Security Games with Limited Surveillance

Bo An, David Kempe
University of Southern California
Los Angeles, CA 90089
{boa,dkempe}@usc.edu

Christopher Kiekintveld
University of Texas, El Paso
El Paso, TX 79968
cdkiekintveld@utep.edu

Eric Shieh
University of Southern California
Los Angeles, CA 90089
eshieh@usc.edu

Satinder Singh
University of Michigan
Ann Arbor, MI 48109
baveja@umich.edu

Milind Tambe
University of Southern California
Los Angeles, CA 90089
tambe@usc.edu

Yevgeniy Vorobeychik*
Sandia National Laboratories
Livermore, CA 94550
yvorobe@sandia.gov

Abstract

Randomized first-mover strategies of Stackelberg games are used in several deployed applications to allocate limited resources for the protection of critical infrastructure. Stackelberg games model the fact that a strategic attacker can surveil and exploit the defender's strategy, and randomization guards against the worst effects by making the defender less predictable. In accordance with the standard game-theoretic model of Stackelberg games, past work has typically assumed that the attacker has perfect knowledge of the defender's randomized strategy and will react correspondingly. In light of the fact that surveillance is costly, risky, and delays an attack, this assumption is clearly simplistic: attackers will usually act on partial knowledge of the defender's strategies. The attacker's imperfect estimate could present opportunities and possibly also threats to a strategic defender.

In this paper, we therefore begin a systematic study of security games with limited surveillance. We propose a natural model wherein an attacker forms or updates a belief based on observed actions, and chooses an optimal response. We investigate the model both theoretically and experimentally. In particular, we give mathematical programs to compute optimal attacker and defender strategies for a fixed observation duration, and show how to use them to estimate the attacker's observation durations. Our experimental results show that the defender can achieve significant improvement in expected utility by taking the attacker's limited surveillance into account, validating the motivation of our work.

Introduction

Stackelberg security games have been used in several deployed applications for allocating limited resources in order to protect critical infrastructure including LAX Airport, US Coast Guard, and the Federal Air Marshals Service (Basilico, Gatti, and Amigoni 2009; Korzhyk, Conitzer, and Parr 2010; Dickerson et al. 2010; Tambe 2011; An et al. 2011b; Pita et al. 2008; An et al. 2011a; Tsai et al. 2009). A Stackelberg security game models an interaction between a defender and an attacker (Kiekintveld et al. 2009). The defender first commits to a security policy (which may be randomized), and the attacker conducts surveillance to learn the defender's policy before launching an attack. A solution to the game yields an optimal randomized strategy for the defender, based on the assumption that the attacker will observe this strategy and respond optimally. Software decision aids based on Stackelberg games have been successfully implemented in several real-world domains, including Los Angeles International Airport (LAX) (Pita et al. 2008), United States Federal Air Marshals Service (FAMS) (Tsai et al. 2009), United States Transportation Security Agency (TSA) (Pita et al. 2011), and the United States Coast Guard (An et al. 2011a; Shieh et al. 2012).

Terrorists conduct surveillance to select potential targets and gain strong situational awareness of targets' vulnerabilities and security operations (Southers 2011). Most existing work on security games, including deployed applications, assumes that the attacker is able to observe the defender's strategy perfectly. (There are some notable exceptions discussed below; our work differs significantly from those.) This assumption is a useful first-level approximation, but it is clearly simplistic. In reality, the attacker may have more limited observation capabilities since surveillance is costly and delays an attack. Attackers may also wish to reduce the number of observations due to the risk of being detected by security forces during surveillance activities (Southers 2011). Therefore, *it is important to consider situations where attackers select targets based on a limited number of observations using explicit belief updates*.

In this paper, we begin a systematic investigation of belief update models for games with limited observation. We make the following contributions: (1) We introduce a natural model of security games with limited surveillance and formulate how the attacker updates his belief based on observations. (2) We provide two formulations for computing the defender's optimal strategy, one using nonconvex programming, and another which is convex, but approximate. (3) We present theoretical analysis and exhibit surprising non-monotonicity phenomena. (4) Our experiments show that the defender can do much better by explicitly modeling the attacker's decision process based on limited observations.

*Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Stackelberg Security Games

A Stackelberg security game (Kiekintveld et al. 2009) has two players, a defender who uses m identical resources to protect a set of targets $T = \{1, 2, \ldots, n\}$ ($m < n$), and an attacker who selects a single target to attack. The defender has N pure strategies $\mathcal{A}$, each a coverage vector representing which m targets are covered. Our model can handle more general security settings in which there may exist scheduling constraints on the assignment of resources (Jain et al. 2010). In that case, $\mathcal{A}$ represents feasible assignments. We write $A_i = 1$ if target i is covered in strategy $A \in \mathcal{A}$, and $A_i = 0$ otherwise. The defender can choose a randomized strategy $\mathbf{x}$, with $x_A \geq 0$ the probability of playing a strategy A. A defender strategy can be represented more compactly using a marginal coverage vector $\mathbf{c}(\mathbf{x}) = \langle c_i(\mathbf{x}) \rangle$ where $c_i(\mathbf{x}) = \sum_{A \in \mathcal{A}} x_A A_i$ is the probability that target i is covered by some defender resource (Kiekintveld et al. 2009). The attacker's strategy is a vector $\mathbf{a} = \langle a_i \rangle$ where a_i is the probability of attacking target i. Since the attacker always has an optimal pure-strategy response, we restrict the attacker's strategies to pure strategies without loss of generality.

The payoffs for each player depend on which target is attacked and the probability that the target is covered. If the attacker attacks target i, there are two cases: If target i is covered, the defender receives a reward R_i^d and the attacker receives a penalty P_i^a. Otherwise, the payoffs for the defender and attacker are P_i^d and R_i^a, respectively. We assume that $R_i^d \geq P_i^d$ and $R_i^a \geq P_i^a$ in order to model that the defender would always prefer the attack to fail, while the attacker would prefer it to succeed. For a strategy profile $\langle \mathbf{c}, \mathbf{a} \rangle$, the expected utilities for both agents are given by:

$$U^d(\mathbf{c}, \mathbf{a}) = \sum_{i \in T} a_i U^d(\mathbf{c}, i), \text{ where } U^d(\mathbf{c}, i) = c_i R_i^d + (1 - c_i) P_i^d$$

$$U^a(\mathbf{c}, \mathbf{a}) = \sum_{i \in T} a_i U^a(\mathbf{c}, i), \text{ where } U^a(\mathbf{c}, i) = c_i P_i^a + (1 - c_i) R_i^a$$

In a Stackelberg game, the defender moves first, choosing $\mathbf{x}$, while the attacker observes $\mathbf{x}$ and plays an optimal response $\mathbf{a}$ to it. The standard solution concept is strong Stackelberg equilibrium (SSE) (von Stengel and Zamir 2004). In an SSE, the defender chooses an optimal strategy $\mathbf{x}$, accounting for the attacker's best response $\mathbf{a}$, under the assumption that the attacker breaks ties in the defender's favor.

Security Games with Limited Observation

We propose to depart from the standard Stackelberg assumption that the attacker has full knowledge of $\mathbf{x}$, and instead model the attacker as a Bayesian decision maker who starts with a prior distribution over the defender's strategies and forms a posterior based on a finite number of observed realizations. We refer to our model as SGLS (Security Game with Limited Surveillance). Throughout most of this paper, we assume that the duration τ of the attacker's observation sequence is determined exogenously, e.g., based on intelligence or expert advice. (At the end, we offer a heuristic approach to estimate τ as an alternative.) The sequence of moves in an SGLS is as follows.

1. First, the defender chooses a strategy. We assume that when choosing a strategy, the defender has knowledge of the attacker's prior beliefs about the defender's strategy and the number of observations the attacker will make.
2. Then, the attacker makes τ observations and selects the optimal target based on his posterior belief about the defender's strategy.

Example 1. We use the LAX airport as an example, based on the ARMOR application (Pita et al. 2008). The police at LAX place m checkpoints on the entrance roads to LAX following a mixed strategy computed by ARMOR. The fact that attackers may engage in surveillance prior to an attack is based on real-world cases and feedback from security experts (Southers 2011), and follows other Stackelberg models deployed in practice and justified elsewhere (Pita et al. 2009).[1] In practice, the attackers will make only a limited number of observations of how the checkpoints are placed before they launch an attack. For example, they might observe placements for 20 days, and then launch an attack a week later after finalizing plans for the attack based on an analysis of the security strategy. A single observation in this domain might involve the attacker driving around the different entrances to the airport in order to determine which ones are covered by checkpoints at any particular time, so each observation gives information about the full strategy of the defender.[2]

We assume that the attacker and the defender have common knowledge of the attacker's prior beliefs over the set of mixed strategies that the defender may execute. *We also assume that the defender does not know the exact times when the attacker will conduct surveillance*, and therefore cannot modify the mixed strategy during the observation sequence. This is realistic if the defender is operating in a steady state, and does not know when or where surveillance operations could take place for planning a specific attack.

In an SGLS, the attacker updates his beliefs about the defender's strategy given his prior and τ observations, labeled $\sigma^1, \ldots, \sigma^\tau$, where each observation is one of the defender's pure strategies. The individual observations are drawn independently from the distribution representing the defender's mixed strategy. Such a sequence of observations σ can be compactly represented by an observation vector

[1] The model in this paper assumes a surveillance phase prior to any actual execution of an attack. In particular, we assume that executing an attack is sufficiently complex that it is prohibitively difficult to observe the pure strategy of the defender and immediately launch an attack against this pure strategy. This assumption is based on real-world cases and feedback from security experts (Southers 2011), and follows other Stackelberg models deployed in practice and justified elsewhere (Pita et al. 2009). One important factor in this is the difficulty of generating and executing complex conditional plans with limited resources.

[2] An alternative model could be developed where the attacker picks one (or a few) targets to observe, and will therefore only learn about a part of the full pure strategy in each observation. This is an interesting direction for future work.

$\mathbf{o} = \langle o_A \rangle$ in which o_A is the number of times pure strategy A is observed. An observation vector $\mathbf{o}$ can represent $\frac{\tau!}{\prod_{A \in \mathcal{A}} o_A!}$ observation sequences. The observation vector space is $\mathcal{O}_\tau = \{\mathbf{o} : o_A \in \{0, \ldots, \tau\}, \sum_{A \in \mathcal{A}} o_A = \tau\}$ when the attacker makes τ observations.

Updating Attacker Beliefs

We assume that the attacker's belief is represented as Dirichlet distributions with support set $\mathcal{S} = \{\mathbf{x} : \sum_{A \in \mathcal{A}} x_A = 1, x_A \geq 0, \forall A \in \mathcal{A}\}$. A Dirichlet distribution $f(\mathbf{x})$ is characterized by a parameter vector $\alpha = \langle \alpha_A \rangle$ with $\alpha_A > -1$ for all $A \in \mathcal{A}$. It assigns probability $\beta \prod_{A \in \mathcal{A}} (x_A)^{\alpha_A}$ to the defender's mixed strategy $\mathbf{x}$, where $\beta = \frac{\Gamma(\sum_{A \in \mathcal{A}} \alpha_A + |\mathcal{A}|)}{\prod_{A \in \mathcal{A}} \Gamma(\alpha_A + 1)}$ is a normalization constant expressed in terms of the gamma function Γ. The prior belief can be represented as follows:

$$f(\mathbf{x}) = \frac{\Gamma(\sum_{A \in \mathcal{A}} \alpha_A + |\mathcal{A}|)}{\prod_{A \in \mathcal{A}} \Gamma(\alpha_A + 1)} \prod_{A \in \mathcal{A}} (x_A)^{\alpha_A}$$

If the defender's mixed strategy is $\mathbf{x}$, the probability that the attacker will observe $\mathbf{o} \in \mathcal{O}_\tau$ is $f(\mathbf{o}|\mathbf{x}) = \frac{\tau!}{\prod_{A \in \mathcal{A}} o_A!} \prod_{A \in \mathcal{A}} (x_A)^{o_A}$. By applying Bayes' rule for observation $\mathbf{o}$, we can calculate the posterior distribution as:

$$f(\mathbf{x}|\mathbf{o}) = \frac{\Gamma(\sum_{A \in \mathcal{A}} \alpha_A + |A| + \tau)}{\prod_{A \in \mathcal{A}} \Gamma(\alpha_A + o_A + 1)} \prod_{A \in \mathcal{A}} (x_A)^{\alpha_A + o_A}$$

Having observed $\mathbf{o}$, the attacker believes that the probability with which the defender chooses pure strategy A is

$$p(A|\mathbf{o}) = \int_{\mathcal{S}} x_A f(\mathbf{x}|\mathbf{o}) \mathrm{d}\mathbf{x} = \frac{\alpha_A + o_A + 1}{\sum_{A \in \mathcal{A}} \alpha_A + |\mathcal{A}| + \tau}.$$

The marginal coverage of target i according to the posterior belief $f(\mathbf{x}|\mathbf{o})$ is

$$c_i^{\mathbf{o}} = \sum_{A \in \mathcal{A}} A_i p(A|\mathbf{o}) = \frac{\sum_{A \in \mathcal{A}} A_i (\alpha_A + o_A + 1)}{\sum_{A \in A} \alpha_A + |A| + \tau}.$$

After making τ observations, the attacker chooses the best target to attack based on the posterior belief $f(\mathbf{x}|\mathbf{o})$, i.e., attacks the target i maximizing $(1 - c_i^{\mathbf{o}}) R_i^a + c_i^{\mathbf{o}} P_i^a$. We denote this target by $\psi(\mathbf{o})$.

Analysis of SGLS

In this section, we explore some general trends among the strategies and payoffs for the attacker and defender. Intuitively, one would expect that more observations will give the attacker more accurate information. In turn, this more accurate information could be exploited to make better (for the attacker) decisions. In zero-sum games, where the two players' utilities are the opposite of each other, this should also lead to lower utility for the defender. Finally, one may expect that more fine-grained knowledge will make the attacker consider a larger set of targets, and the defender may thus have to protect more targets.

Somewhat surprisingly, all of the above intuitions can at times be false, even in the following apparently much more restrictive class of games: the defender has only one resource $m = 1$, so that a pure strategy A consists precisely of protecting a single target i. The attacker's prior has $\alpha_A = 0$ for all strategies A. Furthermore, the game is zero-sum (so $P_i^d = -R_i^a$), and both players' utilities are 0 when the attack fails (so $P_i^a = R_i^d = 0$). In these cases, the game is fully characterized by the target values to the attacker, which we simply write as $R_i := R_i^a$.

The high-level intuition for the failure of the reasoning described above is that when the attacker makes few observations, it imposes a coarse "resolution" on the possible probabilities $\mathbf{x}$ the attacker could obtain as beliefs. After all, the denominator for all x_A is always $\tau + |\mathcal{A}| + \sum_{A \in \mathcal{A}} \alpha_A$. For different values of τ, the particular fractions that can be attained for x_A can interact with the target values R_i in ways that can be surprisingly exploited by the defender. The following example illustrate these effects, showing that increasing τ is not necessarily advantageous for the attacker.

Example 2 (Increasing Defender Utility with Larger τ)**.** There are two targets, with values $R_1 = 1$ and $R_2 = 0.99$. There are two pure defender strategies: protecting target 1 or target 2. Assume that the defender protects target 1 with a probability of $x \in [0, 1]$ and protects target 2 with probability $1 - x$. Consider that the attacker makes one observation and there are two situations: 1) He will observe target 1 being protected with probability x and will then attack target 2; 2) He will observe target 2 being protected with probability $1 - x$ and will attack target 1 in this case. Therefore, the defender's expected utility is $-0.99x^2 - (1 - x)^2$ and her optimal utility is -0.497. If the attacker makes two observations, there are three situations: 1) target 1 is always protected; 2) target 2 is always protected; and 3) target 1 is protected once and target 2 is protected once. We can calculate the optimal defender strategy by hand and the defender's optimal utility is -0.469, which is higher than the defender's utility when the attacker makes one observation.

Intuitively, what happens here is that with one observation, the attacker will always have a belief that makes one target significantly more likely to be protected than the other one. This means that the other target will have to be protected against attack, and the defender needs to defend both targets roughly equally. On the other hand, with two observations, there is a significant chance that the attacker will believe that the protection probabilities are equal. Since target 1 is attacked in that case, the defender can profitably increase coverage on target 1.

Example 3 (Fewer Targets Protected with Larger τ)**.** There are three targets, two of value $R_1 = R_2 = 1.3$, and a third of value $R_3 = 1$.

A slightly tedious but straightforward calculation shows that when $\tau = 2$, the optimal strategy for the defender protects target 3 with probability roughly 16%, and targets 1 and 2 with probability roughly 42% each. On the other hand, if $\tau = 3$, there is no observation sequence under which target 3 is attacked (as either target 1 or target 2 will always be more attractive), so by Proposition 1 below, the

optimal solution is to protect each of targets 1 and 2 with probability 50%. This example thus shows that neither the set nor the number of protected targets needs to be monotone non-decreasing.

These examples raise interesting practical challenges. They show that a defender with very precise information about the observation length τ could possibly exploit subtleties in the attacker's belief update rule to catch the attacker more frequently. Intuitively, such a strategy may not be as robust in practice: it could backfire seriously when the estimate of τ is slightly off. In general, since there may be uncertainty in τ (and we will experimentally show the advantage of our approach despite such uncertainty), an interesting direction for future work is whether such uncertainty could lead to expected monotonicity of behaviors.

Safe Targets

Some targets i have sufficiently low values such that there is *no* observation vector $\mathbf{o}$ for which the attacker would attack i. We call a target i *safe* for observation duration τ if $\psi(\mathbf{o}) \neq i$ for all $\mathbf{o} \in \mathcal{O}_\tau$ where $\psi(\mathbf{o})$ is the target the attacker will attack after observing $\mathbf{o}$. Let Φ_τ denote the set of all targets safe for duration τ. Of course, some targets could be safe even in an SSE.

It is intuitive to believe that safe targets should never be protected in an optimal defender strategy. In other words, the defender never needs to use safe targets as "decoys". This intuition is true, so long as the defender is allowed to leave some resources unused, or the number of unsafe targets is at least as large as the number of resources.

Proposition 1. *Without loss of generality, the optimal defender strategy $\mathbf{x}$ has the property that $c_i(\mathbf{x}) = 0$ for all safe targets i, under the assumption that the defender is allowed to leave resources unused.*

Proof. Define a vector $\mathbf{x}'$ by setting $x'_{A'} = \sum_{A:A\setminus\Phi_\tau=A'} x_A$ for all A'. In other words, when the strategy $\mathbf{x}$ asks the defender to protect a set A (each defender pure strategy $A \in \mathcal{A}$ represents the set of targets the pure strategy A will cover), she instead protects just the unsafe targets in A. We claim that this new strategy $\mathbf{x}'$ does as well as $\mathbf{x}$.

Consider the following specific way in which the defender can generate an action according to $\mathbf{x}$. Partition the unit interval $[0,1]$ into disjoint sets S_A of size $|S_A| = x_A$. Draw a uniformly random number $z \in [0,1]$, and play the unique strategy A with $z \in S_A$. A sequence of τ actions can then be generated by drawing $\mathbf{z} \in [0,1]^\tau$ uniformly at random, and playing the unique action A with $z_j \in S_A$ in round j. The vector $\mathbf{z}$ uniquely determines a sequence of actions for the defender, and thus uniquely determines the observation $\mathbf{o}$. We write $\mathbf{o}(\mathbf{x},\mathbf{z})$ for the unique observation generated when the distribution is $\mathbf{x}$ and the random vector chosen is $\mathbf{z}$. We can now write the defender's expected utility as

$$\mathrm{E}\left[U^d \mid \mathbf{x}\right] = \int_{\mathbf{z}\in[0,1]^\tau} U^d(c_{\psi(\mathbf{o}(\mathbf{z},\mathbf{x}))}(\mathbf{x}), \psi(\mathbf{o}(\mathbf{z},\mathbf{x})))d\mathbf{z}, \quad (1)$$

where $c_i(\mathbf{x})$ is the protection probability of target i under the mixed strategy $\mathbf{x}$.

To generate an action according to $\mathbf{x}'$, we can use a similar strategy. We specifically define the sets $S_{A'} = \bigcup_{A:A\setminus\Phi_\tau=A'} S_A$, which have sizes exactly $x_{A'}$. In other words, to generate an action A' from $\mathbf{x}'$, we draw a set A according to the distribution $\mathbf{x}$, and then protect the set $A' = A \setminus \Phi_\tau$. The result of this specific way of coupling the generation of $\mathbf{o}$ and $\mathbf{o}' = \mathbf{o}(\mathbf{x}',\mathbf{z})$ is that for each vector $\mathbf{z}$, we can obtain $\mathbf{o}' = \mathbf{o}(\mathbf{x}',\mathbf{z})$ from $\mathbf{o} = \mathbf{o}(\mathbf{x},\mathbf{z})$ simply by setting $o'_i = o_i$ for $i \notin \Phi_\tau$, and $o'_i = 0$ for $i \in \Phi_\tau$. (Since we assume that $m = 1$, o_i is defined as o_A if $A_i = 1$.)

To determine $\psi(\mathbf{o}')$ from $\mathbf{o}'$, notice that the attacker does a pairwise comparison between all targets i, based on the utility of the attack and the belief about the protection probability, which is based only on o'_i. Thus, the choice among a set T' of targets will remain the same if $o'_i = o_i$ for all $i \in T'$. Because $\psi(\mathbf{o}') \notin \Phi_\tau$ by definition, and $o'_i = o_i$ for all $i \notin \Phi_\tau$, we get that $\psi(\mathbf{o}(\mathbf{x}',\mathbf{z})) = \psi(\mathbf{o}(\mathbf{x},\mathbf{z}))$ for all $\mathbf{z}$. Plugging this equality into (1), and observing that $c_i(\mathbf{x}') = c_i(\mathbf{x})$ for all $i \notin \Phi_\tau$ now shows that the expected utility is the same under the two strategies. □

Motivated by the study of SSE, one might suspect that the converse is also true, i.e., that every unsafe target must be protected with positive probability. It turns out that the converse is false, as evidenced by the following example:

Example 4 (Not All Unsafe Targets Protected)**.** Consider an instance with 4 targets. The first three have value 1.334 each, while the fourth one has value 1. The attacker makes $\tau = 9$ observations. Expressing the defender's payoff as a function of her probability of protecting target 4 and computing the derivative shows that the optimal strategy never protects target 4, and instead assigns probability $1/3$ to each of targets 1, 2, and 3. This is even though target 4 is not safe: if each of targets 1, 2, 3 is observed protected exactly thrice, the attacker will actually attack target 4.

Intuitively, it does not pay for the defender to divert even a small amount of resources to target 4 to deal with a fairly unlikely (though possible) event. At a higher level, this example shows that a defender should make judicious choices about which targets need to be protected.

Computing the Defender's Optimal Strategy

We now investigate the problem of computing the defender's optimal strategy in general SGLS under the assumption that the number of observations, τ, is known. First, we observe that the attacker's optimal target choice depends only on the observation vector $\mathbf{o}$, and not directly on the defender's strategy. Of course, the defender's strategy $\mathbf{x}$ affects the probability of each observation vector $\mathbf{o}$, and therefore affects the probability that the attacker will choose each target. Nevertheless, we can use this insight to separate the attacker's decision problem, which is discrete, but easy, from the problem of the defender, which is continuous and non-linear. We first describe how we can calculate the optimal attacker response for each observation $\mathbf{o}$. Then, we describe how to compute an optimal defense strategy $\mathbf{x}$.

Attacker's Best Response

Assume that the attacker observes $\mathbf{o} \in \mathcal{O}_\tau$. The attacker's posterior belief about the coverage of target i is $c_i^{\mathbf{o}} = \frac{\sum_{A \in \mathcal{A}} A_i(\alpha_A + o_A + 1)}{\sum_{A \in \mathcal{A}} \alpha_A + |A| + \tau}$. The attacker's expected utility for attacking target i is then $c_i^{\mathbf{o}}(P_i^a - R_i^a) + R_i^a$. If the attacker observes $\mathbf{o}$, the attacker will attack the best target $\psi(\mathbf{o})$ which can give him the highest expected utility: $\psi(\mathbf{o}) = \arg\max_{i \in T} \left(c_i^{\mathbf{o}} P_i^a + (1 - c_i^{\mathbf{o}}) R_i^a\right)$. As in SSE, we assume that the attacker breaks ties in the defender's favor. $\psi(o)$ can be computed simply by going through all targets and comparing their objective values (which are easy to compute).

Exact Formulation

We now introduce an exact (but nonconvex) mathematical program for computing the defender's optimal strategy $\mathbf{x}$, assuming that $\psi(\mathbf{o})$ is pre-computed.

P1:

$$\max \quad \sum_{\mathbf{o} \in \mathcal{O}_\tau} \frac{\tau!}{\prod_{A \in \mathcal{A}} o_A!} \prod_{A \in \mathcal{A}} (x_A)^{o_A} d^{\mathbf{o}} \tag{2}$$

$$\mathbf{s.t.} \quad x_A \in [0,1] \qquad \forall A \in \mathcal{A} \tag{3}$$

$$\sum_{A \in \mathcal{A}} x_A = 1 \tag{4}$$

$$c_i = \sum_{A \in \mathcal{A}} x_A A_i \qquad \forall i \in T \tag{5}$$

$$d^{\mathbf{o}} = c_{\psi(\mathbf{o})} R^d_{\psi(\mathbf{o})} + (1 - c_{\psi(\mathbf{o})}) P^d_{\psi(\mathbf{o})} \qquad \forall \mathbf{o} \in \mathcal{O}_\tau \tag{6}$$

P1 defines the defender's optimal strategy by considering all possible $\mathbf{o} \in \mathcal{O}_\tau$ and evaluating her expected utility for each observation. Equation (2) is the objective function which maximizes the defender's expected payoff $\sum_{\mathbf{o} \in \mathcal{O}_\tau} f(\mathbf{o}|\mathbf{x}) d^{\mathbf{o}}$ where $d^{\mathbf{o}}$ is the defender's expected utility when the attacker's observation is $\mathbf{o}$. Equations (3) and (4) define the legal strategy space for the defender. Equation (5) defines the marginal coverage for each target given the defender's strategy $\mathbf{x}$. Equation (6) defines the defender's expected payoff $d^{\mathbf{o}} = c_{\psi(\mathbf{o})} R^d_{\psi(\mathbf{o})} + (1 - c_{\psi(\mathbf{o})}) P^d_{\psi(\mathbf{o})}$ when the attacker attacks $\psi(\mathbf{o})$ for observation $\mathbf{o}$.

Unfortunately, objective (2) makes this formulation nonconvex so no available solver can guarantee an optimal solution. This motivates us to consider a convex approximation.

Convex Approximation

Taking the log of the defender's objective function (2) does not change the maximizers since log is monotone increasing. However, it keeps the function (2) non-convex, so as an approximation, we move the log inside the summation, which makes it a concave objective $\sum_{\mathbf{o} \in \mathcal{O}_\tau} (\log \frac{\tau!}{\prod_{A \in \mathcal{A}} o_A!} + \sum_{A \in \mathcal{A}} o_A \log x_A + \log d^{\mathbf{o}})$. The value of $d^{\mathbf{o}}$ could be negative, so we cannot safely apply the log operator. By adding a large value to each entry in the payoff matrix we can get an equivalent game in which $d^{\mathbf{o}}$ is always positive. This yields the following convex minimization formulation:

P2:

$$\min \sum_{\mathbf{o} \in \mathcal{O}_\tau} \Big(- \log \frac{\tau!}{\prod_{A \in \mathcal{A}} o_A!} - \sum_{A \in \mathcal{A}} o_A \log x_A - \log d^{\mathbf{o}} \Big) \tag{7}$$

$$\mathbf{s.t.} \quad (3) - (6)$$

Attacker's Number of Observations

In the previous section we assumed that the number of attacker observations was known. As witnessed in many terrorist attacks (Southers 2011), surveillance happens during the terrorist operational planning cycle, and the decision about observation duration is often exogenously made. How would the defender estimate the number τ of observations the attacker would make? This section presents a computational heuristic approach for approximating τ. Specifically, the attacker is supposed to model the defender's best response to his observation duration by assuming that the game is zero-sum. In other words, the attacker assumes that the defender payoffs are $P_i^d = -R_i^a$ and $R_i^d = -P_i^a$ for each target $i \in T$. In general, observations incur an opportunity cost by delaying an attack and increasing the probability that the attackers are captured before an attack can be carried out. We model this opportunity cost as a fixed cost $\lambda > 0$, such that taking τ observations reduces the attacker's expected utility by $\lambda \cdot \tau$.[3] The problem of calculating the optimal observation duration can then be formulated as:

$$\operatorname*{argmax}_{\tau} \Big(\sum_{\mathbf{o} \in \mathcal{O}_\tau} \frac{\tau!}{\prod_{A \in \mathcal{A}} o_A!} \prod_{A \in \mathcal{A}} x_A^*(\tau)^{o_A} k^{\mathbf{o}} - \lambda \cdot \tau \Big)$$

where $\mathbf{x}^*(\tau)$ is the defender's optimal strategy (computed using **P1**) under the assumption of zero-sum games. $k^{\mathbf{o}} = \sum_{A \in \mathcal{A}} x_A^*(\tau) A_{i,\psi(\mathbf{o})} (P^a_{\psi(\mathbf{o})} - R^a_{\psi(\mathbf{o})}) + R^a_{\psi(\mathbf{o})}$ is the attacker's utility when he observes $\mathbf{o}$ and the defender plays strategy $\mathbf{x}^*(\tau)$. In other words, the attacker chooses which target to attack based on his posterior beliefs and the attacker's utility of attacking his best target $\psi(\mathbf{o})$ depends on the attacker's belief about the defender's optimal strategy $\mathbf{x}^*(\tau)$.

We propose a search heuristic to iteratively approximate τ. Intuitively, with a small τ value, the attacker's prior belief has a large impact on his posterior belief, which could be far from the defender's strategy. The defender may be able to exploit this (since we assume the prior is known), leading to a low attacker utility. As τ becomes large, the attacker's belief will converge to the true defender strategy, and the attacker is likely to choose a better response. However, the attacker's utility decreases with increasing τ due to the observation cost λ. Due to these competing factors, it is very likely that the attacker's utility is single-peaked.

We use this structure to approximate the best value of τ using binary search in Algorithm 1. We maintain a lower

[3] If $\lambda = 0$, the attacker will make an infinite number of observations and will be able to completely learn the defender's strategy. The defender's optimal strategy will be the SSE strategy. An alternative way of modeling surveillance cost is using a discount factor.

bound (LB) and an upper bound (UB) on the attacker's optimal observation duration. LB is initialized to 0. Let $U^a(\tau)$ be the attacker's utility if the attacker makes τ observations, which can be computed using $\mathbf{P1}$. We say that $U^a(\tau)$ is increasing iff $U^a(\tau) \leq U^a(\tau+1)$.

Algorithm 1: Estimate optimal observation duration

$LB \leftarrow 0, n \leftarrow 1$;
while $U^a(F_n)$ *is increasing* **do**
  $LB \leftarrow F_n, n++$;
$UB \leftarrow F_n$;
while $UB - LB > 1$ **do**
  $\tau \leftarrow \frac{UB+LB}{2}$;
  if $U^a(\tau)$ *is increasing* **then** $LB \leftarrow \tau + 1$;
  else $UB \leftarrow \tau$;
return $\arg\max_{\tau \in \{LB,UB\}} U^a(\tau)$;

The first stage of Algorithm 1 is estimating the upper bound UB using Fibonacci numbers F_n. If $U^a(F_n)$ is increasing, we will update the lower bound to F_n and continue to check $U^a(F_{n+1})$. If $U^a(F_n)$ is decreasing, we have found a feasible upper bound F_n. In the second stage of Algorithm 1, we use binary search to find the optimal τ. In each iteration, τ is set to be the mean of UB and LB. If $U^a(\tau)$ is increasing, the lower bound is increased, otherwise the upper bound is decreased. The search continues until the upper bound and lower bound are sufficiently close.

Experimental Evaluation

We compare SGLS with the standard SSE model (in which the attacker has full knowledge of the defender's strategy, and the defender plans accordingly) and explore key characteristics of SGLS. We conduct experiments primarily on randomly-generated instances of security games. R_i^d and R_i^a are drawn independently and uniformly from the range $[0, 100]$. P_i^d and P_i^a are drawn from the range $[-100, 0]$. We consider three methods for setting the prior beliefs for the attacker. The first is a uniform prior: the attacker believes that the defender will choose a uniformly random strategy from the space of possible strategies, so $\alpha_A = \nu$ for every $A \in \mathcal{A}$. ν is a parameter capturing the strength of the prior belief. As ν increases, the attacker's posterior belief will give more weight to his prior. The second prior is based on the SSE strategy $\hat{\mathbf{x}}$ for a zero-sum game constructed so that the defender's payoffs are the opposite of the attacker's payoffs. This prior can be computed by an attacker who does not know the defender's payoffs. In the absence of such knowledge, it seems reasonable for an attacker to treat the game as zero-sum. The prior satisfies $\max_A \alpha_A = \nu$ and $\alpha_A / \hat{x}_A = \alpha_{A'} / \hat{x}_{A'}$ for different $A, A' \in \mathcal{A}$. Finally, we consider a hybrid that combines the uniform and SSE using a weighted combination with weights of 0.5 each.

All experiments are averaged over 100 sample games. Unless otherwise specified, we use 5 targets, 1 defender resource, $\lambda = 1$ as the observation cost, and a uniform prior with $\nu = 10$. We use KNITRO version 8.0.0 to solve $\mathbf{P1}$ and $\mathbf{P2}$.

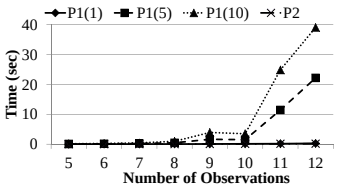

Figure 1: $\mathbf{P1}$ vs $\mathbf{P2}$: runtime

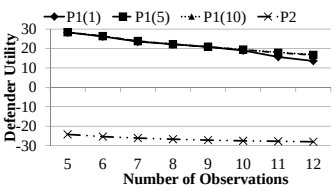

Figure 2: $\mathbf{P1}$ vs $\mathbf{P2}$: utility

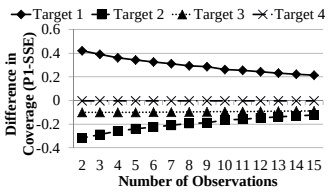

Figure 3: Convergence

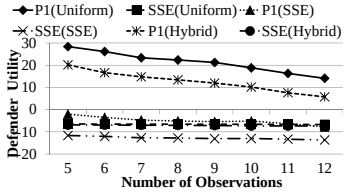

Figure 4:Vary observation length

Comparison of Solution Methods: We begin by evaluating the accuracy and runtime performance of our solution methods, $\mathbf{P1}$ and $\mathbf{P2}$. Formulation $\mathbf{P1}$ is exact but nonconvex, while $\mathbf{P2}$ is approximate but convex. Since existing solvers cannot guarantee exact solutions for nonconvex programs, all of the solutions we find are approximate. However, we can get a sense of the overall solution accuracy by using the multi-start feature of KNITRO when solving the nonconvex formulation. Figure 1 shows that the approximate formulation $\mathbf{P2}$ is faster than the exact formulation, and the runtime for KNITRO increases linearly with additional restarts as expected. In Figure 2, we see that the approximate formulation $\mathbf{P2}$ results in much lower solution quality than $\mathbf{P1}$. We also see that the solution quality for $\mathbf{P1}$ is very similar regardless of the number of restarts. We observe similar results for tests with larger numbers of targets. Based on these data, we use $\mathbf{P1}$ with one starting point in the remaining experiments. In addition, we note that we can use the analytical results from Proposition 1 to speed up $\mathbf{P1}$ by a factor of two.

Convergence Behavior: Intuitively, as τ grows very large, the solution should converge towards the SSE solution. Our experimental results confirm this. Figure 3 shows the difference between the strategy computed using $\mathbf{P1}$ and the SSE solution in random zero-sum games with 4 targets. The payoff for each player is 0 for a failed attack, and we sort the payoffs for the targets so that the values for targets 1, 2, 3, and 4 are in decreasing order. Each data point is the difference between the coverage on the target for the $\mathbf{P1}$ strategy and the SSE strategy. As τ increases, the defender's strategy gradually converges to the SSE strategy (all of the differences converge towards 0).

There is also an interesting pattern in the structure of the solutions when there are few observations: the defender allocates more resources to protect targets with the highest values, and fewer resources to less important targets. In the graph, this is seen in the positive average differences for target 1 and the negative differences for 2 and 3. Intuitively when τ is small, the attacker believes that each target is protected with similar probability, so the important targets

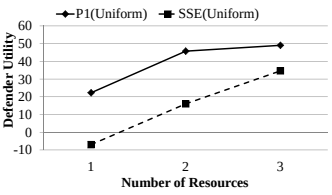

Figure 5: Effect of # of resources

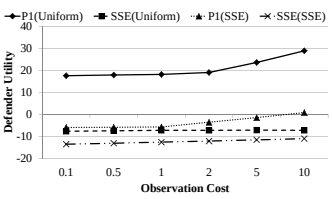

Figure 6: Effect of observe cost

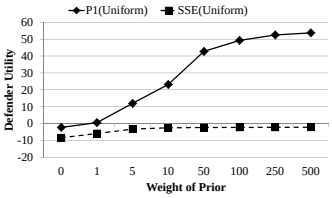

Figure 7: Effect of learning speed

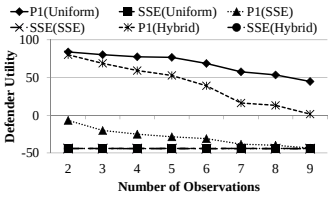

Figure 8: Using LAX data

are relatively underprotected, and vice versa.

SGLS outperforms SSE: We now evaluate the performance of $\mathbf{P1}$ against SSE strategies. In the experiments, the attacker always plays an optimal best response according to the limited surveillance model. Our results presented in Figures 4–7 show that the SGLS solution based on $\mathbf{P1}$ leads to *significantly higher defender utilities than the SSE strategy* across a wide range of different experimental conditions. In Figure 4, the x-axis is the number of observations τ, and the y-axis is the defender's utility for the SGLS solution and SSE. We consider different attacker priors including uniform, zero-sum SSE, and the hybrid prior combining uniform and zero-sum SSE. Figure 4 shows that the strategies computed by $\mathbf{P1}$ always achieved higher utilities than SSE for all priors, though the effect is less dramatic for the SSE-based prior. Additionally, as the number of attacker observation increases, the defender's utility tends to decrease.

Figures 5–7 compare the performance of $\mathbf{P1}$ against SSE strategies when the attacker always chooses the approximately optimal τ (using the methods described previously) for games with varying numbers of resources, observation costs, and initial attacker priors. (We did not include the results for some priors for readability.) Figure 5 shows that the defender's utility increases with more resources, as expected. The defender's utility also increases as the attacker's observation cost increases (Figure 6), which is consistent with our earlier observation since the attacker tends to make fewer observations with increasing cost. Finally, with a higher value of ν, the attacker will update beliefs more slowly, placing greater weight on the prior. The attacker may make more observations to account for this, but must also factor in the utility loss due to observation costs. As shown in Figure 7, the defender's utility is monotone increasing with the decrease of the attacker's learning speed.

Robustness of SGLS: To apply SGLS using $\mathbf{P1}$, the defender needs to estimate the number of observations the attacker will make, which may be difficult in practice. Here we examine the impact that an inaccurate estimate can have on the defender's payoff. In Table 1, each row represents the $\tau^* \in \{1, \ldots, 10\}$ of observations the attacker actually makes before choosing a strategy. Each column represents the defender's estimate of the observation duration $\tau \in \{1, \ldots, 10\}$ which is used in computing the defender strategy; the last column represents the defender using the SSE strategy. Let $U^d(\tau^*, \tau)$ be the defender's utility when the attacker makes τ^* observations but the defender assumes that the attacker will make τ observations. The entry in row τ^* and column τ is $U^d(\tau^*, \tau^*) - U^d(\tau^*, \tau)$, which measures the defender's utility loss for estimating τ^* as τ (or using the SSE strategy in the final column). On the diagonal, $\tau^* = \tau$, so the utility loss is zero by definition.

Table 1: Robustness of SGLS

	1	2	3	4	5	6	7	8	9	10	SSE
1	0.0	4.4	5.9	9.9	11.6	13.6	15.1	17.5	19.2	21.3	49.3
2	7.2	0.0	2.6	5.5	7.4	9.2	10.8	13.0	14.7	16.8	44.7
3	11.8	5.0	0.0	3.7	4.9	6.9	8.3	10.6	12.2	14.2	42.3
4	19.3	12.8	8.6	0.0	2.0	3.1	4.6	6.6	8.2	10.2	38.1
5	23.4	17.1	13.1	4.7	0.0	1.9	2.7	4.8	6.2	8.2	35.9
6	26.0	19.8	15.9	7.8	3.7	0.0	1.4	3.2	4.5	6.3	34.1
7	28.9	22.9	19.2	11.6	7.3	4.0	0.0	2.2	3.1	4.8	32.4
8	34.1	28.4	25.0	17.3	13.4	9.8	6.0	0.0	1.2	2.7	29.8
9	36.4	30.9	27.7	20.8	16.8	13.5	9.5	3.4	0.0	1.8	28.3
10	41.0	35.6	32.6	25.6	21.7	18.6	14.5	8.6	5.4	0.0	26.2

The data in Table 1 show that the utility loss from using an SSE is typically greater than from using an incorrect number of observations. The optimal solutions are fairly robust to small variations in τ, especially as τ grows larger. In addition, we note that the loss for overestimating τ is generally smaller than for a symmetric underestimate of τ; in the table, the values to the upper right (overestimates) are smaller than the values to the lower left (underestimates).

SGLS on Real LAX Game Matrices: In addition to the randomly-generated game instances used in the experiments above, we also ran a comparison of the SGLS solution with the SSE solution using real-world data from the deployed ARMOR system for scheduling canine patrols at the LAX airport. There are eight terminals within the LAX airport and there is one canine unit. The results on the real game matrix shown in Figure 8 are similar to the results for our synthetic examples. (The defender's utilities of SSE strategies with different priors are the same; thus, the corresponding curves overlap.) In particular, the SGLS solution achieves significant improvements in the defender's expected utility compared to SSE.

Summary and Related Work

We present the first systematic study of security games with limited surveillance, making the following contributions: (i) We introduce the SGLS model wherein an attacker forms or updates a belief based on observed actions, and chooses an optimal response; (ii) We investigate SGLS theoretically, providing surprising non-monotonicity phenomena; (iii) We give mathematical programs to compute optimal attacker

and defender strategies for a fixed observation duration, and to estimate the attacker's observation durations; (iv) Our experimental results show that the defender can exploit the limited observation to achieve significantly higher expected utility than what would be achievable by SSE, validating the motivation of our work.

In terms of related work, some recent work has relaxed the perfect observation assumption in security games. Korzhyk et al. (2011) only consider two extreme situations: perfect observation and no observation. Realistically, attackers have partial knowledge of the defender's strategies. RECON (Yin et al. 2011) takes into account possible observation errors by assuming that the attacker's belief is within some distance of the defender's real strategy. It does not address how these errors arise, nor does it explicitly model the process of forming beliefs based on limited observations. The COBRA algorithm (Pita et al. 2010) focuses on human perception of probability distributions by applying support theory (Tversky and Koehler 1994) from psychology. Both RECON and COBRA require hand-tuned parameters to model observation errors, which we avoid in this paper.

Acknowledgments

This research is supported by MURI grant W911NF-11-1-0332 and ONR grant N00014-08-1-0733.

References

An, B.; Pita, J.; Shieh, E.; Tambe, M.; Kiekintveld, C.; and Marecki, J. 2011a. GUARDS and PROTECT: Next generation applications of security games. *SIGECOM* 10:31–34.

An, B.; Tambe, M.; Ordóñez, F.; Shieh, E.; and Kiekintveld, C. 2011b. Refinement of strong Stackelberg equilibria in security games. In *Proc. of the 25th Conference on Artificial Intelligence*, 587–593.

Basilico, N.; Gatti, N.; and Amigoni, F. 2009. Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In *Proc. of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 500–503.

Dickerson, J. P.; Simari, G. I.; Subrahmanian, V. S.; and Kraus, S. 2010. A graph-theoretic approach to protect static and moving targets from adversaries. In *Proc. of The 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 299–306.

Jain, M.; Kardes, E.; Kiekintveld, C.; Ordóñez, F.; and Tambe, M. 2010. Security games with arbitrary schedules: A branch and price approach. In *Proc. of The 24th AAAI Conference on Artificial Intelligence*, 792–797.

Kiekintveld, C.; Jain, M.; Tsai, J.; Pita, J.; Tambe, M.; and Ordóñez, F. 2009. Computing optimal randomized resource allocations for massive security games. In *Proc. of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 689–696.

Korzhyk, D.; Conitzer, V.; and Parr, R. 2010. Complexity of computing optimal Stackelberg strategies in security resource allocation games. In *Proc. of The 24th AAAI Conference on Artificial Intelligence*, 805–810.

Korzhyk, D.; Conitzer, V.; and Parr, R. 2011. Solving Stackelberg games with uncertain observability. In *Proc. of The 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 1013–1020.

Pita, J.; Jain, M.; Western, C.; Portway, C.; Tambe, M.; Ordóñez, F.; Kraus, S.; and Parachuri, P. 2008. Deployed ARMOR protection: The application of a game-theoretic model for security at the Los Angeles International Airport. In *Proc. of The 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 125–132.

Pita, J.; Jain, M.; Ordóñez, F.; Tambe, M.; Kraus, S.; and Magori-Cohen, R. 2009. Effective solutions for real-world Stackelberg games: When agents must deal with human uncertainties. In *Proc. of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 369–376.

Pita, J.; Jain, M.; Tambe, M.; Ordóñez, F.; and Kraus, S. 2010. Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition. *Artificial Intelligence* 174(15):1142–1171.

Pita, J.; Tambe, M.; Kiekintveld, C.; Cullen, S.; and Steigerwald, E. 2011. GUARDS - game theoretic security allocation on a national scale. In *Proc. of The 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 37–44.

Shieh, E.; An, B.; Yang, R.; Tambe, M.; Baldwin, C.; DiRenzo, J.; Maule, B.; and Meyer, G. 2012. PROTECT: A deployed game theoretic system to protect the ports of the United States. In *Proc. of The 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Southers, E. 2011. *LAX - terror target: the history, the reason, the countermeasure*. Cambridge University Press. chapter Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned, 27–50.

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

Tsai, J.; Rathi, S.; Kiekintveld, C.; Ordóñez, F.; and Tambe, M. 2009. IRIS: a tool for strategic security allocation in transportation networks. In *Proc. of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 37–44.

Tversky, A., and Koehler, D. J. 1994. Support thoery: A nonextensional representation of subjective probability. *Psychological Review* 101:547–567.

von Stengel, B., and Zamir, S. 2004. Leadership with commitment to mixed strategies. Technical Report LSE-CDAM-2004-01, CDAM Research Report.

Yin, Z.; Jain, M.; Tambe, M.; and Ordóñez, F. 2011. Risk-averse strategies for security games with execution and observational uncertainty. In *Proc. of The 25th AAAI Conference on Artificial Intelligence (AAAI)*, 758–763.