# Exploring Information Asymmetry in Two-Stage Security Games

**Haifeng Xu**[1], **Zinovi Rabinovich**[2], **Shaddin Dughmi**[1], **Milind Tambe**[1]

[1]University of Southern California, Los Angeles, CA 90007, USA
{haifengx,shaddin,tambe}@usc.edu

[2]Independent Researcher, Jerusalem, Israel
zr@zinovi.net

## Abstract

Stackelberg security games have been widely deployed to protect real-world assets. The main solution concept there is the Strong Stackelberg Equilibrium (SSE), which optimizes the defender's random allocation of limited security resources. However, solely deploying the SSE mixed strategy has limitations. In the extreme case, there are security games in which the defender is able to defend all the assets "almost perfectly" at the SSE, but she still sustains significant loss. In this paper, we propose an approach for improving the defender's utility in such scenarios. Perhaps surprisingly, our approach is to *strategically* reveal to the attacker information about the sampled pure strategy.

Specifically, we propose a two-stage security game model, where in the first stage the defender allocates resources and the attacker selects a target to attack, and in the second stage the defender strategically reveals local information about that target, potentially deterring the attacker's attack plan. We then study how the defender can play optimally in both stages. We show, theoretically and experimentally, that the two-stage security game model allows the defender to achieve *strictly* better utility than SSE.

## Introduction

Security games continue to gain popularity within the research community, and have led to numerous practical applications (Tambe 2011). The basic model is a Security Stackelberg Game (SSG) played between a defender (leader) and an attacker (follower). In the past decade, most research on security games has focused on computing or approximating the Strong Stackelberg Equilibrium (SSE), which optimizes the defender's random allocation of limited resources. A few examples include (Basilico, Gatti, and Amigoni 2009; Jain 2012; An et al. 2012; Vorobeychik and Letchford 2014; Blum, Haghtalab, and Procaccia 2014). However, solely deploying the SSE mixed strategy is insufficient for a good defense in many games. As we will show later, in the extreme case there are security games in which the defender is able to defend all the assets "almost perfectly" at the SSE, but she still sustains significant loss. This raises a natural question: can the defender do better than simply deploying the SSE mixed strategy, and if so can his optimal strategy be computed efficiently? *As our main contribution, we answer both questions in the affirmative in a natural two-stage game model.* Our main technique is to exploit the *information asymmetry* between the defender and attacker — the defender has more information. Specifically, we note that the attacker only observes the deployed mixed strategy by long-term surveillance, but the defender further knows its realization in each deployment. We show that the defender can strictly benefit by revealing such information to the attacker *strategically*.

Optimal information structures have been studied in many contexts, including auctions (Milgrom and Weber 1982; Milgrom 2008; Levin and Milgrom 2010), persuasion (Kamenica and Gentzkow 2009), voting (Alonso and Câmara 2014), and general games (Bergemann and Morris 2013). In security domains, researchers have realized the importance of the information asymmetry between the defender and attacker, however they have focused mainly on whether and how to hide private information by secrecy and deception (Brown et al. 2005; Powell 2007; Zhuang and Bier 2010). A common argument is that more defense is not always beneficial, since it may lead the attacker to suspect the importance of a target. This departs from the Stackelberg security game framework. For example, (Yin et al. 2013) consider optimal allocation of deceptive resources (e.g., hidden cameras), which introduces asymmetric information regarding deployments of resources between the defender and attacker. However, they do not consider strategically revealing such information. Rather, they model the failure of deceptive resources by a probability and feed it to a resource allocation formulation.

In this paper, we initiate the study of strategic information revelation in security games. One particular model relevant to our work is the Bayesian Persuasion (BP) model introduced in (Kamenica and Gentzkow 2009). The basic BP model describes a two-person game between a sender and a receiver with random payoff matrices. The sender can observe the realization of the payoff matrices, while the receiver only knows the prior distribution. The BP model studies how the sender (defender in our case) can design a signaling scheme to strategically reveal this information and convince a rational receiver (attacker in our case) to take a desired action. Back to security games, we observe that there is usually a *timing gap* between the attacker's choice of target and attack execution, and show how the defender can make use of such a timing gap to "persuade" a rational attacker and deter potential attacks. We formalize this as a novel *two-stage* security game model, which combines resource allocation (the first stage) and strategic information revelation (the second stage).

## An Example

To convey the basic idea, let us consider a simple Federal Air Marshal (Tsai et al. 2009) scheduling problem. A defender, against an attacker, aims to schedule $n-1$ air marshals to protect $2n$ *identical* (w.r.t. importance) flights, namely $t_1, ...t_{2n}$. The defender's pure strategies are simply arbitrary subsets of $[2n]$ of size at most $n-1$. The defender gets utility $-2$ $(1)$ if any uncovered (covered) target is attacked (i.e., $U^u_d(t_i) = -2$, $U^c_d(t_i) = 1$); while the attacker gets utility $1$ $(-1)$ if he attacks an uncovered (covered) target (i.e., $U^u_a(t_i) = 1$, $U^c_a(t_i) = -1$), for $i = 1, ..., 2n$. Assume the attacker has an additional option – choose to not attack, in which case both players get utility $0$. As easily observed, the optimal defender strategy is to protect each flight with probability $\frac{n-1}{2n} = 0.5 - \frac{1}{2n}$. The attacker has expected utility $\frac{n+1}{2n} \times 1 + \frac{n-1}{2n} \times (-1) = \frac{1}{n} (> 0)$ by attacking any target. So he attacks a target, resulting in defender utility $\frac{n+1}{2n} \times (-2) + \frac{n-1}{2n} \times 1 = -0.5 - \frac{3}{2n}$.

We have just computed the Strong Stackelberg Equilibrium (SSE) — traditionally we would be done. However, reexamining this game, one might realize the following phenomenon: the defender has done a great job, "almost" stopping the attack at every target. Unfortunately, she lacks just one additional air marshal. Consequently, the defender has to lose at least a constant factor $0.5$, watching the attacker attacking a target and gaining only a tiny payoff of $\frac{1}{n}$. Can we do better? The answer turns out to be YES. Our approach exploits the asymmetric knowledge of the defensive strategy between the defender and attacker — the defender knows more. We show that, surprisingly, the defender can gain *arbitrarily better* utility (in the multiplicative sense) than $-0.5 - \frac{3}{2n}$ by revealing such information.

For any target $t_i$, let $X_c$ $(X_u)$ denote the event that $t_i$ is covered (uncovered). The defender's mixed strategy results in $\mathbb{P}(X_c) = 0.5 - \frac{1}{2n}$ and $\mathbb{P}(X_u) = 0.5 + \frac{1}{2n}$. W.l.o.g, imagine the attacker boards $t_1$ in order to commit an attack. The attacker only knows that $t_1$ is protected with $0.5 - \frac{1}{2n}$ probability, while the defender knows the realization of the current deployment. We design a policy for the defender to reveal this information to the attacker. Specifically, let $\sigma_c$ and $\sigma_u$ be two signals that the defender will ask the captain in flight $t_1$ to announce. The meaning of signals will be clear later, but for now, one may think of them as two messages telling the attacker target $t_1$ is *covered* $(\sigma_c)$ or *uncovered* $(\sigma_u)$[1]. Now, let the defender *commit* to the following public (thus known by the attacker) *signaling scheme*:

$$\mathbb{P}(\sigma_c|X_c) = 1 \qquad \mathbb{P}(\sigma_u|X_c) = 0;$$
$$\mathbb{P}(\sigma_c|X_u) = \frac{0.5 - \frac{1}{2n}}{0.5 + \frac{1}{2n}} \qquad \mathbb{P}(\sigma_u|X_u) = \frac{\frac{1}{n}}{0.5 + \frac{1}{2n}}.$$

In other words, if $t_1$ is protected, the defender will always announce $\sigma_c$; if $t_1$ is not protected, the defender will announce $\sigma_c$ with $\frac{0.5-\frac{1}{2n}}{0.5+\frac{1}{2n}}$ probability and $\sigma_u$ with $\frac{\frac{1}{n}}{0.5+\frac{1}{2n}}$ probability.

Let us analyze this from the attacker's perspective. If he receives signal $\sigma_c$, occurring with probability

$$\mathbb{P}(\sigma_c) = \mathbb{P}(\sigma_c|X_c)\mathbb{P}(X_c) + \mathbb{P}(\sigma_c|X_u)\mathbb{P}(X_u) = 1 - \frac{1}{n},$$

the attacker infers the following posterior, by Bayes' rule: $\mathbb{P}(X_c|\sigma_c) = \frac{\mathbb{P}(\sigma_c|X_c)\mathbb{P}(X_c)}{\mathbb{P}(\sigma_c)} = \frac{1}{2}$ and $\mathbb{P}(X_u|\sigma_c) = \frac{1}{2}$. If he attacks, the attacker's expected utility given $\sigma_c$ is $\frac{1}{2} \times (-1) + \frac{1}{2} \times 1 = 0$, while the defender gains $1 \times \frac{1}{2} - 2 \times \frac{1}{2} = -0.5$. Assume the attacker breaks ties in favor of the defender and chooses to not attack, then both players get utility $0$. On the other hand, if the attacker receives signal $\sigma_u$ (with probability $\frac{1}{n}$), he infers a utility of $1$ and attacks the target, resulting in defender utility $-2$. As a result, in expectation the defender derives utility $-\frac{2}{n}$ on target $t_1$. Multiplicatively, $-\frac{2}{n}$ is arbitrarily better than $-0.5 - \frac{3}{2n}$ as $n \to \infty$. Interestingly, the attacker's expected utility of $\frac{1}{n}$ equals his SSE utility. We will show later that this is actually not a coincidence.

Recalling the concept of a signal, we notice that, signals $\sigma_c, \sigma_u$ have no intrinsic meaning besides the posterior distributions inferred by the attacker based on the signaling scheme and prior information. Intuitively, by designing signals, the defender identifies a "part" of the prior distribution that is "bad" for both players, i.e., the posterior distribution of $\sigma_c$ (the attacker is indifferent at optimality in this example), and signals as much to the attacker, so that the two players can cooperate to avoid it. This is why the defender can do strictly better while the attacker is not worse off.

## Model of Two-Stage Security Games

As observed before, the fact that the defender provides controlled access to information on the realized deployment of security resources can help her gain better utility than just deploying SSE. In this section, we formally model this phenomenon.

At a high level, we propose a two-stage security game model. The first stage is similar to regular security games, in which the defender (randomly) allocates security resources. In the second stage, the defender reveals information about the realized deployment of security resources, using a *signaling scheme*.

Consider a security game with a defender and an attacker. The defender has $K$ resources and needs to protect $T$ targets. Let $\mathcal{S}$ denote the set of all pure strategies and each pure strategy $s \in \mathcal{S}$ is a map $s$: $[K] \to 2^{[T]}$ that assigns each resource $k \in [K]$ to protect a subset of $[T]$. A mixed strategy is a distribution over $\mathcal{S}$, which results in a marginal probabilistic coverage over target set $[T]$. From this perspective,

[1] Physically, $\sigma_c$ could be a sentence like "We are proud to announce air marshal Robinson is on board flying with us today.", while $\sigma_u$ could be just keeping silent.

a marginal coverage vector can also be viewed as a mixed strategy, if it is implementable by a distribution over $\mathcal{S}$. So, instead, we will use $\boldsymbol{z} = (z_1, ..., z_T) \in R^T$ to denote a mixed strategy, where target $t$ is protected with probability $z_t$. Let $U^{c/u}_{d/a}(t)$ be the utility of defender($d$)/attacker($a$) when target $t$, if attacked, is covered($c$)/uncovered($u$). We assume the attacker has the option to not attack, in which case both players get utility 0. [2] Clearly, this is a best response for the attacker if his utility on every target is non-positive. As a standard assumption, we assume $U^c_d(t) > U^u_d(t)$ and $U^c_a(t) < 0 < U^u_a(t)$ for any $t$.

The first stage is similar to regular security games, in which the defender commits to a mixed strategy. We now model the second stage – the signaling procedure. This stage could be viewed as a persuasion procedure (Kamenica and Gentzkow 2009), during which the defender tries to persuade a rational attacker to behave in a desired way. So we call it the *persuasion phase*. Specifically, for any $t \in [T]$ covered with probability $z_t$, let $X = \{X_c, X_u\}$ be the set of events describing whether $t$ is covered ($X_c$) or not ($X_u$) and $\Sigma$ be the set of all possible signals. A signaling scheme, with respect to (w.r.t) target $t$, is a *random* map

$$f_c : X \xrightarrow{rnd} \Sigma.$$

The set of probabilities

$$\{p(x, \sigma) : \ x \in X, \ \sigma \in \Sigma\}$$

completely describes the random map $f$, in which $p(x, \sigma)$ is the probability that event $x \in X$ happens and signal $\sigma \in \Sigma$ is sent. Therefore, $\sum_\sigma p(x, \sigma) = \mathbb{P}(x), \forall x \in X$. On the other hand, upon receiving a signal $\sigma$, the attacker infers a posterior distribution $\mathbb{P}(X_c|\sigma) = \frac{p(X_c,\sigma)}{p(X_c,\sigma)+p(X_u,\sigma)}$ and $\mathbb{P}(X_u|\sigma) = \frac{p(X_u,\sigma)}{p(X_c,\sigma)+p(X_u,\sigma)}$, and makes a decision among two actions: attack or not attack. For every target $t$, the defender seeks a signaling scheme w.r.t. $t$ to maximize her expected utility on $t$.

Mathematically, a signal denotes a posterior distribution on $X$. Thus a signaling scheme splits the prior distribution $(z_t, 1 - z_t)$ into a number $|\Sigma|$ of posteriors to maximize the defender's utility on $t$. However, how many signals are sufficient to design an optimal signaling scheme w.r.t. $t$? It follows from (Kamenica and Gentzkow 2009) that

**Lemma 1.** *Two signals suffice for the defender to design an optimal signaling scheme, w.r.t. target t, with one signal recommending the attacker to attack and another one recommending him to not attack.*

Intuitively, this is because we can always combine any two signals that result in the same consequence. In particular, if the attacker has the same best response on signal $\sigma_1$ and $\sigma_2$, then instead of sending $\sigma_1$ and $\sigma_2$, the defender could have just sent a new signal $\sigma$ with probability $p(x, \sigma) = p(x, \sigma_1) + p(x, \sigma_2), \forall x \in X$. As a result of Lemma 1, a signaling scheme w.r.t. $t$ could be characterized by

$$\begin{aligned} p(X_c, \sigma_c) &= p \qquad & p(X_c, \sigma_u) &= z_t - p; \\ p(X_u, \sigma_c) &= q \qquad & p(X_u, \sigma_u) &= 1 - z_t - q, \end{aligned}$$

in which, $p \in [0, z_t], q \in [0, 1 - z_t]$ are variables. So the attacker infers the following expected utility: $\mathbb{E}(utility|\sigma_c) = \frac{1}{p+q}(pU^c_a + qU^u_a)$ and $\mathbb{E}(utility|\sigma_u) = \frac{1}{1-p-q}((z - p)U^c_a + (1 - z - q)U^u_a)$, where, for ease of notation, we drop the "$t$" in $z_t$ and $U^{c/u}_{d/a}(t)$ when it is clear from context. W.l.o.g, let $\sigma_c$ be a signal recommending the attacker to not attack, i.e., constraining $\mathbb{E}(utility|\sigma_c) \leq 0$, in which case both players get 0. Then the following LP parametrized by coverage probability $z$, denoted as $peLP_t(z)$ (Persuasion Linear Program), computes the optimal signaling scheme w.r.t. $t$:

$$\begin{aligned} \max \quad & (z - p)U^c_d + (1 - z - q)U^u_d \qquad (1) \\ s.t. \quad & pU^c_a + qU^u_a \leq 0 \\ & (z - p)U^c_a + (1 - z - q)U^u_a \geq 0 \\ & 0 \leq p \leq z \\ & 0 \leq q \leq 1 - z. \end{aligned}$$

This yields the attacker utility $\mathbb{P}(\sigma_u)\mathbb{E}(utility|\sigma_u) + \mathbb{P}(\sigma_c) \times 0 = (z - p)U^c_a + (1 - z - q)U^u_a$ and defender utility $(z - p)U^c_d + (1 - z - q)U^u_d$, w.r.t. $t$.

We propose the following two-stage Stackelberg security game model:

- Phase 1 (Scheduling Phase): the defender (randomly) schedules the resources by playing a mixed strategy $\boldsymbol{z} \in [0, 1]^T$, and samples one pure strategy each round.
- Phase 2 (Persuasion Phase): $\forall t \in [T]$, the defender *commits* to an optimal signaling scheme w.r.t. $t$ computed by $peLP_t(z_t)$ before the game starts, and then in each round, sends a signal on each target $t$ according to the commitment.

During the play, the attacker first observes $\boldsymbol{z}$ by surveillance. Then he chooses a target $t_0$ to *approach* or *board* at some round, where the attacker receives a signal and decides whether to attack $t_0$ or not. Note that, the model makes the following three assumptions. First, the defender is able to commit to a signaling scheme, and crucially will also follow the commitment. She is incentivized to do so because otherwise the attacker will not trust the signaling scheme, thus may ignore signals. Then the game becomes a standard Stackelberg game. Second, the attacker breaks ties in favor of the defender. Similar to the definition of SSE, this is without loss of generality since if there is a tie among different choices, we can always make a tiny shift of the probability mass to make the choice, preferred by the defender, $\epsilon$ better than other choices. Third, we assume the attacker cannot distinguish whether a target is protected or not when he approaches it.

With the persuasion phase, both of the defender and the attacker's payoff structures might be changed. Specifically, the defender's utility on any target $t$ is the optimal objective *value* of the linear program $peLP_t(z)$, which is non-linear in $z$. Can the defender always *strictly* benefit by adding the persuasion phase? How can we compute the optimal mixed

[2] Most security game papers incorporate this extra action by adding a fake target with payoff 0 to both players.

strategy in this new model? We answer these questions in the next two sections.

## When to Persuade

In this section, fixing a marginal coverage $z$ on a target $t$, we compare the defender's and attacker's utilities w.r.t. $t$ in the following two different models:

- Model 1: the regular security game model, without persuasion (but the attacker can choose to not attack);
- Model 2: the two-stage security game model, in which the persuasion w.r.t. $t$ is *optimal*.

The following notation will be used frequently in our comparisons and proofs (index $t$ is omitted when it is clear):

$$\begin{aligned} DefU_{1/2}(t) &: \text{ defender's expected utility in Model 1/2;} \\ AttU_{1/2}(t) &: \text{ attacker's expected utility in Model 1/2;} \\ U_{def/att}(t) &: = zU^c_{d/a} + (1-z)U^u_{d/a}, \text{ expected utility of} \\ & \quad \text{defense/attack, } \textit{if attacker attacks } t. \end{aligned}$$

Note that $AttU_1 = \max(U_{att}, 0)$ may not equal to $U_{att}$ since the attacker chooses to not attack if $U_{att} < 0$. Similarly, $DefU_1$ may not equal to $U_{def}$.

### Defender's Utility

First, we observe that the defender will never be worse off in Model 2 than Model 1 w.r.t. $t$.

**Proposition 1.** *For any* $t \in [T]$, $DefU_2 \geq DefU_1$.

*Proof.* If $U_{att} \geq 0$, then $p, q = 0$ is a feasible solution to $peLP_t(z)$ in formula 1, which achieves a defender utility $zU^c_d + (1-z)U^u_d = DefU_1$. So $DefU_2 \geq DefU_1$.

If $U_{att} < 0$, the attacker will choose to not attack in Model 1, so $DefU_1 = 0$. In this case, $p = z, q = 1-z$ is a feasible solution to $peLP_t(z)$, which achieves a defender utility 0. So $DefU_2 \geq 0 = DefU_1$. □

However, the question is, will the defender always *strictly* benefit w.r.t. $t$ from the persuasion phase? The following theorem gives a succinct characterization.

**Theorem 1.** *For any* $t \in [T]$ *with marginal coverage* $z \in [0,1]$, $DefU_2 > DefU_1$, if and only if*:*

$$U_{att}(U^c_d U^u_a - U^c_a U^u_d) < 0. \qquad (2)$$

*Proof.* The inequality Condition 2 corresponds to the following four cases:

1. $U_{att} > 0, U^u_d \geq 0, U^c_d U^u_a - U^c_a U^u_d < 0$;
2. $U_{att} > 0, U^u_d < 0, U^c_d U^u_a - U^c_a U^u_d < 0$;
3. $U_{att} < 0, U^u_d \geq 0, U^c_d U^u_a - U^c_a U^u_d > 0$;
4. $U_{att} < 0, U^u_d < 0, U^c_d U^u_a - U^c_a U^u_d > 0$.

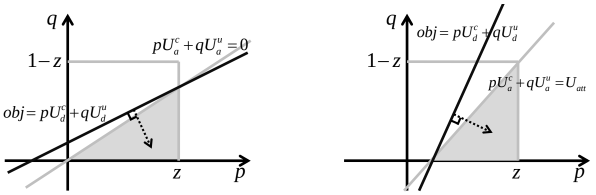

Figure 1: Feasible regions (gray areas) and an objective function gaining strictly better defender utility than SSE for the case $U_{att} > 0$ (Left) and $U_{att} < 0$ (Right).

Case 1 obviously does not happen, since $U^c_d U^u_a - U^c_a U^u_d > 0$ when $U^c_d > U^u_d \geq 0$ and $U^u_a > 0 > U^c_a$. Interestingly, cases 2–4 correspond exactly to all the three possible conditions that make $DefU_2 > DefU_1$. We now give a geometric proof. Instead of $peLP_t(z)$, we consider the following equivalent LP:

$$\begin{aligned} \min \quad & pU^c_d + qU^u_d \\ s.t. \quad & pU^c_a + qU^u_a \leq 0 \\ & pU^c_a + qU^u_a \leq U_{att} \\ & 0 \leq p \leq z \\ & 0 \leq q \leq 1-z, \end{aligned}$$

so that $DefU_2 = U_{def} - Opt$. Figure 1 plots the feasible region for the case $U_{att} > 0$ and $U_{att} < 0$, respectively. Note that, the vertex $(z, 0)$ can never be an optimal solution in either case, since the feasible point $(z - \epsilon, \epsilon)$ for tiny enough $\epsilon > 0$ always achieves strictly smaller objective value, assuming $U^c_d > U^u_d$. When $U_{att} > 0$, the attacker chooses to attack, resulting in $DefU_1 = U_{def}$. So to strictly increase the defender's utility is equivalent to making $Opt < 0$ for the above LP. That is, we only need to guarantee the optimal solution is *not* the origin $(0,0)$ (a vertex of the feasible polytope). This happens when $U^u_d < 0$, and the slope of $obj = pU^c_d + qU^u_d$ is less than the slope of $0 = pU^c_a + qU^u_a$, that is $U^c_d/U^u_d - U^c_a/U^u_a > 0$. These conditions correspond to the case 2. In this case, the defender gains *extra* utility $-Opt = -\frac{z}{U^u_a}(U^u_a U^c_d - U^c_a U^u_d) > 0$ by adding persuasion.

When $U_{att} < 0$, the attacker chooses to not attack, resulting in $DefU_1 = 0$. To increase the defender's utility, we have to guarantee $Opt < U_{def}$. Note that the vertex $(z, 1-z)$ yields exactly an objective $U_{def}$, so we only need to guarantee the optimal solution is the vertex $(\frac{U_{att}}{U^c_a}, 0)$. This happens either when $U^u_d \geq 0$ (corresponding to case 3 in which case $U^c_d U^u_a - U^c_a U^u_d > 0$ holds naturally) or when $U^u_d < 0$ and the slope of $obj = pU^c_d + qU^u_d$ is greater than the slope of $0 = pU^c_a + qU^u_a$. That is, $-U^c_d/U^u_d > -U^c_a/U^u_a$. This corresponds to case 4 above. In such cases, the defender gains *extra* utility $U_{def} - Opt = -\frac{1-z}{U^c_a}(U^u_a U^c_d - U^c_a U^u_d) > 0$ by adding persuasion.

When $U_{att} = 0$, the possible optimal vertices are $(0,0)$ and $(z, 1-z)$, which corresponds to the defender utility 0 and $U_{def}$, respectively. So $DefU_2 = \max\{0, U_{def}\}$ at optimality, which equals to $DefU_1$ assuming the attacker breaks ties in favor of the defender. □

### Interpreting the Condition in Theorem 1

Inequality 2 immediately yields that the defender does not benefit by persuasion in zero-sum security games, since $U_d^c U_a^u - U_a^c U_d^u = 0$ for any target in zero-sum games. Intuitively, this is because there are no posterior distributions, thus signals, where the defender and attacker can cooperate due to the strictly competitive nature of zero-sum games.

One case of the Inequality 2 is $U_{att} > 0$ and $U_d^c U_a^u - U_a^c U_d^u < 0$. To interpret the latter, let us start from a zero-sum game, which assumes $-U_d^u = U_a^u > 0$ and $U_d^c = -U_a^c > 0$. Then the condition $U_d^c U_a^u - U_a^c U_d^u = U_d^c U_a^u - (-U_a^c)(-U_d^u) < 0$ could be achieved by making $-U_d^u > U_a^u$ or $U_d^c < -U_a^c$. That is, the defender values a target more than the attacker ($-U_d^u > U_a^u$), e.g., the damage to a flight causes more utility loss to the defender than the utility gained by the attacker, or the defender values catching the attacker less than the cost to the attacker ($U_d^c < -U_a^c$), e.g., the defender does not gain much benefit by placing a violator in jail but the violator loses a lot. In such games, if the attacker has incentives to attack (i.e., $U_{att} > 0$), the defender can "persuade" him to not attack.

Another case of Condition 2 is $U_{att} < 0$ and $U_d^c U_a^u - U_a^c U_d^u > 0$. In contrast to the situation above, this is when the defender values a target less than the attacker (e.g., a fake target or honey pot) but cares more about catching the attacker. Interestingly, the defender benefits when the attacker does not want to attack (i.e., $U_{att} < 0$), but the defender "entices" him to commit an attack in order to catch him.

### Attacker's Utility

Now we compare the attacker's utilities w.r.t. $t$ in Model 1 and Model 2. Recall that Proposition 1 shows the defender will never be worse off. A natural question is, whether the attacker can be strictly better off? The attacker will never be worse off under *any signaling scheme*. Intuitively, this is because the attacker gets more information about the resource deployment, so he cannot be worse off, otherwise he could just ignore those signals. Mathematically, this holds simply by observing the constraints in $peLP_t(z)$ Formulation 1:

1. when $U_{att} \geq 0$, $AttU_1 = U_{att} = zU_a^c + (1-z)U_a^u$ and $AttU_2 = (z-p)U_a^c + (1-z-q)U_a^u$, so $AttU_1 - AttU_2 = pU_a^c + qU_a^u \leq 0$;
2. when $U_{att} < 0$, $AttU_2 = (z-p)U_a^c + (1-z-q)U_a^u \geq 0 = AttU_1$.

Note that the above conclusion holds without requiring the signaling scheme to be optimal, since the derivation only uses feasibility constraints. Interestingly, if the defender does persuade optimally, then equality holds.

**Theorem 2.** *Given any target* $t \in [T]$ *with marginal coverage* $z \in [0,1]$*, we have* $AttU_1 = AttU_2 = \max(0, U_{att})$.

*Proof.* From $peLP_t(z)$ we know that $AttU_2 = U_{att} - (pU_a^c + qU_a^u)$. The proof is divided into three cases. When $U_{att} > 0$ (left panel in Figure 1), we have $AttU_1 = U_{att}$. As argued in the proof of Theorem 1, the optimal solution can never be the vertex $(z, 0)$. So the only possible optimal vertices are $(0,0)$ and $(z, -z\frac{U_a^c}{U_a^u})$, both of which satisfy $pU_a^c + qU_a^u = 0$. So $AttU_2 = U_{att} - (pU_a^c + qU_a^u) = U_{att} = DefU_1$. When $U_{att} < 0$ (right panel in Figure 1),we have $AttU_1 = 0$. The only possible optimal vertices are $(z, 1-z)$ or $(-\frac{U_{att}}{U_a^c}, 0)$, both of which satisfies $pU_a^c + qU_a^u = U_{att}$. So $AttU_2 = 0 = AttU_1$. For the case $U_{att} = 0$, similar argument holds. To sum up, we always have $AttU_1 = AttU_2$. □

## How to Persuade

As we have seen so far, the defender can strictly benefit by persuasion in the two-stage security game model. Here comes the natural question for computer scientists: how can we compute the optimal mixed strategy? We answer the question in this section, starting with a lemma stating that the defender's optimal mixed strategy in the two-stage model is different from the SSE in its standard security game version.

**Lemma 2.** *There exist security games in which the optimal mixed strategy in Model 2 is different from the SSE mixed strategy in the corresponding Model 1.*

A proof can be found in the online version. We now define the following solution concept.

**Definition 1.** *The optimal defender mixed strategy and signaling scheme in the two-stage Stackelberg security game, together with the attacker's best response, form an equilibrium called the* Strong Stackelberg Equilibrium with Persuasion (peSSE).

Proposition 1 yields that, by adding the persuasion phase, the defender's utility will not be worse off under any mixed strategy, specifically, under the SSE mixed strategy. This yields the following performance guarantee of peSSE.

**Proposition 2.** *Given any security game, defender's utility in peSSE* $\geq$ *defender's utility in SSE.*

Now we consider the computation of peSSE. Note that the optimal signaling scheme can be computed by LP 1 for any target $t$ with given coverage probability $z_t$. The main challenge is about how to compute the optimal mixed strategy in Phase 1. Assume the defender's (leader) mixed strategy, represented as a marginal coverage vector over target set $[T]$, lies in a polytope $\mathcal{P}_d$. [3] With a bit of abuse of notation, let us use $peLP_t(z_t)$ to denote also the optimal objective value of the persuasion LP, as a function of $z_t$. Let

$$U_{att}(t,z) = zU_a^c(t) + (1-z)U_a^u(t)$$

be the attacker's expected utility, if he attacks, as a *linear* function of $z$.

Recall that, given a mixed strategy $\boldsymbol{z} \in [0,1]^T$, the defender's utility w.r.t. $t$ is $peLP_t(z_t)$ and the attacker's utility w.r.t. $t$ is $\max(U_{att}(t,z_t), 0)$ (Theorem 2). Similar to the

[3] Note that a polytope can always be represented by linear constraints (though may need exponentially many). For example, a simple case is the games in which pure strategies are arbitrary subsets $A \subseteq [T]$ with cardinality $|A| \leq k$, $\mathcal{P}_d$ can be represented by $2T+1$ linear inequalities: $\sum_i z_i \leq k$ and $\mathbf{0} \leq \boldsymbol{z} \leq \mathbf{1}$. However, $\mathcal{P}_d$ can be complicated in security games, such that it is NP-hard to optimize a linear objective over $\mathcal{P}_d$ (Xu et al. 2014). Finding succinct representations of $\mathcal{P}_d$ plays a key role in the computation of SSE, but this is not our focus in this paper.

framework in (Conitzer and Sandholm 2006), we define the following optimization problem for every target $t$, denoted as $OPT_t$:

$$\begin{aligned} \max \quad & peLP_t(z_t) && (3)\\ s.t. \quad & \max(U_{att}(t, z_t), 0) \geq \max(U_{att}(t', z_{t'}), 0)\ \forall t' \\ & \boldsymbol{z} \in \mathcal{P}_d, \end{aligned}$$

which computes a defender mixed strategy maximizing the defender's utility on $t$, subject to: 1. the mixed strategy is achievable; 2. attacking $t$ is the attacker's best response. Notice that some of these optimization problems may be infeasible. Nevertheless, at least one of them is feasible. The peSSE is obtained by solving these $T$ optimization problems and picking the best solution among those $OPT_t$'s.

To solve optimization problem 3, we have to deal with non-linear constraints and the specific objective $peLP_t(z_t)$, which is the optimal objective value of another LP. We first simplify the constraints to make them linear. In particular, the following constraints

$$\max(U_{att}(t, z_t), 0) \geq \max(U_{att}(t', z_{t'}), 0), \forall t' \in [T]$$

can be split into two cases, corresponding to $U_{att}(t, z_t) \geq 0$ and $U_{att}(t, z_t) \leq 0$ respectively, as follows,

| CASE 1 | CASE 2 |
|---|---|
| $U_{att}(t, z_t) \geq 0$<br>$U_{att}(t, z_t) \geq U_{att}(t', z_{t'}), \forall t'$ | $U_{att}(t', z_{t'}) \leq 0, \forall t'$ |

Now, the only problem is to deal with the objective function in Formulation 3. Here comes the crux.

**Lemma 3.** *For any $t \in [T]$, $peLP_t(z)$ is increasing on $z$ for any $z \in (0, 1)$.*

*Proof.* For notation simplicity, let $f(z) = peLP_t(z)$. We show that for any sufficiently small $\epsilon > 0$ (so that $z + \epsilon < 1$), $f(z + \epsilon) \geq f(z)$. Fixing $z$, if the optimal solution for $peLP_t(z)$, say $p^*, q^*$, satisfies $q^* = 0$, then we observe that $p^*, q^*$ is also feasible for $peLP_t(z+\epsilon)$. As a result, plugging $p^*, q^*$ in $peLP_t(z+\epsilon)$, we have $f(z+\epsilon) \geq (z-p^*)U_d^c+(1-z-q^*)U_d^u + \epsilon(U_d^c - U_d^c) \geq f(z)$ since $\epsilon(U_d^c - U_d^c) \geq 0$. On the other hand, if $q^* > 0$, then for any small $\epsilon > 0$ (specifically, $\epsilon < q^*$), $p^*+\epsilon, q^*-\epsilon$ is feasible for $peLP_t(z+\epsilon)$. Here the only need is to check the feasibility constraint $(p^*+\epsilon)U_a^c+(q^*-\epsilon)U_a^u = p^*U_a^c+q^*U_a^u+\epsilon(U_a^c-U_a^u) \leq 0$, which holds since $\epsilon(U_a^c - U_a^u) \leq 0$. This feasible solution achieves an objective value equaling to $f(z)$. Therefore, we must have $f(z + \epsilon) \geq f(z)$. □

The intuition behind Lemma 3 is straightforward – the defender should always get more utility by protecting a target more. However, this actually does not hold in standard security games. Simply consider a target with $U_d^c = 2$, $U_d^u = -1$ and $U_a^c = -1$, $U_a^u = 1$. If the target is covered with probability $0.4$, then in expectation both the attacker and defender get $0.2$; however, if the target is covered with probability $0.6$, the attacker will not attack and both of them get $0$. Therefore, the monotonicity in Lemma 3 is really due to the signaling scheme.

Back to the optimization problem 3, here comes our last key observation – the monotonicity property in Lemma 3 reduces the problem to an LP. Specifically, the following lemma is easy to think through.

**Lemma 4.** *Maximizing the* increasing *function $peLP_t(z_t)$ over any feasible region $\mathcal{D}$ reduces to directly maximizing $z_t$ over $\mathcal{D}$ and then plugging in the optimal $z_t$ to $peLP_t(z_t)$.*

To this end, we summarize the main results in this section. The following theorem essentially shows that computing peSSE efficiently reduces to computing SSE [see (Conitzer and Sandholm 2006) for a standard way to compute SSE by multiple LPs]. In other words, adding the persuasion phase does not increase the computational complexity.

**Theorem 3.** *For any security game, the* Strong Stackelberg Equilibrium with Persuasion (peSSE)*, defined in Definition 1, can be computed by multiple LPs.*

*Proof.* According to Lemma 3 and 4, Algorithm 1, based on multiple LPs, computes the peSSE. □

**Algorithm 1** Computing peSSE

```
1: For every target $t \in [T]$, compute the optimal objectives
   for the following two LPs:

      $\max \quad z_t$                                                  (4)
      $s.t. \quad U_{att}(t, z_t) \geq 0$
            $U_{att}(t, z_t) \geq U_{att}(t', z_{t'}), \forall t' \in [T]$
            $\boldsymbol{z} \in \mathcal{P}_d$

   and

      $\max \quad z_t$                                                  (5)
      $s.t. \quad U_{att}(t', z_{t'}) \leq 0, \forall t' \in [T]$
            $\boldsymbol{z} \in \mathcal{P}_d.$

   Let $z^*_{t,1}$, $z^*_{t,2}$ be the optimal objective value for LP 4,
   LP 5 respectively. $z^*_{t,i} = null$ if the corresponding LP
   is infeasible.
2: Choose the non-null $z^*_{t,i}$, denoted as $z^*$, that maximizes
   $peLP_t(z^*_{t,i})$ over $t \in [T]$ and $i = 1, 2$. The optimal
   mixed strategy that achieves $z^*$ in one of the above LPs
   is the peSSE mixed strategy.
```

## Simulations

As expected, our simulation based on more than $20,000$ covariance random security games (Nudelman et al. 2004) shows that peSSE outperforms SSE in terms of the defender utility, and interestingly, performs much better than SSE when the defender has negative SSE utilities. We omit details here due to space constraints and refer the reader to the online version for further information.

## Conclusions and Discussions

In this paper, we studied how the defender can use strategic information revelation to increase defensive effectiveness. The main takeaway is that, besides physical security resources, the defender's extra *information* can also be viewed as a means of defense. This raises several new research questions in security games and beyond, and we list a few: Instead of only observing the signal from the chosen target, what if the attacker simultaneously surveils several targets before deciding which to attack? What about scenarios in which the defender is privy to extra information regarding the payoff structure of the game, such as the vulnerability of various targets and effectiveness of defensive resources, and can strategically reveal such information? Finally, do our results have analogues beyond our two-stage game model, to extensive-form games more broadly?

**Acknowledgment:** This research was supported by MURI grant W911NF-11-1-0332 and NSF grant CCF-1350900.

## References

Alonso, R., and Câmara, O. 2014. Persuading voters. Technical report.

An, B.; Kempe, D.; Kiekintveld, C.; Shieh, E.; Singh, S. P.; Tambe, M.; and Vorobeychik, Y. 2012. Security games with limited surveillance. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, 1242–1248.

Basilico, N.; Gatti, N.; and Amigoni, F. 2009. Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, AAMAS '09, 57–64. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Bergemann, D., and Morris, S. 2013. Bayes correlated equilibrium and the comparison of information structures. *Working paper*.

Blum, A.; Haghtalab, N.; and Procaccia, A. D. 2014. Lazy defenders are almost optimal against diligent attackers. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, 573–579.

Brown, G.; Carlyle, M.; Diehl, D.; Kline, J.; and Wood, K. 2005. A two-sided optimization for theater ballistic missile defense. *Oper. Res.* 53(5):745–763.

Conitzer, V., and Sandholm, T. 2006. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, EC '06, 82–90. New York, NY, USA: ACM.

Jain, M. 2012. Scaling up security games: Algorithms and applications.

Kamenica, E., and Gentzkow, M. 2009. Bayesian persuasion. Working Paper 15540, National Bureau of Economic Research.

Levin, J., and Milgrom, P. 2010. Online advertising: Heterogeneity and conflation in market design. *American Economic Review* 100(2):603–07.

Milgrom, P. R., and Weber, R. J. 1982. A Theory of Auctions and Competitive Bidding. *Econometrica* 50(5):1089–1122.

Milgrom, P. 2008. What the seller won't tell you: Persuasion and disclosure in markets. *Journal of Economic Perspectives* 22(2):115–131.

Nudelman, E.; Wortman, J.; Shoham, Y.; and Leyton-Brown, K. 2004. 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*, 880–887. IEEE Computer Society.

Powell, R. 2007. Allocating defensive resources with private information about vulnerability. *American Political Science Review* 101(04):799–809.

Tambe, M. 2011. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned.* Cambridge University Press.

Tsai, J.; Rathi, S.; Kiekintveld, C.; Ordonez, F.; and Tambe, M. 2009. Iris - a tool for strategic security allocation in transportation networks. In *The Eighth International Conference on Autonomous Agents and Multiagent Systems - Industry Track*.

Vorobeychik, Y., and Letchford, J. 2014. Securing interdependent assets.

Xu, H.; Fang, F.; Jiang, A. X.; Conitzer, V.; Dughmi, S.; and Tambe, M. 2014. Solving zero-sum security games in discretized spatio-temporal domains. In *Proceedings of the 28th Conference on Artificial Intelligence (AAAI 2014), Qubec, Canada*.

Yin, Y.; An, B.; Vorobeychik, Y.; and Zhuang, J. 2013. Optimal deceptive strategies in security games: A preliminary study.

Zhuang, J., and Bier, V. M. 2010. Reasons for secrecy and deception in Homeland-Security resource allocation. *Risk Analysis* 30(12):1737–1743.

# APPENDIX

## A Proof of Lemma 2

**Lemma Statement:** There exist security games, in which the optimal mixed strategy in Model 2 is different from the SSE mixed strategy in the corresponding Model 1.

*Proof.* We prove directly by constructing the following game. Consider a security game with payoff matrix in Table 1.

|  | $U_d^c$ | $U_d^u$ | $U_a^c$ | $U_a^u$ |
|---|---|---|---|---|
| $t_1$ | 1 | -2 | -1 | 1 |
| $t_2$ | 3 | -5 | -3 | 5 |
| $t_3$ | 1 | -4 | -2 | 4 |
| $t_4$ | 0 | -0.5 | -2 | 1 |

Table 1: Payoff

Assume there are two resources, and feasible pure strategies are $A_1 = (t_1, t_2)$, $A_2 = (t_2, t_3)$ and $A_3 = (t_3, t_4)$. Let $\boldsymbol{p} = (p_1, p_2, p_3)$ denote a mixed strategy where $p_i$ is the probability of taking action $A_i$. With a bit calculation, one can find the Strong Stackelberg Equilibrium (SSE) as $\boldsymbol{p} = (\frac{3}{8}, \frac{7}{32}, \frac{13}{32})$ with coverage probability vector $\boldsymbol{z} = (\frac{3}{8}, \frac{19}{32}, \frac{5}{8}, \frac{13}{32})$. The attacker's utility is $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, -\frac{7}{32})$ and defender's utility is $(-\frac{7}{8}, -\frac{1}{4}, -\frac{7}{8}, -\frac{19}{64})$, so the attacker will attack $t_2$.

Now, if we add the persuasion phase as in Model 2, the optimal mixed strategy is $\boldsymbol{p} = (\frac{3}{8}, \frac{3}{8}, \frac{1}{4})$ with coverage probability vector $\boldsymbol{z} = (\frac{3}{8}, \frac{3}{4}, \frac{5}{8}, \frac{1}{4})$. The attacker's utility is $(\frac{1}{4}, -1, \frac{1}{4}, \frac{1}{4})$ and defender's utility is $(-\frac{1}{2}, 1, -\frac{1}{4}, -\frac{1}{8})$, so the attacker will attack $t_4$ in favor of the defender. So the defender's utility changes from $-\frac{1}{4}$ in Model 1 to $-\frac{1}{8}$ in Model 2. □

## Simulations

In this section, we compare SSE and peSSE on randomly generated security games. Our simulations aim to compare the two concepts, SSE and peSSE, in games with various payoff structures.

To generate payoffs, we follow most security game papers and use the covariance random payoff generator (Nudelman et al. 2004), but with a slight modification. Specifically, let $\mu[a, b]$ denote a uniform distribution on interval $[a, b]$, then we randomly generate the following random payoffs: $U_d^c \sim \mu[0, r], U_d^u \sim \mu[-10, 0], U_a^c = aU_d^c \times \frac{10}{r} + b\mu[-10, 0]$ (set $U_d^c \times \frac{10}{r} = 0$ if $r = 0$) and $U_a^u = aU_d^u + b\mu[0, 10]$, where $a = cov, b = \sqrt{1 - a^2}$. Here $cov \in [-1, 0]$ is the covariance parameter between defender's reward (or penalty) and attacker's penalty (or reward). So $cov = 0$ means a totally random payoff structure while $cov = -1$ and $r = 10$ means a zero-sum game. By setting $U_d^c \in [0, r]$ while $U_a^c \in [0, 10]$, we intentionally capture the defender's "overall" value of catching the attacker by parameter $r$. Standard covariance payoff fixes $r = 10$, but Theorem 1 suggests that $r$ may affect the utility difference between SSE and peSSE.

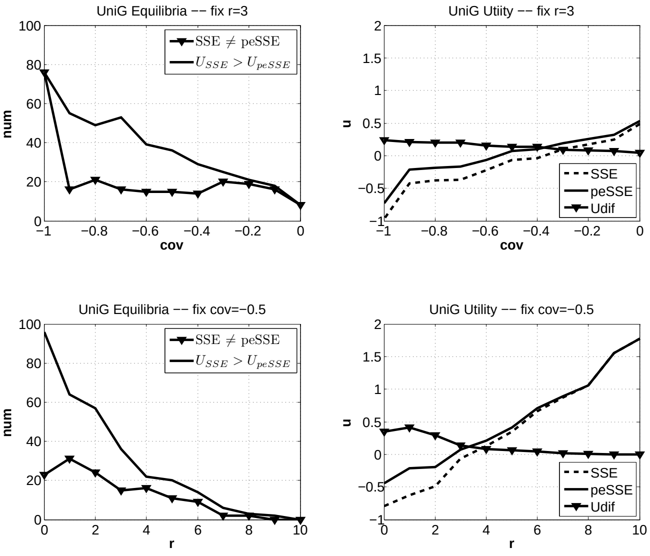

Figure 2: Comparison between SSE and peSSE: fix parameter $r = 3$ (upper) and fix parameter $cov = -0.5$. The trend is similar for different $r$ or $cov$, except the utility scales are different.

In all the simulations, every game has 8 targets and 3 resources, and the attacker has the option to not attack. We simulate two different kinds of pure strategies, which results in two types of games:

1. **Uni**form Strategy **G**ame (**UniG**): in such games, a pure strategy is any subset of targets with cardinality at most 3.
2. **Ran**dom Strategy **G**ame (**RanG**): for each game we randomly generate 6 pure strategies, each of which is a subset of targets with cardinality at most 3. Each target is guaranteed to be covered by at least one pure strategy.

We set $r = 0, 1, ..., 10$ and $cov = 0, -0.1, -0.2, ..., -1$. For each parameter instance, i.e., $r$ and $cov$, *100* random security games are simulated. As a result, in total $2 \times 100 \times 11^2 = 24,200$ (2 types of games, $11^2$ parameter combinations and 100 games per case) random security games are tested in our experiments. We find that the UniG and RanG games have similar experimental performance, except that RanG games have a lower utility at a given parameter instance. This is reasonable since UniG games are relaxations of the RanG games in terms of the set of pure strategies. So we only show results for UniG to avoid repetition.

Figure 2 gives a comprehensive comparison about the difference between SSE and peSSE. All these performances are averaged over 100 games. These figures suggest the following empirical conclusions as expected (note that the trends reflected in the figures are basically similar for different $r$ or $cov$, except the utility scales are different):

- In the left two panels, the line $SSE \neq peSSE$ describes the number of games within 100 simulations that have different SSE and peSSE mixed strategies. This number seems not very sensitive in parameter $cov$ (note games with $cov = -1$ is not zero-sum when $r = 3$), but increases as $r$ decreases. That is, when defender cares less about catching the attacker, then persuading the attacker to not attack benefits the defender more.

- The line $U_{SSE} > U_{peSSE}$ in the left two panels describes how many games have *strictly* greater peSSE utility than SSE utility. This number increases as $cov$ or $r$ decreases. That is, if the defender cares less about catching the attacker or the game becomes more competitive (i.e., $cov$ decreases), then defender benefits more by persuasion. Note that the $Udif$ lines in the right two panels also show the same trend.
- The right two panels show that persuasion usually helps more when the defender's SSE utility is less. Specifically, peSSE can increase the SSE utility by about half when $r$ is small with fixed $cov = -0.5$ (right-lower panel).