_v and $\theta_{v'}$ are both incentivized to report θ_* while $(v, v') \in E$. By our observations above, both θ_v and $\theta_{v'}$ will be induced to play $1^\times$ and obtain utility 0.5 each. Thus, we need the following in order for θ_v and $\theta_{v'}$ to have no incentive to report a different type.

(i) $x^{\theta_{v'}}_{a_{v'}} = 0$, since otherwise θ_v would be better-off reporting $\theta_{v'}$ and obtain utility strictly greater than 0.5.
(ii) $x^{\theta_{v'}}_{b_{v'}} = 0$ and $x^{\theta_{v'}}_{a_{v''}} = 0$ for all $v'' \in \mathcal{N}(v')$, since otherwise $\theta_{v'}$ would be better-off reporting truthfully and obtain utility strictly greater than 0.5.

It follows that, now $\mathbf{x}^{\theta_{v'}}$ can only (and must) pick the last row of the payoff matrix with nonzero probability, so $3^\times$ becomes the only best response of a type-$\theta_{v'}$ player against $\mathbf{x}^{\theta_{v'}}$. Thus, we have $o^{\theta_{v'}} = \langle \mathbf{x}^{\theta_{v'}}, 3^\times \rangle$, in which case, however, $\theta_{v'}$ is able to

obtain utility 1 by reporting truthfully, which contradicts the assumption that $\theta_{v'}$ is incentivized to report θ^*.

Therefore, the number of follower types that are motivated to report θ_* by the optimal policy is exactly to the size of the maximum independent set (and proportional to the overall leader utility). Any $\frac{1}{(|\Theta|-1)^{1-\epsilon}}$-approximation algorithm would also provide a $\frac{1}{|V|^{1-\epsilon}}$-approximation to the MAX-INDEPENDENT-SET problem, which exists unless P = NP. This completes the proof.

Next, we show that the same inapproximability result with OPT-IC. Our reduction will still be from MAX-INDEPENDENT-SET but the underlying idea and the instance constructed are both different: the previous reduction uses the follower types that *violate* IC as an indicator of the independent set, but this approach is infeasible when IC constraints are imposed.

Theorem 3 *For any constant $\epsilon > 0$,* OPT-IC *does not admit a polynomial-time $\frac{1}{|\Theta|^{1-\epsilon}}$-approximation algorithm unless P = NP, even when the follower has only three actions.*

Proof. We show a reduction from the MAX-INDEPENDENT-SET problem. Given an instance of MAX-INDEPENDENT-SET specified by a graph $G = (V, E)$, we construct a game with $|V|$ follower types $\Theta = \{\theta_v : v \in V\}$; each θ_v corresponds to a node v and appears with probability $\frac{1}{|\Theta|}$. The leader has $2|V|$ actions $\{a_v : v \in V\} \cup \{b_v : v \in V\}$; and the follower has three actions $\{1^{\times}, 2^{\times}, 3^{\times}\}$. The payoffs are given below, where $\mathcal{N}(v)$ denotes the set of neighbouring nodes of v, and all the empty entries are 0.

	$1^{\times}$	$2^{\times}$	$3^{\times}$
a_v			
b_v		1	1
$a_{v'} : v' \in \mathcal{N}(v)$	0.5		1
otherwise			1

follower type θ_v

	$1^{\times}$	$2^{\times}$	$3^{\times}$
any i	1		

leader

Now for any independent set V' of size k, consider a leader policy σ with: $o^{\theta_v} = \langle a_v, 1^{\times} \rangle$ if $v \in V'$, and $o^{\theta_v} = \langle b_v, 2^{\times} \rangle$ if $v \notin V'$. In this policy, all types find truthful report to be optimal. Hence, the policy is IC and offers the leader overall utility $\frac{k}{|\Theta|}$.

Conversely, for any IC policy σ, we claim that the corresponding nodes of types that are motivated to respond with $j = 1^{\times}$ always form an independent set. Suppose for the sake of a contradiction that θ_v but $\theta_{v'}$ both respond with $j = 1^{\times}$ and $(v, v') \in E$. Observe that, when θ_v is reported, $j = 1^{\times}$ can be induced only by leader strategy a_v because it is strictly dominated by $j = 3^{\times}$ in all other cases. Thus, $o^{\theta_v} = \langle a_v, 1^{\times} \rangle$, in which case, however, $\theta_{v'}$ would be motivated to report θ_v, which contradicts the assumption that σ is IC.

Therefore, for the optimal IC policy, the number of follower types that respond with $j = 1^{\times}$ (which is proportional to the leader's overall utility) is equal to the size of the maximum independent set in G. Any $\frac{1}{|\Theta|^{1-\epsilon}}$-approximation algorithm would also provide a $\frac{1}{|V|^{1-\epsilon}}$-approximation to the MAX-INDEPENDENT-SET problem.

We show that the above inapproximability bounds in Theorems 2 and 3 are essentially tight by exhibiting an efficient algorithm with matching approximation guarantee.

Theorem 4 OPT *and* OPT-IC *admit a polynomial-time* $\frac{1}{|\Theta|}$*-approximation algorithm.*

Proof. We show an approximation algorithm for OPT. The algorithm for OPT-IC follows a similar procedure.

We enumerate all (true) follower types $\theta \in |\Theta|$ and compute a policy — call it σ_θ — that maximizes only the leader's utility obtained on this specific follower type θ. We show below that (i) such a policy can be computed in polynomial time, and (ii) the best such policy achieves at least $\frac{1}{|\Theta|}$ of the leader utility of the optimal policy.

(i) Polynomial-time computability. To compute σ_θ, the idea is to enumerate all possible reporting strategies $\beta \in \Theta$ of type θ and, for each β, compute the best policy *under the constraint that it is indeed optimal for a type-θ follower to report β*. An observation that helps simplifying the computation is that we only need to guarantee reporting β to be (weakly) preferred by the follower to reporting θ, so it suffices to consider the two associated outcomes in the leader's policy, i.e., $o^\theta = \langle \mathbf{x}^\theta, j^\theta \rangle$ and $o^\beta = \langle \mathbf{x}^\beta, j^\beta \rangle$. This is due to the fact that a policy that plays $\mathbf{x}^\theta$ on all other reported types $\beta' \notin \{\theta, \beta\}$ will trivially make type θ — and hence type β if it is preferred to θ — a better reporting strategy for the follower than all $\beta' \notin \{\theta, \beta\}$ (same as the argument for proving Proposition 1). Now, the problem reduces to solving the following liner program for each pair of possible values of j^θ and j^β, which is in $[n]^2$, so the tractability follows immediately:

$$\begin{aligned}
\max_{\mathbf{x}^\theta, \mathbf{x}^\beta \in \Delta_m} \quad & u^L(\mathbf{x}^\beta, j^\beta) && && (1)\\
\text{s.t.} \quad & u^F_\theta(\mathbf{x}^\theta, j^\theta) \geq u^F_\theta(\mathbf{x}^\theta, j) && \forall\, j \in [n] && (1a)\\
& u^F_\beta(\mathbf{x}^\beta, j^\beta) \geq u^F_\beta(\mathbf{x}^\beta, j) && \forall\, j \in [n] && (1b)\\
& u^F_\theta(\mathbf{x}^\beta, j^\beta) \geq u^F_\theta(\mathbf{x}^\theta, j^\theta) && && (1c)
\end{aligned}$$

Namely, we maximize the leader's utility obtained on a type-θ follower under the condition that the follower is to report β. Here, the first two constraints guarantee that $j^\theta \in \mathit{BR}_\theta(\mathbf{x})$ and $j^\beta \in \mathit{BR}_\beta(\mathbf{x})$ as required by the definition of leader policy; and the third constraint requires that reporting β is indeed a better choice for a type-θ follower than reporting θ.

(ii) Utility guarantee. It remains to show that the best σ_θ computed above offers the desired utility. Let μ^L_θ be the best solution to Program (1) over all $j^\theta, j^\beta \in [n]$; and let $\theta^* = \arg\max_{\theta \in \Theta} \pi_\theta \cdot \mu^L_\theta$. Now consider an arbitrary feasible policy ς with $\langle \mathbf{y}^\theta, k^\theta \rangle$ being the outcome it specifies for each reported type θ. As defined previously, we let $\hat{\theta}(\varsigma)$ denote a type-θ follower's best reporting strategy against policy ς. It holds that $\mu^L_\theta \geq u^L(\mathbf{y}^{\hat{\theta}(\varsigma)}, k^{\hat{\theta}(\varsigma)})$ because $\mathbf{x}^\theta = \mathbf{y}^\theta$ and $\mathbf{x}^\beta = \mathbf{y}^{\hat{\theta}(\varsigma)}$ are feasible under Constraints (1a)–(**??**) when j^θ and j^β are set to k^θ and $k^{\hat{\theta}(\varsigma)}$, respectively. It follows that, for any feasible σ,

$$U^L(\sigma_{\theta^*}) \geq \pi_{\theta^*} \cdot \mu^L_{\theta^*} \geq \tfrac{1}{|\Theta|} \textstyle\sum_\theta \pi_\theta \cdot \mu^L_\theta \geq \tfrac{1}{|\Theta|} \sum_\theta \pi_\theta \cdot u^L(\mathbf{x}^{\hat{\theta}(\varsigma)}, j^{\hat{\theta}(\varsigma)}) = \tfrac{1}{|\Theta|} \cdot U^L(\varsigma).$$

Since the choice of ς is arbitrary, σ_{θ^*} serves as an $\frac{1}{|\Theta|}$-approximation of the optimal policy. The proof is completed.

4.2 An MILP Formulation and Tractability for Small Θ

Given the hardness result, efficient algorithms for computing the optimal policy would seem unlikely. We show an algorithm based on an MILP (mixed integer linear program) as a practical approach to tackle the problem. Besides its practical significance, the MILP formulation also forms the basis of a polynomial-time algorithm for OPT and OPT-IC, which we show in Theorem 5.

First, when IC constraints are not imposed, the following program (not linear yet) computes the optimal leader policy, where μ_θ^L, $\mathbf{x}^\theta$, y_β^θ, and p_j^θ are variables.

$$\max \quad \textstyle\sum_\theta \pi_\theta \cdot \mu_\theta^L \tag{2}$$

$$\text{s.t.} \quad \mu_\theta^L \le \textstyle\sum_j p_j^\beta \cdot u^L(\mathbf{x}^\beta, j) + (1 - y_\beta^\theta) \cdot M \qquad \forall \theta, \beta \in \Theta \tag{2a}$$

$$p_j^\theta \cdot \left[u_\theta^F(\mathbf{x}^\theta, j) - u_\theta^F(\mathbf{x}^\theta, k)\right] \ge 0 \qquad \forall\, \theta \in \Theta,\ j, k \in [n] \tag{2b}$$

$$\textstyle\sum_j p_j^\beta \cdot u_\theta^F(\mathbf{x}^\beta, j) - \sum_j p_j^\gamma \cdot u_\theta^F(\mathbf{x}^\gamma, j) \ge -(1 - y_\beta^\theta) \cdot M \qquad \forall\, \theta, \beta, \gamma \in \Theta \tag{2c}$$

$$\textstyle\sum_i x_i^\theta = \sum_j p_j^\theta = \sum_\beta y_\beta^\theta = 1 \qquad \forall\, \theta \in \Theta \tag{2d}$$

$$\mathbf{x}^\theta \in [0,1]^m,\ \ y_\beta^\theta \in \{0,1\},\ \ p_j^\theta \in \{0,1\} \qquad \forall\, \theta, \beta \in \Theta,\ j \in [n] \tag{2e}$$

Here, the leader policy is represented by variables $\mathbf{x}^\theta$ and p_j^θ, where $\mathbf{x}^\theta$ is the leader mixed strategy prescribed for report θ, and p_j^θ captures the induced follower action. Through Constraint (2b), it is guaranteed that $p_j^\theta = 1$ if j is a best response of type θ to $\mathbf{x}^\theta$ and $p_j^\theta = 0$, otherwise. Similarly, y_β^θ captures the optimal reporting strategy of type θ, i.e., $y_\beta^\theta = 1$ only if reporting β is optimal for θ, and this is guaranteed via Constraint (2c), in which M is a sufficiently large constant. Finally, each μ_θ^L captures the utility the leader obtains from true type θ via Constraint (2a).

This program is not yet an MILP because of the quadratic terms $p_j^\theta \cdot \mathbf{x}^\theta$. To linearize it, we replace these terms with a set of new variables $\tilde{x}_{ji}^\theta$, subject to $0 \le \tilde{x}_{ji}^\theta \le x_i^\theta$ (for all θ, j, i) and $\sum_j \tilde{x}_{ji}^\theta = p_j^\theta$ (for all θ, i). This guarantees $\tilde{x}_{ji}^\theta = p_j^\theta \cdot x_i^\theta$ given that p_j^θ is restricted to be in $\{0,1\}$.

In the situation with IC constraints, we simply force $y_\theta^\theta = 1$ for all $\theta \in \Theta$ in the above MILP. The "windfall" of the MILP formulation is a polynomial-time algorithm for small $|\Theta|$ shown below.

Theorem 5 *When $|\Theta|$ is fixed, both* OPT *and* OPT-IC *can be solved in polynomial time.*

Proof. Following the MILP, when $|\Theta|$ is fixed, the feasible space of the binary variables $\mathbf{y}$ and $\mathbf{p}$ has size $poly(m, n)$, i.e., we have $\langle \mathbf{y}, \mathbf{p} \rangle \in \{0,1\}^{|\Theta|} \times \{0,1\}^{n \cdot |\Theta|}$. We can enumerate every element in this space and solve an MILP with $\mathbf{y}$ and $\mathbf{p}$ fixed to the corresponding values, which is now a linear program with no integer variable.

5 Generalization to Mixed Policy

The leader policy we have considered so far consists of one *outcome* $\langle \mathbf{x}^\theta, j^\theta \rangle$ for each θ. In this section, we expand the space of leader policies, and allow the leader to *randomize*

over the outcomes for each follower type. In other words, the leader's policy will now be *a distribution of* outcomes for each follower type. We call a distribution of outcomes a *mixture*, and call a generalized leader policy consisting of mixtures a *mixed policy*, as opposed to policies in the previous sections, which we will henceforth call *pure policies*.

Now the game proceeds as follows. The leader first commits to a mixed policy. Then the follower observes the mixed policy and reports a type θ. Finally, the leader will sample an outcome $\langle \mathbf{x}^\theta, j^\theta \rangle$ from the mixture, play $\mathbf{x}^\theta$, and induce the follower to respond with j^θ.

Obviously, the leader utility of an optimal mixed policy is at least that of a pure one for all pure policies are also special mixed policies. However, a natural question is why further randomization over the outcomes will benefit the leader since the leader is already playing a mixed strategy in every outcome, and randomizing mixed strategies typically will not bring any extra power for a player. Interestingly, it turns out that this extra randomization will be beneficial for the leader when there is follower deception. The reason is that such randomization can help us tame the incentive constraints for the follower's type reporting, while the randomization in mixed strategies is responsible for inducing desired follower action responses. The following example gives a more concrete illustration.

Example: Strict Improvement of Mixed Policies. We consider again a security game example with two targets 1 and 2. The defender can patrol one of these targets and the attacker chooses one of them to attack. The attacker has three types θ_*, θ_A and θ_B which appear with equal probability $1/3$. The payoffs are given below.

	$1^\times$	$2^\times$
1	1	-1
2	-1	1

defender

	$1^\times$	$2^\times$
	-1	1
	1	-1

attacker type θ_*

	$1^\times$	$2^\times$
	0	0
	1	-2

attacker type θ_A

	$1^\times$	$2^\times$
	-2	1
	0	0

attacker type θ_B

Note that the two targets are of equal importance to the defender and attacker type θ_* (the game against type θ_* is zero-sum). Therefore, θ_* will always attack the less patrolled target, and for an outcome $\langle \mathbf{x}, j \rangle$ corresponding to type θ_*, it always holds that $x_j \leq 1/2$. Suppose w.l.o.g. that $j = 1^\times$. Then type θ_A is able to obtain at least $1/2$ by misreporting θ_*, in which case: if θ_A does misreport his type, the defender obtains at most 0 on θ_A; if θ_A reports his true type, the defender's policy needs to offer him utility at least $1/2$ to incentivize this truthful report, so target 2 needs to be patrolled with probability at least $1/2$ and, again, the defender obtains at most 0. The same argument applies to type θ_B if we suppose $j = 2^\times$. As a result, any pure policy obtains utility more than 0 only on (at most) one of types θ_A and θ_B, in which case the defender's overall utility is at most $1/3$.

Now we consider a *mixed policy* as follows:

- For reported attacker type θ_*, choose outcomes $\langle (1/2, 1/2), 1^\times \rangle$ and $\langle (1/2, 1/2), 2^\times \rangle$, each with probability $1/2$;
- For reported attacker type θ_A, choose $\langle (1, 0), 1^\times \rangle$ with probability 1; and for θ_B, choose $\langle (0, 1), 2^\times \rangle$ with probability 1.

As such, types θ_A and θ_B can only obtain $-1/2$ in expectation if they misreport θ_*. This incentivizes both θ_A and θ_B to report truthfully, yielding expected defender utility $2/3$ — an improvement of at least $1/3$ compared to the optimal pure policy.

Mixtures with Support Size n Suffice. To compute the optimal mixed policy, one challenge is that the mixture for a follower type θ may be supported on a set of infinite size (as the space of feasible $\langle \mathbf{x}^\theta, j^\theta \rangle$'s is infinite). Thus, it is not even clear whether the optimal mixed policy can be represented in polynomial size. Fortunately, the following result implies that it suffices to consider mixtures supported on at most n outcomes; each outcome induces the follower to choose a distinct action in $[n]$.

Proposition 6 *For any feasible mixture prescribed for reported type θ, there is a feasible mixture with support size $\tilde{n} \leq n$ that yields the same utility for the leader and any follower who reports θ.*

Proof (sketch). The proof follows a standard revelation-principle-type argument. Given any mixture, we can merge all outcomes that induce the same follower response into one outcome; this will not change either the leader's or the follower's expected utility.

Following this result, we can represent the mixture for each type θ as a vector $(p_1^\theta, \ldots, p_n^\theta) \in \Delta_n$ together with n mixed strategies $\mathbf{x}_1^\theta, \ldots, \mathbf{x}_n^\theta$; the mixture samples each outcome $\langle \mathbf{x}_j^\theta, j \rangle$ with probability $p_j^\theta \geq 0$. Note that this representation allows outcomes involved to be invalid because it is possible that some follower action j cannot be induced by any leader strategy, i.e., $j \notin BR_\theta(\mathbf{x})$ for all $\mathbf{x} \in \Delta_m$. Thus, a mixture is valid only if $j \in BR_\theta(\mathbf{x}_j^\theta)$ for all j such that $p_j^\theta > 0$.

5.1 Computing the Optimal Mixed Policy

Proposition 6 implies that there exists an optimal mixed policy of polynomial size. Nevertheless, Theorem 7 below shows that the problem of computing the optimal mixed policy (referred to as OPTX) remains to be inapproximable in the case *without* IC constraints. The hardness of inapproximability is essentially tight following a similar argument as in Theorem 4; we state the result in Theorem 8 but omit the proof. Surprisingly, after we impose IC constraints, the problem (OPTX-IC) becomes *tractable*.

Theorem 7 *For any constant $\epsilon > 0$,* OPTX *does not admit any polynomial-time $\frac{1}{(|\Theta|-1)^{1-\epsilon}}$-approximate algorithm unless P = NP, even when the follower has only three actions.*

Proof (sketch). The inapproximability can be shown from a reduction from the same MAX-INDEPENDENT-SET problem as in the proof of Theorem 2.

Theorem 8 OPTX *admits a polynomial-time $\frac{1}{|\Theta|}$-approximation algorithm.*

We note that an MILP formulation for OPTX can be devised, by modifying Program (2) as follows. First, replace every $\mathbf{x}^\theta$ by $\mathbf{x}_j^\theta$, wherever they are associated with p_j^θ. Then relax every p_j^θ to be in $[0, 1]$. The program can be linearized using the same technique applied to Program (1), and similarly this provides a polynomial-time algorithm for OPTX when the number of follower types is fixed (Theorem 9; proof omitted).

If we further impose IC constraints, the optimal mixed policy can be computed in polynomial time via a linear program. This is because the IC constraints further remove the integer variables $y^{\theta}_{\theta'}$'s in the MILP described above (Theorem 10; proof omitted).

Theorem 9 *When $|\Theta|$ is fixed,* OPTX *can be solved in polynomial time.*

Theorem 10 OPTX-IC *can be computed in polynomial time.*

6 Empirical Evaluation

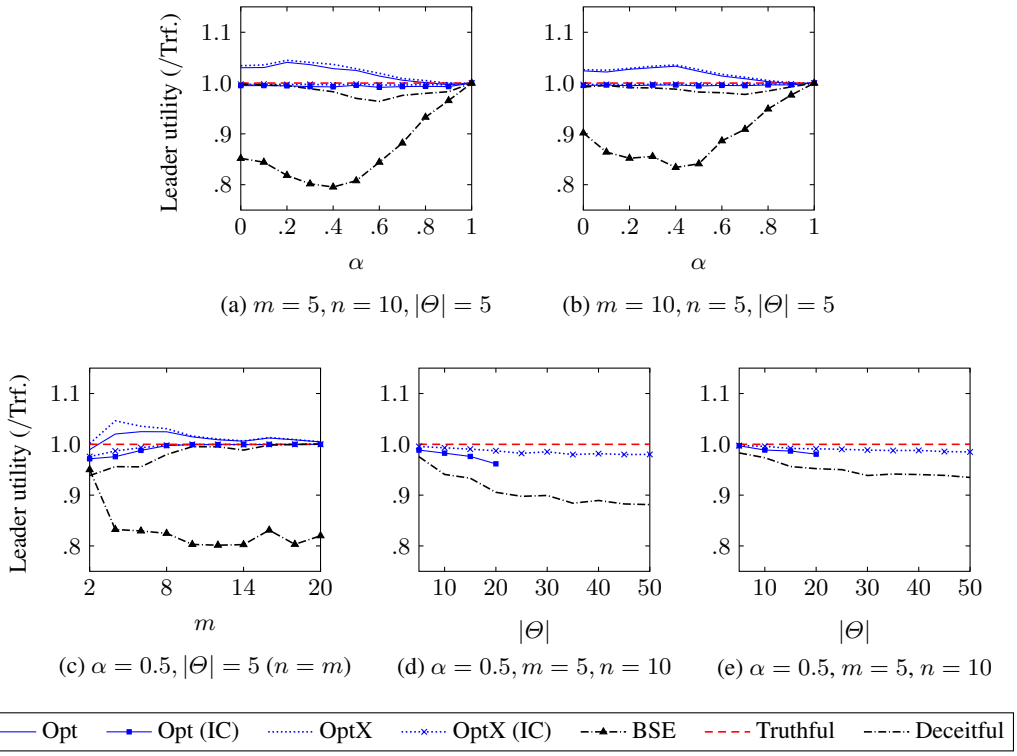
We evaluate our framework with games generated randomly using the well-known covariance game model [13]. We generate the players' payoffs uniformly at random from the range $[0, 1]$, and then adjust the follower's payoffs by blending them with the negative value of the leader's payoffs; the degree of blending is controlled by a parameter $\alpha \in [0, 1]$, i.e., $u^F_\theta \leftarrow (1-\alpha) \cdot u^F_\theta - \alpha \cdot u^L$. As such, the game is zero-sum when $\alpha = 1$; and, when $\alpha = 0$, payoffs of the leader and the follower are completely uncorrelated. We also generate the probabilities π_θ's uniformly at random.

We compare the leader's utility obtained by different policies/approaches proposed in this paper and other existing work, including: (1) Optimal pure policy (labeled *Opt*), with and without IC; (2) Optimal mixed policy (*OptX*), with and without IC; (3) Bayesian Stackelberg Equilibrium (*BSE*); (4) Optimal leader strategy when the follower truthfully reports his type (*Truthful*), and (5) the same strategy but when the follower strategically reports his type (*Deceitful*).

Figure 1 depicts our results. The first two figures, (a) and (b), show leader utility with varying α, from which we see that all our proposed polices (lines in blue) improve the leader's utility significantly upon BSE. Our policies achieve leader utility that is very close to the truthful utility, even when IC constraints are imposed. When IC is not required, the optimal policies perform even better in almost all experiments. Surprisingly, even when the leader naïvely trusts a deceitful follower, the utility obtained may sometimes be higher than the truthful utility. However, this is in line with our observation in Section 3, that follower deception is not always bad for the leader.

The utility differences are the most significant when α is around 0.5. Intuitively, this is because when α approaches 1 the diminishing uncertainty over follower types reduces the effect of follower deception. But when α is close to 0, payoffs of the leader and the follower are less correlated and their objectives less conflicted; hence, when the follower leverages deception to increase his utility, there is a lower chance that this will decrease the leader's utility.

We further investigate the effects of varying number of actions and number of follower types. From Figure 1 (c) we see that utility differences (excluding BSE) quickly diminish when the number of actions increase. Similar to the situation when α is close to 0, the level of randomness of the payoffs increases as the action space grows, which dissolves the negative effect of follower deception. In contrast, as shown in Figure 1 (d) and (e), the effect of follower deception appear to increase with the number of follower types as leader utility in the deceitful setting drops. Utility offered by the optimal mixed IC policy, however, decreases much slower. Hence, the optimal mixed IC policy offers a scalable practical solution when there is a large number of follower types. At a high



(a) $m = 5, n = 10, |\Theta| = 5$ (b) $m = 10, n = 5, |\Theta| = 5$

(c) $\alpha = 0.5, |\Theta| = 5$ $(n = m)$ (d) $\alpha = 0.5, m = 5, n = 10$ (e) $\alpha = 0.5, m = 5, n = 10$

Fig. 1: Leader utility obtained with different approaches, varying with α as in (a) and (b), with m and n as in (c), and with $|\Theta|$ as in (d) and (e). All y-axes represent leader utility as a *ratio to the utility obtained in the truthful situation*. The missing data points are instances that cannot be solved within one hour because of the scalability issue. All data points are averages of 50 runs.

level, these results suggest that uncertainty over follower's payoff information is a major amplifier of the negative effect of follower deception.

7 Conclusion

In this paper, we point out potential occurrences of (imitative) follower deception in Stackelberg games and the risk when they are ignored. We then propose a framework to design leader policies against such deception and provide a systematic study of the computational aspects of this framework. Our results shows that handling follower deception is hard in general, even when the searching of optimal policy is expanded into the less complicated space of mixed policies.

There are a number of potential directions for future work. Perhaps the first ones to investigate are several extended settings: e.g., when there is no prior knowledge about the distribution of follower types, or when follower types fall into a continuous space instead of a finite one as in our model. It would also be interesting to position this work in a specific application area, such as security games where follower deception might be more common, yet more harmful, because of the adversarial nature of the followers.

References

1. Blum, A., Haghtalab, N., Procaccia, A.D.: Learning optimal commitment to overcome insecurity. In: Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14). pp. 1826–1834 (2014)
2. Brown, G., Carlyle, M., Diehl, D., Kline, J., Wood, K.: A two-sided optimization for theater ballistic missile defense. Oper. Res. **53**(5), 745–763 (Sep 2005)
3. Caldentey, R., Haugh, M.: A cournot-stackelberg model of supply contracts with financial hedging and identical retailers. Foundations and Trends® in Technology, Information and Operations Management **11**(1-2), 124–143 (2017)
4. Conitzer, V., Sandholm, T.: Computing the optimal strategy to commit to. In: Proceedings of the 7th ACM Conference on Electronic Commerce (EC'06). pp. 82–90 (2006)
5. Dughmi, S., Kempe, D., Qiang, R.: Persuasion with limited communication. In: Proceedings of the ACM Conference on Economics and Computation (EC'16). pp. 663–680 (2016)
6. Guo, Q., An, B., Bosansky, B., Kiekintveld, C.: Comparing strategic secrecy and stackelberg commitment in security games. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI'17) (2017)
7. Haghtalab, N., Fang, F., Nguyen, T.H., Sinha, A., Procaccia, A.D., Tambe, M.: Three strategies to success: Learning adversary models in security games. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16). pp. 308–314 (2016)
8. Jain, M., Pita, J., Tambe, M., Ordóñez, F., Paruchuri, P., Kraus, S.: Bayesian stackelberg games and their application for security at los angeles international airport. SIGecom Exch. **7**(2), 10:1–10:3 (Jun 2008)
9. Kiekintveld, C., Islam, T., Kreinovich, V.: Security games with interval uncertainty. In: Proceedings of The 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'13). pp. 231–238. International Foundation for Autonomous Agents and Multiagent Systems (2013)
10. Letchford, J., Conitzer, V., Munagala, K.: Learning and approximating the optimal strategy to commit to. In: International Symposium on Algorithmic Game Theory. pp. 250–262. Springer (2009)
11. Li, Y., Conitzer, V.: Game-theoretic question selection for tests. Journal of Artificial Intelligence Research **59**, 437–462 (2017)
12. Nguyen, T.H., Yadav, A., An, B., Tambe, M., Boutilier, C.: Regret-based optimization and preference elicitation for stackelberg security games with uncertainty. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI'14). pp. 756–762 (2014)
13. Nudelman, E., Wortman, J., Shoham, Y., Leyton-Brown, K.: Run the gamut: A comprehensive approach to evaluating game-theoretic algorithms. In: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2. pp. 880–887. IEEE Computer Society (2004)
14. Osborne, M.J., Rubinstein, A.: A course in game theory. MIT press (1994)
15. Page, F.H.: Optimal contract mechanisms for principal-agent problems with moral hazard and adverse selection. Economic Theory **1**(4), 323–338 (1991)
16. Paruchuri, P., Pearce, J.P., Marecki, J., Tambe, M., Ordonez, F., Kraus, S.: Playing games for security: An efficient exact algorithm for solving bayesian stackelberg games. In: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'08). pp. 895–902 (2008)
17. Peng, B., Shen, W., Tang, P., Zuo, S.: Learning optimal strategies to commit to. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19) (2019)
18. Powell, R.: Allocating defensive resources with private information about vulnerability. American Political Science Review **101**(04), 799–809 (2007)

19. Rabinovich, Z., Jiang, A.X., Jain, M., Xu, H.: Information disclosure as a means to security. In: Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'15). pp. 645–653 (2015)
20. Roth, A., Ullman, J., Wu, Z.S.: Watch and learn: Optimizing from revealed preferences feedback. In: Proceedings of the 48th annual ACM symposium on Theory of Computing (STOC'16). pp. 949–962. ACM (2016)
21. Schlenker, A., Thakoor, O., Xu, H., Fang, F., Tambe, M., Tran-Thanh, L., Vayanos, P., Vorobeychik, Y.: Deceiving cyber adversaries: A game theoretic approach. In: Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS'18). pp. 892–900 (2018)
22. Stengel, B.v., Zamir, S.: Leadership with commitment to mixed strategies. CDAM Research Report LSE-CDAM-2004-01, London School of Economics (2004)
23. Tambe, M.: Security and game theory: algorithms, deployed systems, lessons learned. Cambridge University Press (2011)
24. Van Long, N., Sorger, G.: A dynamic principal-agent problem as a feedback stackelberg differential game. Central European Journal of Operations Research **18**(4), 491–509 (2010)
25. Xu, H., Freeman, R., Conitzer, V., Dughmi, S., Tambe, M.: Signaling in bayesian stackelberg games. In: Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16). pp. 150–158 (2016)
26. Xu, H., Rabinovich, Z., Dughmi, S., Tambe, M.: Exploring information asymmetry in two-stage security games. In: Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI'15). pp. 1057–1063 (2015)
27. Yin, Y., An, B., Vorobeychik, Y., Zhuang, J.: Optimal deceptive strategies in security games: A preliminary study (2013)
28. Zhuang, J., Bier, V.M.: Reasons for secrecy and deception in Homeland-Security resource allocation. Risk Analysis **30**(12), 1737–1743 (2010)
29. Zuckerman, D.: Linear degree extractors and the inapproximability of max clique and chromatic number. In: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing. pp. 681–690. ACM (2006)