Passivity-Based Distributed Strategies for Stochastic Stackelberg Security Games

Phillip Lee[1], Andrew Clark[2], Basel Alomair[3], Linda Bushnell[1(✉)], and Radha Poovendran[1]

[1] Network Security Lab, Department of Electrical Engineering, University of Washington, Seattle, WA 98195, USA
{leep3,lb2,rp3}@uw.edu
[2] Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA 01609, USA
aclark@wpi.edu
[3] National Center for Cybersecurity Technology, King Abdulaziz City for Science and Technology (KACST), Riyadh, Saudi Arabia
alomair@kacst.edu.sa

Abstract. Stackelberg Security Games (SSGs) model scenarios where a defender implements a randomized security policy, while an attacker observes the policy and selects an optimal attack strategy. Applications of SSG include critical infrastructure protection and dynamic defense of computer networks. Current work focuses on centralized algorithms for computing stochastic, mixed-strategy equilibria and translating those equilibria into security policies, which correspond to deciding which subset of targets (e.g., infrastructure components or network nodes) are defended at each time step. In this paper, we develop *distributed* strategies for multiple, resource-constrained agents to achieve the same equilibrium utility as these centralized policies. Under our approach, each agent moves from defending its current target to defending a new target with a precomputed rate, provided that the current target is not defended by any other agent. We analyze this strategy via a passivity-based approach and formulate sufficient conditions for the probability distribution of the set of defended targets to converge to a Stackelberg equilibrium. We then derive bounds on the deviation between the utility of the system prior to convergence and the optimal Stackelberg equilibrium utility, and show that this deviation is determined by the convergence rate of the distributed dynamics. We formulate the problem of selecting a minimum-mobility security policy to achieve a desired convergence rate, as well as the problem of maximizing the convergence rate subject to mobility constraints, and prove that both formulations are convex. Our approach is illustrated and compared to an existing integer programming-based centralized technique through a numerical study.

This work was supported by ONR grant N00014-14-1-0029, NSF grant CNS-1446866 and a grant from the King Abdulaziz City for Science and Technology (KACST).

MHR Khouzani et al. (Eds.): GameSec 2015, LNCS 9406, pp. 113–129, 2015.
DOI: 10.1007/978-3-319-25594-1_7

1 Introduction

Intelligent and persistent adversaries typically observe a targeted system and its security policies over a period of time, and then mount efficient attacks tailored to the weaknesses of the observed policies. These attacks have been analyzed within the framework of Stackelberg Security Games (SSG), where the defender (leader) selects a policy in order to maximize its utility under the best response strategy of the adversary (follower) [1,2]. Applications of SSGs include defense of critical infrastructures [3,4] and intrusion detection in computer networks [5]. In both of these applications, the security policy corresponds to defending a set of targets, including ports, checkpoints, or computer network nodes.

The security of the system targeted in an SSG can be further improved through randomized policies, in which the set of nodes or locations that are guarded varies over time with a probability distribution that is chosen by the defender [2–4,6]. An attacker with knowledge of the probability distribution, but not the outcome of the randomized policy at each time step, will have greater uncertainty of the system state and reduced effectiveness of the attack.

Current work in SSGs focuses on centralized computation of the Stackelberg equilibria against different types of attackers, including rational, min-max, and bounded rational [6] attackers, under complete, incomplete, or uncertain information. In scenarios including patrolling and intrusion defense, however, security policies are implemented by distributed agents (e.g., multi-robot patrols, or malware filters in intrusion detection). These agents have limitations on computation, communication, and ability to move between targets. Currently, however, computationally efficient distributed strategies for resource-constrained defenders to achieve the same Stackelberg equilibria as centralized mechanisms are lacking.

In this paper, we developed distributed strategies for multiple defenders that guarantee convergence to a stochastic Stackelberg equilibrium distribution while minimizing the cost of movement. We propose a distributed strategy in which each defender first checks if a neighboring target is undefended, and then transitions to defending that with a certain probability if it is undefended. Since each defender only needs to know whether the neighboring targets are defended, the proposed policy can be implemented with only local communication. We analyze our approach by introducing nonlinear continuous dynamics, where each state variable is equal to the probability that a corresponding target is guarded by at least one defender, that approximate our proposed strategy. We show that, under this mapping, the Stackelberg equilibrium is achieved if and only if the continuous dynamics converge to a fixed point corresponding to the Stackelberg equilibrium. We develop sufficient conditions for convergence of these nonlinear dynamics via a passivity-based approach.

We derive bounds on the utility of an adversary with partial information as a function of the convergence rate of the dynamics, which we characterize as a passivity index. We then formulate the problem of maximizing the convergence rate, subject to mobility constraints, and prove that the formulation is convex, leading to efficient algorithms for computing the optimal policy. Our approach is validated and compared with an existing integer programming-based approach via numerical study.

The paper is organized as follows. In Sect. 2, we review related works on Stackelberg security games. In Sect. 3, the defenders and attacker models are introduced, and a zero-sum game is formulated between multiple defenders and an attacker. In Sect. 4, we propose a distributed defender strategy and prove convergence to the desired Stackelberg equilibrium. Section 5 bounds the utility of the attacker using the convergence rate of the dynamics and presents a convex optimization approach for maximizing the convergence rate. Section 6 presents our simulation results. Section 7 concludes the paper.

2 Related Work

Stackelberg Security Games (SSGs) have been gaining increasing attention in the security community in application including the defense of critical infrastructures such as airports [3,7], large interconnected computer networks [5,8] and protection of location privacy [9,10]. In particular, stochastic Stackelberg games have been used to design randomized security policies instead of deterministic policies that can be learned by the attacker with certainty.

Computing the Stackelberg equilibria has been studied in the existing literatures [11,12]. Computation of mixed-strategy Stackelberg equilibria against a worst-case (minimax or zero-sum) attacker was considered in [7]. Randomized security policies against bounded rational adversaries were proposed in [11]. When the defender has partial or uncertain information on the adversary's goals and capabilities, a repeated Stackelberg framework was proposed to model the learning and adaptation of the defender strategy over time [12]. In [13], a human adversary with bounded rationality was modeled as the quantal response (QR) in which the rationality of the adversary is characterized by a positive parameter λ, with perfect rationality and worst-case (minimax) behavior as the two extremes. Games when the defender is uncertain about the behavioral models of the attacker has been studied. In [6], a monotonic maximin solution was proposed that guarantees utility bound for the defender against a class of QR adversaries. These existing works focus on computing the Stackelberg equilibria, where optimization framework including mixed-integer programming has been used for the computation.

Centralized algorithms for choosing which targets to defend over time to achieve a Stackelberg equilibrium have received significant recent attention [14, 15], leading to deployment in harbor patrols [4] and mass transit security [3,16]. In [14], randomized patrolling of a one-dimensional perimeter by multiple robots was considered, where all robots are governed by a parameter p determining to move forward or back. In [15], a game when the attacker not only has the knowledge of the randomized policy but also the current location of the defender was analyzed, leading to attacker's strategy being function of the defense policy and the previous moves of the defender. In these works, mixed integer linear programming techniques were proposed to compute the defender strategy, which provide guaranteed optimality but require a centralized entity with worst-case exponential complexity in the number of defenders, time steps, and targets.

In the present paper, we instead consider a set of defenders who choose their strategies in a distributed manner in order to approximate the equilibrium of a one-shot SSG.

3 Model and Game Formulation

In this section, we present the defender and adversary models. We then formulate a Stackelberg game modeling the behavior of the adversary and defenders.

3.1 Defender Model

We assume that there are n targets and m defenders where $m \leq n$. The targets are represented as nodes on a complete graph, and each defender is located at one node in the graph at each time t. We model the constrained mobility of defenders and physical distances between nodes by assigning a cost d_{ij} of traversing from target i to target j. The cost of traversing may not be symmetric ($d_{ij} \neq d_{ji}$). Each defender is able to communicate with other defender to obtain information regarding whether any target is currently occupied by another defender. We define S_t to be the set of targets that is defended at time t.

3.2 Adversary Model

We consider an adversary whose goal is to successfully penetrate the system by attacking one or more targets over time. If the adversary attacks target i at time t, the adversary will collect the reward $r_i \geq 0$ if no defender is present at the target at time t. If at least one defender is present at target i at time t, the adversary will pay the cost $c_i \geq 0$. Both reward and cost values are known to the defenders and the adversary.

We consider two types of adversaries with different levels of available information. The first type of adversary is able to observe the fraction of time that a target is occupied by at least one defender for all targets but is unable to observe the current locations of defenders. The second type of adversary is able to observe exact location of one or more defenders at a sequence of times $t_1 < t_2 < \cdots < t_k$ and plan the attack strategy at time $t > t_k$ based on these observations.

3.3 Game Formulation

We consider a Stackelberg game where the defenders first choose the fraction of time that each target will be occupied by at least one defender. The adversary then observes the chosen fraction of time and decides to either attack a specific target, or not attack any target. The goal of the adversary is to maximize its expected utility, defined as the expected reward minus the expected cost of detection. The goal of the defender is to minimize the best-case expected utility of the adversary, leading to a zero-sum formulation.

To formally define the game, we denote x_i as the fraction of time that target i is occupied by at least one defender. If the adversary decides to attack target i, then the expected utility of attacking i, denoted $U_{adv}(i)$, is given as

$$U_{adv}(x_i) = (1 - x_i)r_i - x_i c_i = -(r_i + c_i)x_i + r_i \tag{1}$$

Let z_i be the adversary's chosen probability of attacking target i. Writing $\mathbf{x}$ and $\mathbf{z}$ as the vectors of defender and adversary probabilities, respectively, the expected utility of the adversary can be written as

$$U_{adv}(\mathbf{x}, \mathbf{z}) = -\mathbf{x}^T(C + R)\mathbf{z} + \mathbf{1}^T R\mathbf{z} \tag{2}$$

where C and R are $n \times n$ diagonal matrices with $C_{ii} = c_i$ and $R_{ii} = r_i$. Given $\mathbf{x}$, the adversary obtains the best-response strategy $\mathbf{z}$ by solving the linear program

$$\begin{array}{ll} \text{maximize } -\mathbf{x}^T(C + R)\mathbf{z} + \mathbf{1}^T R\mathbf{z} \\ \mathbf{y} \\ \text{s.t.} \qquad \mathbf{1}^T\mathbf{z} \leq 1, 0 \leq z_i \leq 1,\ i = 1, \ldots, n \end{array} \tag{3}$$

We note that the adversary can maximize its utility by selecting $z_i = 1$ for some i satisfying

$$i \in \arg\max \{(\mathbf{x}^T(C + R) + \mathbf{1}^T R)_j : j = 1, \ldots, n\}$$

and $z_j = 0$ otherwise. Hence, without loss of generality we assume that the adversary selects a best-response strategy $\mathbf{z}^*$ with this structure, implying that the expected utility of the adversary is given by

$$U^*_{adv}(\mathbf{x}) = \max\{\max_{i=1,\ldots,n} \{-(r_i + c_i)x_i + r_i\}, 0\} \tag{4}$$

which is a piecewise linear function in $\mathbf{x}$.

The Stackelberg equilibrium $\mathbf{x}^*$ of the defender can then be obtained as the solution to the optimization problem

$$\begin{array}{ll} \text{minimize } U^*_{adv}(\mathbf{x}) \\ \mathbf{x} \\ \text{s.t.} \qquad \mathbf{1}^T\mathbf{x} \leq m, x_i \in [0, 1] \end{array} \tag{5}$$

where the constraint $\mathbf{1}^T\mathbf{x} \leq m$ reflects the fact that there are m defenders. Equation (5) is a piecewise linear optimization problem, and hence is convex. In the following section, we will discuss how to design the mobility patterns of defenders to achieve the computed $\mathbf{x}^*$ in a distributed manner.

4 Passivity-Based Distributed Defense Strategy

In this section, we present the proposed distributed patrolling strategy of the defenders. We define continuous dynamics that approximate the probability that each target is defended at time t, and show that convergence of the continuous dynamics to the distribution $\mathbf{x}^*$ is equivalent to convergence of the time-averaged defender positions to the Stackelberg equilibrium. We formulate sufficient conditions for convergence of the continuous dynamics via a passivity-based approach.

4.1 Distributed Defender Strategy

Our proposed distributed patrolling strategy is as follows. Each defender decides whether to move to a different target according to an i.i.d. Poisson process with rate γ. At time t, the defender at target i selects a target $j \neq i$ uniformly at random and sends a query message to determine if there is already a defender at target j. If so, then the defender remains at target i. If not, the defender moves to target j with probability p_{ij}.

This defender strategy can be modeled via nonlinear continuous dynamics. Let $x_i(t)$ denote the probability that at least one defender guards target i at time t. For $\delta > 0$ sufficiently small, we then have

$$x_i(t+\delta) = x_i(t) + (1 - x_i(t)) \sum_{j \neq i} \gamma \delta p_{ji} x_j(t) - \sum_{j \neq i} \gamma \delta p_{ij} x_i(t)(1 - x_j(t)).$$

This approximation makes the simplifying assumption that the events $i \in S_t$ and $j \notin S_t$ are independent for $i \neq j$. Dividing by δ and taking the limit as $\delta \to 0$ yields

$$\dot{x}_i(t) = (1 - x_i(t)) \sum_{j \neq i} Q_{ji} x_j(t) - x_i(t) \sum_{j \neq i} Q_{ij}(1 - x_j(t)), \tag{6}$$

where $Q_{ij} = p_{ij}\gamma$. The following lemma establishes that under the dynamics (6), the number total expected number of defended targets is equal to m at each time step, and the probability that each target is defended is within the interval [0,1].

Lemma 1. *If $x_i(0) \in [0,1]$ for all i and $\mathbf{1}^T\mathbf{x}(0) = m$, then $x_i(t) \in [0,1]$ and $\mathbf{1}^T\mathbf{x}(t) = m$ for all $t \geq 0$.*

Proof. To show that $x_i(t) \in [0,1]$ for all $t \geq 0$ when $x_i(0) \in [0,1]$, let

$$t^* = \inf \{t : x_i(t) \notin [0,1] \text{ for some } i\}.$$

By continuity, $x_i(t^*) \in \{0,1\}$ for some i and $x_j(t) \in [0,1]$ for all $j \neq i$. Suppose without loss of generality that $x_i(t^*) = 0$. Then

$$\dot{x}_i(t^*) = \sum_{j \neq i} Q_{ji} x_j(t) \geq 0,$$

implying that $x_i(t) \in [0,1]$ within a neighborhood of t^* and contradicting the definition of t^*. Hence $x_i(t) \in [0,1]$ for all i and $t \geq 0$.

Now, we have that

$$\begin{aligned}
\mathbf{1}^T\dot{\mathbf{x}}(t) &= \sum_{i=1}^{n} \left[(1 - x_i(t)) \sum_{j \neq i} Q_{ji} x_j(t) - x_i(t) \sum_{j \neq i} Q_{ij}(1 - x_j(t)) \right] \\
&= \sum_{i=1}^{n} \left[\sum_{j \neq i} (Q_{ji} x_j(t) - Q_{ij} x_i(t)) + \sum_{j \neq i} (Q_{ij} x_i(t) x_j(t) - Q_{ji} x_i(t) x_j(t)) \right] = 0,
\end{aligned}$$

implying that $\mathbf{1}^T\mathbf{x}(t)$ is constant.

4.2 Passivity-Based Convergence Analysis

We now derive conditions on the matrix Q to ensure that, for any initial distribution $\mathbf{x}(0)$, the dynamics (6) satisfy $\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{x}^*$. If this condition holds, then the time-averaged distribution satisfies $\frac{1}{T}\int_0^T \mathbf{x}(t)\, dt \to \mathbf{x}^*$, and hence the Stackelberg equilibrium is achieved.

By inspection of (6), convergence to $\mathbf{x}^*$ occurs only if

$$(1 - x_i^*) \sum_{j \neq i} Q_{ji} x_j^* = x_i^* \sum_{j \neq i} Q_{ij}(1 - x_j^*)$$

for all i. Defining D^* to be a diagonal matrix with $D_{ii}^* = x_i^*$, this necessary condition can be written in matrix form as

$$(D^*(Q - Q^T) + Q^T)\mathbf{x}^* = D^* Q \mathbf{1}. \tag{7}$$

In order to develop sufficient conditions for convergence to $\mathbf{x}^*$, we introduce a decomposition of the dynamics (6) into a negative feedback interconnection between two passive dynamical systems. Recall that a dynamical system Σ is *output feedback passive* if there exists a positive semidefinite function V such that

$$\dot{V}(t) \leq \rho y(t)^T y(t) + u(t)^T y(t) \tag{8}$$

for all input u and output y for all time t. If $\rho = 0$, then the system is called passive, and the system is called strictly passive if $\rho < 0$. The parameter ρ is defined as the output feedback passivity index of the system [17].

Define $\hat{\mathbf{x}}(t) = \mathbf{x}(t) - \mathbf{x}^*$, and let two input-output dynamical systems be given by

$$(\Sigma_1) \quad \begin{cases} \dot{\hat{x}}_i(t) = -(R_{in}(i) + R_{out}(i))\hat{x}_i(t) + u_i^{(1)}(t) \\ y_i^{(1)}(t) = \hat{x}_i(t) \end{cases} \tag{9}$$

$$(\Sigma_2): \quad \mathbf{y}^{(2)}(t) = -(D^*(Q - Q^T) + Q^T)\mathbf{u}^{(2)}(t) \tag{10}$$

where $R_{in}(i) = \sum_{j\in N(i)} Q_{ji} x_j(t)$ and $R_{out}(i) = \sum_{j\in N(i)} Q_{ij}(1 - x_i(t))$. By inspection, the trajectory of $\hat{x}_j(t)$ in the negative feedback interconnection between (Σ_1) and (Σ_2), shown in Fig. 1, is equivalent to the trajectory of $\hat{x}_j(t)$ under the dynamics (6).

The decomposition of Fig. 1 can be interpreted as follows. The top block represents the change in the probability that each target i is defended, based on the current probability that target i is defended. The input signal from the bottom block can be interpreted as the rate at which defenders from other targets move to target i.

A standard result states that the negative feedback interconnection between two strictly passive systems is globally asymptotically stable [17], which in this case implies that $\mathbf{x}(t)$ converges asymptotically to $\mathbf{x}^*$. Hence, it suffices to derive conditions under which systems (Σ_1) and (Σ_2) are strictly passive. We now present sufficient conditions for strict passivity of (Σ_1) and (Σ_2), starting with (Σ_1).

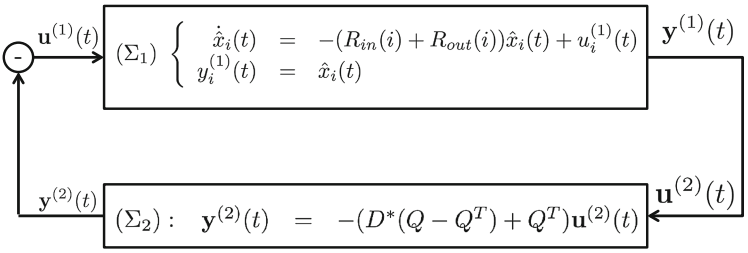


Fig. 1. Decomposition of the patrol dynamics as negative feedback interconnection between passive systems.

Proposition 1. *The system* (Σ_1) *is passive from input* $\mathbf{u}^{(1)}(t)$ *to output* $\mathbf{y}^{(1)}(t)$. *If* $\max_j \{\min\{Q_{ji}, Q_{ij}\}\} > 0$ *for all* i, *then* (Σ_1) *is strictly passive.*

Proof. Consider the storage function $V(\hat{\mathbf{x}}) = \frac{1}{2}\hat{\mathbf{x}}^T\hat{\mathbf{x}}$. We have

$$\dot{V}(\hat{\mathbf{x}}) = -\sum_i (R_{in}(i) + R_{out}(i))\hat{x}_i^2 + (\mathbf{u}^{(1)})^T\hat{\mathbf{x}}.$$

Since the output $\mathbf{y}^{(1)}$ is given by $\mathbf{y}^{(1)}(t) = \hat{\mathbf{x}}$, it suffices to show that $R_{in}(i) + R_{out}(i) > 0$ for all feasible $\mathbf{x}$. We have

$$R_{in}(i) + R_{out}(i) = \sum_i [Q_{ji}x_j + Q_{ij}(1 - x_j)]. \tag{11}$$

Since $x_j \in [0, 1]$, each term of (11) is bounded below by $\min\{Q_{ji}, Q_{ij}\} \geq 0$. Hence the system (Σ_1) satisfies $\dot{V}(\hat{\mathbf{x}}) \leq (\mathbf{u}^{(1)})^T\mathbf{y}$, implying passivity. Furthermore, if the condition $\max_j \{\min\{Q_{ji}, Q_{ij}\}\} =: k > 0$ holds for all i, then

$$\dot{V}(\hat{\mathbf{x}}) < -k\hat{\mathbf{x}}^T\hat{\mathbf{x}} + (\mathbf{u}^{(1)})^T\mathbf{y},$$

implying strict passivity.

The condition $\max_j \{\min\{Q_{ji}, Q_{ij}\}\} > 0$ implies that, for target i, there exists at least one target j such that defenders will transition to target i from target j, and vice versa, with positive probability.

For the system (Σ_2), define matrix $K = (D^*(Q - Q^T) + Q^T)$, so that $\mathbf{y}^{(2)} = -K\mathbf{u}^{(2)}$. If $-\mathbf{u}^T K\mathbf{u} \geq 0$ for all $\mathbf{u}$, then passivity of the bottom block would be guaranteed. On the other hand, since the diagonal entries of K are all 0, the matrix K is neither positive- nor negative-definite. The following proposition gives a weaker sufficient condition.

Proposition 2. *Define* $P = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$. *If* $PKP \leq 0$ *for all* $\mathbf{u}$, *then the system* (Σ_2) *satisfies* $\mathbf{u}^T\mathbf{y} \geq 0$ *for all* $\mathbf{u}$ *satisfying* $\mathbf{1}^T\mathbf{u} = 0$.

Proof. Suppose that $\mathbf{1}^T\mathbf{u} = 0$. Then $P\mathbf{u} = \mathbf{u}$, since P projects any vector onto the subspace orthogonal to $\mathbf{1}$, and hence $\mathbf{u}^T K\mathbf{u} = \mathbf{u}^T PKP\mathbf{u}$. The inequality $PKP \leq 0$ then implies that $\mathbf{u}^T\mathbf{y} = \mathbf{u}^T K\mathbf{u} \leq 0$.

Combining the conditions for passivity of (Σ_1) and (Σ_2) with the fact that $\mathbf{1}^T\hat{\mathbf{x}}(t) = 0$ (Lemma 1) yields the following sufficient condition for convergence to the desired distribution $\mathbf{x}^*$.

Theorem 1. *If the conditions*

$$K\mathbf{x}^* = D^*Q\mathbf{1} \tag{12}$$

$$\max_j \{\min\{Q_{ji}, Q_{ij}\}\} > 0 \ \forall i \tag{13}$$

$$P^T \frac{K + K^T}{2} P \leq 0 \tag{14}$$

hold, then the vector of probabilities $\mathbf{x}(t)$ *converges to* $\mathbf{x}^*$ *as* $t \to \infty$. *There exists at least one realization of* Q *with* $Q_{ij} \geq 0$ *for all* $i \neq j$ *and* $Q_{ii} = 0$ *that satisfies (12)–(14).*

Proof. Condition (12) implies that the equilibrium of the dynamics (6) corresponds to the Stackelberg equilibrium $\mathbf{x}^*$. Conditions (13) and (14) establish strict passivity of (Σ_1) (Proposition 1) and passivity of (Σ_2) (Proposition 2), respectively, when the trajectory satisfies $\mathbf{1}^T\hat{\mathbf{x}}(t) = 0$ and $x_i(t) \in [0, 1]$ for all i and t, which is guaranteed by Lemma 1. Hence the overall system is globally asymptotically stable with equilibrium $\mathbf{x}^*$. It remains to show that there is a feasible matrix Q that satisfies the conditions (12)–(14).

The proof constructs a matrix Q such that $\frac{K+K^T}{2} = \zeta(\frac{1}{n}\mathbf{1}\mathbf{1}^T - I)$ for some $\zeta \geq 0$. By construction, $\frac{1}{2}P(K + K^T)P = -\zeta P^3 \leq 0$, since $P \geq 0$.

For this choice of $\frac{K+K^T}{2}$, the identities $\frac{K+K^T}{2} = \zeta(\frac{1}{n}\mathbf{1}\mathbf{1}^T - I)$ and $K\mathbf{x}^* = D^*Q\mathbf{1}$ are equivalent to

$$x_i^*Q_{ij} + (1 - x_j^*)Q_{ij} + x_j^*Q_{ji} + (1 - x_i^*)Q_{ji} = \zeta \ \forall i \neq j \tag{15}$$

$$\sum_j x_i^*(1 - x_j^*)Q_{ij} = \sum_j x_j^*(1 - x_i^*)Q_{ji} \ \forall i \tag{16}$$

Define

$$\tau_{ij} = \frac{1}{1 - x_j^*} + \frac{1}{x_i^*} + \frac{1}{1 - x_i^*} + \frac{1}{x_j^*},$$

and let $Q_{ij} = \frac{\zeta}{\tau_{ij}x_i^*(1-x_j^*)}$. Substitution of Q_{ij} and Q_{ji} into (15) yields

$$\frac{x_i^*\zeta}{\tau_{ij}x_i^*(1 - x_j^*)} + \frac{(1 - x_j^*)\zeta}{\tau_{ij}x_i^*(1 - x_j^*)} + \frac{x_j^*\zeta}{\tau_{ij}x_j^*(1 - x_i^*)} + \frac{(1 - x_i^*)\zeta}{\tau_{ij}x_j^*(1 - x_i^*)} = \zeta,$$

implying that (15) holds. Furthermore,

$$x_i^*(1 - x_j^*)Q_{ij} = \frac{\gamma}{\tau_{ij}} x_j^*(1 - x_i^*)Q_{ji},$$

and hence (16) holds as well.

Observe that under this choice of Q, $Q_{ij} \geq 0$ for all i, j, and condition (13) is satisfied as well.

While there may be multiple matrices Q satisfying conditions (12)–(14), and hence guaranteeing convergence to $\mathbf{x}^*$, the corresponding dynamics of each defender may lead to a high cost associated with moving between distant targets. The problem of selecting the values of Q that minimize the total movement can be formulated as

$$\begin{aligned} &\text{minimize} \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}Q_{ij}x_i^*(1-x_j^*) \\ &Q, K \\ &\text{s.t.} \quad K = D^*(Q-Q^T)+Q^T \\ &\qquad P(K+K^T)P \leq 0 \\ &\qquad K\mathbf{x}^* = D^*Q\mathbf{1} \\ &\qquad Q_{ij} \geq 0 \ \forall i \neq j, \ Q_{ii} = 0 \ \forall i \\ &\qquad \max_j \{\min\{Q_{ji}, Q_{ij}\}\} > 0 \ \forall i \end{aligned} \tag{17}$$

The objective function $\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}Q_{ij}x_i^*(1-x_j^*)$ can be interpreted as the total movement cost to maintain the Stackelberg equilibrium $\mathbf{x}^*$ once the equilibrium is reached. Equation (17) can be reformulated as a standard-form semidefinite program and solved in polynomial time. Furthermore, the procedure described in Theorem 1 can be used to construct a feasible solution to (17) in $O(n^2)$ time when the number of targets is large.

5 Mitigating Side Information of Adversary

In this section, we analyze the performance of our approach against an adversary with knowledge of the defender positions at a previous time period. We first bound the deviation between the utility of an adversary with partial information and the Stackelberg equilibrium utility. Our bound is a function of the convergence rate of the dynamics (6). We then formulate the problem of maximizing the convergence rate subject to mobility constraints, as well as the problem of selecting the least-costly patrolling strategy to achieve a desired convergence rate.

5.1 Deviation from Stackelberg Equilibrium

An adversary who observes the defender positions at time t' can estimate the probability $x_i(t)$ that target i is defended at time $t > t'$ via the dynamics (6). The adversary then computes the optimal strategy $\mathbf{z}(t)^*$, where $z_i(t)^*$ is the probability of attacking target i at time t, by solving the optimization problem $\max\{-\mathbf{x}(t)^T(C+R)\mathbf{z} + \mathbf{1}^T R\mathbf{z} : \mathbf{1}^T\mathbf{z} = 1, \mathbf{z} \geq 0\}$.

The deviation of the resulting utility from the Stackelberg equilibrium is given by

$$E(t) = \sum_j [z_j(t)^*(c_j x_j(t) + (1-x_j(t))r_j) - z_j^*(x_j^* c_j + (1-x_j^*)r_j)].$$

The following theorem provides an upper bound on $E(t)$ as a function of the convergence rate.

Theorem 2. *The expression $E(t)$ satisfies*

$$E(t) \leq 2 \max_j \{|c_j||x_j(t) - x_j^*| + |r_j||x_j(t) - x_j^*|\} + \max_j |c_j - r_j| \sum_j |x_j(t) - x_j^*|. \tag{18}$$

Proof. Letting $\alpha_j(x_j(t)) = c_j x_j(t) + r_j(1 - x_j(t))$,

$$\begin{aligned} E(t) &= \sum_j [\alpha_j(x_j(t))(z_j(t)^* - z_j^* + z_j^*) - z_j^* \alpha_j(x_j^*)] \\ &= \sum_j [\alpha_j(x_j(t))(z_j(t)^* - z_j^*) + z_j^*(\alpha_j(x_j(t)) - \alpha_j(x_j^*))]. \end{aligned} \tag{19}$$

Considering the two terms of the inner summation in (19) separately, we first have that $\sum_j \alpha_j(x_j(t))(z_j(t)^* - z_j^*)$ is equal to $\alpha_j(x_j(t)) - \alpha_i(x_i(t))$, where j is the target attacked by the adversary in the best-response to distribution $\mathbf{x}(t)$ and i is the target attacked by the adversary in the best-response to $\mathbf{x}^*$. We then have

$$\begin{aligned} \alpha_j(x_j(t)) - \alpha_i(x_i(t)) &= c_j x_j(t) + r_j(1 - x_j(t)) - c_i x_i(t) - r_i(1 - x_i(t)) \\ &= c_j \hat{x}_j(t) - r_j \hat{x}_j(t) - c_i \hat{x}_i(t) + r_i \hat{x}_i(t) \\ &\quad + c_j x_j^* + r_j(1 - x_j^*) - c_i x_i^* - r_i(1 - x_i^*) \\ &\leq c_j \hat{x}_j(t) - r_j \hat{x}_j(t) - c_i \hat{x}_i(t) + r_i \hat{x}_i(t) \qquad (20) \\ &\leq |c_j||x_j - x_j^*| + |r_j||x_j - x_j^*| \qquad (21) \\ &\quad + |c_i||x_i - x_i^*| + |r_i||x_i - x_i^*| \end{aligned}$$

where (20) follows from the fact that i is a best-response to $\mathbf{x}^*$ and (21) follows from the triangle inequality. Taking an upper bound over i and j yields the first term of (18).

Now, consider the second term of $E(t)$. We have

$$\alpha_j(x_j(t)) - \alpha_j(x_j^*) = c_j x_j(t) + (1 - x_j(t)) r_j - c_j x_j^* - r_j(1 - x_j^*) = (c_j - r_j)(x_j(t) - x_j^*).$$

Hence

$$\begin{aligned} \sum_j z_j^*(\alpha_j(x_j(t)) - \alpha(x_j^*)) &= \sum_j z_j^*(c_j - r_j)(x_j(t) - x_j^*) \\ &\leq \max_i |c_i - r_i| \sum_j |x_j(t) - x_j^*|, \end{aligned}$$

the second term of (18).

Theorem 1 implies that the deviation between the optimal adversary utility at time t and the Stackelberg equilibrium is determined by the convergence rate. The convergence rate can be bounded via a Lyapunov-type argument. As a preliminary, we have the following standard result.

Proposition 3. *[17] Let $V(x)$ be a continuously differentiable function such that*

$$c_1||x||^a \leq V(x) \leq c_2||x||^a \tag{22}$$
$$\dot{V}(x) \leq -c_3||x||^a \tag{23}$$

over a domain $D \subset \mathbb{R}^n$. Suppose $\dot{x} = f(x)$ satisfies $f(0) = 0$. Then

$$||x(t)|| \leq \left(\frac{c_2}{c_1}\right)^{1/a} \exp\left(-\frac{c_3}{c_2 a}\right)||x(0)||.$$

A bound on the convergence rate can then be derived via the passivity analysis of Sect. 4.

Proposition 4. *Define $K_p = P^T(\frac{K+K^T}{2})P$, where $P = (I - \frac{1}{n}\mathbf{1}\mathbf{1}^T)$, and suppose that $K_p \leq 0$. Denote the eigenvalues of K_p as $0 \geq -\lambda_1 \geq \cdots \geq -\lambda_{n-1}$ and associated eigenvector of λ_i as q_i. Then, the deviation $||\mathbf{x}(t) - \mathbf{x}^*||_2$ satisfies*

$$||\mathbf{x}(t) - \mathbf{x}^*||_2 \leq \exp(-\lambda_1). \tag{24}$$

Proof. Let $V(\hat{\mathbf{x}}) = \frac{1}{2}\hat{\mathbf{x}}^T\hat{\mathbf{x}}$. In the notation of Proposition 3, we have $a = 2$ and $c_1 = c_2 = \frac{1}{2}$. We will bound $\dot{V}(\hat{\mathbf{x}})$ as a function of $||\hat{\mathbf{x}}||^2$. Any $\hat{\mathbf{x}}$ such that $\mathbf{1}^T\hat{\mathbf{x}} = 0$ satisfies $\hat{\mathbf{x}} = P\hat{\mathbf{x}}$. Then, from the passivity analysis in Proposition 1, we have

$$\dot{V}(\hat{\mathbf{x}}) \leq \hat{\mathbf{x}}^T K \hat{\mathbf{x}} = \hat{\mathbf{x}}^T P^T \frac{K + K^T}{2} P\hat{\mathbf{x}} = \hat{\mathbf{x}}^T K_p \hat{\mathbf{x}}$$

which can be upper bounded as

$$\begin{aligned}
\hat{\mathbf{x}}^T K_p \hat{\mathbf{x}} &\overset{(a)}{=} \sum_{i=1}^{n-1} -\lambda_i (\mathbf{q}_i^T \hat{\mathbf{x}})^2 \leq -\lambda_1 \sum_{i=1}^{n-1} \hat{\mathbf{x}}^T \mathbf{q}_i \mathbf{q}_i^T \hat{\mathbf{x}} \\
&\overset{(b)}{=} -\lambda_1 \sum_{i=1}^{n-1} \hat{\mathbf{x}}^T (I - \frac{1}{n}\mathbf{1}\mathbf{1}^T)\hat{\mathbf{x}} = -\lambda_1 \hat{\mathbf{x}}^T P \hat{\mathbf{x}} \\
&\overset{(c)}{=} -\lambda_1 \hat{\mathbf{x}}^T P^T P \hat{\mathbf{x}} = -\lambda_1 ||\hat{\mathbf{x}}||^2
\end{aligned}$$

where (a) is from eigen decomposition, (b) is from the orthogonality of eigenvectors for symmetric matrices, and (c) is from the idempotent property of the projection matrix. Substituting $-\lambda_1$ as c_3 from Proposition 3, we obtain the desired bound.

The proof of Proposition 4 implies that $\dot{V}(\hat{\mathbf{x}}) \leq -\lambda_1 \hat{\mathbf{x}}^T \hat{\mathbf{x}}$, implying that λ_1 is a *passivity index* [17] for the system (Σ_1). Proposition 4 shows that maximizing over the convergence rate is equivalent to maximizing $|\lambda_1|$, which will be considered in the following section.

5.2 Optimizing the Convergence Rate

The problem of maximizing the convergence rate subject to the mobility constraint can be formulated as

$$\begin{array}{ll} \text{maximize } s & \\ Q, K, s & \\ \text{s.t.} & K = D^*(Q - Q^T) + Q^T \\ & K\mathbf{x}^* = D^*Q\mathbf{1} \\ & Q_{ij} \geq 0 \ \forall i \neq j, \ Q_{ii} = 0 \ \forall i \\ & \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} Q_{ij} \leq d \\ & \max_j \{\min \{Q_{ji}, Q_{ij}\}\} > 0 \ \forall i \\ & P\left(\frac{K+K^T}{2}\right) P + sP \leq 0, s \geq 0 \end{array} \tag{25}$$

The first four constraints are from (17). The last constraint ensures the negative semi-definiteness of the matrix $P(K+K^T)P$ and maximization of $|\lambda_1|$, as shown in the following proposition.

Proposition 5. *Denote the eigenvalues of* $P(K + K^T)P$ *as* $0, \lambda_1, \ldots, \lambda_{n-1}$ *ordered such that* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1}$*, and let* q_i *denote the eigenvector associated with eigenvalue* λ_i*. If* $P(K + K^T)P + sP \leq 0$*, then* $\lambda_1 \leq -s$*.*

Proof. Let $K_P = P(K + K^T)P$. Then the matrix $K_P + sP$ can be rewritten as

$$K_P + sP = PK_PP + sPIP = P(K_P + sI)P \tag{26}$$

by the idempotent property of P. If $P(K_P+sI)P \leq 0$, then $\mathbf{x}^T P(K_P+sI)P\mathbf{x} \leq 0$ for all $\mathbf{x}$. Letting $\hat{\mathbf{x}} = P\mathbf{x}$, we have

$$\hat{\mathbf{x}}^T(K_P + sI)\hat{\mathbf{x}} \leq 0$$

for all $\hat{\mathbf{x}}$ that satisfies $\mathbf{1}^T\hat{\mathbf{x}} = 0$. In particular, choose $\hat{\mathbf{x}} = q_1$, which satisfies the condition $\mathbf{1}^T q_1$ from the orthogonality of eigenvectors of a symmetric matrix. Then $q_1^T(K_P + I)q_1 = \lambda_1 + s \leq 0$, and hence $\lambda_1 \leq -s$.

By Proposition 5, the constraints $P(K + K^T)P + sP$ and $s \geq 0$ ensure the negative semidefiniteness of $P(K + K^T)P$ and maximizing s will result in $s^* = |\lambda_1|$. The formulated optimization problem is a semidefinite program and can be solved efficiently in polynomial time as in the case of (17).

An alternative optimization is minimizing the patrol cost for a given convergence rate λ. This optimization problem can be formulated as

$$\begin{array}{ll} \text{minimize } \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} Q_{ij} x_i^*(1 - x_j^*) & \\ Q, K & \\ \text{s.t.} & K = D^*(Q - Q^T) + Q^T \\ & P\left(\frac{K+K^T}{2}\right) P + \lambda P \leq 0 \\ & K\mathbf{x}^* = D^*Q\mathbf{1} \\ & Q_{ij} \geq 0 \ \forall i \neq j, \ Q_{ii} = 0 \ \forall i \end{array} \tag{27}$$

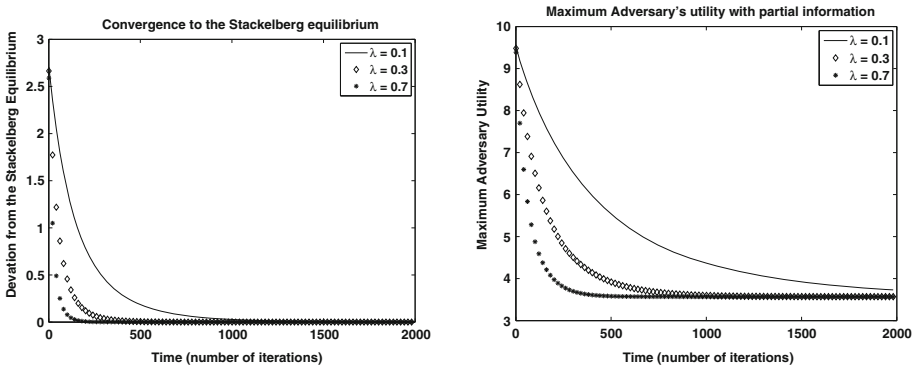
which is also convex. This optimization problem is always feasible by the same argument given in Theorem 1, since given a $\lambda > 0$, one can set $\zeta = \lambda$ in the proof of Theorem 1 and construct a matrix Q that satisfies the constraint of (27). This optimization problem returns the least costly patrolling policy given a security constraint of achieving a desired convergence rate to the Stackelberg equilibrium.

6 Numerical Study

In this section, we conduct a numerical study via Matlab on a patrolling application. The formulated optimization problems were solved using cvx. We consider a network with 30 targets deployed uniformly at random in a square of size 10. The mobility cost d_{ij} was set as the Euclidean distance between target i and j. The number of defenders was set to 5. The diagonal reward and cost matrices R and C were randomly generated where the reward and cost values r_i and c_i were chosen uniformly in the interval $(0, 10)$.

We first obtained a Stackelberg equilibrium $\mathbf{x}^*$ by solving the convex optimization problem (5), and solved for Q for a set of convergence rates λ by solving the optimization problem (27) where the movement cost is minimized for a given convergence rate. The adversary's utility at the Stackelberg equilibrium was 3.56.

Convergence of $\mathbf{x}(t)$ to the Stackelberg equilibrium $\mathbf{x}^*$ under the continuous dynamics (6) is shown in Fig. 2(a). The initial positions were chosen at random among 30 targets. We observe that $\mathbf{x}(t)$ converges to $\mathbf{x}^*$ exponentially with differing convergence rates as shown in Proposition 4. Figure 2(b) shows the maximum



(a) (b)

Fig. 2. (a) Figure illustrating the convergence of $\mathbf{x}(t)$ to $\mathbf{x}^*$. Metric for deviation from the Stackelberg equilibrium was $||\mathbf{x}(t) - \mathbf{x}^*||$ with Q matrices obtained with varying λ by solving optimization problem (27). (b) Maximum adversary's utility with information of the initial locations of defenders. The maximum utility of the adversary decays exponentially, with the maximum utility being the reward value of the target that is not covered by a defender initially.

utility of the adversary over time when the adversary observes the positions of defenders at time $t = 0$. The maximum utility of the adversary at time $t = 0$ is shown to be 9.5 which is the maximum reward value of targets that are not guarded by defender at time $t = 0$. Maximum adversary's utility converges to the defender's utility at Stackelberg equilibrium. The maximum utility of the adversary also decays exponentially with higher convergence rate of (6) offering faster decay of the adversary's utility as observed in Theorem 2.

Our proposed approach is compared with the integer programming-based technique, denoted Raptor, for centralized computation of patrol routes developed in [16] as shown in Fig. 3. Each data point represents an average over 15 independent and random trials with different cost and reward matrices, as well as target locations. The number of defenders was set to 3. For our approach, the minimum patrolling cost was obtained from the optimization problem (27), while the movement cost of Raptor is the minimum cost to transition between two sets of patroller locations sampled randomly with distribution $\mathbf{x}^*$. Our approach is able to achieve comparable mobility cost to Raptor with a convergence rate of $\lambda = 10^{-3}$. We observe that under our approach, as the number of targets increases, the minimum movement cost increases, with the rate of increase proportional to the convergence rate while Raptor's minimum patrolling cost stays relatively constant as the number of targets increase.

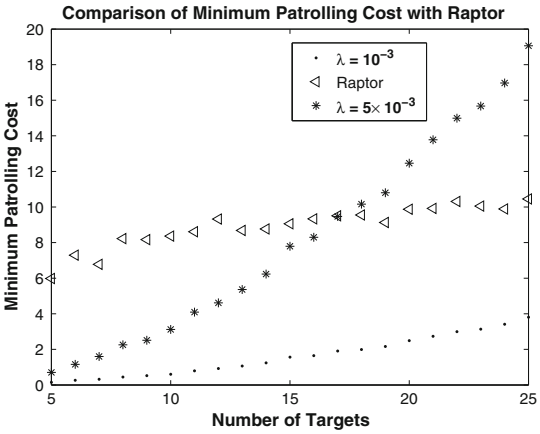


Fig. 3. Minimum patrolling cost with different convergence rate λ and Raptor [16]. The number of defenders was set to 3. It is shown that our approach is able to achieve comparable mobility cost to Raptor with a convergence rate of $\lambda = 10^{-3}$. Under our approach, the minimum movement cost grows in a linear manner as the number of targets grows, and the slope of the line is proportional to the convergence rate λ. Raptor's minimum patrolling cost remains relatively constant as the number of targets grows.

7 Conclusions and Future Work

Stackelberg security games are a modeling framework for scenarios in which a defender chooses a randomized security policy, and an adversary observes the distribution of the randomized policy and selects an attack accordingly. In this paper, we developed a strategy for a team of defenders to implement a stochastic Stackelberg equilibrium security policy. Under our proposed strategy, each defender selects a target according to a precomputed probability distribution at each time step and moves to that target if the target is currently unoccupied. We formulated sufficient conditions, via a passivity-based approach, for a chosen probability distribution to guarantee convergence to the desired Stackelberg equilibrium.

We analyzed the behavior of an intelligent adversary who observes the previous positions of the set of defenders and selects an attack strategy based on these positions and the knowledge of the defender strategies. We proved that the additional impact of the attack provided by knowledge of the defender positions can be bounded as a function of the convergence rate of the defenders to the Stackelberg equilibrum. Under the passivity framework, this convergence rate is interpreted as a passivity index. We formulated the problem of selecting the minimum-cost (in terms of defender movement) strategy to achieve a desired convergence rate, as well as the problem of selecting the fastest-converging defender strategy under mobility constraint, as semidefinite programs, enabling efficient computation of the optimal patrols for each defender. Numerical results verified that both the deviation from the Stackelberg equilibrium and the adversary's utility decayed exponentially over time. The numerical study also suggested that the minimum patrolling cost increased linearly in the number of targets for a fixed number of defenders.

The approach presented in this paper assumes a set of identical defenders that are capable of moving between any two targets within a desired time. A direction of future research is to generalize the approach to heterogeneous defenders who require multiple time steps to move between distant targets, reflecting a deployment over a wide geographical area. We will also extend the proposed approach to arbitrary topologies with mobility constraint of defenders and numerically evaluate the approach with real-world data including the transit network used in [16]. In addition, we will investigate incorporating Bayesian framework where both the defender and the adversary have prior distribution of each other's utility and initial locations and develop approximation algorithms to solve the Bayesian Stackelberg game.

References

1. Paruchuri, P., Pearce, J.P., Marecki, J., Tambe, M., Ordonez, F., Kraus, S.: Playing games for security: an efficient exact algorithm for solving Bayesian Stackelberg games. In: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems, vol. 2, pp. 895–902 (2008)

2. Manshaei, M.H., Zhu, Q., Alpcan, T., Bacşar, T., Hubaux, J.-P.: Game theory meets network security and privacy. ACM Comput. Surv. (CSUR) **45**(3), 25 (2013)
3. Pita, J., Jain, M., Marecki, J., Ordóñez, F., Portway, C., Tambe, M., Western, C., Paruchuri, P., Kraus, S.: Deployed ARMOR protection: the application of a game theoretic model for security at the Los Angeles international airport. In: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems: Industrial Track, pp. 125–132 (2008)
4. Shieh, E., An, B., Yang, R., Tambe, M., Baldwin, C., DiRenzo, J., Maule, B., Meyer, G.: Protect: a deployed game theoretic system to protect the ports of the United States. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, vol. 1, pp. 13–20 (2012)
5. Chen, L., Leneutre, J.: A game theoretical framework on intrusion detection in heterogeneous networks. IEEE Trans. Inf. Forensics Secur. **4**(2), 165–178 (2009)
6. Jiang, A.X., Nguyen, T.H., Tambe, M., Procaccia, A.D.: Monotonic maximin: a robust stackelberg solution against boundedly rational followers. In: Das, S.K., Nita-Rotaru, C., Kantarcioglu, M. (eds.) GameSec 2013. LNCS, vol. 8252, pp. 119–139. Springer, Heidelberg (2013)
7. Brown, G., Carlyle, M., Salmerón, J., Wood, K.: Defending critical infrastructure. Interfaces **36**(6), 530–544 (2006)
8. Zonouz, S., Khurana, H., Sanders, W.H., Yardley, T.M.: Rre: a game-theoretic intrusion response and recovery engine. IEEE Trans. Parallel Distrib. Syst. **25**(2), 395–406 (2014)
9. Shokri, R., Theodorakopoulos, G., Troncoso, C., Hubaux, J.-P., Le Boudec, J.-Y.: Protecting location privacy: optimal strategy against localization attacks. Conf. Comput. Commun. Secure. 617–627 (2012)
10. Shokri, R.: Privacy games: optimal user-centric data obfuscation. Proc. Priv. Enhancing Technol. **2015**(2), 1–17 (2015)
11. Yang, R., Ordonez, F., Tambe, M.: Computing optimal strategy against quantal response in security games. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, vol. 2, pp. 847–854 (2012)
12. Kiekintveld, C., Marecki, J., Tambe, M.: Approximation methods for infinite bayesian stackelberg games: modeling distributional payoff uncertainty. In: The 10th International Conference on Autonomous Agents and Multiagent Systems, vol. 3, pp. 1005–1012 (2011)
13. McKelvey, R.D., Palfrey, T.R.: Quantal response equilibria for normal form games. Games Econ. Behav. **10**(1), 6–38 (1995)
14. Agmon, N., Kraus, S., Kaminka, G. et al.: Multi-robot perimeter patrol in adversarial settings. In: International Conference on Robotics and Automation, pp. 2339–2345 (2008)
15. Vorobeychik, Y., An, B., Tambe, M., Singh, S.: Computing solutions in infinite-horizon discounted adversarial patrolling games. In: International Conference on Automated Planning and Scheduling (2014)
16. Varakantham, P., Lau, H.C., Yuan, Z.: Scalable randomized patrolling for securing rapid transit networks. In: Proceedings of the Twenty-Fifth Innovative Applications of Artificial Intelligence Conference
17. Khalil, H.K.: Nonlinear Systems. Prentice Hall Upper Saddle River (2002)