Imitative Attacker Deception in Stackelberg Security Games *

Thanh Nguyen[1†] and **Haifeng Xu**[2]
[1]University of Oregon
[2]Harvard University
thanhhng@cs.uoregon.edu, hxu@seas.harvard.edu

Abstract

To address the challenge of *uncertainty* regarding the *attacker*'s payoffs, capabilities and other characteristics, recent work in security games has focused on learning the optimal defense strategy from observed attack data. This raises a natural concern that the strategic attacker may mislead the defender by *deceptively* reacting to the learning algorithms. This paper focuses on understanding how such *attacker deception* affects the game equilibrium. We examine a basic deception strategy termed *imitative deception*, in which the attacker simply pretends to have a different payoff assuming his *true* payoff is unknown to the defender. We provide a clean characterization about the game equilibrium as well as optimal algorithms to compute the equilibrium. Our experiments illustrate significant defender loss due to imitative attacker deception, suggesting the potential side effect of learning from the attacker.

1 Introduction

In AI research, there has been an increasing interest in the application of Stackelberg models in addressing security challenges, also known as *Stackelberg security games* (SSGs). This interest is driven in part by a number of high-impact deployed security applications [Tambe, 2011]. In these real-world domains, an important challenge facing the security agency (the *defender*) is her *uncertainty* about the *attacker*'s capabilities, payoffs, and behavior, etc. To address this challenge, various learning techniques have been proposed to learn the attacker's characteristics or the defender's optimal strategy from observed attack data [Letchford *et al.*, 2009; Marecki *et al.*, 2012; Blum *et al.*, 2014; Balcan *et al.*, 2015; Haghtalab *et al.*, 2016; Nguyen *et al.*, 2016; Kar *et al.*, 2017; Gholami *et al.*, 2017; Xu *et al.*, 2016; Peng *et al.*, 2019]. A crucial assumption in these works is that the attacker always responds *honestly* to the defender's algorithm so that the *true* attacker characteristics are learned. However, given the competitive nature of the interaction, this assumption may rarely hold in practice — the strategic attacker may manipulate his reactions to mislead the learning algorithm towards an outcome that favors the attacker. Such concern of *attacker deception* motivates the central research question of this paper:

In SSGs with a defender facing an unknown attacker, how would the attacker's deception affect the equilibrium?

This paper initiates the study of attacker deception in SSGs with *unknown attacker payoff*. We investigate a basic deception strategy termed *imitative deception* in which the attacker simply pretends to have some payoff (which may differ from his true one) and *always* plays consistently to this deceptive payoff. As a result, through learning, the defender eventually commits to an optimal defense strategy against the attacker's deceptive payoff. The attacker aims to find an *optimal deceptive payoff* such that it leads to a defender equilibrium strategy that maximizes the attacker's *true utility*. Such deception can happen when, e.g., the defender seeks to learn the optimal strategy against the attacker [Blum *et al.*, 2014; Gholami *et al.*, 2017; Peng *et al.*, 2019]. Moreover, imitative deception is easy for the attacker to implement. Thus, we believe it serves as a natural and important first step towards a general understanding of attacker deception.

1.1 Results and Implications

We study the attacker's problem of finding the optimal deceptive payoff and the corresponding defender equilibrium strategy, and refer to this game as the *imitative deception game*. When the attacker's deception is *unconstrained* — i.e., he can choose any deceptive payoff — we provide a clean characterization about the attacker's optimal deception and corresponding optimal defender strategy. Interestingly, we prove that the optimal attacker deception in this case is to pretend to have a (deceptive) payoff such that the resulting optimal strategy for the defender is to play the `Maximin` strategy, *regardless of what the attacker's true payoff is*. This result conveys a very interesting conceptual message: in security games, the attacker would always like to pretend to be strictly competitive against the defender regardless what true payoff he possesses. This seemingly irrational behavior is in fact highly strategized and can maximize his true utility. We remark that this is a special property of security games and does not hold in general Stackelberg games.

When the attacker's deception is constrained and cannot be arbitrary payoff, we examine the problem *computation-*

*The two authors contribute equally to the paper.
†Contact Author

ally, and present a general optimization framework to solve the imitative deception game under *constrained attacker deception*. We then show that instantiating our framework for two natural types of constraints results in a compact formulation as a Mixed Integer Linear Program (MILP), which can be solved efficiently as shown by our experiments. Finally, our simulation illustrates significant benefit [loss] of the attacker [defender] in presence of imitative attacker deception.

1.2 Additional Related Work

The study of deception in security domains has a rich literature [Fraunholz *et al.*, 2018]. In SSGs, recent work examines deception from the *defender's side* and study how to mislead the attacker's decision by exploiting the defender's knowledge regarding uncertainties [Zhuang *et al.*, 2010; Xu *et al.*, 2015; Rabinovich *et al.*, 2015; Guo *et al.*, 2017]. However, *attacker deception* — the natural counterpart in this line of research — has not been paid much attention. To our knowledge, the only relevant study is the recent work by [Nguyen *et al.*, 2019], however focusing on *repeated* games and analyzing Bayesian Nash equilibrium, whereas our game is one-shot and is a Stackelberg model with commitment.

Research on adversarial learning has attempted to design various types of attacks to machine learning algorithms [Brückner *et al.*, 2012; Brückner and Scheffer, 2011; Barreno *et al.*, 2010; 2006; Lowd and Meek, 2005] The attacker deception in our game can be viewed as a type of *causative attack* to the defender's learning algorithm, with the goal of maximizing the attacker's utility.

2 Preliminaries and Our Model

We consider a standard SSG where a *defender* allocates K *security resources* to protect N *targets* ($K < N$) against an *attacker's* attack. Let $[N] = \{1, 2, \ldots, N\}$ denote the set of targets. A pure strategy of the defender is an assignment of these K resources to an arbitrary subset of K targets (i.e., no scheduling constraints), and a mixed strategy is thus a probability distribution over these pure strategies. A defender mixed strategy in this setting can be equivalently represented as a vector of *marginal coverage* probabilities $\mathbf{x} = (x_1, x_2, \ldots, x_N)$, where $\sum_j x_j \leq K$ and $x_j \in [0, 1]$ is the probability of protecting target $j \in [N]$ [Kiekintveld *et al.*, 2009]. Let $\mathbf{X}$ denote the set of all these mixed strategies.

If the attacker attacks target j that is not protected by the defender, he obtains a reward of R_j^a while the defender receives a penalty of P_j^d. Conversely, if j is protected, the attacker receives a penalty of $P_j^a (< R_j^a)$ while the defender gets a reward of $R_j^d (> P_j^d)$. The players' expected utilities at j are thus computed as follows:

$$U_j^d(x_j) = x_j R_j^d + (1 - x_j) P_j^d \quad (1)$$

$$U_j^a(x_j) = x_j P_j^a + (1 - x_j) R_j^a \quad (2)$$

Given the payoff structure, for any defense strategy $\mathbf{x} \in \mathbf{X}$, let $\Gamma(\mathbf{x})$ denote the set of targets of the highest expected utility for the attacker. SSG models assume that the attacker is aware of $\mathbf{x}$, and thus the rational attacker will attack some target $j \in \Gamma(\mathbf{x})$ as his best (pure) response [Tambe, 2011]. We call $\Gamma(\mathbf{x})$ the *attack set* with respect to $\mathbf{x}$.

The commonly adopted solution concept in SSGs is the Strong Stackelberg Equilibrium (SSE), which consists of a defender mixed strategy $\mathbf{x}^*$ and an attacker best response $i^* \in \Gamma(\mathbf{x}^*)$. Formally, $(\mathbf{x}^*, i^*)$ is an SSE if:

$$(\mathbf{x}^*, i^*) = \operatorname*{argmax}_{\mathbf{x} \in \mathbf{X}, i \in \Gamma(\mathbf{x})} U_i^d(x_i) \quad (3)$$

Sometimes two action profiles $(\mathbf{x}^*, i^*)$ and $(\mathbf{x}', i')$ are both SSEs of the game. Yet, they must always yield exactly the same utility for both players [Yin *et al.*, 2010]. In such cases, we say $(\mathbf{x}', i')$ is *reducible* to $(\mathbf{x}^*, i^*)$.

Our model assumes that the attacker's true payoff $\{P_j^a, R_j^a\}_{j \in [N]}$ is *unknown* to the defender. The attacker can manipulate his behavior to mislead the defender if that is beneficial. We focus on a basic deception model, which we term *imitative deception*. That is, the attacker simply behaves according to a different payoff $\{\hat{P}_j^a, \hat{R}_j^a\}_{j \in [N]}$, instead of his true payoff, and will do so *consistently throughout the game*. Such deception may happen in many scenarios, especially those where the defender seeks to learn the attacker's payoff. The attacker's goal is to find the optimal *deceptive payoff* so that it leads to a SSE defense strategy that maximizes the attacker's *true utility*. Under imitative attacker deception, the SSE is defined in exactly the same way except that it is now with respect to the attacker's *deceptive payoff*. Yet, to distinguish this deceptive situation from truthful attacker behavior, we call the induced game *imitative deception game* and the corresponding SSE the *deceptive SSE*.

3 Unconstrained Imitative Deception

In this section, we study the case where the attacker's deception is unconstrained. Concretely, the attacker can imitate any payoff $\{\hat{P}_j^a, \hat{R}_j^a\}_{j \in [N]}$ as long as $\hat{P}_j^a < \hat{R}_j^a$. We provide a complete characterization about both players' strategies in this *unconstrained* setting, and prove that the optimal imitative attacker deception is to pretend to have a payoff such that the optimal defender strategy is to play the `Maximin` strategy. Our main theorem is formally stated as follows.

Theorem 1. *For any true attacker payoff $\{P_j^a, R_j^a\}_{j \in [N]}$, the deceptive Strong Stackelberg Equilibrium (SSE) under* optimal *attacker imitative deception is characterized as follows:*

- *The defender's optimal strategy is* `Maximin` $\mathbf{x}^{mm}$*;*
- *The attacker attacks target*

$$i^* = \operatorname{argmax}_{j \in [N]} \left[x_j^{mm} P_j^a + (1 - x_j^{mm}) R_j^a \right];$$

- *The attacker's optimal imitative payoffs can be constructed as follows (this is one possible construction): (i) For any $j \in [N]$, set $\hat{P}_j^a = -R_j^d$ and $\hat{R}_j^a = -P_j^d$; (ii) if $x_{i^*}^{mm} = 0$, then re-set $\hat{R}_{i^*}^a = -U^{mm}$ and $\hat{P}_{i^*}^a = -U^{mm} - 1$ where $U^{mm} = \min_j [x_j^{mm} R_j^d + (1 - x_j^{mm}) P_j^d]$ is the defender's* `Maximin` *utility.*

Theorem 1 provides a complete characterization about both players' optimal strategies under imitative attacker deception. It also illustrates the usefulness of the `Maximin` strategy in handling attacker deception in security domains. Note that

the attacker's optimal deceptive payoffs may *not* be exactly the opposite of the defender's payoffs, i.e., $\hat{P}_j^a = -R_j^d$ and $\hat{R}_j^a = -P_j^d$ — sometimes he needs to treat his "favoriate" target i^* specially to make sure attacking i^* is indeed his best response (i.e., the "if $x_{i^*}^{\text{mm}} = 0$" step in Theorem 1). We remark that Theorem 1 relies on the structure of *security* games and does *not* hold for general Stackelberg games. The following corollary shows that the attacker has no incentive for imitative deception in zero-sum SSGs.

Corollary 1. *In zero-sum games, the attacker's optimal imitative deception strategy is to play truthfully.*

Admittedly, Theorem 1 is somewhat counter intuitive since the defender will always be misled to play the same strategy $\mathbf{x}^{\text{mm}}$ regardless of the attacker's true payoffs. Before giving the full proof of Theorem 1, we provide an illustrative example and attempt to provide some intuition underlying this result. However, we note that its full proof is more involved.

Example 1 (An example and intuitions of Theorem 1)**.** *Consider a security game with* 3 *targets. The defender and attacker payoffs are specified as follows:*

	target 1	*target 2*	*target 3*
R_i^d	*1*	*3*	*1*
P_i^d	*-3*	*-2*	*0*
R_i^a	*2*	*1*	*1*
P_i^a	*-3*	*-1*	*-1*

The defender only has 1 *security guard, which can be allocated to protect any target. We use* (x_1, x_2, x_3) *to denote a defender mixed strategy where* x_i *is the probability of allocating the guard to target* i*. Our model assumes that the attacker payoffs are* unknown *to the defender, and attacker can imitate any other payoff structure if that is beneficial to him.*

If the attacker were honest, then the SSE is $(1/3, 1/3, 1/3)$*. The attacker will be induced to attack target* 3*, resulting in defender utility* $1/3$ *and attacker utility* $1/3$*.*

According to Theorem 1, under optimal attacker imitative deception, the defender will be misled to play the `Maximin` *strategy which is* $(2/3, 1/3, 0)$*. The attacker will attack target* $\text{argmax}_{j\in[3]}\left[x_j^{mm}P_j^a + (1-x_j^{mm})R_j^a\right]$*, which is target* 3 *in this example. This results in defender utility* 0 *and attacker utility* 1 *(larger than the utility of an honest attacker). The optimal deceptive attacker payoff constructed in Theorem 1 is*

$\hat{R}_i^a$	*3*	*2*	*1/3*
$\hat{P}_i^a$	*-1*	*-3*	*-2/3*

It is easy to verify that this deceptive payoff — though make the game non-zero sum at target 3 *— indeed makes the* `Maximin` *strategy* $(2/3, 1/3, 0)$ *the SSE.*

In this example, through deception, the attacker was able to completely "shift" the defender's resource away from target 3 *(i.e., from* $x_3 = 1/3$ *in honest SSE to* $x_3 = 0$ *in* $\mathbf{x}^{mm}$*) and achieve utility* 1 *by attacking the unprotected target* 3*. One might wonder whether the attacker can instead completely shift the defender's resource away from target* 1 *and then attack it, resulting in even higher utility* 2*. It turns out that this is* not *possible because* $(0, x_2, x_3)$ *and* $i^* = 1$ *can never form an SSE — shifting some protection from target* $2, 3$ *to target* 1 *will surely increase the defender's expected utility.*

Some intuitions about why `Maximin`**.** *In the proof of Theorem 1, we will give a full characterization about what kind of* $(\mathbf{x}, i^*)$ *could be a (deceptive) SSE, and identify some "*consistency*" condition. Intuitively, the expected defender utility at* i^* *should be no worse any other target* j *with non-zero protection probability (the "*`Max`*" part) since otherwise she can move some protection from* j *to* i^* *to improve her SSE utility. Now, among all consistent* $(\mathbf{x}, i^*)$*'s, which is the best for the attacker? It shall be the one minimizing the defender's utility (the "*`Min`*" part) since it minimizes* x_{i^*} *and thus maximizes the attacker utility. This is the intuition of why the* `Maximin` *defender strategy shows up at the equilibrium regardless of what the attacker's true payoffs are.*

Example 2 (Failure of Theorem 1 in normal-form games)**.** *Consider a Stackelberg version of the* battle of the sexes *Game, with payoffs as follows, where row player is the leader.*

	`Opera`	`Basketball`
`Opera`	*(2,1)*	*(0,0)*
`Basketball`	*(0,0)*	*(1,2)*

Without deception, the leader should commit to `Opera` *deterministically, resulting in follower best response* `Opera` *and utility* 1*. If the leader plays* `Maximin` *strategy* $(1/3, 2/3)$*, the imitative follower payoff specified by Theorem 1 is to make the game zero-sum in this case. The follower shall take action* `Basketball`*, resulting in follower utility* $4/3$*. However, this is not optimal for the follower — at optimal imitative deception, the follower pretends to not "care about" the leader at all, and* always *have utility* 2 *for* `Basketball` *and* 1 *for* `Opera`*. In this case, the leader will commit to* `Basketball`*, resulting follower utility* 2*. In fact, the follower deception essentially served as a way of* commitment.

Proof of Theorem 1

The main challenge here is that the attacker's deception is to manipulate the space of his payoffs whereas our analysis has to examine the space of (deceptive) SSEs. Unfortunately, the relation between the space of attacker payoffs and the space of SSEs does not admit a clean analytical form. To prove the theorem, we establish various characterizations of SSEs, which we believe might be of independent interest.

Our proofs are divided into three main steps.

Step 1: A Characterization of SSE

As the first step, we provide a characterization of the SSE in security games, which will be crucial for us to analyze what defender mixed strategies can possibly arise in deceptive SSE. Since this characterization may be of independent interest, we state it as Theorem 2. Intuitively, Theorem 2 shows that a strategy profile is an SSE *if and only if* it is reducible to $\{\mathbf{x}^*, i^*\}$ such that: (1) all targets in the attack set have equal attacker utility; (2) i^* has the highest defender utility among all targets in the attack set; (3) any target outside the attack set is covered with a probability of zero; (4) either all the resources are used up or one of the target is covered with a probability of one. We remark that the crucial conditions here are *Condition (3) and (4)*, which are specific to the security game setting, whereas Condition (1) and (2) follow naturally from the definition of SSE.

Theorem 2. *Given any security game, a strategy profile is an SSE if and only if it is* reducible *to* $\{\mathbf{x}^*, i^*\}$ *s.t.*
1) For any $j \in \Gamma(\mathbf{x}^*)$, $P^a_{i^*}x^*_{i^*} + R^a_{i^*}(1-x^*_{i^*}) = P^a_j x^*_j + R^a_j(1-x^*_j)$*; 2) For any* $j \in \Gamma(\mathbf{x}^*)$, $R^d_{i^*}x^*_{i^*} + P^d_{i^*}(1-x^*_{i^*}) \geq R^d_j x^*_j + P^d_j(1-x^*_j)$*; 3) For any* $j \notin \Gamma(\mathbf{x}^*)$, $x^*_j = 0$ *and* $P^a_{i^*}x_{i^*} + R^a_{i^*}(1-x^*_{i^*}) > R^a_j$*; and 4) either (i)* $\sum_j x^*_j = K$ *or (ii)* $\sum_j x^*_j < K$ *and* $x^*_k = 1$ *for some* $k \in \Gamma(\mathbf{x}^*)$.

The full proof of Theorem 2 can be found in the full version. A useful corollary of Theorem 2 is its instantiation to zero-sum games. In particular, we can view the `Maximin` strategy as the Stackelberg equilibrium of a zero-sum security game. Therefore, Theorem 2 gives rise to the following characterization of the defender's `Maximin` strategy, which we denote as $\mathbf{x}^{\text{mm}}$. Here, U^{mm} is the defender's `Maximin` utility and $\Gamma(\mathbf{x}^{\text{mm}})$ denotes the attack set of the `Maximin` strategy.

Lemma 1 (Characterization of `Maximin`)**.** *Any* `Maximin` *defender strategy is reducible to* $\mathbf{x}^{mm}$ *such that: 1) for all* $j \in \Gamma(\mathbf{x}^{mm})$, $x^{mm}_j R^d_j + (1-x^{mm}_j)P^d_j = U^{mm}$*; 2) for all* $j \notin \Gamma(\mathbf{x}^{mm})$, $x^{mm}_j = 0$ *and* $U^{mm} < P^d_j$*; and 3) either (i)* $\sum_j x^{mm}_j = K$ *or (ii)* $\sum_j x^{mm}_j < K$ *and* $x^{mm}_k = 1$ *for some* $k \in \Gamma(\mathbf{x}^{mm})$.

Note that the first two conditions in Theorem 1 are combined as one condition due to $P^a_j = -R^d_j$ and $R^a_j = -P^d_j$.

Step 2: Characterizing the Set of All Deceptive SSE

Our second main step is to characterize the space of all the possible deceptive SSE $\{\mathbf{x}, i^*\}$'s that can possibly arise due to the attacker's imitative deception. Besides revealing useful insights about the SSE, this characterization is also an important step towards analyzing the attacker's imitative deception since the optimal deception strategy is essentially to pick the deceptive SSE that maximizes the attacker's expected utility.

Our characterization of SSE in Theorem 2 will play a key role in this analysis. Naturally, $\mathbf{x}$ must satisfy Condition (4) of Theorem 2 since this condition is out of the attacker's control. Additionally, the i^* cannot be arbitrary. In particular, Condition (3) implies that any j such that $x_j > 0$ must belong to the attack set $\Gamma(\mathbf{x})$. In other words, $\mathbf{supp}(\mathbf{x}) \subseteq \Gamma(\mathbf{x})$ where $\mathbf{supp}$ denotes the support of $\mathbf{x}$. As a result, Condition (2) then implies that $\forall j \in \mathbf{supp}(\mathbf{x})$, we must have $R^d_{i^*}x_{i^*} + P^d_{i^*}(1-x_{i^*}) \geq R^d_j x_j + P^d_j(1-x_j)$. This leads to the following definition of consistency between $\mathbf{x}$ and i^*.

Definition 3.1. *We say* $\{\mathbf{x}, i^*\}$ *is* consistent *if* $R^d_{i^*}x_{i^*} + P^d_{i^*}(1-x_{i^*}) \geq R^d_j x_j + P^d_j(1-x_j), \forall j \in \mathbf{supp}(\mathbf{x})$.

Note that x_{i^*} may equal 0 when $\{\mathbf{x}, i^*\}$ is *consistent*. Interestingly, it turns out that consistency is essentially the only requirement to make $\{\mathbf{x}, i^*\}$ a deceptive SSE.

Lemma 2. *For any* $(\mathbf{x}, i^*)$*, there exists* $\{\hat{R}^a_j, \hat{P}^a_j\}_{j=1}^N$ *to make* $\{\mathbf{x}, i^*\}$ *a deceptive SSE* if and only if $\{\mathbf{x}, i^*\}$ *is consistent and* $\mathbf{x}$ *satisfies Condition (4) of Theorem 2.*

Proof. We have argued about "only if" direction (i.e., the necessity) previously. Here we prove sufficiency.

Given any consistent $(\mathbf{x}, i^*)$ where $\mathbf{x}$ satisfies Condition (4) of Theorem 2, we construct the following deceptive payoff of the attacker: (i) $\forall j$, if $x_j > 0$, let $\hat{R}^a_j = 2$ and $\hat{P}^a_j = 2 - \frac{1}{x_j}$; (ii) $\forall j$, if $x_j = 0$ and $j \neq i^*$, let $\hat{R}^a_j = 0$ and $\hat{P}^a_j = -1$; and (iii) if $x_{i^*} = 0$, set $\hat{R}^a_{i^*} = 1$ and $\hat{P}^a_{i^*} = -1$.

Therefore, $\forall j \neq i^*$, if $x_j > 0$, the attacker's deceptive utility at target j is $x_j\hat{P}^a_j + (1-x_j)\hat{R}^a_j = x_j\left(2 - \frac{1}{x_j}\right) + 2(1-x_j) = 1$; if $x_j = 0$, the attacker's deceptive utility is $\hat{R}^a_j = 0$. For the special target i^*, if $x_{i^*} = 0$, the third step in the above construction sets the deceptive payoff to make the attacker's expected utility at i^* also $1 = \hat{R}^a_{i^*}$. Consequently, we have $\Gamma(\mathbf{x}) = \mathbf{supp}(x) \cup \{i^*\}$. One can easily verify that the constructed $\{\mathbf{x}, i^*\}$ satisfies all the four conditions in Theorem 2: Condition 1 and 2 are satisfied due to the construction of $\{\hat{P}^a_j, \hat{R}^a_j\}_{j=1}^n$ and Condition 3 is satisfied due to consistency assumption. Thus, $\{\mathbf{x}, i^*\}$ is a deceptive SSE. □

Step 3: Completing the Proof

The main part of our final step is to invoke previous characterization results to prove that that under optimal imitative attacker deception, the defender strategy in the deceptive SSE is precisely the `Maximin` strategy $\mathbf{x}^{\text{mm}}$. As a result, by definition of `Maximin`, for any target i we have $R^d_i x^{\text{mm}}_i + P^d_i(1-x^{\text{mm}}_i) \geq R^d_j x^{\text{mm}}_j + P^d_j(1-x^{\text{mm}}_j), \forall j \in \mathbf{supp}(\mathbf{x}^{\text{mm}})$. So $(\mathbf{x}^{\text{mm}}, i)$ is consistent for any $i \in [N]$. By Lemma 2, $(\mathbf{x}^{\text{mm}}, i)$ can be a deceptive SSE and the optimal target i^* for the attacker is then $i^* = \text{argmax}_{j\in[N]}\left[x^{\text{mm}}_j P^a_j + (1-x^{\text{mm}}_j)R^a_j\right]$. In this case, the imitative attacker payoff has an even easier construction due to the special property of $\mathbf{x}^{\text{mm}}$. In particular, we simply let $\hat{P}^a_j = -R^d_j$ and $\hat{R}^a_j = -P^d_j$ for all $j \in\in [N]$. If $x^{\text{mm}}_{i^*} = 0$, reset $\hat{R}^a_{i^*} = -U^{\text{mm}}$, i.e., the negative of the defender's `Maximin` utility, and $\hat{P}^a_{i^*} = -U^{\text{mm}} - 1$. It is easy to verify that in this case the attack set is $\tau(\mathbf{x}^{\text{mm}}) \cup \{i^*\}$.

What remains is to argue that the defender strategy is precisely $\mathbf{x}^{\text{mm}}$, as summarized in the following lemma, whose proof is deferred to the full version. These all together concludes the proof of Theorem 1.

Lemma 3. *Let* $\mathbf{x}^{mm}$ *be the defender's* `Maximin` *strategy. For any* $i^* \in [N]$ *such that* $(\mathbf{x}^{mm}, i^*)$ *is consistent,* $(\mathbf{x}^{mm}, i^*)$ *maximizes the attacker's utility among all consistent* $(\mathbf{x}, i^*)$*'s where* $\mathbf{x}$ *satisfies Condition (4) of Theorem 2.*

4 Constrained Imitative Deception

In the constrained attacker deception scenario, the attacker can pretend his payoff to be different from his true one with some predetermined constraints on his potential lies. For example, the attacker's deceptive rewards and penalties are constrained within some intervals. In this work, we consider two cases: (i) there are value constraints of the attacker's payoff; and (ii) there is a limited number of targets the attacker can report untruthful rewards and penalties.

In value-constrained deception, the attacker can only report its payoff from a *deceptive payoff space* $\Omega \subset \mathbb{R}^N_+ \times \mathbb{R}^N_-$ and the defender follows the corresponding deceptive SSE to play. As a corollary of Theorem 1, the following proposition shows that if imitating a strictly competitive opponent is feasible for the attacker, then doing so is always optimal.

Proposition 1. *If the deceptive payoff space* Ω *includes* $\hat{R}^a_j = -P^d_j, \hat{P}^a_j = -R^d_j, \forall j$*, then the attacker's optimal deceptive strategy is still to make the deception game* zero-sum *and the defender strategy in the deceptive SSE is* `Maximin`.

Next we describe a general framework for computing the optimal imitative deception under *arbitrary* value constraints, and then show how to instantiate this framework for two concrete (and natural) types of value constraints.

As stated previously, the utility of the attacker in the SSE is a decreasing function of the defender's coverage probability at the attacked target. Therefore, our idea is to divide the attacker's deceptive payoff space into N sub-spaces such that for all the attacker payoffs from the same sub-space, the attacker attacks the same target in the deceptive SSE. Based on Theorem 2, the problem of computing the optimal deceptive SSE in each sub-space can then be represented as the two Mixed Integer Non-Linear Programs (MINLPs); each corresponds to either Condition 4(i) or 4(ii):

The following MINLP is with respect to the Condition 4(i):

$$\min\ x_i \text{ s.t.} \tag{4}$$

$$x_i\hat{P}^a_i + (1-x_i)\hat{R}^a_i \geq x_j\hat{P}^a_j + (1-x_j)\hat{R}^a_j, \forall j \tag{5}$$

$$x_i\hat{P}^a_i + (1-x_i)\hat{R}^a_i \leq x_j\hat{P}^a_j + (1-x_j)\hat{R}^a_j + (1-h_j)M, \forall j \tag{6}$$

$$x_i\hat{P}^a_i + (1-x_i)\hat{R}^a_i \geq \hat{R}^a_j - h_jM + \epsilon, \forall j \tag{7}$$

$$x_iR^d_i + (1-x_i)P^d_i \geq x_jR^d_j + (1-x_j)P^d_j - (1-h_j)M, \forall j \tag{8}$$

$$h_i = 1, h_j \in \{0,1\}, x_j \leq h_j, \forall j \tag{9}$$

$$\sum\nolimits_j x_j = K, x_j \in [0,1], \forall j \tag{10}$$

$$\{\hat{P}^a_j, \hat{R}^a_j\} \in \Omega \tag{11}$$

where $\hat{P}^a_j, \hat{R}^a_j, x_j, h_j$ are variables. This MINLP is non-linear because it has product of variables, i.e., $x_j\hat{P}^a_j$. The MINLP minimizes the defender's coverage probability at the attacked target i, or equivalently, to maximize the attacker's true expected utility. In particular, h_j is a binary variable which indicates if target j belongs to the attack set ($h_j = 1$) or not ($h_j = 0$). Constraint (5) ensures that target i has the highest deceptive expected utility for the attacker. Constraint (8) forces the defender's utility at i to be the highest among targets in the attack set. In other words, constraints (5) and (8) guarantee that i is the attacked target. Constraints (5) and (6) force the deceptive expected utility for the attacker at every target in the attack set to be equal to the one at i. Constraints (7) and (9) force $x_j = 0$ and the attacker deceptive utility at i is strictly greater than R^a_j if $j \notin \Gamma(\mathbf{x})$. Constraint (10) satisfies Condition 4(i) of Theorem 2. In summary, constraints (5–10) guarantee that the outcome of this MINLP is a deceptive SSE. Finally, constraint (11) forces the attacker's deceptive payoff to be in the space Ω. Note that, M and ϵ are very large and small constants, respectively.

A similar MINLP with respect to Condition 4(ii) of Theorem 2 can be formulated, simply by substituting Constraint (10) by the following constraints, with additional binary variable q_j indicating whether $x_j = 1$ ($q_j = 1$) or not ($q_j = 0$).

$$\sum\nolimits_j x_j \leq K, \sum\nolimits_j q_j = 1, x_j \geq q_j, q_j \in \{0,1\}, \forall j \tag{12}$$

4.1 Value-Bound Constraints

We now instantiate the above framework with value-bounded constraints. That is, the space Ω can be represented as a set of separate lower and upper bound constraints on the attacker's rewards and penalties. We can convert the aforementioned non-linear optimization problems into Mixed Integer Linear Programs (MILPs). In particular, assume that Ω can be represented as follows:

$$\Omega_j = \{(\hat{P}^a_j, \hat{R}^a_j) :\ l^p_j \leq \hat{P}^a_j \leq u^p_j, l^r_j \leq \hat{R}^a_j \leq u^r_j\}, \forall j$$

where $(l^r_j, u^r_j, l^p_j, u^p_j)$ are constants. We introduce new variables $y^p_j = \hat{P}^a_j x_j$ and $y^r_j = \hat{R}^a_j(1-x_j)$. We now can reformulate the problem (4–11) as the following MILP:

$$\min\ x_i \tag{13}$$

$$\text{s.t. } y^p_i + y^r_i \geq y^p_j + y^r_j, \forall j \tag{14}$$

$$y^p_i + y^r_i \leq y^p_j + y^r_j + (1-h_j)M \tag{15}$$

$$y^p_i + y^r_i \geq y^p_j + y^r_j - h_jM + \epsilon \tag{16}$$

Constraints (8–10)

$$l^p_j x_j \leq y^p_j \leq u^p_j x_j, \forall j \tag{17}$$

$$l^r_j(1-x_j) \leq y^r_j \leq u^r_j(1-x_j), \forall j \tag{18}$$

where constraints (14–16) correspond to (5–7). Constraint (17–18) correspond to (11). Similarly, we also obtain a MILP w.r.t. Condition 4(ii). These two MILPs compute the optimal deceptive payoff given that i is the attacked target in the corresponding deceptive SSE. Finally, the best deceptive payoff is chosen as to provide the highest true expected utility for the attacker among all these choices of the attacked target.

4.2 Target-Limited Constraints

Now we show how the above framework can be applied to target-limited constraints. That is, the attacker can lie for up to $L < T$ targets. Here, we introduce a binary variable z_j which indicates if the attacker lie at target j ($z_j = 1$) or not ($z_j = 0$). We now can then formulate a new program with the following additional constraints:

$$R^a_j - z_jM \leq \hat{R}^a_j \leq R^a_j + z_jM, \forall j \tag{19}$$

$$P^a_j - z_jM \leq \hat{P}^a_j \leq P^a_j + z_jM, \forall j \tag{20}$$

$$\sum\nolimits_j z_j \leq L, z_j \in \{0,1\} \tag{21}$$

where $\{R^a_j, P^a_j\}_{j\in[N]}$ is the true payoff of the attacker.

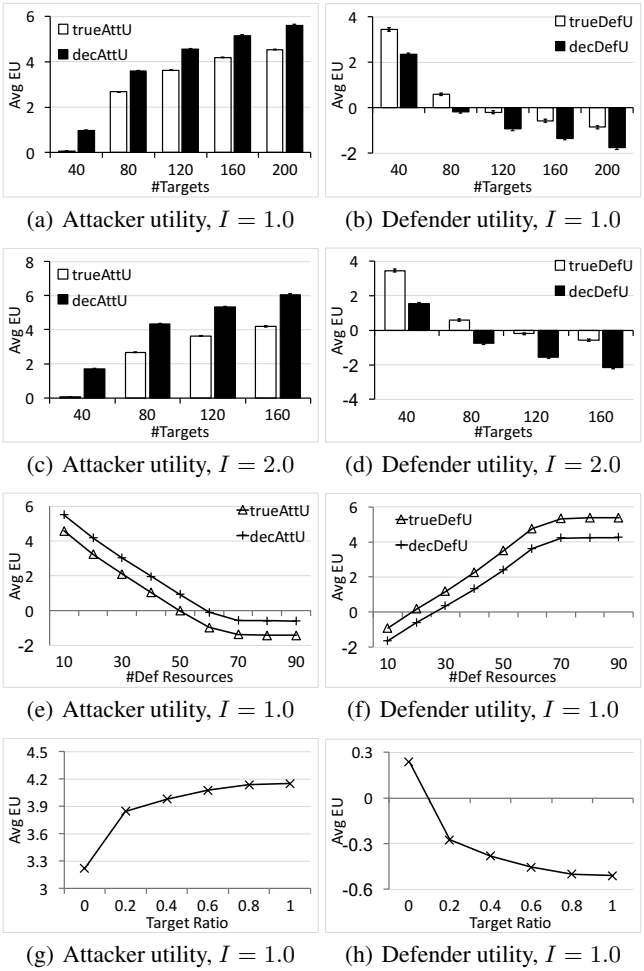
When we additionally have value-bound constraints, then the corresponding target-limited constraints for y^p_j and y^r_j are:

$$R^a_j(1-x_j) - z_jM \leq y^r_j \leq R^a_j(1-x_j) + z_jM \tag{22}$$

$$P^a_j x_j - z_jM \leq y^p_j \leq P^a_j x_j + z_jM \tag{23}$$

5 Experiments

We evaluate the solution quality of our proposed deceptive algorithm. We aim at empirically analyzing the benefit [loss] of the attacker [defender] in terms of expected utility in the presence of the attacker's deception. In our experiments, the players' rewards and penalties are generated in the ranges $[1, 10]$



(a) Attacker utility, $I = 1.0$ (b) Defender utility, $I = 1.0$

(c) Attacker utility, $I = 2.0$ (d) Defender utility, $I = 2.0$

(e) Attacker utility, $I = 1.0$ (f) Defender utility, $I = 1.0$

(g) Attacker utility, $I = 1.0$ (h) Defender utility, $I = 1.0$

Figure 1: Solution evaluation

and $[-10, 1]$ using the covariance game generator, GAMUT (http://gamut.stanford.edu/). The covariance value r governs the correlation between the players' payoffs. If $r = -1.0$, the generated games are zero-sum. Since the attacker always play truthfully in zero-sum games (Corollary 1), we only choose r within $[-0.8, 0.0]$ with a step size of 0.2. Each data point in our results is averaged over 250 different games (50 games per covariance value). Finally, we consider two scenarios: (i) small deceptive payoff space with an interval size of $I = 1.0$; and (ii) large space with $I = 2.0$. Our evaluations are based on various game settings with varying number of deceivable targets, number of targets, and number of defender resources.

In Figure 1, we evaluate the attacker and the defender's average expected utility in two cases: (i) the attacker plays truthfully; and (ii) the attacker is rationally deceptive. We name the attacker and defender's utilities (`trueAttU`, `trueDefU`) and (`decAttU`, `decDefU`), respectively. In the deception case, `decAttU` is computed based on the optimal deceptive SSE and the attacker's true payoff.

In Figures 1(a,d), the x-axis represents the number of targets and the y-axis is the average expected utility of the attacker [defender]. Overall, when the attacker is rationally deceptive, the attacker's [defender's] utility is roughly an increasing concave [decreasing convex] function of the number of targets, which is similar to the case of a truthful attacker. This makes sense, as when the number of target increases, the defender has less protection on the targets, leaving targets more vulnerable to attacks. Furthermore, the `decAttU` [`decDefU`] is quantitatively higher [lower] than the `trueAttU` [`trueDefU`]. This result shows a significant benefit [loss] of the attacker [defender] in the presence of the attacker's deception. We also see an increase [decrease] in `decAttU` [`decDefU`] when the interval size increases ($I = 1.0$ vs $I = 2.0$). This reflects the growth in options for deception (i.e., deceptive payoff space), as well as increased potential benefit for the attacker to play deceptively.

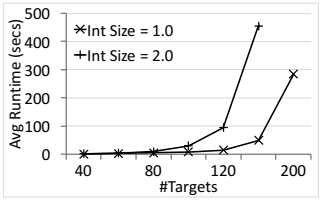


Figure 2: Runtime evaluation

Figures 1(e,f) show that the attacker's [defender's] utility in both cases is decreasing [increasing] in the number of defender resources K, reflecting the increased coverage probabilities of the defender over the targets. In Figures 1(g,h), the *target ratio* is the proportion of the target set at which the attacker can lie about his payoff. When the ratio is 0, the attacker plays truthfully. When the ratio is 1.0, the attacker can manipulate the whole target set. The attacker's [defender's] utility is shown to be roughly an increasing concave [decreasing convex] function of the *target ratio*. Similar to the increase in the interval size, this result reflects the growth in options for deception of the attacker, and increased benefit [increased loss] for the attacker [defender].

Our last experiment evaluates the runtime performance. In Figure 2, the x-axis represents the number of targets. The y-axis is the average runtime in seconds. Figure 2 shows that the runtime grows exponentially when N increases. Nevertheless, the algorithm can easily scale up to $N = 160$ targets and interval size $I = 2.0$ (solved within ≈ 455 seconds). We also see an increase in the runtime when $I = 2.0$ compared to $I = 1.0$, reflecting an increased deceptive payoff space to search for an optimal deceptive SSE.

6 Summary

We studied a basic attacker deception strategy termed *imitative deception* motivated by security contexts where the defender needs to learn the unknown attacker payoffs from observed attack data. We show that the optimal *unconstrained attacker deception* is to make the defender play `Maximin` in the deceptive game. We also present a general optimization framework to solve the game under *constrained deception*. Our experiments illustrate the significant benefit [loss] of the attacker [defender] in the presence of the imitative deception, suggesting potential side effects of learning from the attacker.

References

[Balcan *et al.*, 2015] Maria-Florina Balcan, Avrim Blum, Nika Haghtalab, and Ariel D. Procaccia. Commitment without regrets: Online learning in Stackelberg security games. In *16th ACM Conference on Economics and Computation*, pages 61–78, 2015.

[Barreno *et al.*, 2006] Marco Barreno, Blaine Nelson, Russell Sears, Anthony D. Joseph, and J. Doug Tygar. Can machine learning be secure? In *ACM Symposium on Information, Computer and Communications Security*, 2006.

[Barreno *et al.*, 2010] Marco Barreno, Blaine Nelson, Anthony D. Joseph, and J. D. Tygar. The security of machine learning. *Machine Learning*, 81(2):121–148, 2010.

[Blum *et al.*, 2014] Avrim Blum, Nika Haghtalab, and Ariel D. Procaccia. Learning optimal commitment to overcome insecurity. In *Advances in Neural Information Processing Systems*, pages 1826–1834, 2014.

[Brückner and Scheffer, 2011] Michael Brückner and Tobias Scheffer. Stackelberg games for adversarial prediction problems. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2011.

[Brückner *et al.*, 2012] Michael Brückner, Christian Kanzow, and Tobias Scheffer. Static prediction games for adversarial learning problems. *Journal of Machine Learning Research*, 13:2617–2654, 2012.

[Fraunholz *et al.*, 2018] D. Fraunholz, S. D. Anton, C. Lipps, D. Reti, D. Krohmer, F. Pohl, M. Tammen, and H. D. Schotten. Demystifying deception technology: A survey. *arXiv preprint arXiv:1804.06196*, 2018.

[Gholami *et al.*, 2017] S. Gholami, B. Ford, F. Fang, A. Plumptre, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, M. Nsubaga, and J. Mabonga. Taking it for a test drive: a hybrid spatio-temporal model for wildlife poaching prediction evaluated through a controlled field test. In *European Conference on Machine Learning*, 2017.

[Guo *et al.*, 2017] Qingyu Guo, Bo An, Branislav Bosansky, and C. Kiekintveld. Comparing strategic secrecy and Stackelberg commitment in security games. In *26th International Joint Conference on Artificial Intelligence*, 2017.

[Haghtalab *et al.*, 2016] Nika Haghtalab, Fei Fang, Thanh Hong Nguyen, Arunesh Sinha, Ariel D. Procaccia, and Milind Tambe. Three strategies to success: Learning adversary models in security games. In *25th International Joint Conference on Artificial Intelligence*, pages 308–314, 2016.

[Kar *et al.*, 2017] D. Kar, B. Ford, S. Gholami, F. Fang, A. Plumptre, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, and M. Nsubaga. Cloudy with a chance of poaching: Adversary behavior modeling and forecasting with real-world poaching data. In *16th International Conference on Autonomous Agents and Multi-Agent Systems*, 2017.

[Kiekintveld *et al.*, 2009] Christopher Kiekintveld, Manish Jain, Jason Tsai, James Pita, Fernando Ordóñez, and Milind Tambe. Computing optimal randomized resource allocations for massive security games. In *8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 689–696, 2009.

[Letchford *et al.*, 2009] Joshua Letchford, Vincent Conitzer, and Kamesh Munagala. Learning and approximating the optimal strategy to commit to. In *International Symposium on Algorithmic Game Theory*, pages 250–262, 2009.

[Lowd and Meek, 2005] Daniel Lowd and Christopher Meek. Adversarial learning. In *ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, pages 641–647, 2005.

[Marecki *et al.*, 2012] Janusz Marecki, Gerry Tesauro, and Richard Segal. Playing repeated Stackelberg games with unknown opponents. In *11th International Conference on Autonomous Agents and Multiagent Systems*, 2012.

[Nguyen *et al.*, 2016] T. H. Nguyen, A. Sinha, S. Gholami, A. Plumptre, L. Joppa, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, R. Critchlow, et al. Capture: A new predictive anti-poaching tool for wildlife protection. In *15th International Conference on Autonomous Agents and Multi-Agent Systems*, pages 767–775, 2016.

[Nguyen *et al.*, 2019] Thanh H Nguyen, Yongzhao Wang, Arunesh Sinha, and Michael P Wellman. Deception in finitely repeated security games. In *33th AAAI Conference on Artificial Intelligence*, 2019.

[Peng *et al.*, 2019] B. Peng, Weiran Shen, Pingzhong Tang, and Song Zuo. Learning optimal strategies to commit to. In *33th AAAI Conference on Artificial Intelligence*, 2019.

[Rabinovich *et al.*, 2015] Zinovi Rabinovich, Albert Xin Jiang, Manish Jain, and Haifeng Xu. Information disclosure as a means to security. In *14th International Conference on Autonomous Agents and Multi-Agent Systems*, pages 645–653, 2015.

[Tambe, 2011] Milind Tambe. *Security and game theory: algorithms, deployed systems, lessons learned*. Cambridge University Press, 2011.

[Xu *et al.*, 2015] Haifeng Xu, Zinovi Rabinovich, Shaddin Dughmi, and Milind Tambe. Exploring information asymmetry in two-stage security games. In *29th AAAI Conference on Artificial Intelligence*, pages 1057–1063, 2015.

[Xu *et al.*, 2016] Haifeng Xu, Long Tran-Thanh, and Nicholas R. Jennings. Playing repeated security games with no prior knowledge. In *15th International Conference on Autonomous Agents and Multi-Agent Systems*, pages 104–112, 2016.

[Yin *et al.*, 2010] Zhengyu Yin, Dmytro Korzhyk, Christopher Kiekintveld, Vincent Conitzer, and Milind Tambe. Stackelberg vs. Nash in security games: Interchangeability, equivalence, and uniqueness. In *9th International Conference on Autonomous Agents and Multi-Agent Systems*, pages 1139–1146, 2010.

[Zhuang *et al.*, 2010] Jun Zhuang, Vicki M. Bier, and Oguzhan Alagoz. Modeling secrecy and deception in a multi-period attacker-defender signaling game. *European Journal of Operational Research*, 203:409–418, 2010.