
Manipulating a Learning Defender and Ways to Counteract

Jiarui Gan

University of Oxford
Oxford, UK
jiarui.gan@cs.ox.ac.uk

Qingyu Guo

Nanyang Technological University
Singapore
qguo005@e.ntu.edu.sg

Long Tran-Thanh

University of Southampton
Southampton, UK
l.tran-thanh@soton.ac.uk

Bo An

Nanyang Technological University
Singapore
boan@ntu.edu.sg

Michael Wooldridge

University of Oxford
Oxford, UK
mjw@cs.ox.ac.uk

Abstract

In Stackelberg security games when information about the attacker’s type (i.e., payoff parameters) is uncertain, algorithms have been proposed to *learn* the optimal defender commitment by interacting with the attacker and observing their best responses. In this paper, we show that, however, these algorithms can be easily manipulated if the attacker responds untruthfully. As a key finding, attacker manipulation normally leads to the defender learning a *maximin strategy*, which effectively renders the learning attempt meaningless as to compute a maximin strategy requires no additional information about the other player at all. We then apply a game-theoretic framework at a higher level to counteract such manipulation, in which the defender commits to a *policy* that specifies her strategy commitment according to the learned information. We provide a polynomial-time algorithm to compute the optimal such policy, and in addition, a heuristic approach inspired by the quantal response model, that applies even when the attacker’s type space is infinite or completely unknown. Empirical evaluation shows that our approaches improve the defender’s utility significantly as compared to the situation when attacker manipulation is ignored.

1 Introduction

Stackelberg security games (SSGs) are Stackelberg game models developed for deriving optimal security resource allocation in strategic scenarios. In the AI community, a line of work applying SSG models forms the algorithmic basis of resource scheduling systems, that are in use by the Los Angeles Airport, the US Coast Guard, the Federal Air Marshal Service, etc, to assist in protecting high-profile infrastructures, and public and natural resources [21].

The standard solution concept of SSG, the *strong Stackelberg equilibrium* (SSE) captures the situation where a defender (the leader) commits to her optimal strategy, assuming that an attacker (the follower) will respond optimally to her commitment. There are many algorithms designed to compute SSEs in different SSG models when complete information about the attacker’s type (i.e., his payoff parameters) is provided. While payoff information may be incomplete in many real environments, algorithms were also designed for the defender to *learn* the optimal commitment through interacting with the attacker: by committing to a series of carefully chosen defender strategies and observing the attacker’s best responses to these strategies [11, 4, 8, 19, 17]. The optimality of the learned commitment thus relies crucially on the assumption of a truthful attacker, one who responds to the defender’s commitment

optimally according to their actual payoffs. Unfortunately, when there is no guarantee that the attacker will indeed be truthful, a strategic attacker can easily manipulate the learning algorithm by using fake best responses — typically by imitating the responses of a different attacker type. The defender will then learn a commitment that is optimal with respect to the imitated type but, very likely, suboptimal with respect to the true attacker type. As we will show in this paper, the attacker is often incentivized to imitate a type that makes the game zero-sum; a credulous defender would then only learn a *maximin strategy* (i.e., the optimal commitment in a zero-sum game). Effectively, the learning attempt now becomes meaningless: to compute a maximin strategy, one needs no additional information about the other player’s payoffs at all!

Driven by this issue, we study what can be done to reduce the defender’s loss due to attacker manipulation. We apply a game-theoretic framework at a higher level. In the framework, the defender commits to a *policy* that specifies her strategy commitment according to the learned information. A strategic attacker then takes into account the defender’s policy, choosing optimally what he wants the defender to learn so that the policy outputs a strategy that benefits him the most. We make several other contributions under this framework. (i) We propose a novel quality measure of the defender’s policy and argue why it is a reasonable choice in the context of SSGs. (ii) We develop a polynomial-time algorithm to compute the optimal policy with respect to this quality measure, as well as (iii) a heuristic approach which applies even when the attacker type space is infinite or completely unknown. The heuristic approach is inspired by the famous *quantal response* model that was initially proposed to model bounded rationality of human players. It suggests a rationality of behaving in a “boundedly rational” manner in the presence of attacker manipulation. (iv) Finally, through empirical evaluation we show that our approaches can improve the defender’s utility significantly in randomly generated games, as compared to the situation when attacker manipulations are ignored.

Our work follows some recent research effort on understanding manipulation of leader learning algorithms in general Stackelberg games [7] and shed light on the same issue in SSGs. The SSG model offers us an appropriate level of specification that enables us to derive a richer set of results than from a general model (consider, e.g., when the interests of the leader and the follower completely align, there is simply no incentive for the follower to manipulate the leader), while it also captures sufficiently many real-world applications of significant practical value.

Additional Related Work Apart from [7], manipulation of leader learning algorithms remains largely an under-explored topic, though there are many papers focusing on the design learning algorithms for the leader. In addition to the aforementioned efforts to learn the optimal leader commitment against a fixed but unknown follower type [11, 4, 8, 19, 17], a couple of papers also take the regret-minimization perspective and design online learning algorithms for the leader to use in the adversarial setting [1, 22]. Our work can be seen as a middle-ground approach between the overoptimistic assumption of a truthful follower adopted by the former line of work and the pessimistic assumption of a worst-case opponent by the latter. Our approach to deal with attacker manipulation is, in a nutshell, to reduce information uncertainty by acquiring additional information while bearing in mind that information acquired may be manipulated by the attacker. There are also approaches to deal with uncertainties without additional information retrieval attempts; they are immune to manipulation as a result. In particular, algorithms were designed to compute robust leader strategies when the leader can bound the follower’s payoffs in certain intervals [11, 10, 14], or knows a probability distribution of the follower’s type [6, 16, 9, 18]. Follower manipulation is not a concern in applying these approaches but leader strategies yielded are weaker. In a broader sense, our work is also related to poisoning attacks in adversarial machine learning, where an attacker manipulates the training data (in our model, their payoffs) to undermine the performance of learning algorithms; see, e.g., pioneering work in this area [2, 3] and some recent surveys [5, 12].

2 SSG Preliminaries

An SSG is played between a *defender* (the leader) and an *attacker* (the follower). The defender allocates m security resources to a set of targets $T = \{1, \dots, n\}$ ($n > m$ without loss of generality), and the attacker chooses a target to attack. In the pure strategy setting, a defender strategy is a resource allocation. An attack on a target i is unsuccessful as long as one resource is allocated to i , in which case the attacker receives a penalty p_i^a and the defender a reward r_i^d . Otherwise, i.e., when no resource is allocated to i , the attack is successful, in which case the attacker receives a reward r_i^a

and the defender a penalty p_i^d . We say that a target is *protected*, or *covered*, if at least one resource is assigned to it; and *unprotected*, or *uncovered*, otherwise. It is assumed that $r_i^a > p_i^a$ and $r_i^d > p_i^d$ for all i , so the attacker always prefers a successful attack and the defender prefers the opposite.

The defender can further randomize the resource allocation and commit to a mixed strategy, i.e., a probability distribution over pure strategies. The structure of SSGs allows a defender mixed strategy to be represented more compactly as a *coverage (vector)* $\mathbf{c} = (c_i)_{i \in T}$, with each c_i representing the probability that target i is protected. We will stick to this representation and use the terms *coverage* and *defender mixed strategy* interchangeably. Under the constraint that the defender can use at most m resources, the space of feasible mixed strategies is $\mathcal{C} = \{\mathbf{c} \in \mathbb{R}^n : 0 \leq \mathbf{c} \leq 1, \sum_{i \in T} c_i \leq m\}$; provably, any coverage vector in \mathcal{C} can be implemented as a distribution of pure strategies each involving at most m resources, and any such distribution results in a coverage vector in \mathcal{C} . Let $u^d(\mathbf{c}, i)$ and $u^a(\mathbf{c}, i)$ be the expected utilities of the defender and the attacker, respectively, when the defender plays a mixed strategy \mathbf{c} and the attacker attacks target i . With slight abuse of notation, we write

$$u^d(\mathbf{c}, i) = u^d(c_i, i) = c_i \cdot r_i^d + (1 - c_i) \cdot p_i^d, \quad (1)$$

$$u^a(\mathbf{c}, i) = u^a(c_i, i) = (1 - c_i) \cdot r_i^a + c_i \cdot p_i^a. \quad (2)$$

It is worth noting that the defender's utility function is strictly increasing with respect to c_i , and attacker's utility function strictly decreasing. By the standard assumption, in an SSG the attacker is able to observe the defender's mixed strategy through surveillance before he launches an attack, but the instantiated pure strategy is not observable.

The *strong Stackelberg equilibrium* (SSE) is the standard solution concept of SSGs. In an SSE, the defender commits to an optimal mixed strategy, taking into account that the attacker will observe this strategy and respond optimally. It is assumed that ties are broken in favor of the defender when the attacker has multiple best responses, so without loss of generality we can assume that the attacker always responds by playing a *pure* strategy. The assumption is justified by the fact that the defender can often induce the attacker's strict preference to the favored target by reducing the coverage of this target by an infinitesimal amount.¹ Formally, a strategy profile $(\hat{\mathbf{c}}, \hat{i})$ forms an SSE if:

$$(\hat{\mathbf{c}}, \hat{i}) \in \arg \max_{\mathbf{c} \in \mathcal{C}, i \in \text{BR}(\mathbf{c})} u^d(\mathbf{c}, i),$$

where $\text{BR}(\mathbf{c})$ is the set of the attacker's best responses to \mathbf{c} , i.e., $\text{BR}(\mathbf{c}) = \arg \max_{i \in T} u^a(\mathbf{c}, i)$. An SSE always exists and can be computed in polynomial time, e.g., using a multiple-LP approach [6].

In this paper, we will refer to a full set of attacker payoff parameters $(\mathbf{r}^a, \mathbf{p}^a)$ as an *attacker type*. To distinguish, we extend Eq. (2) and denote by u_θ^a the utility function of a type- θ attacker, i.e., for $\theta = (\mathbf{r}, \mathbf{p})$, we have $u_\theta^a(\mathbf{c}, i) = (1 - c_i) \cdot r_i + c_i \cdot p_i$. Definition of the best response set is extended likewise, with $\text{BR}_\theta(\mathbf{c}) = \arg \max_{i \in T} u_\theta^a(\mathbf{c}, i)$. We will refer to an SSE in a game where the attacker's true type is θ as *SSE on attacker type θ* .

Example 1. Consider an SSG where the defender allocates one security guard to protect two targets A and B . The defender has three pure strategies: to assign the guard to protect A or B , or to send the guard on vacation; the corresponding mixed strategy space is $\mathcal{C} = \{\mathbf{c} \in \mathbb{R}^2 : 0 \leq \mathbf{c} \leq 1, c_1 + c_2 \leq 1\}$. The attacker can choose to attack A or B . In this game, the targets are of equal importance to the defender, such that a successful attack on either target $i \in \{A, B\}$ results in utility $p_i^d = -1$ for the defender, and an unsuccessful one results in $r_i^d = 0$. For the attacker, the payoffs are $r_A^a = 3, r_B^a = 1$, and $p_A^a = p_B^a = 0$. The bi-matrix representation of the game is shown below, with the defender and the attacker being the row and column players, respectively.

	attack A	attack B
protect A	0, 0	-1, 1
protect B	-1, 3	0, 0
protect \emptyset	-1, 3	-1, 1

The SSE of this game can be found using the indifference rule, i.e., by identifying a point where the attacker is indifferent of attacking A and B , while the defender cannot improve coverage of the targets any further (however, not in every game can an SSE be found in this way). In the only SSE

¹See [20] and [21] (Chapter 8) for more discussions on the SSG and the solution concepts.

of this game, the defender protects (A, B) with probabilities $\hat{c} = (\frac{3}{4}, \frac{1}{4})$ (which is equivalent to a mixed strategy $\mathbf{x} = (\frac{3}{4}, \frac{1}{4}, 0)$ as in the bi-matrix representation), to which the attacker finds his best responses to be $\text{BR}(\hat{c}) = \{A, B\}$ and, by assumption, breaks the tie in favor of the defender by attacking A . The defender gets utility $u^d(\hat{c}, A) = -\frac{1}{4}$ and the attacker gets $u^a(\hat{c}, A) = \frac{3}{4}$.

3 Manipulating a Learning Defender

We investigate how attacker manipulation can take place and begin with a warm-up example below.

Example 2. Consider now the attacker in Example 1 pretends to have payoff $r_A^a = 1$ (all other parameters remain the same) and “best” responds to queries of the defender’s learning algorithm according to this fake parameter. Let the resultant fake attacker type be β . The defender will unconsciously be misled into learning an SSE strategy against β , which is $\hat{c} = (\frac{1}{2}, \frac{1}{2})$ in this case. We have $\text{BR}_\beta(\hat{c}) = \{A, B\}$. The attacker can respond (still with ties broken in favor of the defender) by attacking A , and this results in the attacker’s utility to increase to $\frac{3}{2}$, but the defender’s utility to drop to $-\frac{1}{2}$. There is a loss of $\frac{1}{4}$ for the defender compared to the truthful situation! Note that the attacker behaves consistently even after having misled the defender into learning the fake information. Hence, the defender is unable to distinguish him from a truthful type- β attacker.

In the above example, the attacker actually lies to the defender that the game they are playing is zero-sum. It turns out that this is not a coincidence specific to this example but a general phenomenon in SSGs. We show next that it is always optimal for the attacker to mislead the defender into playing her *maximin strategy*. A maximin strategy is one that maximizes the defender’s utility against a worst possible attacker type, i.e., $\bar{c} \in \arg \max_{\mathbf{c}} \min_i u^d(\mathbf{c}, i)$; it is exactly also the defender’s optimal commitment in a zero-sum game (see Lemma 13 in the full paper).

A couple of “disclaimers” would be appropriate before our analysis: First, we only consider the players’ utilities in the fake SSE the defender learns. The cost incurred for both players during the learning process is omitted as we expect the learning algorithm to run efficiently and the learned SSE to repeat in sufficiently many rounds. This simplification is in line with previous work on leader learning algorithm design (e.g., [11, 4, 17]). Without loss of generality, we view the learning process as a reporting step in which the attacker simply *reports* his type to the defender. To manipulate the defender, the attacker can report a fake type, and we refer to this as his *reporting strategy*.

Second, we assume that the attacker behaves according to the reported (fake) type throughout the game. This means that the attacker may be playing a fake best response — hence, a suboptimal one — in the learned SSE. Consequently, the attacker may now exploit the defender for an even higher utility by switching back to his true best response. Nevertheless, since such a change in the attacker’s behavior will inevitably make the defender aware of the manipulation and further complicates the interaction, we ignore the possibility of such a behavior change and adopt this cleaner model to capture the essence of the manipulation problem.

Optimal Attacker Report The following program computes the optimal reporting strategy of a type- θ attacker.

$$\max_{\beta, \mathbf{z}, t} \quad u_\theta^a(\mathbf{z}, t) \quad (3)$$

$$\text{s.t.} \quad (\mathbf{z}, t) \in \arg \max_{\mathbf{c} \in \mathcal{C}, i \in \text{BR}_\beta(\mathbf{c})} u^d(\mathbf{c}, i) \quad (3a)$$

$$\beta \in \Theta \quad (3b)$$

Here $\Theta = \{(\mathbf{r}, \mathbf{p}) \in \mathbb{R}^{n \times n} : r_i > p_i \text{ for all } i \in T\}$ is the set of types that adhere to the basic assumption that an attacker always prefers a successful attack. In the program, the attacker reports a fake type β that results in the defender to learn a strategy \mathbf{z} which, by Eq. (3a), forms an SSE (on type β) along with a best response t (of type β). An optimal solution thus yields a reporting strategy for a type- θ attacker, that maximizes his *true* utility as specified in the objective function.

Based on the program, our next key result Theorem 3 reveals that it is always optimal for the attacker to mislead the defender into playing her maximin strategy. This is surprising as the defender essentially learns no information: to compute the maximin strategy requires no additional knowledge about the other player’s payoffs at all! We present a proof sketch for Theorem 3. Omitted proofs can be found in the full version of this paper.

Theorem 3. *There exists an optimal solution (β, \mathbf{z}, t) of Program (3) such that \mathbf{z} is a maximin strategy of the defender, i.e., $\mathbf{z} \in \arg \max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^d(\mathbf{c}, i)$.*

Proof sketch. Let $\bar{\mathbf{c}}$ be a maximin strategy of the defender and \bar{u} be her maximin utility. Consider a solution (β, \mathbf{z}, t) such that: $z_i = \max \left\{ 0, \frac{\bar{u} - p_i^d}{r_i^d - p_i^d} \right\}$ for all $i \in T$; $t \in \text{BR}_\theta(\mathbf{z})$; and $\beta = (\mathbf{r}, \mathbf{p})$ is such that $r_i = \begin{cases} -p_i^d, & \text{if } i \neq t \\ -\min\{p_t^d, \bar{u}\}, & \text{if } i = t \end{cases}$, and $p_i = -r_i^d$ for all $i \in T$. It can be verified that \mathbf{z} is indeed also a maximin strategy of the defender and (β, \mathbf{z}, t) is an optimal solution. \square

Theorem 3 can be further strengthened under the assumption that the defender’s maximin strategy $\bar{\mathbf{c}}$ is fully mixed, i.e., $0 < \bar{c}_i < 1$ for all i . (In this case the maximin strategy is also unique; see Lemma 15 in the full paper.) The assumption is mild as it is normally expected that no target would be too worthless to the extent that the defender would leave it wide open for the attacker to attack, while on the other hand resources are normally insufficient to allow any target to be fully protected. The strengthening is two-fold: (i) under the additional assumption, the defender’s maximin strategy is her *only* SSE strategy induced by any optimal attacker report, so the equilibrium selection issue that arises when a reported type induces multiple SSEs is avoided; (ii) one optimal attacker report, in particular, is the type that makes the game zero-sum, so our result holds even for a more stringent specification of Θ (e.g., when the defender has more precise knowledge of the attacker’s possible types) as long as Θ contains this zero-sum type (indeed, it is very natural for an attacker to have the zero-sum type given the adversarial nature of SSGs). We state the result in Theorem 4.

Theorem 4. *Suppose $\bar{\mathbf{c}}$ is a maximin defender strategy and it is fully mixed, i.e., $0 < \bar{c}_i < 1$ for all $i \in T$. Let (β, \mathbf{z}, t) be an arbitrary optimal solution of Program (3). For every SSE $(\hat{\mathbf{c}}, \hat{i})$ on type β , it holds that $\hat{\mathbf{c}} = \bar{\mathbf{c}}$. In addition, there exists an optimal solution $(\beta', \mathbf{z}', t')$ such that $\beta' = (-\mathbf{p}^d, -\mathbf{r}^d)$.*

4 Handling Attacker Manipulation — A New Playbook

We explore ways to handle attacker manipulation. Recall our analysis. The key to the success of the attacker’s trick is the naive playbook the defender follows — to always play the learned optimal commitment *as is*. It appears that the defender can be more strategic. Consider Example 2. Suppose the defender tweaks her strategy slightly, playing instead $\tilde{\mathbf{c}} = (\frac{1}{2}, \frac{49}{100})$ even when she learns that $(\frac{1}{2}, \frac{1}{2})$ is the optimal commitment. The attacker, who imitates a type β that makes the game zero-sum, would then have to attack B as now the best response set of β is $\text{BR}_\beta(\tilde{\mathbf{c}}) = \{B\}$. The attacker only obtains utility $\frac{1}{2}$, which is even lower than his utility $\frac{3}{4}$ in the truthful situation. Therefore, if the defender commits to playing, e.g., $(c_1, c_2 - \frac{1}{100})$ whenever she learns that (c_1, c_2) is the optimal commitment, the attacker will at least lose the incentive to mislead the defender into playing her maximin strategy. The question then becomes: what is the best the defender can achieve by revising her playbook in similar ways? We formalize this as finding an optimal *policy* to commit.

Committing to a Policy Formally, a policy is a function $\pi : \Theta \rightarrow \mathcal{C} \times T$ that maps a reported attacker type to an *outcome* $(\mathbf{c}, i) \in \mathcal{C} \times T$. An outcome (\mathbf{c}, i) is a strategy profile consisting of a defender strategy \mathbf{c} and a best response $i \in \text{BR}_\theta(\mathbf{c})$ of the reported attacker type θ .² As an example, the way the defender plays when she ignores attacker manipulation can itself be viewed as a policy that maps every reported type θ to an SSE on θ ; we will refer to this policy as the *SSE policy*.

We assume that a policy can be observed or learned by the attacker through constant interaction with the defender, or the defender can simply announce it to the attacker. In response, a strategic attacker of true type θ chooses to report an optimal type $\beta^* \in \arg \max_{\beta \in \Theta} u_\theta^a(\pi(\beta))$ that will maximize his utility in the outcome of the policy. At a higher level, this can be seen as a Stackelberg game in which the defender commits to a policy and the attacker reports optimally in response to this commitment.

To find the optimal policy, we need a good measure for the quality of a policy. When there is no other prior information about the attacker’s type, the worst-case analysis seems to be appropriate and a

²The reason that we also specify an attacker response in the output of a policy is to avoid complication caused by the tie-breaking issue. Same as in an SSE, the defender can implicitly induce the attacker to respond in a specific way through an infinitesimal deviation.

straightforward choice of quality measure is the utility the defender obtains when playing against the worst attacker type. However, as Proposition 5 suggests, this measure disallows us to well distinguish the quality of many policies. Specifically, when playing against the zero-sum attacker type, no feasible policy can achieve anything better than the maximin utility, so if we take the worst-case utility as the measure, the quality of all policies would be hindered at this attacker type, and the SSE policy, which achieves exactly the maximin utility in the worst case, would immediately be the best policy we can hope for. Essentially, there will be no room for any improvement other than letting the attacker lie.

Proposition 5. *Let $\bar{\mathbf{c}}$ be a maximin strategy of the defender, and $\bar{u} = \min_{i \in T} u^d(\bar{\mathbf{c}}, i)$ be the maximin utility. For any policy π , let $\gamma \in \Theta$ be the corresponding optimal report of an attacker of true type $\beta = (-\mathbf{p}^d, -\mathbf{r}^d)$, i.e., $\gamma \in \arg \max_{\theta \in \Theta} u_{\beta}^a(\pi(\theta))$. Then $u^d(\pi(\gamma)) \leq \bar{u}$.*

This is unreasonable: even in the truthful situation, it is impossible for the defender to achieve more than the maximin utility when the attacker has the zero-sum type, so it would be unfair to underrate a policy simply because it underperforms against the zero-sum type. For this reason, we propose an alternative measure, termed *the efficiency of a policy (EoP)*, which takes into consideration the hardness of playing against each attacker type in the truthful setting. As in Definition 6, the EoP is the worst-case ratio between the utility the defender obtains and what she should have obtained had the attacker been truthful. A higher EoP indicates a smaller loss due to attacker manipulation, and the value of the EoP always lies in between 0 and 1 according to Proposition 7. For the EoP to be meaningful, we shift all payoffs to be non-negative. Without loss of generality, we will hereafter also assume Θ — previously defined as the set of types the attacker is allowed to report — to also be the set of possible (true) attacker types, which is common knowledge to both players.

Definition 6 (EoP). For each $\theta \in \Theta$, let $\beta_{\theta}^{\pi} = \arg \max_{\beta \in \Theta} u_{\theta}^a(\pi(\beta))$ be the attacker’s optimal reporting strategy in response to a policy π (tie-breaking in favor of the defender). The efficiency of π on attacker type θ is $\text{EoP}_{\theta}(\pi) = \frac{u^d(\pi(\beta_{\theta}^{\pi}))}{\hat{u}(\theta)}$, where $\hat{u}(\theta) = \max_{\mathbf{c} \in \mathcal{C}, i \in \text{BR}_{\theta}(\mathbf{c})} u^d(\mathbf{c}, i)$ is the defender’s utility in an SSE on type θ . The (overall) efficiency of π is $\text{EoP}(\pi) = \min_{\theta \in \Theta} \text{EoP}_{\theta}(\pi)$.

Proposition 7. $\text{EoP}(\pi) \in [0, 1]$ for any feasible defender policy π .

Another challenge we face is the representation of a defender policy, which is a function to be optimized. For a computationally feasible formulation of our optimization task, we follow a modeling approach in the literature and consider a discrete version of the problem where the set Θ of attacker types is finite. This approach has been widely adopted to model Bayesian games (e.g., [6, 16, 9, 18]). A finite type set can be seen as an approximation to the continuous type space, while in some scenarios attacker types might also be discrete by nature. For example, in defense against poaching, payoffs of the attacker (poachers) may depend on the type of animal products they are interested in, which falls in a finite set. In addition to this approach, we also propose a heuristic policy that applies to an infinite or even an unknown set of attacker types. We present these approaches next.

5 Computing the Optimal Policy

5.1 Optimal Policy for Finite Attacker Types

When Θ is a finite set, a defender policy can be represented as a list of $\lambda = |\Theta|$ outcomes; we will therefore also write a policy as $\pi = (\mathbf{c}^{\theta}, i^{\theta})_{\theta \in \Theta}$, meaning that $\pi(\theta) = (\mathbf{c}^{\theta}, i^{\theta})$ for each $\theta \in \Theta$. Our analysis reveals that to compute the EoP maximizing policy is NP-hard in general Stackelberg games (see Section D in the full paper), but thanks to the special utility structure of SSGs, the problem admits a polynomial-time algorithm when the underlying game is an SSG.

We consider the decision version of the optimization problem: for a given value ξ , decide whether any defender policy π achieves $\text{EoP}(\pi) \geq \xi$. Trivially, once we have an efficient algorithm for this decision problem, the best EoP can be found efficiently using binary search (in particular, we already know that the value always lies in $[0, 1]$). Our algorithm for this decision problem, presented as Algorithm 1, is constructive and produces a satisfying policy when there exists one. In the remainder of this section, we will let $\Theta = \{\theta_1, \dots, \theta_{\lambda}\}$ such that $\theta_1, \dots, \theta_{\lambda}$ are ordered by the SSE utility they offer the defender in the truthful setting, i.e., $\hat{u}(\theta_1) \geq \hat{u}(\theta_2) \geq \dots \geq \hat{u}(\theta_{\lambda})$; the ordered types can be obtained efficiently given that the SSEs can be computed in polynomial time. We call a policy ℓ -compatible if truthful report is incentivized for every attacker type θ_j , $j \leq \ell$; see Definition 8.

Algorithm 1: Decide if there exists a policy π such that $\text{EoP}(\pi) \geq \xi$.

1. For each $\theta \in \Theta$, compute an SSE $(\hat{\mathbf{c}}^\theta, \hat{i}^\theta)$ on type θ . Let $\hat{u}(\theta) = u^d(\hat{\mathbf{c}}^\theta, \hat{i}^\theta)$.
 2. Sort attacker types in Θ by $\hat{u}(\theta)$, so that $\hat{u}(\theta_1) \geq \hat{u}(\theta_2) \cdots \geq \hat{u}(\theta_\lambda)$.
 3. For each $\ell = 1, \dots, \lambda$, let $\pi(\theta_\ell) = (\mathbf{z}, t)$, where $z_i = \min\{\hat{c}_i^{\theta_\ell}, h_i\}$, $t = \text{BR}_{\theta_\ell}(\mathbf{h})$, and
$$h_i = \max \left\{ 0, \frac{\xi \cdot \hat{u}(\theta_\ell) - p_i^d}{r_i^d - p_i^d}, \max_{\theta \in \{\theta_1, \dots, \theta_{\ell-1}\}} \frac{u_\theta^a(\pi(\theta)) - r_i^\theta}{p_i^\theta - r_i^\theta} \right\}.$$
 4. If $\text{EoP}(\pi) \geq \xi$, return π as a satisfying policy; Otherwise, claim that no such policy exists.
-

The correctness of Algorithm 1 is shown via Theorem 10. Briefly speaking, Algorithm 1 can be viewed as a process of repeatedly replacing the ℓ -th outcome of a satisfying policy (suppose we are given one) with the outcome generated in the ℓ -th iteration of Step 3. The observation in Lemma 9 ensures that the new policy obtained after every replacement will still be a satisfying one. Hence, eventually, we will obtain a satisfying policy that consists of outcomes generated through the algorithm, and this means that we do not need to actually know a satisfying policy to begin with. Interestingly, the policy generated by Algorithm 1 is also *incentive compatible (IC)* (λ -compatible as in Definition 8); it always incentivizes the attacker to report their true type.

Definition 8. A policy π is ℓ -compatible ($0 \leq \ell \leq \lambda$), if in response to π , it is optimal for every attacker type $\theta \in \{\theta_1, \dots, \theta_\ell\}$ to report truthfully, i.e., $u_\theta^a(\pi(\theta)) \geq u_\theta^a(\pi(\beta))$ for all $\beta \in \Theta$.

Lemma 9. Let π be the policy generated in Step 3 of Algorithm 1. Suppose that there exists an $(\ell - 1)$ -compatible policy π^* , $\text{EoP}(\pi^*) \geq \xi$. Then the policy $\tilde{\pi}$, such that
$$\tilde{\pi}(\theta) = \begin{cases} \pi^*(\theta), & \text{if } \theta \in \Theta \setminus \{\theta_\ell\} \\ \pi(\theta), & \text{if } \theta = \theta_\ell \end{cases},$$
 is feasible and ℓ -compatible, and $\text{EoP}(\tilde{\pi}) \geq \xi$.

Theorem 10. In time polynomial in m , n , and $|\Theta|$, Algorithm 1 either outputs a policy π with $\text{EoP}(\pi) \geq \xi$, or decides correctly that no such policy exists. The policy generated is IC.

Proof. The algorithm terminates in λ iterations and the polynomial runtime is readily seen.

Given how \mathbf{z}^θ and t^θ are set in Step 3, Algorithm 1 always constructs a feasible policy, so when no feasible policy can achieve $\text{EoP} \xi$, we have $\text{EoP}(\pi) < \xi$ in Step 4 and the algorithm will decide correctly that no satisfying policy exists.

Suppose that there exists a policy π^* , $\text{EoP}(\pi^*) \geq \xi$. Let $\tilde{\pi}^0 = \pi^*$, and for each $\ell = 1, \dots, \lambda$, we construct a policy $\tilde{\pi}^\ell$ by replacing the outcome $\tilde{\pi}^{\ell-1}(\theta_\ell)$ in $\tilde{\pi}^{\ell-1}$ with $\pi(\theta_\ell)$ generated in Step 3 of Algorithm 1; thus, $\tilde{\pi}^\lambda = \pi$. Trivially, $\tilde{\pi}^0$ is 0-compatible as is any feasible policy. Applying Lemma 9 iteratively, we can conclude that $\tilde{\pi}^\ell$ is ℓ -compatible and $\text{EoP}(\tilde{\pi}^\ell) \geq \xi$ for every ℓ ; in particular, $\tilde{\pi}^\lambda$ is λ -compatible and $\text{EoP}(\tilde{\pi}^\lambda) \geq \xi$. Algorithm 1 will output $\tilde{\pi}^\lambda$ as a satisfying policy in Step 4. \square

5.2 Beyond Finite Attacker Types

The above approach only applies to a finite type set, we now present a heuristic approach to deal with a continuous or even unknown attacker type set. The approach is inspired by the *quantal response (QR)* model that is developed to study bounded rationality of human players [13]. In a QR equilibrium, players are assumed to play not only their optimal pure strategy, but also every other strategy with some probability. The probability of a strategy being chosen is positively related to the utility the player gets from playing that strategy.

The QR policy imitates the irrational behavior in a QR equilibrium. Recall that in an SSE, a rational defender commits to an optimal strategy \mathbf{c} and induces a type- θ attacker to choose a response $i^* \in \text{BR}_\theta(\mathbf{c})$ that maximizes the defender's utility. The QR policy, however, induces the attacker's tie-breaking choice in an "irrational" way. It induces the attacker to choose not only i^* , but also every target in $\text{BR}_\theta(\mathbf{c})$ with some probability; the probability a target being chosen is positively related to the defender's utility when this target is attacked. The idea is to add some uncertainty in the outcome, so that the attacker cannot benefit from the defender inducing him to choose a particular response, which is crucial for his manipulation. This encourages truthful report to some extent: a truthful attacker, who reports his true type θ , is indifferent of which response he is induced to choose in $\text{BR}_\theta(\mathbf{c})$ and hence immune to such uncertainty. The QR policy is as follows.

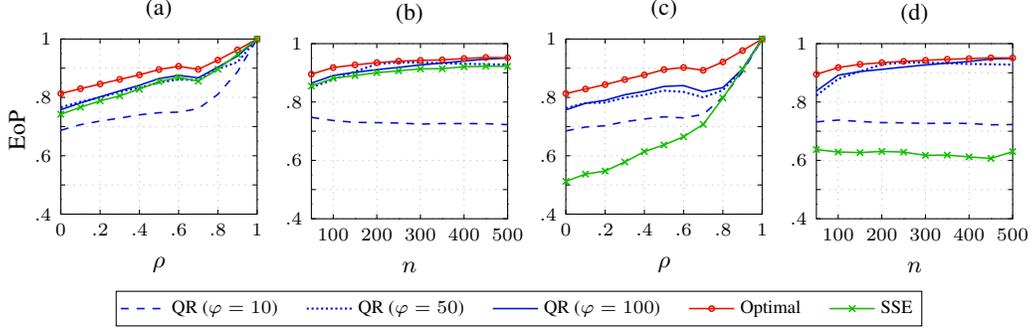


Figure 1: Empirical evaluation of the EoP. In (a), results are obtained with other parameters set to $\lambda = 100$, $m = 10$, and $n = 50$; and in (b) with $m = n/5$, $\rho = 0.5$, and $\lambda = 100$. Figures (c) and (d) repeat (a) and (b), respectively, with the difference that the results are obtained when the zero-sum attacker type is always included in Θ in addition to other randomly generated types.

Definition 11 (QR policy). For each type θ , let \hat{c}^θ be the defender strategy in an SSE on attacker type θ . A QR policy π^{QR} maps a reported type θ to a distribution σ over outcomes in $\{(\hat{c}^\theta, i) : i \in \text{BR}_\theta(\hat{c}^\theta)\}$, such that the probability $\sigma(i)$ of each outcome (\hat{c}^θ, i) is

$$\sigma(i) = \frac{e^{\varphi \cdot u^d(\hat{c}^\theta, i)}}{\sum_{j \in \text{BR}_\theta(\hat{c}^\theta)} e^{\varphi \cdot u^d(\hat{c}^\theta, j)}},$$

where $\varphi > 0$ is a parameter that represents a player’s rationality level in the QR model.³

A defender who uses π^{QR} then samples an outcome from $\pi^{\text{QR}}(\theta)$ to implement when θ is reported. The players are now concerned with their expected utility over the outcome distribution, e.g., for the defender: $u^d(\pi^{\text{QR}}(\theta)) = \sum_{i \in \text{BR}_\theta(\hat{c}^\theta)} \sigma_\theta(i) \cdot u^d(\hat{c}^\theta, i)$. The EoP is defined accordingly with the expected utility. Since \hat{c}^θ and σ_θ are independent of the type set Θ , a QR policy can be implemented on-the-fly for the reported type, and is thus able to handle infinite or unknown type sets.

The QR policy strikes a balance between two unaligned aspects of playing against attacker manipulation: 1) it adds some noise to discourage attacker manipulation; meanwhile, 2) the softmax function that defines $\sigma_\theta(i)$ loosely strings the induced attacker response to the optimal one for the defender, so that the cost of achieving 1) is kept away from being too high. This reveals the rationality of behaving in a “irrational” manner in the presence of attacker manipulation. We empirically evaluate the performance of the QR policy in randomly generated games in the next section.

6 Empirical Evaluation

In our evaluations, attacker types are randomly generated using the covariance model [15], with a parameter $\rho \in [0, 1]$ controlling the closeness of the generated game to a zero-sum game. That is, we shift each payoff parameter x towards the corresponding one y of a zero-sum attacker type and let $x \leftarrow (1 - \rho) \cdot x + \rho \cdot y$. Thus, when $\rho = 1$ the game generated is exactly zero-sum, and when $\rho = 0$ all payoffs are generated uniformly at random. All evaluations are conducted on finite type sets since it is unclear to us how the optimal attacker report can be obtained on infinite type sets. For our QR policy, such a setting still simulates situations with an unknown (but finite) type set, though situations with infinite type sets requires more advanced approaches to evaluate. All results shown are the average of at least 50 runs.

We compare the EoP achieved by our optimal and heuristic policies, with the SSE policy as a benchmark (i.e., the situation when attacker manipulation is ignored). The results are shown in Figure 1. The first set of results in (a) and (b) shows the variance of the EoP with respect to ρ and the scale of the game. Except for the QR policy with $\varphi = 10$, the performance of all other approaches is

³When $\varphi \rightarrow 0$, a player behaves completely irrationally, playing each strategy uniformly at random; when $\varphi \rightarrow +\infty$, a player becomes fully rational, choosing the optimal strategy with probability approaching 1.

very close to each other in all these figures, though there is still a discernable gap between the optimal policy and the SSE policy. In general, in these results, the loss due to ignoring attacker manipulation appears to be insignificant; the improvement brought by our approaches is marginal.

A more interesting set of results is shown (c) and (d), in which we slightly tweak the randomly generated type set, by always adding a zero-sum attacker type in it. This small change leads to a very different pattern in the results. There is a very significant gap between the optimal and the SSE policies, and the QR policies normally rest in between them. The results corroborate our theoretical analysis, that all attacker types will be incentivized to report the zero-sum type when they are allowed to, which undermines the performance of the SSE policy significantly. The optimal policy, however, is able to achieve very high EoP, close to recovering the defender’s truthful utility ($EoP = 1$) in many results. The QR policies also exhibit good performance in this setting, and the performance tends to be robust with respect to the problem size.

7 Conclusion

In this paper, we investigate manipulation of algorithms that are designed to learn the optimal strategy to commit to in Stackelberg security games, and aim at remedying the overoptimistic assumption of a truthful attacker adopted by these algorithms. We build a novel game framework for this purpose, and propose exact and heuristic approaches to reduce the loss due to manipulation in this framework. The effectiveness of our approaches are evaluated both theoretically and empirically. It would be interesting to look at similar problems in other variants of Stackelberg games, where our framework and approaches may apply.

Acknowledgments

Jiarui Gan was supported by the EPSRC International Doctoral Scholars Grant EP/N509711/1.

References

- [1] Maria-Florina Balcan, Avrim Blum, Nika Haghtalab, and Ariel D Procaccia. Commitment without regrets: Online learning in stackelberg security games. In *Proceedings of the 16th ACM conference on economics and computation (EC’15)*, pages 61–78, 2015.
- [2] Battista Biggio, Blaine Nelson, and Pavel Laskov. Support vector machines under adversarial label noise. In *Proceedings of the Asian Conference on Machine Learning (ACML’11)*, volume 20, pages 97–112, 2011.
- [3] Battista Biggio, Blaine Nelson, and Pavel Laskov. Poisoning attacks against support vector machines. In *Proceedings of the 29th International Conference on International Conference on Machine Learning (ICML’12)*, pages 1467–1474, 2012.
- [4] Avrim Blum, Nika Haghtalab, and Ariel D. Procaccia. Learning optimal commitment to overcome insecurity. In *Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS’14)*, pages 1826–1834, 2014.
- [5] Anirban Chakraborty, Manaar Alam, Vishal Dey, Anupam Chattopadhyay, and Debdeep Mukhopadhyay. Adversarial attacks and defences: A survey. *CoRR*, abs/1810.00069, 2018.
- [6] Vincent Conitzer and Tuomas Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM Conference on Electronic Commerce (EC’06)*, pages 82–90, 2006.
- [7] Jiarui Gan, Haifeng Xu, Qingyu Guo, Long Tran-Thanh, Zinovi Rabinovich, and Michael Wooldridge. Imitative follower deception in stackelberg games. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC’19)*, pages 639–657, 2019.
- [8] Nika Haghtalab, Fei Fang, Thanh H. Nguyen, Arunesh Sinha, Ariel D. Procaccia, and Milind Tambe. Three strategies to success: Learning adversary models in security games. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI’16)*, pages 308–314, 2016.

- [9] Manish Jain, James Pita, Milind Tambe, Fernando Ordóñez, Praveen Paruchuri, and Sarit Kraus. Bayesian Stackelberg games and their application for security at Los Angeles International Airport. *SIGecom Exch.*, 7(2):10:1–10:3, June 2008.
- [10] Christopher Kiekintveld, Towhidul Islam, and Vladik Kreinovich. Security games with interval uncertainty. In *Proceedings of The 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'13)*, pages 231–238, 2013.
- [11] 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.
- [12] Qiang Liu, Pan Li, Wentao Zhao, Wei Cai, Shui Yu, and Victor CM Leung. A survey on security threats and defensive techniques of machine learning: A data driven view. *IEEE access*, 6:12103–12117, 2018.
- [13] Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6 – 38, 1995.
- [14] Thanh Hong Nguyen, Amulya Yadav, Bo An, Milind Tambe, and Craig Boutilier. Regret-based optimization and preference elicitation for Stackelberg security games with uncertainty. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI'14)*, pages 756–762, 2014.
- [15] Eugene Nudelman, Jennifer Wortman, Yoav Shoham, and Kevin Leyton-Brown. 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*, pages 880–887, 2004.
- [16] Praveen Paruchuri, Jonathan P Pearce, Janusz Marecki, Milind Tambe, Fernando Ordonez, and Sarit Kraus. 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)*, pages 895–902, 2008.
- [17] Binghui Peng, Weiran Shen, Pingzhong Tang, and Song Zuo. Learning optimal strategies to commit to. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI'19)*, 2019.
- [18] James Pita, Manish Jain, Fernando Ordóñez, Milind Tambe, Sarit Kraus, and Reuma Magori-Cohen. Effective solutions for real-world Stackelberg games: When agents must deal with human uncertainties. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'09)*, pages 369–376, 2009.
- [19] Aaron Roth, Jonathan Ullman, and Zhiwei Steven Wu. Watch and learn: Optimizing from revealed preferences feedback. In *Proceedings of the 48th annual ACM symposium on Theory of Computing (STOC'16)*, pages 949–962, 2016.
- [20] Bernhard von Stengel and Shmuel Zamir. Leadership with commitment to mixed strategies. *CDAM Research Report LSE-CDAM-2004-01*, London School of Economics, 2004.
- [21] Milind Tambe. *Security and game theory: algorithms, deployed systems, lessons learned*. Cambridge University Press, 2011.
- [22] Haifeng Xu, Long Tran-Thanh, and Nicholas R. Jennings. Playing repeated security games with no prior knowledge. In *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16)*, pages 104–112, 2016.

A Optimal Attacker Report

Lemma 12. *Suppose $(\hat{\mathbf{c}}, \hat{i})$ is an SSE on attacker type θ . The following holds:*

- (i) *If $\hat{c}_i < 1$ for all $i \in T$, then $\{i \in T : \hat{c}_i > 0\} \subseteq \text{BR}_\theta(\hat{\mathbf{c}})$ and $\sum_{i \in T} \hat{c}_i = m$.*
- (ii) *If $\hat{c}_i = 1$ for some $i \in T$, then there exists $j \in \text{BR}_\theta(\hat{\mathbf{c}})$ such that $\hat{c}_j = 1$.*

Proof. By definition, $\hat{\mathbf{c}}$ forms an SSE with some $\hat{i} \in \text{BR}_\theta(\hat{\mathbf{c}})$, i.e., $u^d(\hat{\mathbf{c}}, \hat{i}) \geq u^d(\mathbf{c}, i)$ for all $\mathbf{c} \in \mathcal{C}$, $i \in \text{BR}_\theta(\mathbf{c})$. We claim that $\hat{\mathbf{c}}$ is the optimal solution to the following linear program:

$$\max_{\mathbf{c}} \quad u^d(\mathbf{c}, \hat{i}) \quad (4)$$

$$\text{s.t.} \quad u_\theta^a(\mathbf{c}, i) \leq u_\theta^a(\mathbf{c}, \hat{i}) \quad \forall i \in T \setminus \{\hat{i}\} \quad (4a)$$

$$\sum_{i \in T} c_i \leq m \quad (4b)$$

$$0 \leq c_i \leq 1 \quad \forall i \in T \quad (4c)$$

To see this, note that, first, $\hat{\mathbf{c}}$ satisfies all the constraints above: Eq. (4a) is equivalent to $\hat{i} \in \text{BR}_\theta(\mathbf{c})$, and Eqs. (4b) and (4c) combined are equivalent to $\mathbf{c} \in \mathcal{C}$. Thus, it is a feasible solution. Second, the assumption that $\hat{\mathbf{c}}$ is not optimal would imply the existence of another feasible solution $\mathbf{z} \neq \hat{\mathbf{c}}$, such that $u^d(\mathbf{z}, \hat{i}) > u^d(\hat{\mathbf{c}}, \hat{i})$; this contradicts the assumption that $(\hat{\mathbf{c}}, \hat{i})$ forms an SSE.

By the Karush-Kuhn-Tucker (KKT) conditions, $\hat{\mathbf{c}}$ is an optimal solution only if there exists constants (i.e., KKT multipliers) $\alpha_i, \beta, \gamma_i$ and δ_i , each corresponding to an inequality constraint in Eqs. (4a)–(4c), such that for all $i \in T$ (let $w_i^a = p_i^\theta - r_i^\theta$ and $w_i^d = r_i^d - p_i^d$ for each i below):

$$\begin{cases} -w_i^a \cdot \alpha_i - \beta + \gamma_i - \delta_i = 0 & (\text{if } i \neq \hat{i}) \\ w_i^d & -\beta + \gamma_i - \delta_i = 0 \end{cases} \quad (\text{by stationarity}^4) \quad (5)$$

$$\alpha_i, \beta, \gamma_i, \delta_i \geq 0 \quad (\text{by dual feasibility}) \quad (6)$$

$$\begin{cases} \alpha_i \cdot (u_\theta^a(\hat{\mathbf{c}}, i) - u_\theta^a(\hat{\mathbf{c}}, \hat{i})) = 0 \\ \beta \cdot (\sum_{j \in T} \hat{c}_j - m) = 0 \\ \gamma_i \cdot \hat{c}_i = 0 \\ \delta_i \cdot (\hat{c}_i - 1) = 0 \end{cases} \quad (\text{by complementary slackness}^5) \quad (7)$$

Now we show (i) and (ii) in the statement of the lemma separately.

Part (i). Since $\hat{c}_i < 1$ for all $i \in T$, we have $\delta_i = 0$ for all $i \in T$ by the last equation in Eq. (7). Suppose towards a contradiction that $\sum_{i \in T} \hat{c}_i < m$. We would have $\beta = 0$ by the second equation in Eq. (7); and further, by the second equation in Eq. (5), $w_i^d + \gamma_i = 0$, which is a contradiction as $w_i^d = r_i^d - p_i^d > 0$ and $\gamma_i \geq 0$. Thus, $\beta > 0$ and $\sum_{i \in T} \hat{c}_i = m$.

Similarly, if we suppose $\hat{c}_t > 0$ for some $t \in T$, but $t \notin \text{BR}_\theta(\hat{\mathbf{c}})$, we would have $u_\theta^a(\hat{\mathbf{c}}, t) < \max_{i \in T} u_\theta^a(\hat{\mathbf{c}}, i) = u_\theta^a(\hat{\mathbf{c}}, \hat{i})$. Thus, $\gamma_t = 0$ and $\alpha_t = 0$ by the third and the first equations in Eq. (7), and further, $\beta + \delta_t = 0$ by the first equation in Eq. (5) (note that $t \neq \hat{i}$ since $t \notin \text{BR}_\theta(\hat{\mathbf{c}})$), which contradicts $\beta > 0$ and $\delta_i = 0$ for all $i \in T$ as we have shown above.

Part (ii). Suppose that $\hat{c}_t = 1$, $t \in T$, but $\hat{c}_i < 1$ for all $i \in \text{BR}_\theta(\mathbf{c})$ (in particular, $\hat{c}_i < 1$). Thus, $t \notin \text{BR}_\theta(\hat{\mathbf{c}})$, so $u_\theta^a(\hat{\mathbf{c}}, t) < \max_{i \in T} u_\theta^a(\hat{\mathbf{c}}, i) = u_\theta^a(\hat{\mathbf{c}}, \hat{i})$. We have $\gamma_t = 0$ and $\alpha_t = 0$ by Eq. (7), which implies $\beta + \delta_t = 0$ by Eq. (5); thus $\beta = 0$. In addition, $\hat{c}_i < 1$ implies $\delta_i = 0$ by Eq. (7). Again, by Eq. (5), we end up with the contradiction that $w_i^d + \gamma_i = 0$. \square

⁴A solution x satisfies stationarity if $\nabla f(x) = \lambda_1 \cdot \nabla g_1(x) + \dots + \lambda_\ell \cdot \nabla g_\ell(x)$, where f is the objective function (minimization), each g_i corresponds to an inequality constraint (in the form $g_i(x) \leq 0$), and each λ_i is a KKT multiplier.

⁵A solution x satisfies complementary slackness if $\lambda_i \cdot g_i(x) = 0$ for each KKT multiplier λ_i and their corresponding inequality constraint function g_i .

Lemma 13. *Suppose $\bar{\mathbf{c}}$ is a maximin strategy of the defender, i.e., $\bar{\mathbf{c}} \in \arg \max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^d(\mathbf{c}, i)$. Then $(\bar{\mathbf{c}}, i)$ forms an SSE in a zero-sum game for any $i \in \text{BR}_\beta(\bar{\mathbf{c}})$, where $\beta = (-\mathbf{p}^d, -\mathbf{r}^d)$.*

Proof. Suppose $(\bar{\mathbf{c}}, j)$, $j \in \text{BR}_\beta(\bar{\mathbf{c}})$ is not an SSE. Thus, there exists $\mathbf{z} \in \mathcal{C}$ and $t \in \text{BR}_\beta(\mathbf{z})$, such that $u^d(\mathbf{z}, t) > u^d(\bar{\mathbf{c}}, j)$; equivalently, $u_\beta^a(\mathbf{z}, t) < u_\beta^a(\bar{\mathbf{c}}, j)$ as β makes the game zero-sum. Since $j \in \text{BR}_\beta(\bar{\mathbf{c}})$ and $t \in \text{BR}_\beta(\mathbf{z})$ are the attacker's best responses, we have $u_\beta^a(\bar{\mathbf{c}}, j) = \max_{i \in T} u_\beta^a(\bar{\mathbf{c}}, i)$ and $u_\beta^a(\mathbf{z}, t) = \max_{i \in T} u_\beta^a(\mathbf{z}, i)$ by definition; thus, $\max_{i \in T} u_\beta^a(\mathbf{z}, i) < \max_{i \in T} u_\beta^a(\bar{\mathbf{c}}, i)$. This leads to the following contradiction:

$$\max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^d(\mathbf{c}, i) = \min_{i \in T} u^d(\bar{\mathbf{c}}, i) = -\max_{i \in T} u_\beta^a(\bar{\mathbf{c}}, i) < -\max_{i \in T} u_\beta^a(\mathbf{z}, i) = \min_{i \in T} u^d(\mathbf{z}, i). \quad \square$$

Corollary 14. *Suppose $\bar{\mathbf{c}}$ is a maximin strategy of the defender and $\bar{\mathbf{c}}$ is fully mixed, i.e., $0 < \bar{c}_i < 1$ for all $i \in T$. Then $(\bar{\mathbf{c}}, i)$ forms an SSE in a zero-sum game for all $i \in T$, and $u^d(\bar{\mathbf{c}}, i) = \min_{j \in T} u^d(\bar{\mathbf{c}}, j)$ for all $i \in T$.*

Proof. By Lemma 13, $(\bar{\mathbf{c}}, i)$ forms an SSE on attacker type $\beta = (-\mathbf{p}^d, -\mathbf{r}^d)$ for all $i \in \text{BR}_\beta(\bar{\mathbf{c}})$. Since $\bar{\mathbf{c}}$ is fully mixed, $T = \{i \in T : \bar{c}_i > 0\}$, and by Lemma 12 (i), $T \subseteq \text{BR}_\beta(\bar{\mathbf{c}}) \subseteq T$. Thus, $\text{BR}_\beta(\bar{\mathbf{c}}) \subseteq T$, so $(\bar{\mathbf{c}}, i)$ forms an SSE on attacker type β (i.e., in a zero-sum game) for all $i \in T$; we have $u_\beta^a(\bar{\mathbf{c}}, i) = \max_{j \in T} u_\beta^a(\bar{\mathbf{c}}, j)$. Further, it follows that $u^d(\bar{\mathbf{c}}, i) = -u_\beta^a(\bar{\mathbf{c}}, i) = -\max_{j \in T} u_\beta^a(\bar{\mathbf{c}}, j) = \min_{j \in T} u^d(\bar{\mathbf{c}}, j)$. \square

Lemma 15. *Suppose $\bar{\mathbf{c}}$ is a maximin strategy of the defender and $\bar{\mathbf{c}}$ is fully mixed, i.e., $0 < \bar{c}_i < 1$ for all $i \in T$. Then $\bar{\mathbf{c}}$ is the only maximin strategy of the defender.*

Proof. Suppose $\mathbf{z} = \arg \max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^d(\mathbf{c}, i)$ is a maximin strategy and $\mathbf{z} \neq \bar{\mathbf{c}}$. Thus, either: (i) $z_i \geq \bar{c}_i$ for all $i \in T$, and this is strictly satisfied by some i ; or (ii) $z_j < \bar{c}_j$ for some $j \in T$. We show either of them leads to a contradiction.

Since $\bar{\mathbf{c}}$ is a maximin strategy, it is also an SSE defender strategy by Lemma 13; and by Lemma 12 (i), $\sum_{i \in T} \bar{c}_i = m$. Thus, in the former case, it follows immediately that $\sum_{i \in T} z_i > \sum_{i \in T} \bar{c}_i = m$, which contradicts $\mathbf{z} \in \mathcal{C}$. In the latter case, we have $u^d(\mathbf{z}, j) < u^d(\bar{\mathbf{c}}, j)$ by monotonicity of $u^d(\cdot, j)$, which implies $\min_{i \in T} u^d(\mathbf{z}, i) \leq u^d(\mathbf{z}, j) < u^d(\bar{\mathbf{c}}, j) = \min_{i \in T} u^d(\bar{\mathbf{c}}, i)$, where the last equality follows by Corollary 14. This contradicts the assumption that \mathbf{z} is a maximin strategy. \square

A.1 Proof of Theorem 3

Proof. Let $\bar{\mathbf{c}}$ be a maximin strategy of the defender and \bar{u} be the corresponding maximin utility, i.e., $\bar{\mathbf{c}} \in \arg \max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^d(\mathbf{c}, i)$ and $\bar{u} = \min_{i \in T} u^d(\bar{\mathbf{c}}, i)$. Consider the following solution (β, \mathbf{z}, t) :

- $z_i = \max \left\{ 0, \frac{\bar{u} - p_i^d}{r_i^d - p_i^d} \right\}$ for all $i \in T$; (8)

- $t \in \text{BR}_\theta(\mathbf{z})$ is an arbitrary best response of a type- θ attacker; (9)

- $\beta = (\mathbf{r}, \mathbf{p})$, where $r_i = \begin{cases} -p_i^d, & \text{if } i \neq t \\ -\min\{p_t^d, \bar{u}\}, & \text{if } i = t \end{cases}$ and $p_i = -r_i^d$ for all $i \in T$. (10)

We show that (i) \mathbf{z} is a maximin defender strategy, (ii) (β, \mathbf{z}, t) is a feasible solution of Program (3) and (iii) it is optimal.

We first focus on the case when $p_t^d \leq \bar{u}$, and will show how the proof can be modified to show the same results when $p_t^d > \bar{u}$. When $p_t^d \leq \bar{u}$ we have $(r_i, p_i) = (-p_i^d, -r_i^d)$ for all $i \in T$ by the specification in (10), so for any $\mathbf{c} \in \mathcal{C}$,

$$u_\beta^a(\mathbf{c}, i) = -u^d(\mathbf{c}, i). \quad (11)$$

Now we show (i)–(iii).

(i) Maximin. For all $i \in T$, since $z_i = \max \left\{ 0, \frac{\bar{u} - p_i^d}{r_i^d - p_i^d} \right\} \geq \frac{\bar{u} - p_i^d}{r_i^d - p_i^d}$, we have

$$u^d(\mathbf{z}, i) \geq u^d \left(\frac{\bar{u} - p_i^d}{r_i^d - p_i^d}, i \right) = \frac{\bar{u} - p_i^d}{r_i^d - p_i^d} \cdot r_i^d + \left(1 - \frac{\bar{u} - p_i^d}{r_i^d - p_i^d} \right) \cdot p_i^d = \bar{u}. \quad (12)$$

It follows that

$$\min_{i \in T} u^d(\mathbf{z}, i) \geq \bar{u} \geq \min_{i \in T} u^d(\mathbf{c}, i)$$

for all $\mathbf{c} \in \mathcal{C}$, so \mathbf{z} is indeed a maximin strategy. We still need to make sure that \mathbf{z} is feasible, i.e., $\mathbf{z} \in \mathcal{C}$. Observe that when $z_i > 0$, (12) becomes an equality, so we have $u^d(\mathbf{z}, i) = \bar{u} \leq u^d(\bar{\mathbf{c}}, i)$ which implies $z_i \leq \bar{c}_i$ by monotonicity. Thus, $\sum_{i \in T} z_i \leq \sum_{i \in T} \bar{c}_i \leq m$. It remains to show that $0 \leq z_i \leq 1$ for all i . Trivially, by (8), $z_i \geq 0$ for all i . To see that $z_i \leq 1$, it suffices to show that $\frac{\bar{u} - p_i^d}{r_i^d - p_i^d} \leq 1$. Indeed, this holds as $\bar{u} = \min_{j \in T} u^d(\bar{\mathbf{c}}, j) \leq u^d(\bar{\mathbf{c}}, i) \leq r_i^d$.

(ii) Feasibility We show that (β, \mathbf{z}, t) satisfies all the constraints of Program (3). Clearly, (3b) is satisfied because $r_i \geq -p_i^d > -r_i^d = p_i$ for all $i \in T$ by (10). To see that it also satisfies (3a), first observe that when $p_t^d \leq \bar{u}$ we have

$$u_\beta^a(\mathbf{z}, t) = u_\beta^a \left(\frac{\bar{u} - p_t^d}{r_t^d - p_t^d}, t \right) = \left(1 - \frac{\bar{u} - p_t^d}{r_t^d - p_t^d} \right) \cdot (-p_t^d) + \frac{\bar{u} - p_t^d}{r_t^d - p_t^d} \cdot (-r_t^d) = -\bar{u}. \quad (13)$$

Combining this with (11) and (12) gives, for all i ,

$$u_\beta^a(\mathbf{z}, t) = -\bar{u} \geq -u^d(\mathbf{z}, i) = u_\beta^a(\mathbf{z}, i). \quad (14)$$

Thus, $t \in \text{BR}_\beta(\mathbf{z})$.

Now that $t \in \text{BR}_\beta(\mathbf{z})$ and in (i) we have shown that $\mathbf{z} \in \mathcal{C}$, if we suppose (3a) is not satisfied, we would have $u^d(\mathbf{z}', t') > u^d(\mathbf{z}, t)$ for some $\mathbf{z}' \in \mathcal{C}$ and $t' \in \text{BR}_\beta(\mathbf{z}')$. Applying (11), we find the following for all i :

$$u^d(\mathbf{z}', i) = -u_\beta^a(\mathbf{z}', i) \geq -u_\beta^a(\mathbf{z}', t') = u^d(\mathbf{z}', t'),$$

where the inequality is due to the fact that $t' \in \text{BR}_\beta(\mathbf{z}')$. Thus,

$$u^d(\mathbf{z}', i) \geq u^d(\mathbf{z}', t') > u^d(\mathbf{z}, t) \geq \min_{i \in T} u^d(\mathbf{z}, i) = \bar{u}. \quad (15)$$

It follows that

$$\min_{i \in T} u^d(\mathbf{z}', i) > \bar{u} = \max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^d(\mathbf{c}, i),$$

which is a contradiction given that $\mathbf{z}' \in \mathcal{C}$.

(iii) Optimality. Suppose that (β, \mathbf{z}, t) is not optimal. Thus, there exists a feasible solution $(\beta', \mathbf{z}', t')$ such that $u_\theta^a(\mathbf{z}', t') > u_\theta^a(\mathbf{z}, t)$. By 9, $t \in \text{BR}_\theta(\mathbf{z})$, so we have $u_\theta^a(\mathbf{z}, t') \leq u_\theta^a(\mathbf{z}, t) < u_\theta^a(\mathbf{z}', t')$, which implies $z_{t'} < z_{t'}$ by monotonicity. Since it is defined $z_{t'} = \max \left\{ 0, \frac{\bar{u} - p_{t'}^d}{r_{t'}^d - p_{t'}^d} \right\}$, now that $z_{t'} > z_{t'} \geq 0$, it must be that $z_{t'} < z_{t'} = \frac{\bar{u} - p_{t'}^d}{r_{t'}^d - p_{t'}^d}$. Substituting this into the defender's utility function gives

$$\begin{aligned} u^d(\mathbf{z}', t') &< u^d \left(\frac{\bar{u} - p_{t'}^d}{r_{t'}^d - p_{t'}^d}, t' \right) = \frac{\bar{u} - p_{t'}^d}{r_{t'}^d - p_{t'}^d} \cdot r_{t'}^d + \left(1 - \frac{\bar{u} - p_{t'}^d}{r_{t'}^d - p_{t'}^d} \right) \cdot p_{t'}^d \\ &= \bar{u} = \max_{\mathbf{c} \in \mathcal{C}} \min_{i \in T} u^d(\mathbf{c}, i) \leq \max_{\mathbf{c} \in \mathcal{C}, i \in \text{BR}_\beta(\mathbf{c})} u^d(\mathbf{c}, i). \end{aligned}$$

Thus, $(\beta', \mathbf{z}', t')$ violates (3a), contradicting the assumption that $(\beta', \mathbf{z}', t')$ is a feasible solution.

It remains to deal with the case when $p_t^d > \bar{u}$. The only difference in this case is that now $r_t = -\bar{u} > -p_t^d$ by (10), so (11) only holds for $i \neq t$. In our proof above, the arguments that rely on the assumption that $p_t^d \leq \bar{u}$ and (11) are the equations in (13), (14), and (15), where the latter two now only hold for all $i \neq t$. However, observe the following:

- When $p_t^d > \bar{u}$, we have $\frac{\bar{u}-p_t^d}{r_t^d-p_t^d} < 0$ and thus $z_t = 0$ by (8). It follows that $u_\beta^a(\mathbf{z}, t) = u_\beta^a(0, t) = r_t = -\bar{u}$, so (13) holds as well.
- (14) holds trivially for $i = t$.
- When $p_t^d > \bar{u}$, we have $u^d(\mathbf{z}', t) \geq p_t^d > \bar{u}$, thus establishing (15) for $i = t$ as well.

Therefore, (i)–(iii) hold when $p_t^d > \bar{u}$ and the proof is completed. \square

A.2 Proof of Theorem 4

Proof. In Theorem 3, we have shown the existence of an optimal solution containing a maximin strategy of the defender. By Lemma 15, $\bar{\mathbf{c}}$ is the only maximin strategy if it is fully mixed. Thus, there exists an optimal solution that contains $\bar{\mathbf{c}}$. We fix $\bar{\mathbf{c}}$ in Program (3), and show that in the resultant program, an optimal solution (β', t') is such that $\beta' = (-\mathbf{p}^d, -\mathbf{r}^d)$ and $t' \in \arg \max_{i \in T} u_\theta^a(\bar{\mathbf{c}}, i)$. In fact, now that $\bar{\mathbf{c}}$ is fixed, the optimality of (β', t') follows directly from the specification $t' \in \arg \max_{i \in T} u_\theta^a(\bar{\mathbf{c}}, i)$. It remains to show that (β', t') is feasible. Since the game with attacker type β' is a zero-sum game, by Lemma 13, $(\bar{\mathbf{c}}, t')$ forms an SSE on type β' . Thus, Constraint (3a) is satisfied, and (β', t') is feasible. Therefore, $(\beta', \bar{\mathbf{c}}, t')$ is an optimal solution to Program (3).

Now we show the first part of the theorem, i.e., the uniqueness of the induced defender strategy. Suppose for a contradiction that there exists an SSE $(\hat{\mathbf{c}}, \hat{i})$ on attacker type β , such that $\hat{\mathbf{c}} \neq \bar{\mathbf{c}}$. Consider the two possibilities under this condition.

Case (i). $\hat{c}_i \leq \bar{c}_i$ for all $i \in T$, and this is strictly satisfied for some i . It follows that $\sum_{i \in T} \hat{c}_i < \sum_{i \in T} \bar{c}_i \leq m$, which contradicts Lemma 12 (i).

Case (ii). $\hat{c}_j > \bar{c}_j$ for some $j \in T$. If it is also the case that $\hat{c}_i < 1$ for all $i \in T$, by Lemma 12 (i), we have $j \in \text{BR}_\beta(\hat{\mathbf{c}})$; If otherwise $\hat{c}_i = 1$ for some $i \in T$, by Lemma 12 (ii), there exists $j' \in \text{BR}_\beta(\hat{\mathbf{c}})$ such that $\hat{c}_{j'} = 1 > \bar{c}_{j'}$. In both cases, we find some $j \in \text{BR}_\beta(\hat{\mathbf{c}})$ such that $\hat{c}_j > \bar{c}_j$; hence, $u^d(\hat{\mathbf{c}}, j) > u^d(\bar{\mathbf{c}}, j)$ by monotonicity of $u^d(\cdot, j)$. We have

$$u^d(\hat{\mathbf{c}}, \hat{i}) = \max_{i \in \text{BR}_\beta(\hat{\mathbf{c}})} u^d(\hat{\mathbf{c}}, i) \geq u^d(\hat{\mathbf{c}}, j) > u^d(\bar{\mathbf{c}}, j) = u^d(\bar{\mathbf{c}}, \hat{i}),$$

where the last equality follows by Corollary 14. Thus, by the monotonicity, we have $\hat{c}_i > \bar{c}_i$ and, in turn, $u_\theta^a(\hat{\mathbf{c}}, \hat{i}) < u_\theta^a(\bar{\mathbf{c}}, \hat{i})$. This gives

$$u_\theta^a(\hat{\mathbf{c}}, \hat{i}) < u_\theta^a(\bar{\mathbf{c}}, \hat{i}) \leq \max_{i \in T} u_\theta^a(\bar{\mathbf{c}}, i) = u_\theta^a(\bar{\mathbf{c}}, t'),$$

so $(\beta', \bar{\mathbf{c}}, t')$ is a better solution than (β, \mathbf{z}, t) and this contradicts the assumption in the statement of the theorem.

Both cases lead to contradictions. This completes the proof. \square

B The EoP Measure

B.1 Proof of Proposition 5

Proof. Let $(\mathbf{z}, j) = \pi(\beta)$, which is the outcome a type- β attacker would get if he reports truthfully. By definition, $j \in \text{BR}_\beta(\mathbf{z})$. By Lemma 13, $\bar{\mathbf{c}}$ as the defender's maximin strategy is exactly her SSE strategy in a zero-sum game and $(\bar{\mathbf{c}}, t)$ forms an SSE for any $t \in \text{BR}_\beta(\bar{\mathbf{c}})$. Thus, $u^d(\mathbf{z}, j) \leq u^d(\bar{\mathbf{c}}, t)$ and $u_\beta^a(\bar{\mathbf{c}}, t) = \max_{i \in T} u_\beta^a(\bar{\mathbf{c}}, i)$. Since β makes the game zero-sum, $u_\beta^a(\mathbf{c}, i) = u^d(\mathbf{c}, i)$ for any \mathbf{c} and i . It follows that

$$u_\beta^a(\mathbf{z}, j) = -u^d(\mathbf{z}, j) \geq -u^d(\bar{\mathbf{c}}, t) = u_\beta^a(\bar{\mathbf{c}}, t) = \max_{i \in T} u_\beta^a(\bar{\mathbf{c}}, i) = -\min_{i \in T} u^d(\bar{\mathbf{c}}, i) = -\bar{u}.$$

Now suppose towards a contradiction that $u^d(\pi(\gamma)) > \bar{u}$. Then we have

$$u_\beta^a(\pi(\gamma)) = -u^d(\pi(\gamma)) < -\bar{u} \leq u_\beta^a(\mathbf{z}, j) = u_\beta^a(\pi(\beta)),$$

so the attacker would be strictly better-off reporting β . This contradicts the assumption that γ is an optimal reporting strategy of a type- β attacker. \square

B.2 Proof of Proposition 7

Proof. Clearly, $\text{EoP}(\pi) \geq 0$ as the payoffs are shifted to be non-negative. We show that $\text{EoP}(\pi) \leq 1$.

Let $\hat{u}(\theta) = \max_{\mathbf{c} \in \mathcal{C}, i \in \text{BR}_\theta(\mathbf{c})} u^d(\mathbf{c}, i)$ denote the defender's utility in an SSE on attacker type θ , and let β be an attacker type that provides the best defender utility in an SSE, i.e., $\beta \in \arg \max_{\theta \in \Theta} \hat{u}(\theta)$. Consider the best reporting strategy γ of a type- β attacker in response to π . For the outcome $(\mathbf{z}, t) = \pi(\gamma)$ to be feasible, we must have $t \in \text{BR}_\gamma(\mathbf{z})$; thus, $u^d(\pi(\gamma)) = u^d(\mathbf{z}, t) \leq \max_{\mathbf{c} \in \mathcal{C}, i \in \text{BR}_\gamma(\mathbf{c})} u^d(\mathbf{c}, i) = \hat{u}(\gamma) \leq \hat{u}(\beta)$. It follows that

$$\text{EoP}(\pi) = \min_{\theta \in \Theta} \text{EoP}_\theta(\pi) \leq \text{EoP}_\beta(\pi) = \frac{u^d(\pi(\gamma))}{\hat{u}(\beta)} \leq 1. \quad \square$$

C Correctness of Algorithm 1

Lemma 16. *Let $(\hat{\mathbf{c}}, \hat{i})$ be an arbitrary SSE on an attacker type $\theta \in \Theta$. For any policy π , truthful report guarantees a type- θ attacker his SSE utility, i.e., $u_\theta^a(\pi(\theta)) \geq u_\theta^a(\hat{\mathbf{c}}, \hat{i})$.*

Proof. Suppose towards a contradiction that $u_\theta^a(\pi(\theta)) < u_\theta^a(\hat{\mathbf{c}}, \hat{i})$. Let $\pi(\theta) = (\mathbf{z}, t)$. For π to be feasible, t must be a best response of a type- θ attacker to \mathbf{z} . Thus, $u_\theta^a(\mathbf{z}, t) \geq u_\theta^a(\mathbf{z}, i)$ for all $i \in T$; in particular, $u_\theta^a(\mathbf{z}, t) \geq u_\theta^a(\mathbf{z}, \hat{i})$. We have

$$u_\theta^a(\hat{\mathbf{c}}, \hat{i}) > u_\theta^a(\pi(\theta)) = u_\theta^a(\mathbf{z}, t) \geq u_\theta^a(\mathbf{z}, \hat{i}).$$

Since $u_\theta^a(\mathbf{c}, \hat{i})$ changes continuously with respect to c_i , the above inequality implies the existence of a number $\phi \in (\hat{c}_i, z_i]$, such that $u_\theta^a(\phi, \hat{i}) = u_\theta^a(\mathbf{z}, t)$.

Consider a defender strategy \mathbf{z}' with $z'_i = \phi$ and $z'_i = z_i$ for all $i \in T \setminus \{i\}$. We have $0 \leq z_i \leq 1$ and $\sum_{i \in T} z'_i \leq \sum_{i \in T} z_i \leq m$, so $\mathbf{z}' \in \mathcal{C}$. In addition, $u_\theta^a(\mathbf{z}, i) = u_\theta^a(\mathbf{z}', i)$ for all $i \in T \setminus \{i\}$. Thus,

$$u_\theta^a(\mathbf{z}', \hat{i}) = u_\theta^a(\phi, \hat{i}) = u_\theta^a(\mathbf{z}, t) \geq u_\theta^a(\mathbf{z}, i) = u_\theta^a(\mathbf{z}', i),$$

which means \hat{i} is a best response of a type- θ attacker to \mathbf{z}' , i.e., $\hat{i} \in \text{BR}_\theta(\mathbf{z}')$. This gives rise to the following contradiction:

$$\max_{\mathbf{c} \in \mathcal{C}, i \in \text{BR}_\theta(\mathbf{c})} u^d(\mathbf{c}, i) \geq u^d(\mathbf{z}', \hat{i}) = u^d(\phi, \hat{i}) > u^d(\hat{\mathbf{c}}, \hat{i}) = \max_{\mathbf{c} \in \mathcal{C}, i \in \text{BR}_\theta(\mathbf{c})} u^d(\mathbf{c}, i),$$

where $u^d(\phi, \hat{i}) > u^d(\hat{\mathbf{c}}, \hat{i})$ since $\phi \in (\hat{c}_i, z_i]$. \square

C.1 Proof of Lemma 9

Proof. We will write $\pi = (\mathbf{c}^\theta, i^\theta)_{\theta \in \Theta}$. Let $\beta \in \Theta$ be an optimal report of a type- θ_ℓ attacker in response to π , i.e., $u_{\theta_\ell}^a(\pi(\beta)) \geq u_{\theta_\ell}^a(\pi(\beta'))$ for all $\beta' \in \Theta$. We first show a couple of useful observations.

Claim 1. $u_{\theta_\ell}^a(\mathbf{h}, i^\beta) \geq u_{\theta_\ell}^a(\pi(\beta))$, i.e., a type- θ_ℓ attacker (weakly) prefers outcome (\mathbf{h}, i^β) to $\pi(\beta)$.

Proof of Claim 1. suppose towards a contradiction that $u_{\theta_\ell}^a(\mathbf{h}, i^\beta) < u_{\theta_\ell}^a(\pi(\beta)) = u_{\theta_\ell}^a(\mathbf{c}^\beta, i^\beta)$. By monotonicity of $u_{\theta_\ell}^a(\cdot, i^\beta)$, we have $c_{i^\beta}^\beta < h_{i^\beta} = \max \left\{ 0, \frac{\xi \cdot \hat{u}(\theta_\ell) - p_{i^\beta}^d}{r_{i^\beta}^d - p_{i^\beta}^d}, \max_{\theta \in \{\theta_1, \dots, \theta_{\ell-1}\}} \frac{u_\theta^a(\pi(\theta)) - r_{i^\beta}^\theta}{p_{i^\beta}^\theta - r_{i^\beta}^\theta} \right\}$. Since $c_{i^\beta}^\beta \geq 0$, we have $h_{i^\beta} > 0$, so either (i) $c_{i^\beta}^\beta < h_{i^\beta} = \frac{\xi \cdot \hat{u}(\theta_\ell) - p_{i^\beta}^d}{r_{i^\beta}^d - p_{i^\beta}^d}$, or (ii) $c_{i^\beta}^\beta < h_{i^\beta} = \frac{u_\theta^a(\pi(\theta)) - r_{i^\beta}^\theta}{p_{i^\beta}^\theta - r_{i^\beta}^\theta}$ for some $\theta \in \{\theta_1, \dots, \theta_{\ell-1}\}$. We show that both cases lead to contradictions.

Case (i). $c_{i^\beta}^\beta < \frac{\xi \cdot \hat{u}(\theta_\ell) - p_{i^\beta}^d}{r_{i^\beta}^d - p_{i^\beta}^d}$. It follows by monotonicity of $u^d(\cdot, i^\beta)$, that $u^d(\pi(\beta)) = u^d(\mathbf{c}^\beta, i^\beta) < u^d \left(\frac{\xi \cdot \hat{u}(\theta_\ell) - p_{i^\beta}^d}{r_{i^\beta}^d - p_{i^\beta}^d}, i^\beta \right) = \xi \cdot \hat{u}(\theta_\ell)$; thus, $\text{EoP}(\pi) \leq \text{EoP}_{\theta_\ell}(\pi) = \frac{u^d(\pi(\beta))}{\hat{u}(\theta_\ell)} < \xi$, which contradicts the assumption that π is a satisfying policy.

Case (ii). $c_{i^\beta}^\beta < \frac{u_\theta^a(\pi(\theta)) - r_{i^\beta}^\theta}{p_{i^\beta}^\theta - r_{i^\beta}^\theta}$ for some $\theta \in \{\theta_1, \dots, \theta_{\ell-1}\}$. It follows by monotonicity (decreasing) of $u^a(\cdot, i^\beta)$, that $u_\theta^a(\pi(\beta)) = u_\theta^a(\mathbf{c}^\beta, i^\beta) > u_\theta^a\left(\frac{u_\theta^a(\pi(\theta)) - r_{i^\beta}^\theta}{p_{i^\beta}^\theta - r_{i^\beta}^\theta}, i^\beta\right) = u_\theta^a(\pi(\theta))$, so a type- θ attacker would be strictly better-off reporting type β in response to π , contradicting the assumption that π is $(\ell - 1)$ -compatible. \square

Claim 2. $h_t \leq \hat{c}_t^{\theta_\ell}$; and hence, $z_t = \min\{\hat{c}_t^{\theta_\ell}, h_t\} = h_t$.

Proof of Claim 2. By definition, $t \in \text{BR}_{\theta_\ell}(\mathbf{h})$, so

$$u_{\theta_\ell}^a(\mathbf{h}, t) = \max_{i \in T} u_{\theta_\ell}^a(\mathbf{h}, i) \geq u_{\theta_\ell}^a(\mathbf{h}, i^\beta). \quad (16)$$

Since β is the optimal report of a type- θ_ℓ attacker, we have $u_{\theta_\ell}^a(\pi(\beta)) \geq u_{\theta_\ell}^a(\pi(\theta_\ell))$; further, by Lemma 16, $u_{\theta_\ell}^a(\pi(\theta_\ell)) \geq u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell})$; thus,

$$u_{\theta_\ell}^a(\pi(\beta)) \geq u_{\theta_\ell}^a(\pi(\theta_\ell)) \geq u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell}). \quad (17)$$

Combining (16), Claim 1, and (17) gives:

$$u_{\theta_\ell}^a(\mathbf{h}, t) \geq u_{\theta_\ell}^a(\mathbf{h}, i^\beta) \geq u_{\theta_\ell}^a(\pi(\beta)) \geq u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}^{\theta_\ell}). \quad (18)$$

It follows that $u_{\theta_\ell}^a(\mathbf{h}, t) \geq u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}) = \max_{i \in T} u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, i) \geq u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, t)$. By monotonicity of $u_{\theta_\ell}^a(\cdot, t)$, we have $h_t \leq \hat{c}_t^{\theta_\ell}$. \square

Next, we show the following parts to complete this proof: (i) (\mathbf{z}, t) is indeed feasible as an outcome prescribed for report θ_ℓ , i.e., $\mathbf{z} \in \mathcal{C}$ and $t \in \text{BR}_{\theta_\ell}(\mathbf{z})$; (ii) $\tilde{\pi}$ is ℓ -compatible; (iii) $\text{EoP}(\tilde{\pi}) \geq \xi$.

Part (i). Since $(\hat{\mathbf{c}}^{\theta_\ell}, i^{\theta_\ell})$ is an SSE, by definition, $\hat{\mathbf{c}}^{\theta_\ell} \in \mathcal{C}$ and $\sum_{i \in T} \hat{c}_i^{\theta_\ell} \leq m$. Since $z_i = \min\{\hat{c}_i^{\theta_\ell}, h_i\} \leq \hat{c}_i^{\theta_\ell}$ for all $i \in T$, we have $0 \leq z_i \leq 1$ and $\sum_{i \in T} z_i \leq \sum_{i \in T} \hat{c}_i^{\theta_\ell} \leq m$. Thus, $\mathbf{z} \in \mathcal{C}$. To see that $t \in \text{BR}_{\theta_\ell}(\mathbf{z})$, suppose towards a contradiction that it does not hold. Thus, $u_{\theta_\ell}^a(\mathbf{z}, i^*) > u_{\theta_\ell}^a(\mathbf{z}, t)$ for some $i^* \in T$. By Claim 2, $z_t = h_t$, so we have

$$u_{\theta_\ell}^a(\mathbf{z}, i^*) > u_{\theta_\ell}^a(\mathbf{z}, t) = u_{\theta_\ell}^a(\mathbf{h}, t) = \max_{i \in T} u_{\theta_\ell}^a(\mathbf{h}, i) \geq u_{\theta_\ell}^a(\mathbf{h}, i^*),$$

which implies that $z_{i^*} < h_{i^*}$ by monotonicity of $u_{\theta_\ell}^a(\cdot, i^*)$. Hence, $z_{i^*} = \min\{\hat{c}_{i^*}^{\theta_\ell}, h_{i^*}\} = \hat{c}_{i^*}^{\theta_\ell}$, and

$$u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}) = \max_{i \in T} u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, i) \geq u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, i^*) = u_{\theta_\ell}^a(\mathbf{z}, i^*) > u_{\theta_\ell}^a(\mathbf{z}, t).$$

This leads to the following contradiction:

$$u_{\theta_\ell}^a(\mathbf{z}, t) = u_{\theta_\ell}^a(\mathbf{h}, t) \geq u_{\theta_\ell}^a(\hat{\mathbf{c}}^{\theta_\ell}, \hat{i}) > u_{\theta_\ell}^a(\mathbf{z}, t),$$

where the first two (in)equalities follow by Claim 2 and 18, respectively.

Part (ii). We show that it is optimal for every type- θ attacker, $\theta \in \{\theta_1, \dots, \theta_\ell\}$, to report truthfully in response to $\tilde{\pi}$.

First we consider the case for a type- θ_ℓ attacker and show that $u_{\theta_\ell}^a(\tilde{\pi}(\theta_\ell)) \geq u_{\theta_\ell}^a(\tilde{\pi}(\beta'))$ for all $\beta' \in \Theta \setminus \{\theta_\ell\}$. Observe the following:

$$u_{\theta_\ell}^a(\mathbf{z}, t) = u_{\theta_\ell}^a(\mathbf{h}, t) \geq u_{\theta_\ell}^a(\mathbf{h}, i^\beta) \geq u_{\theta_\ell}^a(\pi(\beta)) \geq u_{\theta_\ell}^a(\pi(\beta')),$$

where the first three (in)equalities follow by Claim 2, (16), and Claim 1, respectively; and the last is due to the assumption that β is the optimal reporting strategy of a type- θ_ℓ attacker. By definition, $\tilde{\pi}(\theta_\ell) = (\mathbf{z}, t)$ and $\tilde{\pi}(\beta') = \pi(\beta')$ for all $\beta' \in \Theta \setminus \{\theta_\ell\}$, so $u_{\theta_\ell}^a(\tilde{\pi}(\theta_\ell)) = u_{\theta_\ell}^a(\mathbf{z}, t)$ and $u_{\theta_\ell}^a(\tilde{\pi}(\beta')) = u_{\theta_\ell}^a(\pi(\beta'))$. It follows that

$$u_{\theta_\ell}^a(\tilde{\pi}(\theta_\ell)) = u_{\theta_\ell}^a(\mathbf{z}, t) \geq u_{\theta_\ell}^a(\pi(\beta')) = u_{\theta_\ell}^a(\tilde{\pi}(\beta')),$$

which is the desired result.

Next, consider the case for each type $\theta \in \{\theta_1, \dots, \theta_{\ell-1}\}$. By Claim 2, we have $z_t = h_t \geq \max_{\theta \in \{\theta_1, \dots, \theta_{\ell-1}\}} \frac{u_\theta^a(\pi(\theta)) - r_t^\theta}{p_t^\theta - r_t^\theta}$. Thus, we have

$$u_\theta^a(\tilde{\pi}(\theta)) = u_\theta^a(\pi(\theta)) = u_\theta^a\left(\frac{u_\theta^a(\pi(\theta)) - r_t^\theta}{p_t^\theta - r_t^\theta}, t\right) \geq u_\theta^a(\mathbf{z}, t) = u_\theta^a(\tilde{\pi}(\theta_\ell)).$$

Since π is $(\ell - 1)$ -compatible, $u_\theta^a(\pi(\theta)) \geq u_\theta^a(\pi(\beta'))$ for all $\beta' \in \Theta$, so for those $\beta' \neq \theta_\ell$ we have

$$u_\theta^a(\tilde{\pi}(\theta)) = u_\theta^a(\pi(\theta)) \geq u_\theta^a(\pi(\beta')) = u_\theta^a(\tilde{\pi}(\beta')).$$

Therefore, $u_\theta^a(\tilde{\pi}(\theta)) \geq u_\theta^a(\tilde{\pi}(\beta'))$ holds for all $\beta' \in \Theta$; it is optimal for a type- θ attacker to report truthfully.

Part (iii). We show that $\text{EoP}_\theta(\tilde{\pi}) \geq \hat{\xi}$ for every type $\theta \in \Theta$, which will imply $\text{EoP}(\tilde{\pi}) = \min_{\theta \in \Theta} \text{EoP}_\theta \geq \hat{\xi}$ and complete the proof.

Since we have shown that $\tilde{\pi}$ is ℓ -compatible, truthful report is incentivized for every type in $\{\theta_1, \dots, \theta_\ell\}$. Thus, for every $\theta \in \{\theta_1, \dots, \theta_\ell\}$, we have $\text{EoP}_\theta(\tilde{\pi}) \geq \frac{u^d(\tilde{\pi}(\theta))}{\hat{u}(\theta)}$. (The reason we have an inequality here is due to the optimistic tie-breaking assumption in Definition 6.) For type θ_ℓ , since $z_t = h_t \geq \frac{\xi \cdot \hat{u}(\theta_\ell) - p_t^d}{r_t^d - p_t^d}$ by Claim 2, we have

$$\text{EoP}_{\theta_\ell}(\tilde{\pi}) \geq \frac{u^d(\tilde{\pi}(\theta_\ell))}{\hat{u}(\theta_\ell)} = \frac{u^d(\tilde{\pi}(\mathbf{z}, t))}{\hat{u}(\theta_\ell)} \geq \frac{u^d\left(\frac{\xi \cdot \hat{u}(\theta_\ell) - p_t^d}{r_t^d - p_t^d}, t\right)}{\hat{u}(\theta_\ell)} = \frac{\xi \cdot \hat{u}(\theta_\ell)}{\hat{u}(\theta_\ell)} = \xi.$$

For types $\theta \in \{\theta_1, \dots, \theta_{\ell-1}\}$, we have

$$\text{EoP}_\theta(\tilde{\pi}) \geq \frac{u^d(\tilde{\pi}(\theta))}{\hat{u}(\theta)} = \frac{u^d(\pi(\theta))}{\hat{u}(\theta)} = \text{EoP}_\theta(\pi) \geq \text{EoP}(\pi) \geq \xi.$$

For the other types $\theta' \in \{\theta_{\ell+1}, \dots, \theta_\lambda\}$, if their optimal report remains the same as under π , for the same argument above, $\text{EoP}_{\theta'}(\tilde{\pi}) = \text{EoP}_{\theta'}(\pi) \geq \text{EoP}(\pi) \geq \xi$. Otherwise, since $\tilde{\pi}$ and π differs only in the outcomes prescribed for type θ_ℓ , if the attacker's optimal reporting strategy changes under $\tilde{\pi}$, it will only change to θ_ℓ , in which case we have

$$\text{EoP}_{\theta'}(\tilde{\pi}) = \frac{u^d(\tilde{\pi}(\theta_\ell))}{\hat{u}(\theta')} \geq \frac{u^d(\tilde{\pi}(\theta_\ell))}{\hat{u}(\theta_\ell)} = \text{EoP}_{\theta_\ell}(\tilde{\pi}) \geq \xi,$$

where the first inequality holds because Algorithm 1 orders attacker types in a way such that $\hat{u}(\theta') \leq \hat{u}(\theta_\ell)$ for all $\theta' \in \{\theta_{\ell+1}, \dots, \theta_\lambda\}$. \square

D Complexity of Computing Optimal Policy in General Stackelberg Games

We show the complexity of computing the optimal leader policy in *general* Stackelberg games where payoff parameters of each player (or player type) are given by a matrix, with no restriction on the values. We let $u^L \in \mathbb{R}^{m \times n}$ denote the leader's payoff matrix, and $u_\theta^F \in \mathbb{R}^{m \times n}$ denote a type- θ follower's payoff matrix for each follower type $\theta \in \Theta$, where m and n denote the numbers of the leader's and the follower's actions (i.e., pure strategies), respectively. The entries $u^d(i, j)$ and $u_\theta^a(i, j)$ are, respectively, the utilities of the leader and a type- θ follower, when the leader plays her i -th action and the follower plays his j -th action. In an SSE, the leader plays a mixed strategy $\mathbf{x} \in \Delta_m$ and the follower best responds to \mathbf{x} with a pure strategy j , yielding leader utility $u^L(\mathbf{x}, j) = \sum_{i=1}^m x_i \cdot u^L(i, j)$ and follower utility $u^F(\mathbf{x}, j) = \sum_{i=1}^m x_i \cdot u^F(i, j)$. All other definitions and notation are the same as in Section 2. In contrast to the tractability of computing the optimal defender policy in an SSG, the problem is hard in general Stackelberg games.

Theorem 17. *It is NP-complete to decide whether there exists a leader policy π with $\text{EoP}(\pi) \geq \xi$.*

Proof. The NP membership of the problem is straightforward as for any given policy π , we can efficiently verify whether $\text{EoP}(\pi) \geq \xi$. For the NP-hardness, we show a reduction from the VERTEX

		1 ^F	2 ^F	3 ^F	
∀i	1	0.5			
		leader			
		1 ^F	2 ^F	3 ^F	
∀i ∈ {a _{v₁} , a _{v₂} }	0.9	0.9	1		
a ₀	0.4		0.4		
otherwise		0.9	1		
		flw. type θ _e (e = {v ₁ , v ₂ } ∈ E)			
		1 ^F	2 ^F	3 ^F	
∀i	1				
		flw. type θ _ℓ (ℓ = 1, ..., k)			
		1 ^F	2 ^F	3 ^F	
a ₀			1	1	
∀i ≠ a ₀		1	1		
		flw. type θ ₀			

Figure 2: Payoff parameters (blank entries are all 0).

COVER problem, which is well-known to be NP-complete. A *vertex cover* V' of an undirected graph $G = (V, E)$ is a subset of V such that $v_1 \in V'$ or $v_2 \in V'$ for every edge $\{v_1, v_2\} \in E$. An instance of the VERTEX COVER problem is given by a graph $G = (V, E)$ and an integer $k \leq |V|$. It is a yes-instance if there exists a vertex cover of G of size at most k .

For a VERTEX COVER instance, we construct the following game and show that the VERTEX COVER instance is a yes-instance if and only if there exists a leader policy π , $\text{EoP}(\pi) \geq 1$. In the game, the leader has $|V| + 1$ actions $\{a_v : v \in V\} \cup \{a_0\}$. The follower has three actions $\{1^F, 2^F, 3^F\}$. The set of possible follower types is $\Theta = \{\theta_1, \dots, \theta_k\} \cup \{\theta_e : e \in E\} \cup \{\theta_0\}$. The payoffs are given in Figure 2.

We first make several observations about the SSEs of this game (in the truthful situation). Below we let $\hat{u}^L(\theta)$ denote the leader's utility in an SSE when she plays against a type- θ follower, i.e., $\hat{u}^L(\theta) = \max_{\mathbf{x} \in \Delta_m, j \in \text{BR}(\mathbf{x})} u^L(\mathbf{x}, j)$.

- The leader's utility only depends on the follower's action, with 3^F being the most detrimental action the leader would anyhow avoid the follower to choose, followed by 2^F , and 1^F is the most preferred follower action.
- The *only* SSE strategy of the leader when she plays against a type- θ_e follower, $e \in E$, is the pure strategy a_0 . When a_0 is played, the follower finds his best responses to be $\text{BR}_{\theta_e}(a_0) = \{3^F, 1^F\}$ and breaks the tie in favor of the leader, playing 1^F ; the leader obtains $u^L(a_0, 1^F) = 1$, which is obviously the highest possible utility she can obtain; hence, $(a_0, 1^F)$ forms an SSE and $\hat{u}^L(\theta_e) = 1$. To see that this is the only SSE, observe that if the leader plays any other pure strategy with some probability, the follower would strictly prefer 3^F to 1^F and would not respond by playing 1^F , in which case the leader cannot get utility 1.
- Every leader strategy is an SSE strategy when the follower she plays against has type θ_ℓ , $\ell = 1, \dots, k$, because the follower will always respond by playing 1^F irrespective of the strategy the leader plays, which always gives the leader utility 1. We have $\hat{u}^L(\theta_\ell) = 1$.
- Every mixed leader strategy over pure strategies $i \neq a_0$ is an SSE strategy of the leader when the follower she plays against has type θ_0 . When such a strategy is played, the follower finds his best response set to be $\{3^F, 2^F\}$ and breaks the tie in favor of the leader, playing 2^F ; this is the best the leader can hope for because a type- θ_0 follower will never play 1^F as it is strictly dominated by 3^F . We have $\hat{u}^L(\theta_0) = 0.5$.

Suppose that there exists a vertex cover $V' = \{v'_1, \dots, v'_k\} \subseteq V$ of size k . The following leader policy π achieves EoP 1.

$$\pi(\theta) = \begin{cases} (a_0, 1^F), & \text{for each } \theta \in \{\theta_e : e \in E\}; \\ (a_{v'_\ell}, 1^F), & \text{for each } \theta = \theta_\ell \in \{\theta_1, \dots, \theta_k\}; \\ (a_{v'_1}, 2^F), & \text{for } \theta = \theta_0. \end{cases}$$

Clearly, all the outcomes prescribed are feasible, so π is a feasible policy. Further, it can be verified that when the leader commits to π , the optimal reporting strategy of the follower is the following.

- For every type- θ_e follower, $e \in E$, it is optimal to report a type θ_ℓ with $v'_\ell \in V'$ being an end point of e . Such a θ_ℓ always exists given that V' is a vertex cover. The leader obtains utility 1 when θ_ℓ is reported, so $\text{EoP}_{\theta_e}(\pi) \geq \frac{1}{\hat{u}(\theta_e)} = 1$.
- For every type- θ_ℓ follower, $\ell \in \{1, \dots, k\}$, it is optimal to report truthfully. The leader obtains utility 1 when $\theta - \ell$ is reported, so $\text{EoP}_{\theta_\ell}(\pi) \geq \frac{1}{\hat{u}(\theta_\ell)} = 1$.
- For a type- θ_0 follower, it is optimal to report truthfully. The leader obtains utility 0.5 when θ_0 is reported, so $\text{EoP}_{\theta_0}(\pi) \geq \frac{0.5}{\hat{u}(\theta_0)} = 1$.

Therefore, $\text{EoP}(\pi) = \min_{\theta \in \Theta} \text{EoP}_\theta(\pi) = 1$. (We have $\text{EoP}(\pi) \leq 1$ by Proposition 7; in fact, $\text{EoP}_\theta(\pi)$ is also upper-bounded by 1 for types θ_e and θ_ℓ above.)

Conversely, suppose that there exists a policy π with $\text{EoP}(\pi) = 1$. We show that there exists a vertex cover of G of size at most k . For $\text{EoP}(\pi) = 1$, we need $\text{EoP}_\theta(\pi) \geq 1$ for all $\theta \in \Theta$. Thus, the actual utility the leader obtains must be: at least 1 on each θ_e and θ_ℓ ($\ell \neq 0$), and at least 0.5 on θ_0 .

Now consider the reporting strategy of a type- θ_0 follower in response to π ; let $\beta \in \Theta$ be the optimal reporting strategy of a type- θ_0 follower, and let $\pi(\beta) = (\mathbf{x}^\beta, j^\beta)$. For the leader to obtain actual utility at least 0.5 on type θ_0 , we need $j^\beta \in \{1^F, 2^F\}$. Observe that a type- θ_0 follower gets utility 0 if $j^\beta = 1^F$, in which case he would be better-off reporting truthfully to avoid being induced to “best” respond by playing 1^F . Thus, the only possibility is $j^\beta = 2^F$. Now that $j^\beta = 2^F$, it must also be that $x_{a_0}^\beta = 0$ since otherwise the follower obtains less than 1 by reporting β and would, again, be better-off reporting truthfully (in which case he is guaranteed utility 1 by the best response 3^F).

Given this, a type- θ_e follower, $e = \{v_1, v_2\} \in E$, is able to obtain utility 0.9 by reporting β . Let γ be the type a type- θ_e follower is incentivized to report, and $\pi(\gamma) = (\mathbf{x}^\gamma, j^\gamma)$; we therefore have $u_{\theta_e}^F(\mathbf{x}^\gamma, j^\gamma) \geq 0.9$. For the leader to obtain utility at least 1 on type θ_e , we need $j^\gamma = 1^F$, in which case $u_{\theta_e}^F(\mathbf{x}^\gamma, j^\gamma) \geq 0.9$ only if $x_{a_{v_1}}^\gamma + x_{a_{v_2}}^\gamma = 1$ (i.e., only the first row of the payoff matrix of θ_e is chosen), so we have $x_{a_0}^\gamma = 0$. Therefore, $\gamma \notin \{\theta_{e'} : e' \in E\}$, because this would lead to $\text{BR}_\gamma(\mathbf{x}^\gamma) = \{3^F\} \not\geq 1^F$ given that $x_{a_0}^\gamma = 0$. For the same reason, we also have $\gamma \neq \theta_0$, so the remaining possibility is that $\gamma = \theta_\ell$ for some $\ell \in \{1, \dots, k\}$.

Let $V^\gamma = \{v \in V : x_{a_v}^\gamma \geq 0\}$, and let v'_γ be the first vertex in V^γ in lexicographical order. Since $x_{a_{v_1}}^\gamma + x_{a_{v_2}}^\gamma = 1$, we have $v'_\gamma \in e$, and moreover, $\{v'_{\theta_\ell} : \ell = 1, \dots, k\} \cap e \neq \emptyset$ given that $\gamma = \theta_\ell$ for some ℓ . This holds for all $e \in E$. Thus, $V' = \{v'_{\theta_\ell} : \ell = 1, \dots, k\}$ forms a vertex cover of G , and $|V'| \leq k$. \square

E Additional Experiment Results

EoP Comparison Figures 3 and 4 show additional results of the EoP comparison. In both figures, (a)–(c) show the variance of EoP with respect to ρ , with type sets of different scales ($\lambda = 10, 100$, and 1000, respectively); (d)–(f) show the variance of the EoP with respect to the scale of the game, under different target-resource ratios ($\frac{n}{m} = 10, 5$, and 2, respectively). In Figure 3, attacker types are generated with the covariance model, while in Figure 4, the zero attacker type is always included in Θ in addition to types generated by the covariance model.

Algorithm Runtime Figure 5 shows results of the runtime test of our algorithms. All results are obtained on a platform with a 2.60 GHz CPU and a 8.0 GB memory. The time for computing SSEs is excluded in the results as this is handled by an existing algorithm, the performance of which is not our focus. Both our algorithms for computing the optimal policy and the QR policy exhibit good scalability, capable of solving problems of 5000 attacker types and 500 targets in a reasonable amount of time. The computation of QR policy is extremely efficient thanks to its simplicity.

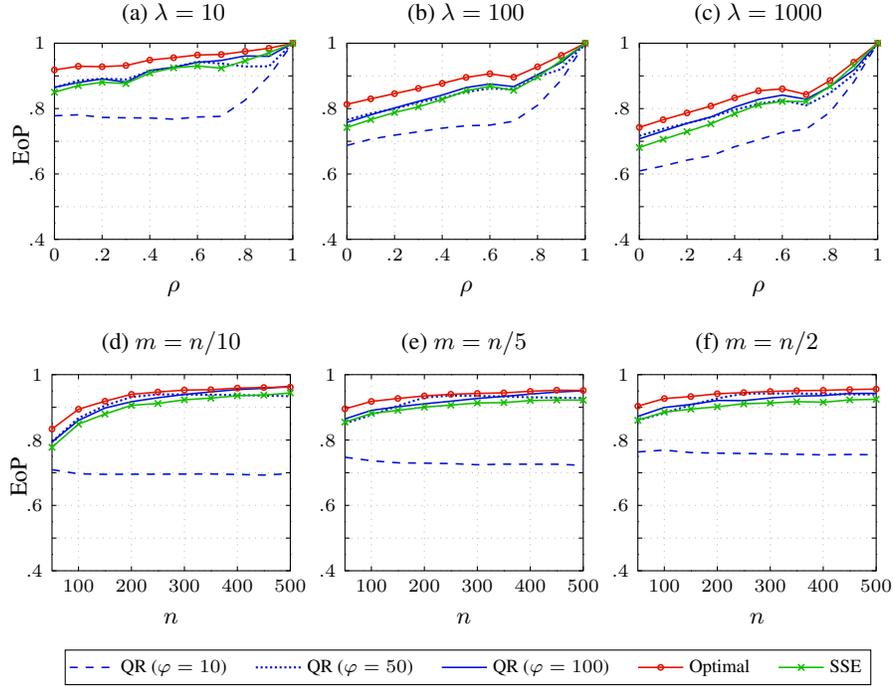


Figure 3: Comparison of the EoP. Other parameters are set to $n = 50, m = 10$ in (a)–(c); and $\rho = 0.5, \lambda = 100$ in (d)–(f).

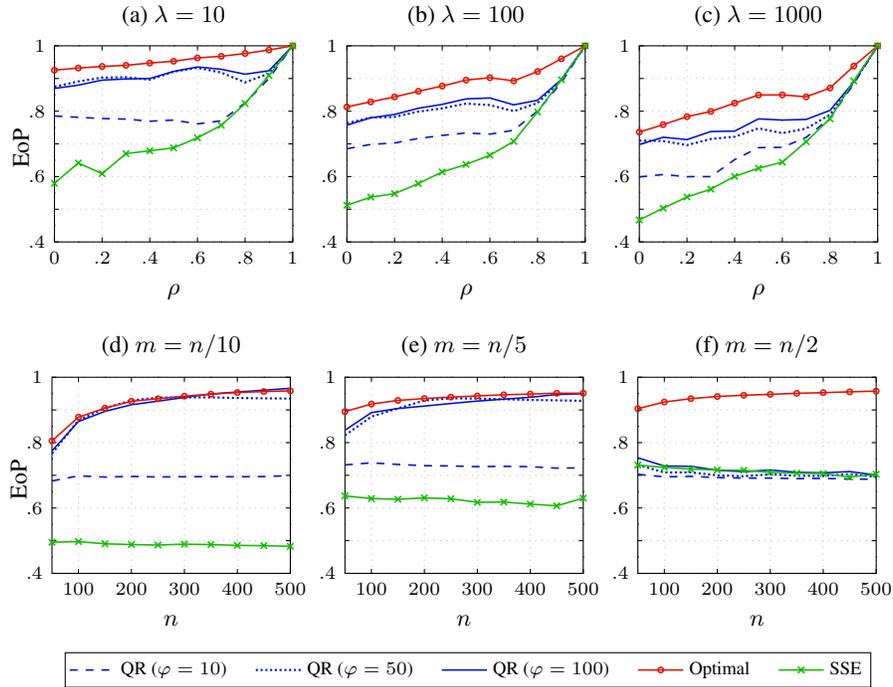


Figure 4: Comparison of the EoP when the zero-sum attacker type is always included in Θ in addition to types generated by the covariance model. Other parameters are specified in the same way as in Figure 3.

(a) Optimal policy						(b) QR policy					
	$n : 100$	200	300	400	500		$n : 100$	200	300	400	500
$\lambda : 1000$	3.54	7.45	10.65	14.26	17.96	$\lambda : 1000$	0.00	0.02	0.02	0.03	0.03
2000	14.84	30.81	46.84	75.99	78.75	2000	0.01	0.03	0.04	0.06	0.10
3000	32.42	63.50	96.92	126.96	161.37	3000	0.02	0.05	0.07	0.10	0.12
4000	57.47	115.05	171.00	228.20	288.19	4000	0.03	0.06	0.09	0.12	0.16
5000	88.84	184.77	273.51	365.91	480.54	5000	0.04	0.08	0.10	0.16	0.20

Figure 5: Algorithm runtime (seconds).