Security Games with Limited Surveillance: An Initial Report

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
eug.vorobey@gmail.com

Abstract

Stackelberg games have been used in several deployed applications of game theory to make recommendations for allocating limited resources for protecting critical infrastructure. The resource allocation strategies are randomized to prevent a strategic attacker from using surveillance to learn and exploit patterns in the allocation. An important limitation of previous work on security games is that it typically assumes that attackers have *perfect* surveillance capabilities, and can learn the exact strategy of the defender. We introduce a new model that explicitly models the process of an attacker observing a sequence of resource allocation decisions and updating his beliefs about the defender's strategy. For this model we present computational techniques for updating the attacker's beliefs and computing optimal strategies for both the attacker and defender, given a specific number of observations. We provide multiple formulations for computing the defender's optimal strategy, including non-convex programming and a convex approximation. We also present an approximate method for computing the optimal length of time for the attacker to observe the defender's strategy before attacking. Finally, we present experimental results comparing the efficiency and runtime of our methods.

Introduction

Stackelberg games have been used in several deployed applications of game theory to make recommendations for allocating limited resources for protecting critical infrastructure (Basilico, Gatti, and Amigoni 2009; Korzhyk, Conitzer, and Parr 2010; Dickerson et al. 2010; Tambe 2011; An et al. 2011b). A Stackelberg security game models an interaction between an attacker and a defender (Kiekintveld et al. 2009). The defender first commits to a security policy (which may be randomized), and the attacker is able to use surveillance to learn about 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 implemented in several real-world domains, including LAX (Los Angeles International Airport) (Pita et al. 2008), FAMS (United States Federal Air Marshals Service) (Tsai et al. 2009), TSA (United States Transportation Security Agency) (Pita et al. 2011), and the United States Coast Guard (An et al. 2011a).

Most of the existing work on security games (including the methods used in the deployed applications listed above) assumes that the attacker is able to observe the defender's strategy perfectly. In reality, the attacker may have more limited observation capabilities, and our goal in this research is to develop models that capture some of these limitations in a more realistic way. Terrorists conduct surveillance to select potential targets and gain strong situational awareness of targets' vulnerabilities and security operations (Southers 2011). One important limitation is the number of observations an attacker can make; it is not possible to conduct surveillance for an infinite period of time. 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 numbers of observations using explicit belief updates.

There has been some recent work that relaxes the perfect observation assumption in security games. RECON (Yin et al. 2011) takes into account possible observation errors by assuming that the attacker's observation is within some distance from the defender's real strategy, but does not address how these errors arise or 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 observations errors, which we avoid in this paper. Yin et. al (2010) prove the equivalence of Stackelberg equilibria and Nash equilibria for some classes of security games. In general, however, Stackelberg and Nash equilibria may differ in security games, and the optimal strategy in cases with limited surveillance may be different than both. There also has been some work on understanding the value of commitment for the leader in general Stackelberg games where observations are limited or costly (Bagwell 1995;

*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.

Morgan and Vardy 2007; van Damme and Hurkens 1997).

The important difference between previous work and the methods we develop in this paper is that we consider a more detailed model of how attackers conduct surveillance operations and update their beliefs about the defender's strategy. We make the following contributions to this line of work: (1) We introduce a model of security games with strategic surveillance and formulate how the attacker updates his belief given limited observations. (2) We provide multiple formulations for computing the defender's optimal strategy, including non-convex programming and a convex approximation. (3) We provide an approximate approach for computing the optimal number of observations for the attacker. (4) We present experimental results comparing the efficiency and runtime of the methods we develop.

Stackelberg Security Games

A Stackelberg security game has two players, a defender who decides how to use m identical resources to protect a set of targets $T = \{t_1, t_2, \ldots, t_{|T|}\}$ ($m < |T|$), and an attacker who selects a single target to attack. The defender's pure strategies are all possible *feasible* assignments of the security resources to targets, with at most m targets from T protected by a single resource each. The defender's mixed strategies consist of all probability distributions over these pure strategies. The attacker's pure strategies coincide with the set of targets that can be attacked (T). In a Stackelberg game we assume that the attacker is able to (perfectly) observe the defender's mixed strategy before selecting a target to attack.

Let $A = \{A_i\}$ be a set of feasible resource assignments, where A_i is the defender's i^{th} pure strategy. If $A_{ij} = 1$, target t_j is covered by the defender in assignment A_i, and $A_{ij} = 0$ otherwise. We denote a mixed strategy for the defender by $\mathbf{x} = \langle x_i \rangle$ where x_i is the probability of choosing A_i. In many cases, we can use a compact representation for this mixed strategy (Kiekintveld et al. 2009). The strategy is represented using a marginal coverage vector $\mathbf{c} = \langle c_j \rangle$ where $c_j = \sum_{A_i \in A} x_i A_{ij}$ is the probability that target t_j is covered by some defender resource. The attacker's mixed strategy is a vector $\mathbf{a} = \langle a_j \rangle$ where a_j is the probability of attacking target t_j.

The payoffs for each player depend on which target is attacked and the probability that the target is covered by the defender. Given a target t_j, the defender receives payoff R_j^d if the adversary attacks t_j and it is covered; otherwise, the defender receives payoff P_j^d. The attacker receives payoff P_j^a in the former case and payoff R_j^a in the latter case. We assume that $R_j^d > P_j^d$ and $R_j^a > P_j^a$, so adding resources to cover a target hurts the attacker and helps the defender. For a strategy profile $\langle \mathbf{c}, \mathbf{a} \rangle$, the expected utilities for both agents are given by (notations are listed in Table 1):

$$U_d(\mathbf{c}, \mathbf{a}) = \sum_{t_j \in T} a_j U_d(\mathbf{c}, t_j), \text{ where } U_d(\mathbf{c}, t_j) = c_j R_j^d + (1 - c_j) P_j^d$$

$$U_a(\mathbf{c}, \mathbf{a}) = \sum_{t_j \in T} a_j U_a(\mathbf{c}, t_j), \text{ where } U_a(\mathbf{c}, t_j) = c_j P_j^a + (1 - c_j) R_j^a$$

In a Stackelberg model, the defender chooses a strategy first, and the attacker chooses a strategy after observing the defender's strategy. The standard solution concept for Stackelberg games is Strong Stackelberg Equilibrium (SSE) (Breton, Alg, and Haurie 1988; Leitmann 1978; von Stengel and Zamir 2004). An SSE requires that the attacker will choose his best target(s) in response to the defender's strategy, with ties broken optimally for the defender if there are multiple best responses for the attacker. Since there always exists an optimal pure-strategy response for the attacker, we restrict the attacker's strategies to pure strategies without loss of generality in this case.

We now introduce a new model that moves away from the Stackelberg model of perfect observation for security games. We call this class of games as 'security games with strategic surveillance (SGSS)'. In our model, the attacker makes a limited number of observations. The attacker may decide the number of observations to make strategically, considering the cost of conducting surveillance. The sequence of moves in an SGSS is as follows.

1. The attacker first decides how many observations to make (denoted by τ), considering the subsequent game and the cost incurred while making observations.
2. Next, the defender chooses a strategy considering the attacker's prior beliefs about the defender's strategy and the number of observations the attacker will make.
3. Finally, the attacker makes τ observations and selects the optimal target based on his posterior belief about the defender's strategy.

We assume that the attacker and the defender have common prior beliefs over the set of mixed strategies that the defender may execute. In addition, we introduce a discount factor to model the cost of surveillance. We also assume that the defender does not know the exact times when the attacker will observe the strategy being executed, and therefore cannot strategically change the strategy during times when it could be observed. 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 the rest of the paper we apply a backwards induction approach to analyze SGSS. First we model how the attacker updates his belief and chooses the best target to attack. Then we formulate an optimization problem for the defender's optimal strategy, given that the attacker will make a known number of observations. Finally, we discuss how the attacker can make a decision on how many observations to make.

Updating Attacker Beliefs

In an SGSS, the attacker updates his beliefs about the defender's strategy given his prior and τ observations, labeled $O^0, \ldots, O^{\tau-1}$, where each observation corresponds to one of the defender's pure strategies. The individual observations are drawn independently from the distribution representing the defender's mixed strategy. We can imagine the belief update proceeding sequentially, with an updated belief calculated after each observation. The attacker begins with a prior belief over the defender's mixed strategies,

Variable	Definition
m	Number of defender resources
T	Set of targets
A	Set of defender strategies
Φ	All ways of allocating m resources on T
$\mathbf{x}$	Defender mixed strategy x_i
$\mathbf{c}$	Defender coverage c_j
$\mathbf{a}$	Attacker coverage a_j
α	Parameter of attacker's prior belief
τ	Number of observations
$f^0(\mathbf{x})$	PDF of attacker's prior belief
$f^\tau(\mathbf{x}\|o)$	PDF of attacker's posterior belief given o
a^o	Attacker's strategy when his observation is o
c_j^o	Attacker's updated belief about t_j's coverage given o
Z	Huge positive constant
γ	Attacker's utility discount rate

Table 1: Notations used in this paper

represented by the probability density function $f^0(\mathbf{x})$ which represents the probability that the defender's mixed strategy is $\mathbf{x}$. We assume this prior is common knowledge. Given the first observation O^0, the attacker applies Bayes' rule to calculate the posterior distribution $f^1(\mathbf{x}|O^0)$ over the defender's mixed strategies $\mathbf{x}$. The posterior distribution $f^1(\mathbf{x})$ is then used as the prior belief distribution for observation O^1. After making τ observations, the attacker attacks the target with the highest expected valued with respect to the final posterior distribution $f^\tau(\mathbf{x}|O^0,\ldots,O^{\tau-1})$.

Example 1. We use the LAX airport as an example, based on the ARMOR application. The police at LAX place m checkpoints on the entrance roads to LAX following a mixed strategy computed using the ARMOR system (Assistant for Randomized Monitoring over Routes) (Pita et al. 2008). Attackers may engage in surveillance prior to an attack.[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 analysis of the security strategy. A single observation in this domain might involve the attacker driving around the different entrances to the airport 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]

[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 learn about only part of the full pure strategy in each observation. We consider the simpler case in this work where there is no decision about which targets to observe, only how many observations to make.

For simplicity, we assume in this work that the attacker's beliefs can be represented as a Dirichlet distribution, which is a conjugate prior for the multinomial distribution. Specifically, the support for the prior distribution $f^0(\mathbf{x})$ is the simplex $\mathcal{S} = \{\mathbf{x} : \sum_{A_i\in\Phi} x_i = 1, x_i \geq 0, \forall A_i \in \Phi\}$, where Φ is the enumeration of all possible ways of allocating m resources to cover the targets in T.[3] We can consider more general security settings in which there may exist scheduling constraints on the assignment of resources, e.g., resources have restrictions on which sets of targets they can cover (Jain et al. 2010). In this case, it follows that $A \subseteq \Phi$. If we assume that the attacker has no knowledge of the defender's scheduling/resource constraints, the attacker will have priors and update beliefs on the set of pure strategies Φ.

The Dirichlet distribution for $f^0(\mathbf{x})$ is of the form $f^0(\mathbf{x}) = \beta\prod_{A_i\in\Phi}(x_i)^{\alpha_i}$ where $\alpha = \langle\alpha_i\rangle$ is a parameter of the Dirichlet distribution and $\alpha_i > 0$. By solving the integral $\beta\int_{\mathcal{S}} f^0(\mathbf{x})\mathrm{d}\mathbf{x} = 1$, we have $\beta = \frac{(\sum_{A_i\in\Phi}\alpha_i+|\Phi|-1)!}{\prod_{A_i\in\Phi}\alpha_i!}$. The prior belief can then be represented as follows:

$$f^0(\mathbf{x}) = \frac{(\sum_{A_i\in\Phi}\alpha_i + |\Phi| - 1)!}{\prod_{A_i\in\Phi}\alpha_i!}\prod_{A_i\in\Phi}(x_i)^{\alpha_i}$$

The probability that the defender will choose pure strategy A_i given the attacker's prior belief $f^0(\mathbf{x})$ is

$$f^0(x_i) = \int_{\mathcal{S}} x_i f^0(\mathbf{x})\mathrm{d}\mathbf{x} = \frac{\alpha_i + 1}{\sum_{A_i\in\Phi}\alpha_i + |\Phi|}$$

The marginal coverage of target t_j given prior belief $f^0(\mathbf{x})$ is

$$p^0(j) = \sum_{A_i\in\Phi} A_{ij} f^0(x_i) = \frac{\sum_{A_i\in\Phi} A_{ij}(\alpha_i + 1)}{\sum_{A_i\in\Phi}\alpha_i + |\Phi|}$$

If $\alpha_i = \alpha_k$ for every i, k, $f^0(x_i) = \frac{1}{|\Phi|}$ for any strategy $A_i \in \Phi$. That is, (from the attacker's perspective) the defender chooses each strategy with the same probability. The probability of strategy A_i will increase with the increase of α_i.

Next we discuss how the attacker updates his belief given his prior belief and the sequence of observations $O = \{O^0,\ldots,O^{\tau-1}\}$ where $O^k \in \Phi$. Let $o_i(O)$ (or o_i for short) be the number of times each pure strategy A_i is executed, with $\sum_{A_i\in\Phi} o_i = \tau$. If the defender's mixed strategy is $\mathbf{x}$, the probability that the attacker will observe $o = \langle o_i\rangle$ is $f(o|\mathbf{x}) = \frac{\tau!}{\prod_{A_i\in\Phi} o_i!}\prod_{A_i\in\Phi}(x_i)^{o_i}$. After the first observation $O^0 = A_i$, the attacker's posterior distribution $f^1(\mathbf{x}|O^0)$ can be computed by applying Bayes' rule as follows:

$$f^1(\mathbf{x}|O^0) = \frac{x_i f^0(\mathbf{x})}{\int_{\mathcal{S}} x_i f^0(\mathbf{x})\mathrm{d}\mathbf{x}} = \frac{(\sum_{A_k\in\Phi}\alpha_k + |\Phi|)!}{\prod_{A_k\in\Phi}\alpha_k!(\alpha_i+1)}\prod_{A_k\in\Phi}(x_k)^{\alpha_i}x_i$$

[3]We assume that the attacker has prior knowledge about the probability distribution $f^0(\mathbf{x})$ over the defender's pure strategies. It is also possible that the attacker has prior belief on targets T's marginal coverage (say $f^0(\mathbf{c})$). In that case, we can convert $f^0(\mathbf{c})$ to $f^0(\mathbf{x})$ by solving a set of linear functions $c_j = \sum_{A_i\in\Phi} x_i A_{ij}, \forall t_j \in T$.

After applying Bayes' rule for all τ observations, we can can calculate the posterior distribution as:

$$f^{\tau}(\mathbf{x}|o) = \frac{(\sum_{A_i \in \Phi} \alpha_i + |\Phi| + \tau - 1)!}{\prod_{A_i \in \Phi}(\alpha_i + o_i)!} \prod_{A_i \in \Phi} (x_i)^{\alpha_i + o_i}$$

The marginal coverage of target t_j given the posterior belief $f^{\tau}(\mathbf{x})$ is

$$p^{\tau}(j) = \sum_{A_i \in \Phi} A_{ij} f^{\tau}(x_i) = \frac{\sum_{A_i \in \Phi} A_{ij}(\alpha_i + o_i + 1)}{\sum_{A_i \in \Phi} \alpha_i + |\Phi| + \tau}$$

After calculating these belief updates for all of the observations, the attacker chooses the best target to attack based on the final posterior belief $f^{\tau}(\mathbf{x}|o)$. The defender's real strategy $\mathbf{x}$ can affect the probability of the attacker's observations and therefore affect the attacker's choice of target.

Computing the Defender's Optimal Strategy

In this section we consider the problem of computing the defender's optimal strategy $\mathbf{x}$ given (1) the attacker's prior belief $f^0(\mathbf{x})$ represented as a Dirichlet distribution with parameter $\alpha = \langle \alpha_i \rangle$, and (2) the fact that the attacker will make a known and fixed number of observations (τ) before launching his attack.

Attacker's Optimal Strategy

We first discuss the problem of calculating the optimal attacker strategy in response to a defender strategy. Let $\mathcal{O}_\tau$ be the space of possible observations when the attacker makes τ observations, represented as $\mathcal{O}_\tau = \{o : o_i \in \{0, \ldots, \tau\}, \sum_{A_i \in A} o_i = \tau\}$. The space $\mathcal{O}_\tau$ is finite and independent of the defender's strategy $\mathbf{x}$.

One feature of SGSS is that the attacker's decision about which target to attack is determined by his prior belief and his observation o. Therefore, we can compute offline the attacker's optimal strategy $\mathbf{a}^o$ for each observation o by solving the following linear program (LP): $\mathbf{P1}$:

$$\max \quad d^o \tag{1}$$

$$a_j^o \in \{0, 1\} \quad \forall t_j \in T \tag{2}$$

$$\sum_{t \in T} a_j^o = 1 \tag{3}$$

$$d^o - c_j^o(R_j^d - P_j^d) - P_j^d \leq (1 - a_j^o)Z \quad \forall t_j \in T \tag{4}$$

$$c_j^o = \frac{\sum_{A_i \in \Phi} A_{ij}(\alpha_i + o_i + 1)}{\sum_{A_i \in \Phi} \alpha_i + |\Phi| + \tau} \quad \forall t_j \in T \tag{5}$$

$$0 \leq k^o - c_j^o(P_j^a - R_j^a) - R_j^a \leq (1 - a_j^o)Z \quad \forall t_j \in T \tag{6}$$

The formulation $\mathbf{P1}$ is similar to the MILP formulations for security games presented in (Kiekintveld et al. 2009). Equation (1) is the objective function which maximizes the defender's expected payoff from the attacker's perspective. As in a Strong Stackelberg equilibrium, we still assume that the attacker breaks ties in favor of the defender. Equations (2) and (3) force the attacker vector to assign a single target probability 1 for each observation o. Equation (4) defines the defender's payoff from the defender's perspective. Equation (6) defines the optimal response for attacker. Equation (5) defines the attacker's updated belief about the coverage of each target given the observation o. a^o represents the attacker's strategy when his observation is o. Z is a huge positive constant. c_j^o is the attacker's updated belief about the coverage of target t_j if his observation is o. k^o is the attacker's expected utility (from the attacker's perspective) when his observation is o.

In the rest of this section, we provide three mathematical programming formulations for computing the defender's optimal strategy $\mathbf{x}^*$ when the number τ of observations is known. Throughout, we assume that a^o is known for each potential observation o.

Non-convex Optimization Formulation

The formulation $\mathbf{P2}$ provides a straightforward approach for computing the defender's optimal strategy. Equation (7) is the objective function which maximizes the defender's expected payoff $\sum_{o \in \mathcal{O}_\tau} f(o|\mathbf{x}) d^o$ where d^o is the defender's utility when the attacker's observation is o. Equations (8) and (9) restrict the defender's strategy space $\mathbf{x}$. Equation (10) computes each target's marginal coverage given the defender's strategy $\mathbf{x}$. Equation (11) defines the defender's expected payoff d^o when the attacker's observation is o. The constraint places an upper bound $c_j(R_j^d - P_j^d) + P_j^d$ on the defender's expected utility d^o when t_j is attacked. $\mathbf{P2}$:

$$\max \quad \sum_{o \in \mathcal{O}_\tau} \frac{\tau!}{\prod_{A_i \in A} o_i!} \prod_{A_i \in A} (x_i)^{o_i} d^o \tag{7}$$

$$x_i \in [0, 1] \quad \forall A_i \in A \tag{8}$$

$$\sum_{A_i \in A} x_i = 1 \tag{9}$$

$$c_j = \sum_{A_i \in A} x_i A_{ij} \quad \forall t_j \in T \tag{10}$$

$$d^o - c_j(R_j^d - P_j^d) - P_j^d \leq (1 - a_j^o)Z \quad \forall t_j, o \in T \times \mathcal{O}_\tau \tag{11}$$

Convex Approximation

The objective function (7) in formulation $\mathbf{P2}$ is not convex, and no existing solver can guarantee finding the optimal solution. One approach in this case is to fall back to approximation. In this case, we can approximate the original problem by taking the log inside the summation for the objective function, changing equation (7) to $\sum_{o \in \mathcal{O}_\tau} \left(\log \frac{\tau!}{\prod_{A_i \in A} o_i!} + \sum_{A_i \in A} o_i \log(x_i) + \log d^o \right)$. However, the value of d^o could be negative, so we cannot safely apply the log operator. This issue can be resolved by adding a large value to each entry in the payoff matrix so d^o will always be positive. Since the equation $\sum_{o \in \mathcal{O}_\tau} \left(\log \frac{\tau!}{\prod_{A_i \in A} o_i!} + \sum_{A_i \in A} o_i \log(x_i) + \log d^o \right)$ is concave, we can convert this to a convex minimization problem as

P3:

$$\min \sum_{o \in \mathcal{O}_\tau} \left(-\log \frac{\tau!}{\prod_{A_i \in A} o_i!} - \sum_{A_i \in A} o_i \log(x_i) - \log d^o \right) \quad (12)$$

follows:

$$(8) - (11) \quad (13)$$

We have conducted initial experiments to evaluate the above two formulations **P2** and **P3**. In all the experiments, there is one defender resource, a varying numbers of targets, and randomly-generated payoffs satisfying the constraint that rewards are higher than penalties. R_j^d and R_j^a are drawn uniformly from the range $[100, 200]$. P_j^d and P_j^a are drawn uniformly from the range $[0, 100]$. The results were averaged over 250 trials.

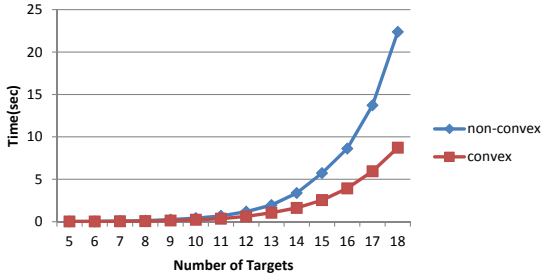

Figure 1: Runtime with different number of targets($\tau = 5$)

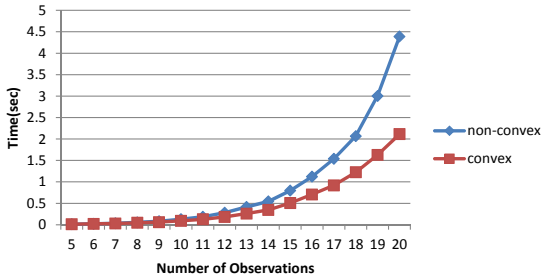

Figure 2: Runtime with different observation lengths ($|T| = 5$)

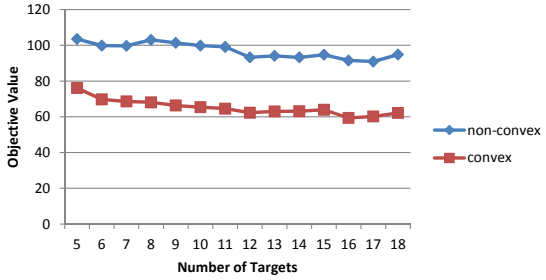

Figure 3: Utility with different number of targets ($\tau = 5$)

Figures 1 and 2 compare the runtime performance of formulations **P2** and **P3**. The x-axis is the size of the game (in terms of the number of targets or observations), and the y-axis is runtime in seconds. Initial results show that the convex optimization formulation **P3** is faster compared to formulation **P2** and the advantage increases with the increase of the scale of the game.

Figure 3 and 4 compare the expected defender utilities for formulations **P2** and **P3**. The x-axis is the size of the game

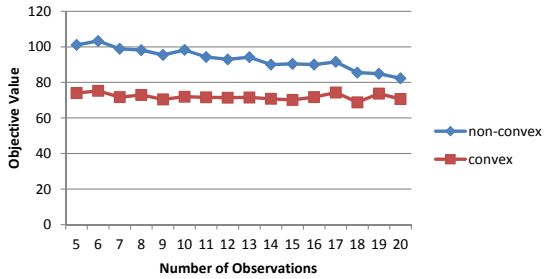

Figure 4: Utility with different observation lengths ($|T| = 5$)

(in terms of the number of targets or observations), and the y-axis is runtime in seconds. We can find that the approximate approach (**P3**) achieved lower expected defender utility with different number of targets and observations.

The Optimal Number of Observations

This section discusses how the attacker decides the number of observations to make, considering the the cost of surveillance. We model surveillance cost by introducing a discount factor $\lambda \in (0, 1)$ for the attacker. We can then formulate the attacker's optimization problem with the discount factor as a bilevel optimization problem **P4** by extending the formulation **P2**:

P4:

$$\max_{\tau} \lambda^\tau \sum_{o \in \mathcal{O}_\tau} \frac{\tau!}{\prod_{A_i \in A} o_i!} \prod_{A_i \in A} (x_i)^{o_i} k^o \quad (14)$$

$$\tau \in \mathbb{N} \quad (15)$$

$$\mathbf{x} = \text{argmax} \sum_{o \in \mathcal{O}_\tau} \frac{\tau!}{\prod_{A_i \in A} o_i!} \prod_{A_i \in A} (x_i)^{o_i} d^o \quad (16)$$

$$x_i \in [0, 1] \quad \forall A_i \in A \quad (17)$$

$$\sum_{A_i \in A} x_i = 1 \quad (18)$$

$$c_j = \sum_{A_i \in A} x_i A_{ij} \quad \forall t_j \in T \quad (19)$$

$$d^o - c_j(R_j^d - P_j^d) - P_j^d \le (1 - a_j^o)Z \quad \forall t_j, o \in T \times \mathcal{O}_\tau \quad (20)$$

$$0 \le k^o - c_j^o(P_j^a - R_j^a) - R_j^a \le (1 - a_j^o)Z \quad \forall t_j, o \in T \times \mathcal{O}_\tau \quad (21)$$

In formulation **P4**, Equation (14) is the objective function which maximizes the attacker's expected payoff $\lambda^\tau \sum_{o \in \mathcal{O}_\tau} f(o|\mathbf{x}) k^o$ when the attacker makes τ observations, and the defender takes strategy $\mathbf{x}$. Equation (15) restricts the possible number of observations the attacker can make. Equations (16)-(21) maximize the defender's expected utility when τ is known. k^o is the attacker's utility when 1) the attacker makes τ observations and 2) the defender takes strategy $\mathbf{x}$.

Bilevel optimization problems are intrinsically hard, and **P4** is even more difficult to solve since both the upper-level problem and the second-level problem are not convex. One approach is to try different values of τ and solve the defender's optimization problems using the methods described previously. Intuitively, due to the existence of discount factor λ, the attacker's utility will decrease as τ increases for sufficiently large values of τ. Therefore, we may be able to use some form of intelligent search to find the optimal value of τ.

Conclusion

This paper explicitly models the attacker's belief update and strategic surveillance decisions in security games, and presents efficient solution techniques to compute agents' optimal strategies. Our primary contributions are as follows: (1) We model the security games with strategic surveillance and formulate how the attacker updates his belief given limited observations. (2) We provide multiple formulations for computing the defender's optimal strategies, including non-convex programming and convex approximation. (3) We provide an approximate approach for computing the attacker's optimal surveillance length. (4) We present initial experimental results comparing the efficiency and runtime of our algorithms.

Our future work will focus on designing more efficient algorithms for computing the optimal strategy. Since solving bilevel optimization problems is very difficult, we will also look at some heuristic algorithm such as penalty function methods and trust-region methods. We also plan to conduct more extensive experiments to explore the implications of limited observation on both the strategies and outcomes in security games.

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.; Ordonez, 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.

Bagwell, K. 1995. Commitment and observability in games. *Games and Economic Behavior* 8:271–280.

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.

Breton, M.; Alg, A.; and Haurie, A. 1988. Sequential stackelberg equilibria in two-person games. *Optimization Theory and Applications* 59(1):71–97.

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.; Ordonez, 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 Ordonez, 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.

Leitmann, G. 1978. On generalized stackelberg strategies. *Optimization Theory and Applications* 26(4):637–643.

Morgan, J., and Vardy, F. 2007. The value of commitment in contests and tournaments when observation is costly. *Games and Economic Behavior* 60(2):326–338.

Pita, J.; Jain, M.; Western, C.; Portway, C.; Tambe, M.; Ordonez, 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)*.

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)*.

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.; Ordonez, 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.

van Damme, E., and Hurkens, S. 1997. Games with imperfectly observable commitment. *Games and Economic Behavior* 21(1-2):282–308.

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

Yin, Z.; Korzhyk, D.; Kiekintveld, C.; Conitzer, V.; ; and Tambe, M. 2010. Stackelberg vs. nash in security games: interchangeability, equivalence, and uniqueness. In *Proc.*

of The 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 1139–1146.

Yin, Z.; Jain, M.; Tambe, M.; and Ordonez, 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.