Submitted to *Interfaces*
manuscript (Please, provide the mansucript number!)

A Deployed Quantal Response Based Patrol Planning System for the US Coast Guard

Bo An
boa@usc.edu
University of Southern California, USA

Fernando Ordóñez
fordon@dii.uchile.cl
Universidad de Chile, Chile *and* University of Southern California, USA

Milind Tambe, Eric Shieh, Rong Yang
{tambe, eshieh, yangrong}@usc.edu
University of Southern California, USA

Craig Baldwin, Joseph DiRenzo, Ben Maule, Garrett Meyer, Kathryn Moretti
{Craig.W.Baldwin, Joseph.DiRenzo, Ben.J.Maule, Garrett.R.Meyer, Kathryn.A.Moretti}@uscg.mil
United States Coast Guard, USA

In this paper we describe the model, theory developed and deployment of PROTECT, a game-theoretic system in use by the United States Coast Guard (USCG) in the Port of Boston for scheduling patrols. The USCG evaluated the deployment of PROTECT in the Port of Boston as a success and is currently evaluating the system in the Port of New York, with the potential for nationwide deployment.

The PROTECT system is premised on an attacker-defender Stackelberg game model but its development and implementation required both theoretical contributions and detailed evaluations. In this paper we describe the work required in the deployment which we group into five key innovations. First, we propose a compact representation of the defender's strategy space, by exploiting equivalence and dominance, that makes PROTECT efficient enough to solve real world sized problems. Second, this system does not assume that adversaries are perfectly rational, a regular assumption in previous game theoretic models for security. Instead, PROTECT relies on a quantal response (QR) model of the adversary's behavior — to the best of our knowledge, this is the first real-world deployment of a QR model. Third, we develop specialized solution algorithms that are able to solve this problem for real-world instances and give theoretical guarantees. Fourth, our experimental results illustrate that PROTECT's QR model handles real-world uncertainties more robustly than a perfect rationality model. Finally, this paper presents real-world evaluation of PROTECT by: (i) a comparison of human-generated vs PROTECT security schedules, and (ii) results from an Adversarial Perspective Team's (human mock attackers) analysis.

Key words: Game Theory, Security, Applications, Stackelberg Games

Introduction

The United States Coast Guard (USCG) continues to face challenges from potential terrorists within the maritime environment, which includes both the Maritime Global Commons and the ports and waterways that make up the United States Maritime Transportation System. The former Director of National Intelligence, Dennis Blair noted in 2010 a persistent threat "from al-Qa'ida and potentially others who share its anti-Western ideology. A major terrorist attack may emanate from either outside or inside the United States" (Blair 2010, p.8). This threat was reinforced in May of 2011 following the raid on Osama Bin Laden's home, where a large trove of material was uncovered, including plans to attack an oil tanker. "There is an indication of intent, with operatives seeking the size and construction of tankers, and concluding it's best to blow them up from the inside because of the strength of their hulls" (Dozier 2011). These oil tankers transit the US Maritime Transportation System. The USCG plays a key role in the security of this system and the protection of seaports to support the economy, environment, and way of life in the US (Young and Orchard 2011). These threats, coupled with challenging economic times, force USCG to operate as effectively as possible, achieving maximum benefit from every hour spent on patrol.

Recent work has successfully deployed game theory models to help plan patrols for certain real-world security applications. Examples of such work are the deployed systems ARMOR, IRIS and GUARDS. The system ARMOR is used by the Los Angeles International Airport Police (Pita et al. 2008) to decide the location and times of road checkpoints and canine patrols. The software IRIS helps the US Federal Air Marshal Service (Tsai et al. 2009) to schedule air marshals on international flights. Finally, the system GUARDS (Pita et al. 2011) is under evaluation by the US Transportation Security Administration to allocate the resources available for airport protection.

This paper presents a new game-theoretic security application to aid the United States Coast Guard (USCG), called *Port Resilience Operational/Tactical Enforcement to Combat Terrorism* (PROTECT). The USCG's mission includes maritime security of the US coasts, ports, and inland waterways; a security domain that faces increased risks in the context of threats such as terrorism and drug trafficking. The USCG conducts patrols, as part of its Ports, Waterways, and Coastal Security (PWCS) mission, to protect the different critical infrastructure and possible entry locations present at every port that may be targeted by an adversary. Planning PWCS patrols is complicated by the fact that there are limited security resources, which imply that USCG patrols cannot provide 24/7 coverage at any, let alone all of the critical infrastructure, and in addition the adversary has the opportunity to observe the security patrol patterns. To assist the USCG in allocating its patrolling

resources, similar to previous applications (Pita et al. 2008, 2011, Tsai et al. 2009), PROTECT uses an attacker-defender Stackelberg game framework, with USCG as the defender against terrorist adversaries that conduct surveillance before potentially launching an attack. PROTECT's solution is to provide a mixed strategy, i.e., randomized patrol patterns taking into account the importance of different targets, and the adversary's surveillance and anticipated reaction to USCG patrols. Each patrol is a sequence of patrol areas and associated defensive activities at each patrol area, and are constrained by a maximum patrol time. The output of PROTECT is a schedule of patrols that includes when the patrols are to begin, what critical infrastructure to visit for each patrol, and what activities to perform at each critical infrastructure.

While PROTECT builds on previous work on deployed security games, it extends this work and provides five key contributions, which we present in this paper. Three contributions correspond to modeling and algorithmic developments, while the last two have to do with the evaluations of the model and deployed system. First, PROTECT represents the security game efficiently by using a compact formulation of defender strategies through dominance and equivalence analysis. Experimental results show the significant benefits of this compact representation which enable the solution of real-world sized problems.

The second and most important contribution is PROTECT's departure from the assumption of perfectly rational human adversaries (used in previous applications). The assumption of perfect rationality is well-recognized as a limitation of classical game theory, and several research directions have been proposed to address this limitation, giving rise to the field of behavioral game-theory (Camerer 2003). From this literature we borrow the quantal response equilibrium model and adapt it to our security domain. Quantal response models have emerged as a promising approach to model human bounded rationality (Camerer 2003, McKelvey and Palfrey 1995, Wright and Leyton-Brown 2010) including recent results that illustrate the benefits of the quantal response (QR) model in the context of security games (Yang et al. 2011). The PROTECT system is built using a QR model of a human adversary in a Stackelberg game to plan USCG patrols. To the best of our knowledge, this is the first time that a QR model has been used in a real-world security application.

Third, we develop specialized solution algorithms that are able to solve this problem for real-world instances and give theoretical guarantees. This USCG patrolling problem can be approximated arbitrarily by a mixed integer linear programming (MILP) formulation. This approach

allows us to compute efficiently an arbitrary approximation of the global optimal defender strategy with MILPs whose size depend on the accuracy of the approximation.

Regarding evaluation of the deployed system, our contributions are as follows: Fourth, this paper presents a detailed simulation analysis of PROTECT's robustness to uncertainty that may arise in the real-world. Our results show that PROTECT's quantal-response-based approach leads to significantly improved robustness when compared to an approach that assumes full attacker rationality. PROTECT has been in use at the Port of Boston since April 2011 (see Figure 1) and under evaluation by the USCG during this time. This real-world evaluation provides the final key contribution of this paper: we provide actual real-world data providing a head-to-head comparison of human-generated schedules with those generated via a game-theoretic algorithm. We also provide results from an Adversarial Perspective Team's (APT) analysis and comparison of patrols before and after the use of the PROTECT system from a viewpoint of an attacker. Again, to the best of our knowledge, this is the first time that results are given on a real-world evaluation of a game-theoretic security system. Given the success of PROTECT in Boston, PROTECT is currently being tested in the Port of New York, and based on the outcome there, it may potentially be extended to other ports in the US.

(a) PROTECT is being used in Boston

(b) Extending PROTECT to NY

Figure 1 USCG boats are patrolling the ports of Boston and NY

Partial results of this work have appeared in a number of conference papers, including the performance of QR models for stackelberg security games against human adversaries in (Yang et al. 2011), algorithms for computing QR based stackelberg security games in (Yang et al. 2012), and a preliminary description of the PROTECT system (Shieh et al. 2012). In addition to the archival benefit of an article that presents this work as a whole, the current paper expands on the challenges

and lessons learned during the deployment of the PROTECT system and presents more details of the real-world evaluation.

The rest of this paper is organized as follows. We start with a discussion of related work and a detailed description of the concept of Stackelberg equilibrium. Then we discuss how to model the real-world maritime patrolling problem of PWCS patrols as a Stackelberg game and its efficient compact representation. We then present the quantal response based security game developed for the PROTECT system. In particular we present its mathematical formulation and an efficient solution algorithm. Later we describe the details of the implementation of the PROTECT system with the USCG. We also present the empirical and real world evaluation of the PROTECT system, respectively. Finally, we conclude the paper, outline some future research directions and summarize lessons learned from applying the PROTECT system into practice.

Related Work

This section reviews related work on operations research methods for security. We also introduce the concept of Stackelberg games, which is the foundation of our PROTECT system for USCG.

OR Models for Security

There are mainly four lines of related work. The first applies optimization techniques to model the security domain, but does not address the strategic aspects of the problem. These methods provide a randomization strategy for the defender, but they do not take into account the fact that the adversaries can observe the defender's actions and then adjust their behavior. Examples of such approaches include (Paruchuri et al. 2006, Ruan et al. 2005) which are based on learning, Markov Decision Processes (MDPs) and Partially Observable Markov Decision Processes (POMDPs). As part of this work, the authors model the patrolling problem with locations and varying incident rates in each of the locations and solve for optimal routes using an MDP framework. Other examples are algorithms for perimeter patrolling in arbitrary topologies (Basilico et al. 2009), maritime patrols in simulations for deterring pirate attacks (Vanek et al. 2011), and in research looking at the impact of uncertainty in adversarial behavior (Agmon et al. 2009). Another example is the "Hypercube Queueing Model" (Larson 1974) which is based on queueing theory and depicts the detailed spatial operation of urban police departments and emergency medical services. It has found application in police beat design, in allocation of patrolling time, etc. Such frameworks can address many of the problems we raise, including different target values and increasing uncertainty by using many possible patrol routes. However, they fail to account for the possibility that an intelligent

attacker will observe and exploit patterns in the security policy. If a policy is based on the historical frequency of attacks, it is essentially a reactive policy and an intelligent attacker will always be one step ahead.

A second set of work uses Stackelberg games to model a variety of security domains. Bier (2007) gives a strong endorsement of this type of modeling for security problems. Game-theoretic models have been applied in a variety of homeland security settings, such as protecting critical infrastructure (Brown et al. 2006, Nie et al. 2007, Pita et al. 2008). Wein (2008) applies Stackelberg games in the context of screening visitors entering the US. In their work, they model the US Government as the leader who specifies the biometric identification strategy to maximize the detection probability using finger print matches, and the follower is the terrorist who can manipulate the image quality of the finger print. They have also been used for studying missile defense systems (Brown et al. 2005a) and for studying the development of an adversary's weapon systems (Brown et al. 2005b). A family of Stackelberg games known as inspection games is closely related to the security games we are interested in and includes models of arms inspections and border patrols (Avenhaus et al. 2002). Another recent work is on randomized security patrolling using Stackelberg games for generic "police and robbers" scenarios (Basilico et al. 2009) and perimeter patrols (Agmon et al. 2008). Our work differs from this line of work in two main aspects. First, we use a new, more efficient game representation and MILP for modeling and solving the Stackelberg games to enable systems to scale to complex real-world situations. Second, we model the game with defender actions that incorporate the domain constraints (e.g., scheduling constraints) to more accurately model the specific games we are interested in.

The third area of related work is the application of game theoretic techniques that are not based on Stackelberg games to security applications. Security problems are increasingly studied using game-theoretic analysis, ranging from computer network security (Srivastava et al. 2005, wei Lye and Wing 2005) to terrorism (Sandler and Arce 2003). Babu et al. (2006) have worked on modeling passenger security system at US airports using linear programming approaches, however, their objective is to classify the passengers in various groups and then screen them based on the group they belong to. Thus, although game theory has been used in security domains in the past, our work focuses on overcoming the challenges that arise from its application in the real-world.

All the above three lines of related work studying the security domain focus on theoretic analysis of hypothetical scenarios. In contrast, the fourth line of related work focuses on developing tools based on Stackelberg games for use in real-world security operations and addresses many practical

aspects of the problem that only arise in fielded applications. (This paper also belongs to this line of research.) A Stackelberg security game models an interaction between a defender and an attacker in which the defender first commits to a security policy and the attacker conducts surveillance to learn the defender's policy before launching an attack. Software decision aids based on Stackelberg games have been successfully implemented in several real-world domains. ARMOR (Assistant for Randomized Monitoring Over Routes) (Pita et al. 2008), the first application of this game-theoretic framework, was successfully deployed at the Los Angeles International Airport in 2007, and has been in use since. The second application, IRIS (Intelligent Randomization In Scheduling), in use by the US Federal Air Marshal Service since 2009 randomizes deployment of air marshals on US air carriers (Tsai et al. 2009), while a third application, GUARDS (Game-theoretic Unpredictable and Randomly Deployed Security), is being evaluated by the US Transportation Security Administration for national deployment across over 400 US airports (Pita et al. 2011). The main differences between PROTECT and existing applications (i.e., ARMOR (Pita et al. 2008), IRIS (Tsai et al. 2009), GUARDS (Pita et al. 2011)) are as follows. The maritime security domain introduces new modeling changes and new scheduling constraints, which make it more difficult to compute the optimal allocation of resources. For instance, the patrolling actions of the USCG are geographically constrained and each patrol has to be finished within certain amount of time. In addition, there are multiple actions for protecting each target, but with different effectiveness. While existing deployed systems assume fully rational attackers, PROTECT relaxes the unrealistic assumption and provides the first real-world deployment of a quantal response (QR) model of the adversary's behavior. Additionally, this paper for the first time provides real-world data: (i) comparison of human-generated vs PROTECT security schedules, and (ii) results from an Adversarial Perspective Team's (human mock attackers) analysis.

Stackelberg Equilibrium

PROTECT builds on Stackelberg games to reason about the interaction between the USCG and the adversary to provide a randomized security policy. Before introducing how we cast the USCG security scenario to a Stackelberg (security) game in next section, we first explain the Stackelberg equilibrium concept.

A generic Stackelberg game has two players, a *leader*, and a *follower* (Fudenberg and Tirole 1991). A leader commits to a strategy first, and then a follower optimizes his reward, considering the action chosen by the leader (von Stengel and Zamir 2004). The two players in a Stackelberg game need not represent individuals, but could also be groups that cooperate to execute a joint

strategy, such as a police force or a terrorist organization. Each player has a set of possible *pure strategies*, or the actions that they can execute. A *mixed strategy* allows a player to play a probability distribution over pure strategies. Payoffs for each player are defined over all possible pure-strategy outcomes for both the players. The payoff functions are extended to mixed strategies by taking the expectation over pure-strategy outcomes. The follower can observe the leader's strategy, and then act in a way to optimize his own payoffs.

To see the advantage of being the leader in a Stackelberg game, consider the game with the payoff as shown in Table 1. The leader is the row player and the follower is the column player. The only pure-strategy Nash equilibrium for this game is when the leader plays a and the follower plays c which gives the leader a payoff of 3; in fact, for the leader, playing b is strictly dominated.

However, in the simultaneous game if the leader can commit to playing b before the follower chooses his strategy, then the leader will obtain a payoff of 4, since the follower would then play d to ensure a higher payoff for himself. If the leader commits to a mixed strategy of playing a and b with equal (0.5) probability, then the follower will play d, leading to a higher expected payoff for the leader of 4.5. As we can see from this example, the equilibrium strategy in the Stackelberg game can be in fact different from the Nash equilibria.

	c	d
a	3,1	5, 0
b	2,0	4, 2

Table 1 Payoff table for example Stackelberg game.

Stackelberg games are used to model the attacker-defender strategic interaction in security domains and this class of Stackelberg games (with certain restrictions on payoffs (Yin et al. 2010)) is called *Stackelberg security games*. In the Stackelberg security game framework, the security force (defender) is modeled as the leader and the terrorist adversary (attacker) is in the role of the follower. The defender commits to a mixed (randomized) strategy, whereas the attacker conducts surveillance of these mixed strategies and responds with a pure strategy of an attack on a target. Thus, the Stackelberg game framework is a natural approximation of the real-world security scenarios. In contrast, the surveillance activity of the attacker cannot be modeled in the simultaneous move games with the Nash equilibrium solution concept. The objective is to find the optimal mixed strategy for the defender.

The standard solution concept is a strong Stackelberg equilibrium (SSE) (Breton et al. 1988, Leitmann 1978, von Stengel and Zamir 2004). In an SSE, the defender chooses an optimal strategy, accounting for the attacker's best response, under the assumption that the attacker breaks ties in the defender's favor. Strong Stackelberg equilibria are known to exist in all Stackelberg games (Basar and Olsder 1995). A standard argument for the tie-breaking assumption of SSEs suggests that the leader is often able to induce the attacker by selecting a strategy arbitrarily close to the equilibrium which causes the follower to strictly prefer the defender's desired strategy (von Stengel and Zamir 2004). The strong Stackelberg equilibrium solution concept is commonly used in the related literature and is also used in all the deployed systems (Pita et al. 2008, 2011, Tsai et al. 2009, An et al. 2011a,b,c, 2012).

A Security Game Model for USCG Patrols

In this section, we begin by discussing the USCG domain. Next we discuss how to practically cast this real-world maritime patrolling problem of PWCS patrols as a Stackelberg game. We also show how to reduce the number of defender strategies.

Stackelberg games have been well established in the literature (Conitzer and Sandholm 2006, Korzhyk et al. 2011, Fudenberg and Tirole 1991) and PROTECT models the PWCS patrol problem as a Stackelberg game with USCG as the leader (defender) and the terrorist adversaries in the role of the follower. In this Stackelberg game framework, the defender commits to a mixed (randomized) strategy of patrols, which is known to the attacker. This is a reasonable approximation of the practice since the attacker conducts surveillance to learn the mixed strategies that the defender carries out, and responds with a pure strategy of an attack on a target. The optimization objective is to find the optimal mixed strategy for the defender.

To model the USCG patrolling domain as a Stackelberg game, we need to define (i) the set of attacker strategies, (ii) the set of defender strategies, and (iii) the payoff function. These strategies and payoffs center on the targets in a port — ports, such as the Port of Boston, have a significant number of potential targets (critical infrastructure).

Player Strategies

The attacks an attacker can launch on different possible targets are considered as his pure strategies. However, the definition of defender strategies is not as straightforward. Patrols last for some fixed duration during the day as specified by USCG, e.g., 4 hours. We generate defender strategies by grouping nearby targets into patrol areas (in real world scenarios such as the Port of Boston, some

targets are very close to each other and it is thus natural to group targets together according to their geographic locations). The presence of patrol areas led the USCG to redefine the set of defensive activities to be performed on patrol areas to provide a more accurate and expressive model of the patrols. Activities that take a longer time provide the defender a higher payoff compared to activities that take a shorter time to complete. This impacts the final patrol schedule as one patrol may visit fewer areas but conduct longer duration defensive activities at the areas, while another patrol may have more areas with shorter duration activities.

To generate all the permutations of patrol schedules, a graph is created with the patrol areas as vertices and adjacent areas connected via edges. Using the graph of patrol areas, PROTECT generates all possible patrol schedules, each of which is a closed walk in the graph that starts and ends at the patrol area b, the base patrol area for the USCG. Each patrol schedule is a sequence of patrol areas and associated defensive activities at each patrol area, and are constrained by a maximum patrol time τ. (Note that even when the defender just passes by a patrol area, this is treated as an activity.) The defender may visit a patrol area multiple times in a schedule due to geographic constraints and the fact that each patrol is a closed walk. For instance, the defender in each patrol should visit the base patrol area at least twice since she needs to start the patrol from the base and finally come back to the base to finish the patrol.

The graph along with the constraints b and τ are used to generate the defender strategies (patrol schedules). Given each patrol schedule, the total patrol schedule time is calculated (this also includes traversal time between areas, but we ignore it in the following for expository purposes); we then verify that the total time is less than or equal to the maximum patrol time τ. After generating all possible patrol schedules, a game is formed where the set of defender strategies is composed of patrol schedules and the set of attacker strategies is the set of targets. The attacker's strategy was based on targets instead of patrol areas because an attacker will choose to attack a single target.

Table 2 gives an example, where the rows correspond to the defender's strategies and the columns correspond to the attacker's strategies. In this example, there are two possible defensive activities, activity k_1 and k_2, where k_2 provides more effective protection (also takes more time) for the defender than k_1. Suppose that the time bound disallows more than two k_2 activities (given the time required for k_2) within a patrol. Patrol area 1 has two targets (target 1 and 2) while patrol areas 2 and 3 each have one target (target 3 and 4 respectively). In the table, a patrol schedule is composed of a sequence of patrol areas and a defensive activity in each area. The patrol schedules are ordered so that the first patrol area in the schedule denotes which patrol area the defender needs

to visit first. In this example, patrol area 1 is the base patrol area, and all of the patrol schedules begin and end at patrol area 1. For example, the patrol schedule in row 2 first visits patrol area 1 with activity k_2, then travels to patrol area 2 with activity k_1, and returns back to patrol area 1 with activity k_1.

Patrol Schedule	Target 1	Target 2	Target 3	Target 4
(1:k_1), (2:k_1), (1:k_1)	50,-50	30,-30	15,-15	-20,20
(1:k_2), (2:k_1), (1:k_1)	100,-100	60,-60	15,-15	-20,20
(1:k_1), (2:k_1), (1:k_2)	100,-100	60,-60	15,-15	-20,20
(1:k_1), (3:k_1), (2:k_1), (1:k_1)	50,-50	30,-30	15,-15	10,-10
(1:k_1), (2:k_1), (3:k_1), (1:k_1)	50,-50	30,-30	15,-15	10,-10

Table 2 Portion of a simplified example of a game matrix

Payoff Matrix

For the payoffs, if a target i is the attacker's choice and the attack fails, then the defender would gain a reward R_i^d while the attacker would receive a penalty P_i^a, else the defender would receive a penalty P_i^d and the attacker would gain a reward R_i^a. Furthermore, let G_{ij}^d be the payoff for the defender if the defender chooses patrol j and the attacker chooses to attack target i. G_{ij}^d can be represented as a linear combination of the defender reward/penalty on target i and A_{ij}, the effectiveness probability of the defensive activity performed on target i for patrol j, as described by Equation 1. A_{ij} depends on the most effective activity on target i in patrol j. The value of A_{ij} is 0 if target i is not in patrol j. If patrol j only includes one activity in a patrol area that covers target i, then we determine its payoff using the following equation (any additional activity may provide an additional incremental benefit in that area and we discuss this in the following section).

$$G_{ij}^d = A_{ij} R_i^d + (1 - A_{ij}) P_i^d \tag{1}$$

In the USCG problem, rewards and penalties are based on an analysis completed by a contracted company of risk analysts that looked at the targets in the Port of Boston and assigned corresponding values for each one. The types of factors taken into consideration for generating these values include economic damage and injury/loss of life. Meanwhile, the effectiveness probability, A_{ij}, for different defensive activities are decided based on the duration of the activities. Longer activities lead to a higher possibility of capturing the attackers.

While loss of life and property helps in assessing damage in case of a successful attack, assessing payoffs requires that we determine whether the loss is viewed symmetrically by the defender and attacker. Similarly, whether the payoffs are viewed symmetrically for the attacker and defender also holds for the scenario when there is a failed attack. These questions to go the heart of determining whether security games should be modeled as zero-sum games (Tambe 2011). Past work in security games (e.g., ARMOR (Pita et al. 2008), IRIS (Tsai et al. 2009), GUARDS (Pita et al. 2011)) has used non-zero-sum game models, e.g., one assumption made is that the attacker might view publicity of a failed attack as a positive outcome. However, non-zero-sum games require further knowledge acquisition efforts to model the asymmetry in payoffs. For simplicity, as the first step PROTECT starts with the assumption of a zero-sum game. However, the algorithm used in PROTECT is not restricted to zero-sum games and the USCG has proposed to relax this assumption in the future. It is also important to note that while Table 2 shows point estimates of payoffs, we recognize that estimates may not be accurate. To that end, in the experimental results section, we evaluated the robustness of our approach when there is payoff noise, observation noise, and execution error.

Compact Representation

In our game, the number of defender strategies, i.e., patrol schedules, grows combinatorially, generating a scale-up challenge. To achieve scale-up, PROTECT uses a compact representation of the patrol schedules using two ideas: (i) combining equivalent patrol schedules and; (ii) removal of dominated patrol schedules.

With respect to equivalence, different permutations of patrol schedules provide identical payoff results. Furthermore, if an area is visited multiple times with different activities in a schedule, we only consider the activity that provides the defender the highest payoff, not the incremental benefit due to additional activities. This decision is made in consideration of the tradeoff between modeling accuracy and efficiency. On the one hand, the additional value of more activities is small. Currently, the patrol time of each schedule is relatively short (e.g., 1 hour) and the defender may visit a patrol area more than once within the short period and will conduct an activity each time. For instance, the defender may pass by a patrol area 10 minutes after conducting a more effective activity at the same patrol area. The additional value of the pass by activity given the more effective activity is therefore very small. On the other hand, it leads to significant computational benefits which are described in this section if we just consider the most effective activity in each patrol.

Therefore, many patrol schedules are equivalent if the set of patrol areas visited and the most effective defensive activities in each patrol area in the schedules are the same even if their order differs. Such equivalent patrol schedules are combined into a single compact defender strategy, represented as a set of patrol areas and defensive activities (and minus any ordering information). The idea of combining equivalent actions is similar to action abstraction for solving large scale dynamic games (Gilpin 2009). Table 3 presents a compact version of Table 2, which shows how the game matrix is simplified by using equivalence to form compact defender strategies, e.g., the patrol schedules in the rows 2-3 from Table 2 are represented as a compact strategy Γ_2 = $\{(1,k_2)$, $(2,k_1)\}$ in Table 3.

Compact Strategy	Target 1	Target 2	Target 3	Target 4
Γ_1 = $\{(1{:}k_1)$, $(2{:}k_1)\}$	50,-50	30,-30	15,-15	-20,20
Γ_2 = $\{(1{:}k_2)$, $(2{:}k_1)\}$	100,-100	60,-60	15,-15	-20,20
Γ_3 = $\{(1{:}k_1)$, $(2{:}k_1)$, $(3{:}k_1)\}$	50,-50	30,-30	15,-15	10,-10

Table 3 Example compact strategies and game matrix

Next, the idea of dominance is illustrated using Table 3 and noting the difference between Γ_1 and Γ_2 is the defensive activity on patrol area 1. Since activity k_2 gives the defender a higher payoff than k_1, Γ_1 can be removed from the set of defender strategies because Γ_2 covers the same patrol areas while giving a higher payoff for patrol area 1. To generate the set of compact defender strategies, a naive approach would be to first generate the full set of patrol schedules and then prune the dominated and equivalent schedules. Instead, PROTECT generates compact strategies in the following manner: we start with patrols that visit the most patrol areas with the least effective activities within the patrol time limit; these activities take a shorter amount of time but we can cover more areas within the given time limit. Then we gradually consider patrols visiting less patrol areas but with increasingly effective activities. This process will stop when we have considered all patrols in which all patrol areas are covered with the most effective activities and cannot include any additional patrol area.

A QR model of Human Adversary

This section presents the mathematical formulation and solution algorithm to solve the Stackelberg security game with a quantal response model for follower behavior.

Problem Formulation

The Quantal Response model is an important model in behavioral game theory (McKelvey and Palfrey 1995, Rogers et al. 2009). It suggests that instead of strictly maximizing utility, individuals respond stochastically in games: the chance of selecting a non-optimal strategy increases as the cost of such an error decreases. Recent work (Wright and Leyton-Brown 2010) shows that Quantal Level-k (Stahl and Wilson 1994) is best suited for predicting human behavior in simultaneous move games. (We applied the QR model instead of Quantal Level-k because in Stackelberg security games the attacker observes the defender's strategy, so level-k reasoning is not applicable.) We present results later that show that QR based Stackelberg security models perform better than perfectly rational models or prospect theory models in empirical results.

The QR model asumes that humans will choose better actions at a higher frequency, but with noise added to the decision making process. In the case of the Stackelberg security game we consider, only the follower has a quantal response model. This follower selects targets based on the attacker utility of selecting each target. The attacker will choose a target with a higher utility at a higher frequency. The defender aims to maximize the defender's expected utility given that the adversary attacks a target following the QR model. The problem of computing the optimal defender strategy given a QR model of the adversary can be formulated as a non-linear and non-convex optimization problem. The formal formulation of the QR model and the mathematical formulation of the defender's optimization problem `P1` is in Appendix. A.

PASAQ Algorithm for Solving the QR Model

We need to solve `P1` to compute the optimal defender strategy, which requires optimally solving a non-convex problem which is in general an NP-hard problem (Vavasis 1995). There are two principal difficulties with `P1`: it has a fractional objective and the objective is composed of non-linear functions. We present the PASAQ algorithm in parts: we first show that a binary search method can be used to handle the fractional objective function by successively solving non-linear problems and then we present a transformation of these non-linear optimization problems to MILP using piece-wise linear functions.

Binary Search Method The key idea of the binary search method is to iteratively bound the optimal value of the fractional objective function of `P1` by solving related optimization problems `CF-OPT` that do not have a fractional objective. The binary search algorithm first initializes the upper bound U_0 and lower bound L_0 of the optimal objective function value. Then, in each iteration,

we solve a related optimization problem CF−OPT and use the result to increase of lower bound or decrease the upper bound. The search continues until the upper-bound and lower-bound are sufficiently close. Appendix. B shows the details of the binary search method. However, the difficulty is that the objective function in the related optimization problem is still non-convex, therefore, solving it directly is still a hard problem. We propose the Piecewise linear Approximation of optimal Strategy Against Quantal response (PASAQ) algorithm to address this.

PASAQ: Algorithm 1 + Linear Approximation PASAQ is an algorithm to compute the approximate optimal defender strategy. The key idea in PASAQ is to use a piecewise linear function to approximate the nonlinear objective function in CF−OPT, and thus convert it into a Mixed-Integer Linear Programming (MILP) problem. Such a problem can easily include assignment constraints giving an approximate solution for an SSG against a QR-adversary with assignment constraints.

PASAQ first uniformly divides the range $[0, 1]$ into multiple pieces (segments) and the details of the piecewise linear approximation approach is in Appendix. C. This piecewise linear approximation leads to a mixed integer programming problem. Furthermore, given a game instance, the solution quality of PASAQ is bounded linearly by the binary search threshold and the piecewise linear accuracy (Theorem 1 in Appendix. D). Therefore the PASAQ solution can be made arbitrarily close to the optimal solution with sufficiently small ϵ and sufficiently large K.

Implementation of the PROTECT Model in USCG

We now describe in detail the implementation of the PROTECT model in USCG. The end users of PROTECT are security officers, and the system must be simple enough for them to be comfortable using it on a regular basis. In particular, the systems are designed to hide as much of the complexity of the game-theoretic models as possible. PROTECT is a stand-alone desktop Java application. Due to security concerns, PROTECT is run on machines that are not connected to any network. The underlying solution methods use the CPLEX software to solve the necessary mixed-integer programs. The core architecture can be divided into three modules, which we describe in detail in the rest of this section.

1. **Input:** Various parameters and domain knowledge.
2. **Back-end:** Inputs are translated into a game model, which is passed to the Stackelberg game solver and then to a final process that generates a specific sample schedule based on the computed probabilities.
3. **Output:** The final schedule is presented to the user.

User input

We rely on the users and domain experts to provide the knowledge required to specify the game model. The basic inputs that PROTECT requires fall into five categories: (1) the number of available resources (i.e., boats), different possible actions at each location and their capabilities (effectiveness), (2) the set of targets to be protected, (3) payoff values for each target, (4) different types of scheduling constraints (e.g., time constraints and geographic constraints), and (4) supplemental data (e.g., geographic information). The application allows users to save and re-use this information across multiple executions. All the basic inputs are provided by the Coast Guard and risk analysts.

In addition to the above basic inputs, we need to decide the value of parameter λ in the adversary's QR model. Clearly, a λ value of 0 (uniform random) and ∞ (fully rational) are not reasonable. Comparing the solutions obtained for the payoff data for Boston, we observe that an attacker's strategy with $\lambda = 4$ starts approaching a fully rational attacker — the probability of attack focuses on a single target. In addition, an attacker's strategy with $\lambda = 0.5$ is similar to a fully random strategy that uniformly chooses a target to attack. USCG experts (with expertise in terrorist behavior modeling) suggested that we could use a broad range for representing possible λ values used by the attacker. Combining the above observations, it was determined that the attacker's strategy is best modeled with a λ value that is in the range [0.5, 4], rather than a single point estimate. A discrete sampling approach was used to determine a λ value that gives the highest average defender expected utility across attacker strategies within this range to get $\lambda = 1.5$. Specifically, the defender considers different assumptions of the attacker's λ value and for each assumption about the λ value, the defender computes her expected utility against the attacker with different λ values within the range [0.5, 4]. We find that when the defender assumes the attacker is using the QR model with $\lambda = 1.5$, the defender's strategy leads to the highest defender expected utility when the attacker follows the QR model with a λ value uniformly randomly chosen from the range of [0.5, 4]. Selecting an appropriate value for λ remains a complex issue however, and it is a key agenda item for future work.

Back-end: Game generation and solving

Based on all of the data provided by domain experts, we generate a Stackelberg game as described before. However, the definition of defender strategies is not as straightforward. Patrols last for some fixed duration during the day as specified by USCG, e.g., 4 hours. Our first attempt was to model each target as a node in a graph and allow patrol paths to go from each individual target

to (almost all) other targets in the port, generating an almost complete graph on the targets. This method yields the most flexible set of patrol routes that would fit within the maximum duration, covering any permutation of targets within a single patrol. This method unfortunately faced significant challenges: (i) it required determining the travel time for a patrol boat for each pair of targets, a daunting knowledge acquisition task given the hundreds of pairs of targets; (ii) it did not maximize the use of port geography whereby boat crews could observe multiple targets at once and; (iii) it was perceived as micro-managing the activities of the USCG boat crews, which was undesirable. We group nearby targets (according to their geographic locations) into patrol areas and generate defender strategies based on patrol areas.

Once the explicit game model has been generated it is passed as input to the PASAQ algorithm. Note that the PASAQ algorithm will only solve the compactly represented games and the output of the PASAQ algorithm is a probability distribution over compact strategies. Then we generate the probability of full patrol schedules. Figure 2 shows a high level view of the steps of the algorithm using the compact representation. The compact strategies are used instead of full patrol schedules to generate the game matrix. Once the optimal probability distribution is calculated for the compact strategies, the strategies with a probability greater than 0 are expanded to a complete set of patrol schedules.

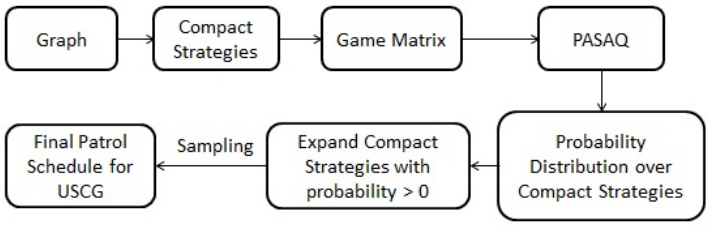


Figure 2 Flow chart of the PROTECT system

In this expansion from a compact strategy to a full set of patrol schedules, we need to determine the probability of choosing each patrol schedule, since a compact strategy may correspond to multiple patrol schedules. The focus here is to increase the difficulty for the attacker to conduct surveillance by increasing unpredictability, which we achieve by randomizing uniformly over all expansions of the compact defender strategies. (Creating optimal Stackelberg defender strategies that increase the attacker's difficulty of surveillance is an open research issue in the literature; here we choose to maximize unpredictability as the first step.) The uniform distribution provides the maximum entropy (greatest unpredictability). Thus, all the patrol schedules generated from a

Day	Hour: 0000 - 2300	Patrol
Day: 1	Hour: 1500	Patrol: [(1:A), (5:C), (6:A), (8:A), (9:B), (8:B), (6:A), (5:A), (1:A)]
Day: 2	Hour: 0300	Patrol: [(1:A), (5:A), (6:A), (8:A), (9:A), (8:A), (6:A), (5:C), (1:A), (2:A), (1:A)]
Day: 3	Hour: 1700	Patrol: [(1:A), (2:C), (4:B), (2:A), (1:B), (2:B), (1:A)]
Day: 4	Hour: 1600	Patrol: [(1:A), (2:B), (4:B), (2:A), (1:B)]
Day: 5	Hour: 1800	Patrol: [(1:A), (5:A), (6:A), (8:A), (9:B), (8:A), (6:A), (5:B), (1:A)]
Day: 6	Hour: 2300	Patrol: [(1:A), (5:A), (6:A), (8:A), (7:C), (5:A), (1:A), (2:A), (4:B), (2:B), (1:A)]
Day: 7	Hour: 0200	Patrol: [(1:A), (2:B), (4:B), (2:A), (1:B)]
Day: 8	Hour: 1400	Patrol: [(1:A), (5:C), (6:A), (8:A), (9:B), (8:A), (6:A), (5:B), (1:A)]
Day: 9	Hour: 0600	Patrol: [(1:A), (5:A), (6:C), (8:B), (9:B), (8:A), (6:A), (5:A), (1:A)]
Day: 10	Hour: 1900	Patrol: [(1:A), (5:C), (6:A), (8:A), (9:B), (8:A), (6:C), (5:B), (1:A)]

Table 4 Sample schedules for 10 days

single compact strategy are assigned a probability of v_i/w_i where v_i is the probability of choosing a compact strategy i and w_i is the total number of expanded patrol schedules for compact strategy i. The complete set of patrol schedules and the associated probabilities are then sampled and provided to the USCG, along with the start time of the patrol generated via uniform random sampling.

Output

From the randomized schedule, we generate a sample schedule for the coast guard. This sample schedule specifies exactly when each boat leaves the base, the sequence of patrol areas to be visited, and the action at each patrol area. The final schedules conform to the domain constraints input by the coast guard. The user can then review the schedule and accept it as is, or add additional constraints and run the scheduling process again. Figure 4 shows one coast guard's sample schedules for 10 days. Each row in the table represents one patrol for one day. The second column corresponds to the starting time of each patrol and the third column corresponds to the detailed patrol schedule which forms a closed walk with different actions at each patrol area. For instance, the patrol on day 1 starts from 3PM. The patrol starts from base patrol area 1, with patrol action A. Then the boat goes to patrol area 5 and conducts patrol action C. After that, the boat goes to patrol area 6 and takes patrol action A. Finally, the boat returns to base patrol area 1.

Experimental Results

This section presents our evaluations on the model and algorithm based on planned experiments. These experiments explore the following questions: the efficiency of a QR based security game when facing human adversaries, the computational difficulty of solving the model with the PASAQ algorithm, sensitivity analysis of the results for the USCG patroll planning problem. The first set of experiments presented consider a simple abstract security game that can be easily explained to human participants. The rest of the experiments, including the payoff values and graph (composed

of 9 patrol areas), were based off the Port of Boston. The solutions for all experiments are run on a machine with an Intel Dual Core 1.4 GHz processor and 2 GB of RAM.

Human Subject Experiment

We conducted empirical tests with human subjects playing an online game to evaluate the performance of different defender strategies. For this experiment we consider a simple security game domain which is easy to explain and requires little information for human participants. In this domain, which corresponds to the problem considered in (Pita et al. 2008), a human attacker selects one of eight targets and the defender decides, without additional side constraints, where to place up to three security resources that perfectly detect an attack (i.e., the corresponding $A_{ij} = 1$ for all targets i that are covered in strategy j).

To test defender strategies human subjects play the role of adversaries. The human subject is able to observe the leader's mixed strategy and the complete payoff matrix. The subject also knows the number of resources the defender can use and with this information selects a target to attack. Subjects are rewarded based on the reward/penalty shown for each target and the probability that a guard was behind the target (i.e., the exact randomized strategy of the defender). To motivate the subjects they could earn or lose money based on whether or not they succeeded in attacking a target; if the subject opened a gate not protected by the guards, they won; otherwise, they lost. Subjects start with $\$8$ and each point won or lost in a game instance was worth $\$0.1$. On average, subjects earned about $\$14.1$ in cash.

We considered seven different payoff matrices: four are representative payoff matrices for a clustering based classification of randomly generated payoff matrices, see (Yang et al. 2011). The other three payoffs are originally from (Pita et al. 2009). For each payoff structure we tested the mixed strategies generated by five algorithms:

1. BRPT: is an MILP formulation for the optimal leader strategy against players whose response follows a Prospect Theory (PT) model (Kahneman and Tvesky 1979).

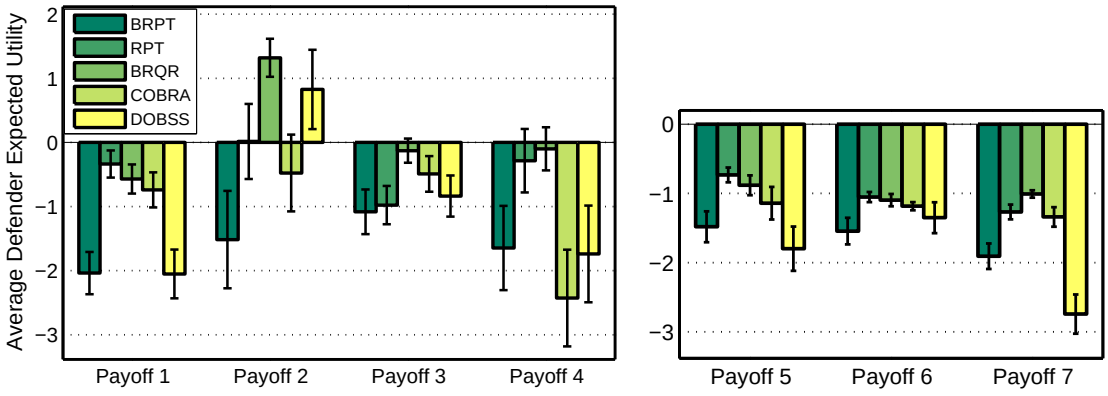
2. RPT: a modified BRPT method that takes into account the uncertainty present in the adversaries' selection, caused (for example) by imprecise computations (Simon 1956).

3. BRQR: the approximate solution of (P1) by solving PASAQ.

4. DOBSS: a Stackelberg security game that assumes perfectly rational adversaries.

5. COBRA: a modified DOBSS that takes into account that the follower might deviate within $\varepsilon > 0$ and assumes humans exhibit an anchoring biases to protect against limited observation conditions.

The specific details of each of these algorithms can be found in the conference articles that introduced them. Specifically BRPT, RPT and BRQR appear in (Yang et al. 2011), DOBSS in (Paruchuri et al. 2008), and COBRA in (Pita et al. 2010).



(a) Payoffs from (Yang et al. 2011) (b) Payoffs from (Pita et al. 2010)

Figure 3 Average expected utility of defender

For each payoff matrix and the five optimal defender strategies, Figure 3 displays the average and standard deviaton of the defender's expected utility with responses coming from 40 different adversaries. The performance of the strategies is closer in payoffs 5∼7 than in payoffs 1∼4. The main reason is that strategies are not very different in payoffs 5∼7 in terms of the Kullback-Leibler divergence. We evaluate the statistical significance of our results using the bootstrap-t method (Wilcox 2003). The comparison is summarized below:

- BRQR outperforms COBRA in all seven payoff structures. The result is statistically significant in three cases ($p<0.005$) and borderline ($p=0.05$) in payoff 3 ($p<0.06$). BRQR also outperforms DOBSS in all cases, with statistical significance in five of them ($p<0.02$).
- RPT outperforms COBRA except in payoff 3. The difference is statistically significant in payoff 4 ($p<0.005$). In payoff 3, COBRA outperforms RPT ($p>0.07$). Meanwhile, RPT outperforms DOBSS in five payoff structures, with statistical significance in four of them ($p<0.05$). In the other two cases, DOBSS has better performance ($p>0.08$).
- BRQR outperforms RPT in three payoff structures with statistical significance ($p<0.005$). They have very similar performance in the other four cases.
- BRPT is outperformed by BRQR in all cases with statistical significance ($p<0.03$). It is also outperformed by RPT in all cases, with statistical significance in five of them ($p<0.02$) and one

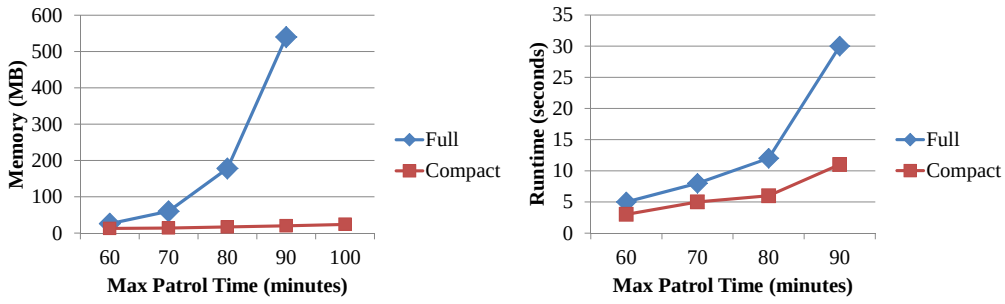
borderline ($p<0.06$). BRPT's failure to perform better (and even worse than COBRA) is a surprising outcome.

In summary, the QR based BRQR algorithm achieved higher defender utilities than algorithms based on the PT model and perfect rationality. Based on this observation from human subject experiments in the security domain, we decide to use the QR based model as the decision making function for the adversaries in the PROTECT system.

Memory and Run-time Analysis

The defender's payoff values have a range of [-10,5] while the attacker's payoff values have a range of [-5,10]. The game was modeled as a zero-sum game in which the attacker's loss or gain is balanced precisely by the defender's gain or loss. For PASAQ, the defender's strategy is computed assuming that the attacker follows the QR model with $\lambda = 1.5$ as justified in the previous section.

This section presents the results based on simulation to show the efficiency in memory and run-time of the compact representation versus the full representation. In Figure 4(a), the x-axis is the maximum patrol time allowed and the y-axis is the memory needed to run PROTECT. In Figure 4(b), the x-axis is the maximum patrol time allowed and the y-axis is the run-time of PROTECT. The maximum patrol time allowed determines the number of combinations of patrol areas that can be visited — so the x-axis indicates a scale-up in the number of defender strategies. When the maximum patrol time is set to 90 minutes, the full representation takes 30 seconds and uses 540 MB of memory while the compact representation takes 11 seconds to run and requires 20 MB of memory. Due to the exponential increase in the memory and run-time that is needed for the full representation, it cannot be scaled up beyond 90 minutes.



(a) Memory comparison (b) Runtime comparison

Figure 4 Comparison of full vs. compact representation

Utility Analysis

It is useful to understand whether PROTECT using PASAQ with $\lambda = 1.5$ provides an advantage when compared to: (i) a uniform random defender's strategy; (ii) a mixed strategy with the assumption of the attacker attacking any target uniformly at random ($\lambda = 0$) or; (iii) a mixed strategy assuming a fully rational attacker ($\lambda = \infty$). The previously existing DOBSS algorithm was used for $\lambda = \infty$ (Paruchuri et al. 2008). Additionally, comparison with the $\lambda = \infty$ approach is important because of the extensive use of this assumption in previous applications (for our zero-sum case, DOBSS is equivalent to minimax but the utility does not change). Typically, we may not have an estimate of the exact value of the attacker's λ value, only a possible range. Therefore, ideally we would wish to show that PROTECT (using $\lambda = 1.5$ in computing the optimal defender strategy) provides an advantage over a range of λ values assumed for the attacker (not just over a point estimate) in his best-response, justifying our use of the PASAQ algorithm. In other words, we are distinguishing between (i) the actual λ value employed by the attacker in best-responding, and (ii) the λ assumed by PASAQ in computing the defender's optimal mixed strategy. The point is to see how sensitive the choice of (ii) is, with respect to prevailing uncertainty about (i).

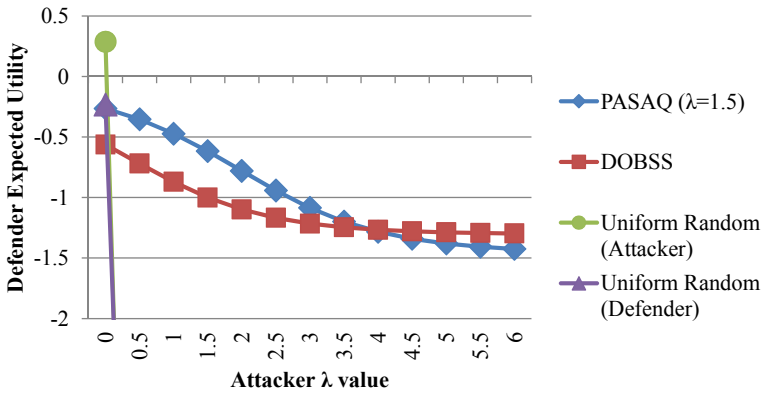


Figure 5 Defender's Expected Utility when varying λ for attacker's strategy

To achieve this, we compute the average defender utility of the four approaches above as the λ value of the attacker's strategy changes from [0, 6], which subsumes the range [0.5, 4] of reasonable attacker strategies. In Figure 5, the y-axis represents the defender's expected utility and the x-axis is the λ value that is used for the attacker's strategy. Both uniform random strategies perform well when the attacker's strategy is based on $\lambda = 0$. However, as λ increases, both strategies quickly drop to a very low defender expected utility. In contrast, the PASAQ strategy with $\lambda = 1.5$ provides a higher expected utility than that assuming a fully rational attacker over a range of attacker λ values (and indeed over the range of interest), not just at $\lambda = 1.5$.

Robustness Analysis

In the real world, observation, execution, and payoffs, are not always perfect due to the following: noise in the attacker's surveillance of the defender's patrols, the many tasks and responsibilities of the USCG where the crew may be pulled off a patrol, and limited knowledge of the attacker's payoff values. Our hypothesis is that PASAQ with $\lambda = 1.5$ is more robust to such noise than a defender strategy which assumes full rationality of the attacker such as DOBSS (Paruchuri et al. 2008), i.e., PASAQ's expected defender utility will not degrade as much as DOBSS over the range of attacker λ of interest. This is illustrated by comparing both PASAQ and DOBSS against observation, execution, and payoff noise (Kiekintveld et al. 2011, Korzhyk et al. 2011, Yin et al. 2011). Intuitively, the QR model is more robust than models assuming perfect rationality since the QR model assumes that the attacker may attack multiple targets with positive probabilities, rather than attacking a single target in the model assuming perfect rationality of the adversaries. Such intuition has been verified in other contexts (Rogers et al. 2009). (A comparison of the uniform random strategies was not included due to its poor performance shown in Figure 5.) All experiments were run generating 200 samples with added noise and averaging over all the samples. For Figures 6, 7, and 8, the y-axis represents the defender's expected utility and the x-axis is the attacker's λ value.

The first experiment considers observational noise, which means that the attacker has noise associated with observing the defender's patrol strategy as shown in Figure 6. In this scenario, if the defender covered a target with probability p, the attacker may perceive the probability to be in $[p - \omega, p + \omega]$ where ω is the noise. The low observation error corresponds to $\omega = 0.1$ while for high error $\omega = 0.2$. Contrary to expectation, observation error leads to an increase in defender expected utility in PASAQ, but a potential decrease (or no change) in DOBSS — thus PASAQ ends up dominating DOBSS by a larger margin over bigger ranges of λ, further consolidating the reason to use PASAQ rather than a full-rationality model.

An example illustrates PASAQ's unexpected behavior. Suppose there are two targets 1 and 2 and the defender's marginal coverage on the two targets is x. Given x, defender expected utilities for targets 1 and 2 are $U_1^d(x) = -2$ and $U_2^d(x) = -1$, with the attacker's expected utility $U^a(x)$ being the opposite because this is a zero-sum game. For an attacker strategy with a higher λ, the adversary will choose to attack target 1 and the defender would get a utility of -2. When observation noise is added, increases in the coverage of target 1 results in decreases in $U_1^a(x')$ so the attacker might choose to attack target 2 instead, giving the defender a higher utility than when noise is absent. If

the coverage of target 1 decreases, $U_1^a(x')$ will increase and the attacker will still choose to attack target 1, but $U_1^d(x)$ will remain the same as when there was no noise.

The reason there is a different trend for DOBSS is because DOBSS minimizes the maximum attacker's expected utility or, in our situation, also maximizes the minimum defender's expected utility. This results in multiple targets with the same minimum defender's utility; these targets are referred to as an *attack set* (Jain et al. 2010). Typically, when the coverage over the attack set varies due to observation error, some of the targets have less and some have more coverage, but the attacker ends up attacking the targets in the attack set regardless, giving the defender almost no change in its expected utility.

For the second experiment, noise is added to the execution phase of the defender. These results, presented in Figure 7 show the performance of different strategies while considering execution noise. The y-axis represents the defender's expected utility and the x-axis is the attacker's λ value. If the defender covered a target with probability p, this probability now changes to be in $[p-\omega, p+\omega]$ where ω is the noise. The low execution error corresponds to $\omega = 0.1$ whereas high error corresponds to $\omega = 0.2$. In the experiments, the attacker best-responds to the mixed strategy with added noise. The key takeaway here is that execution error leads to PASAQ dominating DOBSS over all tested values of λ, further strengthening the reason to use PASAQ rather than a full-rationality model. For both algorithms, the defender's expected utility decreases as more execution error is added because the defender's strategy is impacted by the additional error. When execution error is added, PASAQ dominates DOBSS because the latter seeks to maximize the minimum defender's expected utility so multiple targets will have the same minimum defender utility. For DOBSS, when execution error is added, there is a greater probability that one of these targets will have less coverage, resulting in a lower defender's expected utility. For PASAQ, typically only one target has the minimum defender expected utility. As a result changes in coverage do not impact it as much as DOBSS. As execution error increases, the advantage in the defender's expected utility of PASAQ over DOBSS increases even more.

In the third experiment shown in Figure 8, payoff noise is added by aggregating mean-0 Gaussian noise to the attacker's original payoff values (similar to (Kiekintveld et al. 2011)). As more noise is added to the payoffs, both defenders' strategies result in an increase in the defender's expected utility because the game is no longer zero-sum. The low payoff noise corresponds to a standard deviation of 1 while a high payoff noise corresponds to a standard deviation of 1.5. Similar to the previous experiments, when payoff noise is added, DOBSS is dominated by PASAQ, indicating

the robustness of PASAQ. As noise is added to the attacker's payoff but not the defender's payoff, the attacker's strategy may no longer result in the lowest possible defender expected utility. For example, with no payoff noise, target 1 gives the attacker the highest utility and the defender the lowest utility. When noise is added to the attacker's payoffs, target 1 may no longer give the attacker the highest utility; instead, he will choose to attack target 2, and the defender receives a higher utility than target 1. In essence, with a zero-sum game, the defender has planned a conservative strategy, based on maximin, and as such any change in the attacker is to the defender's benefit in this case.

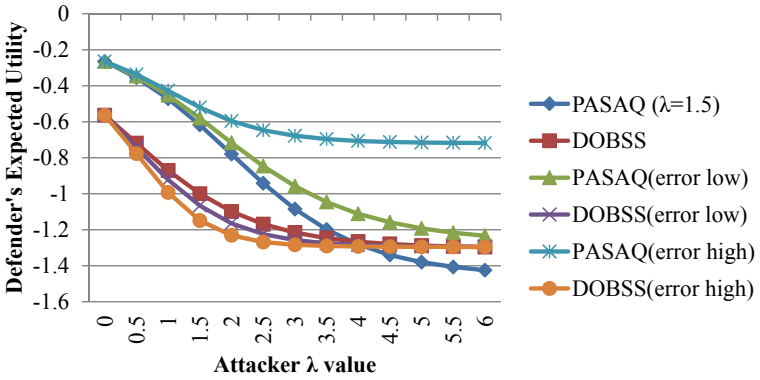


Figure 6 Defender's expected utility in consideration of observation noise

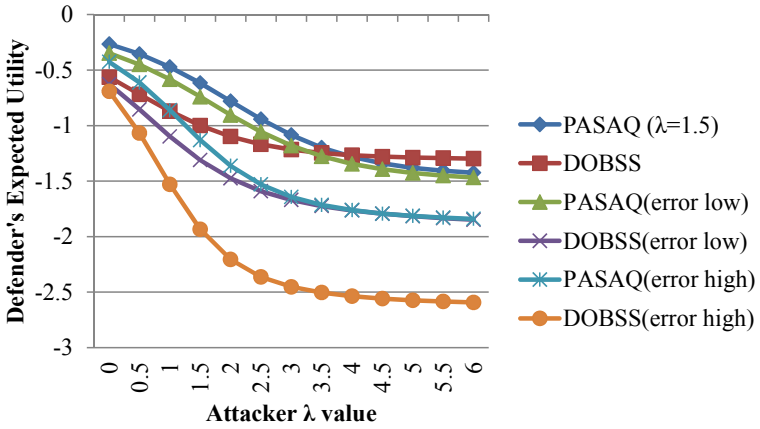


Figure 7 Defender's expected utility in consideration of execution noise

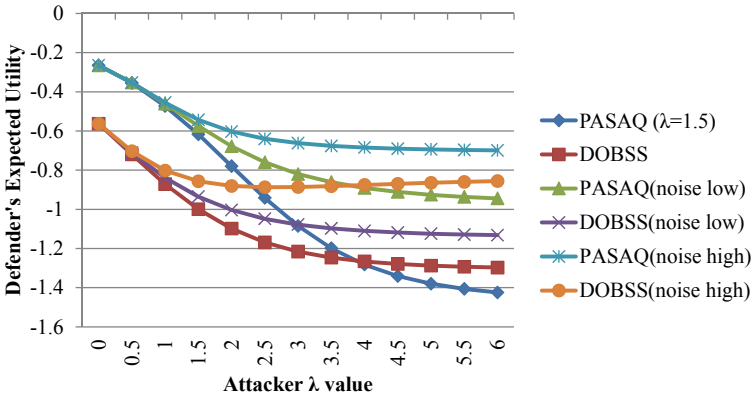


Figure 8 Defender's expected utility in consideration of payoff noise

USCG Real-World Evaluation

In addition to the results obtained from our planned experiments, the USCG conducted its own real-world evaluation of PROTECT. With permission, some aspects of the evaluation are presented in this paper.

Real-world scheduling data: Unlike prior publications of real-world applications of game theory for security, a key novelty of this paper is the inclusion of actual data from USCG patrols before and after the deployment of PROTECT at the Port of Boston. Figure 9 and Figure 10 show the frequency of visits by USCG to different patrol areas over a number of weeks. Figure 9 shows pre-PROTECT patrol visits per day by area and Figure 10 shows post-PROTECT patrol visits per day by area. The x-axis is the day of the week, and the y-axis is the number of times a patrol area is visited for a given day of the week. The y-axis is intentionally blurred for security reasons as this is real data from Boston. There are more lines in Figure 9 than in Figure 10 because during the implementation of PROTECT, new patrol areas were formed which contained more targets and thus fewer patrol areas in the post-PROTECT figure. Figure 9 depicts a definite pattern in the patrols. While there is a spike in patrols executed on Day 5, there is a dearth of patrols on Day 2. Besides this pattern, the lines in Figure 9 intersect, indicating that some days, a higher value target was visited more often while on other days it was visited less often. This means that there was not a consistently high frequency of coverage of higher value targets before PROTECT.

In Figure 10, we notice that the pattern of low patrols on Day 2 (from Figure 9) disappears. Furthermore, lines do not frequently intersect, i.e., higher valued targets are visited consistently across the week. The top line in Figure 10 is the base patrol area and is visited at a higher rate than all other patrol areas.

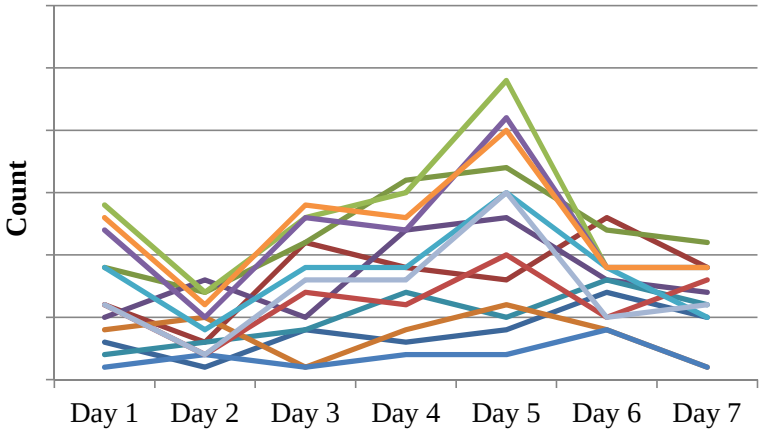


Figure 9 Patrol visits per day by area - pre-PROTECT, one line per area

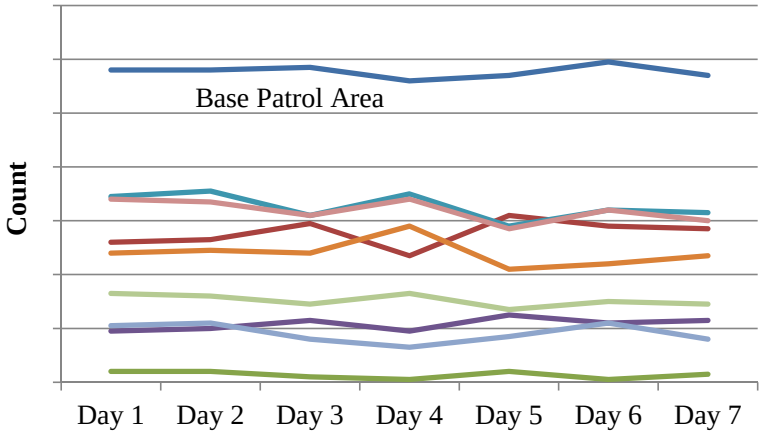


Figure 10 Patrol visits per day by area - post-PROTECT, one line per area

Adversary Perspective Teams(APT): To obtain a better understanding of how the adversary views the potential targets in the port, the USCG created the Adversarial Perspective Team (APT), a mock attacker team. The APT provides assessments from the terrorist perspective and as a secondary function, assesses the effectiveness of the patrol activities before and after deployment of PROTECT. In their evaluation, the APT incorporates the adversary's known intent, capabilities, skills, commitment, resources, and cultural influences. In addition, it screens attack possibilities and assists in identifying the level of deterrence projected at and perceived by the adversary. For the purposes of this research, the adversary is defined as an individual(s) with ties to al-Qa'ida or its affiliates.

The APT conducted a pre- and post-PROTECT assessment of the system's impact on an adversary's deterrence at the Port of Boston. This analysis uncovered a positive trend where the effectiveness of deterrence increased from the pre- to post- PROTECT observations.

Additional Real-world Indicators: The use of PROTECT and APT's improved guidance given to boat crews on how to conduct the patrol jointly provided a noticeable increase in the quality and effectiveness of the patrols. Prior to implementing PROTECT, there were no documented reports of illicit activity. After implementation, USCG crews, reported more illicit activities within the port (therefore justify the effectiveness of the PROTECT system) and provided a noticeable "on the water" presence with industry port partners commenting, "the Coast Guard seems to be everywhere, all the time." With no actual increase in the number of resources applied, and therefore no increase in capital or operating costs, these outcomes support the practical application of game theory in the maritime security environment.

Conclusions and Lessons Learned

This paper reports on PROTECT, a game-theoretic system deployed by the USCG in the Port of Boston since April 2011 for scheduling their PWCS patrols. USCG has deemed the deployment of PROTECT in Boston a success and efforts are underway to deploy PROTECT in the Port of New York, and to other ports in the United States. PROTECT uses an attacker-defender Stackelberg game model, and includes five key innovations.

First, to improve PROTECT's efficiency, we generate a novel compact representation of the defender's strategy space, exploiting equivalence and dominance. Second, PROTECT moves away from the assumption of perfect adversary rationality seen in previous work, relying instead on a quantal response (QR) model of the adversary's behavior. While the QR model has been extensively studied in the realm of behavioral game theory, to the best of our knowledge, this is its first real-world deployment. Third, this work presents a new algorithm, PASAQ, to overcome the difficulties in computing the best defender strategy assuming a QR adversary, including solving a nonlinear and non-convex optimization problem and handling constraints on assigning security resources in designing defender strategies. PASAQ provides efficient computation of the defender strategy with nearly-optimal solution quality. Fourth, we provide experimental results illustrating that PROTECT's QR model of the adversary is better able to handle real-world uncertainties than a perfect rationality model. Finally, for the first time in a security application evaluation, we use real-world data: (i) providing a comparison of human-generated security schedules versus those generated via a game-theoretic algorithm and; (ii) results from an APT's analysis of the impact of the PROTECT system. As a result, PROTECT has advanced the state of the art beyond previous applications of game theory for security. Building on this initial success of PROTECT, we hope to

deploy it at more and much larger-sized ports. In so doing, in the future, we will consider significantly more complex attacker strategies, including potential real-time surveillance and coordinated attacks. We will also consider 1) more complex defender strategies due to larger ports/water ways and heterogeneous defense resources, 2) better models of human behavior in security games, and 3) better algorithms.

Developing the PROTECT system was a collaborative effort involving university researchers and USCG personnel representing decision makers, planners and operators. Building on the lessons reported in (Pita et al. 2011) for working with security organizations, we informed the USCG of (i) the assumptions underlying the game-theoretic approaches, e.g., full adversary rationality, and strengths and limitations of different algorithms — rather than pre-selecting a simple heuristic approach; (ii) the need to define and collect correct inputs for model development and; (iii) a fundamental understanding of how the inputs affect the results. As a result of this project we gained three new insights involving real-world applied research:

(i) Unforeseen positive benefits because security agencies were compelled to reexamine their assumptions. During the course of the project, the USCG was compelled to reassess their operational assumptions as a result of working through the research problem. A positive result of this reexamination prompted USCG to develop new PWCS mission tactics, techniques and procedures. Through the iterative development process, USCG reassessed the reasons why boat crews performed certain activities and whether they were sufficient. For example, instead of "covered" vs "not covered" as the only two possibilities at a patrol point, there are now multiple sets of activities at each patrol point.

(ii) Requirement to work with multiple teams in a security organization at multiple levels of their hierarchy. Applied research requires the research team to collaborate with planners and operators on the multiple levels of a security organization to ensure the model accounts for all aspects of a complex real world environment. Initially when we started working on PROTECT, the focus was on patrolling each individual target. This appeared to micromanage the activities of boat crews, and it was through their input that individual targets were grouped into patrol areas associated with a PWCS patrol. On the other hand, input from USCG headquarters and the APT mentioned earlier, led to other changes in PROTECT, e.g., departing from a fully rational model of an adversary to a QR model.

(iii) The need to prepare answers to end-user practical questions not always directly related to the "meaty" research problems. One example of the need to explain results involved the user

citing that one patrol area was being repeated and hence, randomization did not seem to occur. After assessing this concern, we determined that the cause for the repeated visits to a patrol area was its high reward — order of magnitude greater than the rarely visited patrol areas. PROTECT correctly assigned patrol schedules that covered the more "important" patrol areas more frequently. In another example, the user noted that PROTECT did not assign any patrols to start at 4:00 AM or 4:00 PM over a 60 day test period. They expected patrols would be scheduled to start at any hour of the day, leading them to ask if there was a problem with the program. This required us to develop a layman's briefing on probabilities, randomness, and sampling. With 60 patrol schedules, a few start hours may not be chosen given our uniform random sampling of the start time. These practitioner-based issues demonstrate the need for researchers to not only be conversant in the algorithms and math behind the research, but also be able to explain from a user's perspective how solutions are accurate. An inability to address these issues would result in a lack of real-world user confidence in the model.

Acknowledgments

We thank the USCG offices, and particularly sector Boston, for their exceptional collaboration. The views expressed herein are those of the author(s) and are not to be construed as official or reflecting the views of the Commandant or of the United States Coast Guard. This research was supported by the United States Department of Homeland Security through the National Center for Risk and Economic Analysis of Terrorism Events (CREATE) under award number 2010-ST-061-RE0001.

Appendix. A. The Quantal Response Model Formulation

The QR model asumes that humans will choose better actions at a higher frequency, but with noise added to the decision making process. The model assumes that a player will select action i with probability q_i given by

$$q_i = \frac{e^{\lambda G_i^a(x_i)}}{\sum_{j=1}^{T} e^{\lambda G_j^a(x_j)}}. \tag{2}$$

Here, the parameter $\lambda \in [0, \infty)$ represents the amount of noise in the attacker's strategy. When $\lambda = 0$ the probabilities q_i correspond to a uniform probability distribution over the set of possible actions. As the value λ increases the probabilities q_i become closer to a pure strategy indicating the action with largest value of $G_j^a(x_j)$. We will show our approach for determining the appropriate λ value in the Implementation section.

In the case of the Stackelberg security game we consider, only the follower has a quantal response model. This follower selects targets corresponding to the probability q_i, where $G_j^a(x_j)$ now corresponds to the attacker utility of selecting target j. Using the notation described in the previous section and letting R_j^a and P_j^a (or R_j^d and P_j^d) correspond to the reward or penalty to the adversary (or defender) of selecting target j, respectively, we have

$$\begin{aligned} G_j^a(x_j) &= x_j P_j^a + (1 - x_j) R_j^a = R_j^a - x_j (R_j^a - P_j^a) \\ G_j^d(x_j) &= x_j R_j^a + (1 - x_j) P_j^a = P_j^a + x_j (R_j^a - P_j^a) \, . \end{aligned}$$

The defender aims to maximize the defender's expected utility given that the adversary attacks target i with probability q_i. Given T targets, the coverage vector x, the defender's expected utility against a QR-adversary is:

$$U^d(x) = \sum_{i=1}^{T} q_i(x) U_i^d(x_i) = \sum_{i=1}^{T} q_i(x)(x_i R_i^d + (1-x_i)P_i^d)$$

Therefore, given J compact strategies and effectiveness matrix A_{ij}, the problem of computing the optimal defender strategy given a QR model of the adversary can be formulated as the following non-linear and non-convex optimization problem `P1`:

$$\texttt{P1:} \begin{cases} \max\limits_{x,a} \dfrac{\sum_{i=1}^{T} e^{\lambda R_i^a} e^{-\lambda(R_i^a - P_i^a)x_i}((R_i^d - P_i^d)x_i + P_i^d)}{\sum_{i=1}^{T} e^{\lambda R_i^a} e^{-\lambda(R_i^a - P_i^a)x_i}} \\ x_i = \sum\limits_{j=1}^{J} a_j A_{ij}, \quad \forall i \\ \sum\limits_{j=1}^{J} a_j = 1 \\ 0 \le a_j \le 1, \quad \forall j \end{cases}$$

Recall that A_{ij} corresponds to the effectiveness on target i of the joint compact strategy j. For ease of reference we summarize this notation in Table 5.

T	Number of Targets
J	Total number of compact strategies
R_i^d	Defender reward on covering target i if it's attacked
P_i^d	Defender penalty on not covering target i if it's attack
R_i^a	Attacker reward on attacking target i if it's not covered
P_i^a	Attacker penalty on attacking target i if it's covered
λ	Noise parameter in quantal response model
A_{ij}	Effectiveness probability of compact strategy j on target i
a_j	Probability of choosing compact strategy j
x_i	Marginal coverage on target i

Table 5 PASAQ notation as applied to PROTECT

The objective function of the problem corresponds to the computation of the defender's expected utility resulting from a combination of Equations 1 and 2. Unlike previous application (Jain et al. 2010, Kiekintveld et al. 2011, Paruchuri et al. 2008), x_i in this case not just summarizes presence or absence on a target, but also the effectiveness probability A_{ij} on the target as well. That is, the second line computes the marginal coverage on the targets based on the effectiveness factor A_{ij} and the probability of choosing compact strategy j, denoted as a_j.

Appendix. B. Binary Search Method

For simplicity of the notation, let's define $\theta_i := e^{\lambda R_i^a} > 0$, $\beta_i := \lambda(R_i^a - P_i^a) > 0$, and $\alpha_i := R_i^d - P_i^d > 0$. Using this notation we can express the objective function of P1 as $N(x)/D(x)$, where

- $N(x) = \sum_{i=1}^{T} \theta_i \alpha_i x_i e^{-\beta_i x_i} + \sum_{i=1}^{T} \theta_i P_i^d e^{-\beta_i x_i}$
- $D(x) = \sum_{i=1}^{T} \theta_i e^{-\beta_i x_i} > 0$.

Let also $\mathscr{X}_f$ be the feasible region of P1 and p^* the optimal value. Therefore P1 can be written as: $p^* = \max_{(x,a)\in\mathscr{X}_f} \frac{N(x)}{D(x)}$. A binary search method can be used to handle the fractional objective of this type of problem.

The key idea of the binary search method is to iteratively bound the optimal value (p^*) of the fractional objective function of P1 by solving related optimization problems that do not have a fractional objective. Given a real value r, we define the optimization problem:

$$\texttt{CF-OPT:} \qquad \delta_r^* = \min_{x\in\mathscr{X}_f} \; rD(x) - N(x)$$

The following result shows that $r \le p^*$ is equivalent to $\delta_r^* \le 0$.

LEMMA 1. *Let $N(x)$, $D(x)$ be continuous functions defined on a closed bounded set $\mathscr{X}_f$. Let $D(x) > 0\ \forall x \in \mathscr{X}_f$. If $p^* = \max_{x\in\mathscr{X}_f} \frac{N(x)}{D(x)}$, $r \in \mathscr{R}$, and δ_r^* as defined in CF-OPT, then $r \le p^* \iff \delta_r^* \le 0$.*

Proof: '$\Rightarrow$': Since p^* is the optimal solution value of a continuous objective over a closed bounded set, then there exists an optimal solution x^* such that $p^* = \frac{N(x^*)}{D(x^*)} \ge r$. By rearranging $p^* = \frac{N(x^*)}{D(x^*)} \ge r$, we can get the result.
'$\Leftarrow$': Similarly, there exists $\bar{x}$ such that $\delta_r^* = rD(\bar{x}) - N(\bar{x}) \le 0$, which means that $r \le \frac{N(\bar{x})}{D(\bar{x})} \le p^*$; $\square$

Therefore, by solving this related optimization problem and checking if $\delta_r^* \le 0$, we can answer if a given r is larger or smaller than the global maximum. Algorithm 1 presents the binary search method for problem P1 using as inputs the tolerance ϵ, the payoff matrix (P_M) and the total number of security resources ($numRes$).

Algorithm 1 Binary Search

Input: ϵ, P_M and $numRes$
$(U_0, L_0) \leftarrow \texttt{EstimateBounds}(P_M, numRes)$
$(U, L) \leftarrow (U_0, L_0)$
while $U - L \ge \epsilon$ **do**
  $r \leftarrow \frac{U+L}{2}$
  Solve CF-OPT, let x^r, δ_r^* be the optimal solution and optimal solution value
  if $\delta_r^* \le 0$ **then**
    $L \leftarrow r$
  else
    $U \leftarrow r$
  end if
end while

Algorithm 1 first initializes the upper bound (U_0) and lower bound (L_0) of the optimal objective function value on Line 2 (`EstimateBounds`$(P_M, numRes)$). Then, in each iteration, r is set to be the mean of U and L. Line 6 solves `CF-OPT` to checks whether the current $r \leq p^*$ or not. If so the lower-bound of the binary search needs to be increased and this process also returns a valid strategy x^r. Otherwise, $p^* < r$ and the upper-bound of the binary search should be decreased. The search continues until the upper-bound and lower-bound are sufficiently close, i.e., $U - L < \epsilon$. The number of iterations in Algorithm 1 is bounded by $O(\log(\frac{U_0 - L_0}{\epsilon}))$. Specifically for SSGs we can estimate the upper and lower bounds (corresponding to `EstimateBounds`$(P_M, numRes)$ on Line 2) as follows:

Lower bound: Let s_u be any feasible defender strategy. The defender utility based on using s_u against a adversary's quantal response is a lower bound of the optimal solution of `P1`. A simple example of s_u is the uniform strategy.

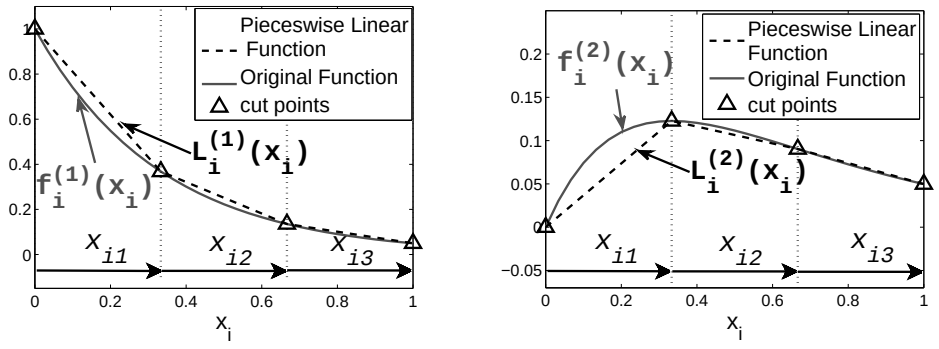
Upper bound: Since $P_i^d \leq U_i^d \leq R_i^d$ we have $U_i^d \leq \max_{i=1}^T R_i^d$. The defender's utility is computed as $\sum_{i=1}^T q_i U_i^d$, where U_i^d is the defender utility on target i and q_i is the probability that the adversary attacks target i. Thus, the maximum R_i^d serves as an upper bound of U_i^d.

Appendix. C. Piecewise Linear Approximation

In order to demonstrate the piecewise approximation in PASAQ, we first rewrite the nonlinear objective function of `CF-OPT` as:

$$\sum_{i=1}^{T} \theta_i (r - P_i^d) e^{-\beta_i x_i} - \sum_{i=1}^{T} \theta_i \alpha_i x_i e^{-\beta_i x_i}$$

The goal is to approximate the two nonlinear function $f_i^{(1)}(x_i) = e^{-\beta_i x_i}$ and $f_i^{(2)}(x_i) = x_i e^{-\beta_i x_i}$ as two piecewise linear functions in the range $x_i \in [0, 1]$, for each $1 \leq i \leq T$.



(a) approximation of $f_i^{(1)}(x_i)$ (b) approximation of $f_i^{(2)}(x_i)$

Figure 11 Piecewise linear approximation

We first uniformly divide the range $[0, 1]$ into K pieces (segments). Simultaneously, we introduce a set of new variables $\{x_{ik}, k = 1..K\}$ to represent the portion of x_i in each of the K pieces, $\{[\frac{k-1}{K}, \frac{k}{K}], k = 1..K\}$. Therefore, $x_{ik} \in [0, \frac{1}{K}], \forall k = 1..K$ and $x_i = \sum_{k=1}^K x_{ik}$. In order to ensure that $\{x_{ik}\}$ is a valid partition of x_i, all x_{ik} must satisfy: $x_{ik} > 0$ only if $x_{ik'} = \frac{1}{K}, \forall k' < k$. In other words, x_{ik} can be non-zero only when all the previous pieces are completely filled. Figures 11(a) and 11(b) display two examples of such a partition.

Thus, we can represent the two nonlinear functions as piecewise linear functions using $\{x_{ik}\}$. Let $\{(\frac{k}{K}, f_i^{(1)}(\frac{k}{K})), k = 0..K\}$ be the $K+1$ cut-points of the linear segments of function $f_i^{(1)}(x_i)$, and $\{\gamma_{ik}, k = 1..K\}$ be the slopes of each of the linear segments. Starting from $f_i^{(1)}(0)$, the piecewise linear approximation of $f_i^{(1)}(x_i)$, denoted as $L_i^{(1)}(x_i)$:

$$L_i^{(1)}(x_i) = f_i^{(1)}(0) + \sum_{k=1}^{K} \gamma_{ik} x_{ik} = 1 + \sum_{k=1}^{K} \gamma_{ik} x_{ik}$$

Similarly, we can obtain the piecewise linear approximation of $f_i^{(2)}(x_i)$, denoted as $L_i^{(2)}(x_i)$:

$$L_i^{(2)}(x_i) = f_i^{(2)}(0) + \sum_{k=1}^{K} \mu_{ik} x_{ik} = \sum_{k=1}^{K} \mu_{ik} x_{ik}$$

where, $\{\mu_{ik}, k = 1..K\}$ is the slope of each linear segment. and $\{(\frac{K}{K}, f_i^{(2)}(\frac{K}{K})), k = 0, \ldots, L\}$ be the $L+1$ endpoints of linear segments of function $f_i^{(2)}(x_i)$. In order to represent the objective function of `CF-OPT` as a piecewise linear function, we divide each variable x_i into L parts $\{x_{ik}, k = 1, \ldots, L\}$, with each part related to the l^{th} linear segments of the function. Therefore, we could write the objective function of `CF-OPT` as

$$\sum_{i}^{T} \theta_i (k - P_i^d)(1 + \sum_{k=1}^{K} \gamma_{ik} x_{ik}) - \sum_{i}^{T} \theta_i \alpha_i \sum_{k=1}^{K} \mu_{ik} x_{ik}$$

where, γ_{ik} represent the slope of the l^{th} line segment of function $e^{-\beta_i x_i}$ and μ_{ik} represent the slope of l^{th} linear segment of function $x_i e^{-\beta_i x_i}$.

PASAQ consists of Algorithm 1, but with `CF-OPT` rewritten as follows:

$$\min_{x,z,a} \sum_{i=1}^{T} \theta_i (r - P_i^d)(1 + \sum_{k=1}^{K} \gamma_{ik} x_{ik}) - \sum_{i=1}^{T} \theta_i \alpha_i \sum_{k=1}^{K} \mu_{ik} x_{ik}$$

$$\text{s.t. } 0 \le x_{ik} \le \frac{1}{K}, \quad \forall i, \quad k = 1 \ldots K \tag{3}$$

$$z_{ik} \frac{1}{K} \le x_{ik}, \quad \forall i, \quad k = 1 \ldots K-1 \tag{4}$$

$$x_{i(k+1)} \le z_{ik}, \quad \forall i, \quad k = 1 \ldots K-1 \tag{5}$$

$$z_{ik} \in \{0, 1\}, \quad \forall i, \quad k = 1 \ldots K-1 \tag{6}$$

$$\sum_{k=1}^{K} x_{ik} = \sum_{A_j \in \mathcal{A}} a_j A_{ij}, \quad \forall i \tag{7}$$

$$\sum_{A_j \in \mathcal{A}} a_j = 1 \tag{8}$$

$$0 \le a_j \le 1, \quad \forall j \tag{9}$$

Let's refer to the above MILP formulation as `PASAQ-MILP`.

The solution provided by PASAQ is in the feasible region of `P1` as shown in Lemma 2.

LEMMA 2. *The feasible region for* $x = \langle x_i = \sum_{k=1}^{K} x_{ik}, 1 \le i \le T \rangle$ *of* `PASAQ-MILP` *is equivalent to that of* `P1`.

JUSTIFICATION. The auxiliary integer variable z_{ik} indicates whether or not $x_{ik} = \frac{1}{K}$. Equation (4) enforces that $z_{ik} = 0$ only when $x_{ik} < \frac{1}{K}$. Simultaneously, Equation (5) enforces that $x_{i(k+1)}$ is positive only if $z_{ik} = 1$. Hence,$\{x_{ik}, k = 1..K\}$ is a valid partition of x_i and $x_i = \sum_{k=1}^{K} x_{ik}$ and that $x_i \in [0, 1]$. Thus, the feasible region of `PASAQ-MILP` is equivalent to `P1`.

$\underline{\theta} := \min_{i=1}^{T} \theta_i$	$\overline{R^d} := \max_{i=1}^{T} \lvert R_i^d \rvert$	$\overline{\beta} := \max_{i=1}^{T} \beta_i$
$\overline{\theta} := \max_{i=1}^{T} \theta_i$	$\overline{P^d} := \max_{i=1}^{T} \lvert P_i^d \rvert$	$\overline{\alpha} := \max_{i=1}^{T} \alpha_i$

Table 6 Notations for error bound proof

However, PASAQ approximates the minimum value of `CF-OPT` by using `PASAQ-MILP`, and furthermore solves `P1` approximately using binary search. Hence, we need to show an error bound on the solution quality of PASAQ.

We define two constants which are decided by the game payoffs: $C_1 = (\overline{\theta}/\underline{\theta})e^{\overline{\beta}}\{(\overline{R^d} + \overline{P^d})\overline{\beta} + \overline{\alpha}\}$ and $C_2 = 1 + (\overline{\theta}/\underline{\theta})e^{\overline{\beta}}$ (the notation used is defined in Table 6). In the following, we are interested in obtaining a bound on the difference between p^* (the global optimal obtained from `P1`) and $Obj_{P1}(\tilde{x}^*)$, where $\tilde{x}^*$ is the strategy obtained from PASAQ.

THEOREM 1. *Let $\tilde{x}^*$ be the defender strategy computed by* PASAQ*, p^* is the global optimal defender expected utility,*

$$0 \leq p^* - Obj_{P1}(\tilde{x}^*) \leq 2C_1 \frac{1}{K} + (C_2 + 1)\epsilon$$

The proof of Theorem 1 is in Appendix. D.

Appendix. D. Proof of Theorem 1

For simplicity, let's first define the following notations:

- $F^{(r)}(x)$, the objective function of the `CF-OPT` problem associated with a given estimation value r:

$$F^{(r)}(x) = \sum_{i=1}^{T} \theta_i (r - P_i^d) e^{-\beta_i x_i} - \sum_{i=1}^{T} \theta_i \alpha_i x_i e^{-\beta_i x_i}$$

- $\nu^{(r)} = arg\min_x F^{(r)}(x)$
- $\tilde{F}^{(r)}(x)$, the objective function of the `PASAQ-MILP` problem associated with a given estimation value r:

$$\tilde{F}^{(r)}(x) = \sum_{i=1}^{T} \theta_i (r - P_i^d)(1 + \sum_{k=1}^{K} \gamma_{ik} x_{ik}) - \sum_{i=1}^{T} \theta_i \alpha_i \sum_{k=1}^{K} \mu_{ik} x_{ik}$$

- $\tilde{\nu}^{(r)} = arg\min_x \tilde{F}^{(r)}(x)$

Also, we define the game constants decided by the payoff in Table 6.

Next, we prove a number of lemmas that will be used to prove Theorem 1.

LEMMA 3. *Let $\tilde{N}(x) = \sum_{i=1}^{T} \theta_i \alpha_i L_i^{(2)}(x_i) + \sum_{i=1}^{T} \theta_i P_i^d L_i^{(1)}(x_i)$ and $\tilde{D}(x) = \sum_{i=1}^{T} \theta_i L_i^{(1)}(x_i) > 0$ be the piecewise linear approximation of $N(x)$ and $D(x)$ respectively. Then, $\forall x \in \mathscr{X}_f$*

$$|N(x) - \tilde{N}(x)| \leq (\overline{\theta}\overline{\alpha} + \overline{P^d \theta \beta}) \frac{T}{K}$$

$$|D(x) - \tilde{D}(x)| \leq \overline{\theta\beta} \frac{T}{K}$$

Proof: Let $f_i(x_i)$ be the original function, and $L_i(x_i)$ be the corresponding piecewise linear approximation function. The following proof holds for both $f_i(x_i) = e^{-\beta_i x_i}$ and $f_i(x_i) = x_i e^{-\beta_i x_i}$:

$$\max_{0\le x_i\le 1} |f_i(x_i) - L_i(x_i)| = \max_{k=1}^{K} \max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} |f_i(x_i) - L_i(x_i)| \tag{10}$$

We now prove the bound on $\max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} |f_i(x_i) - L_i(x_i)|$ in three steps:

1. Assuming $f_i(x_i) \ge L_i(x_i)$, $\frac{k-1}{K} \le x_i \le \frac{k}{K}$

$$\begin{aligned}
&\max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} |f_i(x_i) - L_i(x_i)| \le \max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i) - \min_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} L_i(x_i)\\
&= \max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i) - \min\{L_i(\frac{k-1}{K}), L_i(\frac{k}{K})\}\\
&= \max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i) - \min\{f_i(\frac{k-1}{K}), f_i(\frac{k}{K})\}\\
&\le \max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i) - \min_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i) \le \frac{1}{K} \max_{0\le x_i\le 1} |f_i'(x_i)|
\end{aligned}$$

2. Assuming $f_i(x_i) \le L_i(x_i)$, $\frac{k-1}{K} \le x_i \le \frac{k}{K}$

$$\begin{aligned}
&\max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} |f_i(x_i) - L_i(x_i)| \le \max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} L_i(x_i) - \min_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i)\\
&= \max\{L_i(\frac{k-1}{K}), L_i(\frac{k}{K})\} - \min_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i)\\
&= \max\{f_i(\frac{k-1}{K}), f_i(\frac{k}{K})\} - \min_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i)\\
&\le \max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i) - \min_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} f_i(x_i) \le \frac{1}{K} \max_{0\le x_i\le 1} |f_i'(x_i)|
\end{aligned}$$

3. If $f_i(x_i)$ and $L_i(x_i)$ get across in $[\frac{k-1}{K}, \frac{k}{K}]$, we could partition the range into small regions such that within each sub partition, the two functions do not get across. We then can apply (a) or (b) within each partition.

Combining the above three conditions, we have

$$\max_{\frac{k-1}{K}\le x_i\le \frac{k}{K}} |f_i(x_i) - L_i(x_i)| \le \frac{1}{K} \max_{0\le x_i\le 1} |f_i'(x_i)| \tag{11}$$

where $f_i'(x_i)$ is the first order derivative of function $f_i(x_i)$.

Combining with Equation (10), we have

$$\max_{0\le x_i\le 1} |f_i(x_i) - L_i(x_i)| \le \frac{1}{K} \max_{0\le x_i\le 1} |f_i'(x_i)| \tag{12}$$

Hence, the approximation error bound is decided by the maximum absolute value of the first order derivative. It can be shown that

$$\max_{0\le x_i\le 1} |\frac{d(e^{-\beta_i x_i})}{dx_i}| = |\frac{d(e^{-\beta_i x_i})}{dx_i}|_{x_i=0} = \beta_i \tag{13}$$

$$\max_{0\le x_i\le 1} |\frac{d(x_i e^{-\beta_i x_i})}{dx_i}| = |\frac{d(x_i e^{-\beta_i x_i})}{dx_i}|_{x_i=0} = 1 \tag{14}$$

Combining Equation (12)-(14) gives the following result:

$$|e^{-\beta_i x_i} - L_i^{(1)}(x_i)| \leq \frac{\beta_i}{K}, 0 \leq x_i \leq 1, \quad 1 \leq i \leq T \tag{15}$$

$$|x_i e^{-\beta_i x_i} - L_i^{(2)}(x_i)| \leq \frac{1}{K}, 0 \leq x_i \leq 1, \quad 1 \leq i \leq T \tag{16}$$

Given Equation (15)-(16) and the definitions of $N(x)$, $\tilde{N}(x)$, $D(x)$, and $\tilde{D}(x)$, it follows that

$$|N(x) - \tilde{N}(x)| \leq (\overline{\theta}\overline{\alpha} + \overline{P^d\theta\beta})\frac{T}{K}$$

$$|D(x) - \tilde{D}(x)| \leq \overline{\theta\beta}\frac{T}{K} \quad \square$$

LEMMA 4. *The difference between the objective funciton of* `P1`*,* $Obj_{P1}(x)$*, and its corresponding piecewise linear approximation,* $\tilde{Obj}_{P1}(x)$*, is less than* $C_1\frac{1}{K}$.

Proof: Let $\tilde{N}(x) = \sum_{i=1}^{T}\theta_i\alpha_i L_i^{(2)}(x_i) + \sum_{i=1}^{T}\theta_i P_i^d L_i^{(1)}(x_i)$ and $\tilde{D}(x) = \sum_{i=1}^{T}\theta_i L_i^{(1)} > 0$ be the piecewise linear approximation of the numerator and denominator of Obj_{P1} respectively.

$$\begin{aligned}
&|Obj_{P1}(x) - \tilde{Obj}_{P1}(x)| = |\frac{N(x)}{D(x)} - \frac{\tilde{N}(x)}{\tilde{D}(x)}| \\
&= |\frac{N(x)}{D(x)} - \frac{N(x)}{\tilde{D}(x)} + \frac{N(x)}{\tilde{D}(x)} - \frac{\tilde{N}(x)}{\tilde{D}(x)}| \\
&\leq |\frac{N(x)}{D(x)}\frac{\tilde{D}(x) - D(x)}{\tilde{D}(x)}| + |\frac{N(x) - \tilde{N}(x)}{\tilde{D}(x)}| \\
&= \frac{1}{\tilde{D}(x)}(|Obj_{P1}(x)| \cdot |D(x) - \tilde{D}(x)| + |N(x) - \tilde{N}(x)|)
\end{aligned}$$

Based on Lemma 3,

$$|N(x) - \tilde{N}(x)| \leq \sum_{i=1}^{T}\theta_i\alpha_i\frac{1}{K} + \sum_{i=1}^{T}\theta_i|P_i^d|\frac{\beta_i}{K} \leq (\overline{\theta}\overline{\alpha} + \overline{P^d\theta\beta})\frac{T}{K}$$

$$|D(x) - \tilde{D}(x)| \leq \sum_{i=1}^{T}\theta_i\frac{\beta_i}{K} \leq (\overline{\theta}/\underline{\theta})\beta\frac{T}{K}$$

At the same time, $|Obj_{P1}(x)| \leq \overline{R^d}$ and $\tilde{D}(x) \geq T\underline{\theta}e^{-\overline{\beta}}$. Hence,

$$|Obj_{P1}(x) - \tilde{Obj}_{P1}(x)| \leq (\overline{\theta}/\underline{\theta})e^{\overline{\beta}}\overline{\beta}\{\overline{R^d} + \overline{P^d} + \frac{\overline{\alpha}}{\overline{\beta}}\} \cdot \frac{1}{K} = C_1\frac{1}{K} \quad \square$$

LEMMA 5. *Let* $\tilde{L}^*$ *and* $\tilde{U}^*$ *be the final lower and upper bounds of* PASAQ*, and* $\tilde{x}^*$ *is the defender strategy returned by* PASAQ. *Then,*

$$\tilde{L}^* \leq \tilde{Obj}_{P1}(\tilde{x}^*) \leq \tilde{U}^*$$

Proof: When the algorithm stops, we have $F^{(\tilde{L}^*)}(\tilde{x}^*) \leq 0 \Rightarrow \tilde{L}^* \leq \frac{N(\tilde{x}^*)}{D(\tilde{x}^*)} = Obj_{P1}(\tilde{x}^*)$ At the same time, $F^{(\tilde{U}^*)}(\tilde{x}^*) > 0, \forall\tilde{x} \Rightarrow \tilde{U}^* > \frac{N(\tilde{x}^*)}{D(\tilde{x}^*)} = Obj_{P1}(\tilde{x}^*)$. $\square$

LEMMA 6. $\forall x \in \mathscr{X}_f$*, the following condition holds*

$$|F^{(r)}(x) - \tilde{F}^{(r)}(x)| \leq (|r| + \overline{P^d})\overline{\theta}\sum_{i=1}^{T}\frac{\beta_i}{K} + \overline{\alpha}\overline{\theta}\frac{T}{K} \tag{17}$$

Proof: Let $L_i^{(1)}(x_i) = 1 + \sum_{k=1}^{K} a_{ik}x_{ik}$ be the piecewise linear approximations of function $e^{-\beta_i x_i}$, and $L_i^{(2)}(x_i) = \sum_{k=1}^{K} b_{ik}x_{ik}$ be that of function $x_i e^{-\beta_i x_i}$. We have

$$\begin{aligned}
&|F^{(r)}(x) - \tilde{F}^{(r)}(x)| \\
&\leq |\sum_{i=1}^{T} \theta_i (r - P_i^d) e^{-\beta_i x_i} - \sum_{i=1}^{T} \theta_i (r - P_i^d) L_i^{(1)}(x_i)| + |\sum_{i=1}^{T} \theta_i \alpha_i x_i e^{-\beta_i x_i} - \sum_{i=1}^{T} \theta_i \alpha_i L_i^{(2)}(x_i)| \\
&\leq \sum_{i=1}^{T} \theta_i |r - P_i^d| \cdot |e^{-\beta_i x_i} - L_i^{(1)}(x_i)| + \sum_{i=1}^{T} \theta_i \alpha_i |x_i e^{-\beta_i x_i} - L_i^{(2)}(x_i)| \\
&\leq (|r| + P^d)\overline{\theta} \sum_{i=1}^{T} |e^{-\beta_i x_i} - L_i^{(1)}(x_i)| + \overline{\alpha}\overline{\theta} \sum_{i=1}^{T} |x_i e^{-\beta_i x_i} - L_i^{(2)}(x_i)| \qquad (18)
\end{aligned}$$

Combining Equation (15) and (16), we have

$$|F^{(r)}(x) - \tilde{F}^{(r)}(x)| \leq (|r| + P^d)\overline{\theta} \sum_{i=1}^{T} \frac{\beta_i}{K} + \overline{\alpha}\overline{\theta}\frac{T}{K} \quad \square$$

LEMMA 7. *Let $\tilde{L}^*$ be the estimated maximum of $Obj_{P1}(x)$ by running* PASAQ*, then*

$$\tilde{F}^{(\tilde{L}^*)}(x) \geq -\epsilon\overline{\theta}T, \quad \forall x \in \mathscr{X}_f \qquad (19)$$

Proof: Let U^* and L^* be the upper and lower bound when the algorithm stops. According to Line 3 in Algorithm 1, $U^* - L^* \leq \epsilon$. Furthermore, $U^* > \tilde{L}^* \geq L^*$. Therefore we know$\tilde{L}^* + \epsilon \geq U^*$, so the optimal solution value of `CF-OPT` with $r = \tilde{L}^* + \epsilon$ must satisfy $\delta^*_{\tilde{L}^*+\epsilon} > 0$. In other words,

$$\tilde{F}^{(\tilde{L}^*+\epsilon)}(x) > 0, \quad \forall x \in \mathscr{X}_f \qquad (20)$$

On the other hand,

$$\tilde{F}^{(\tilde{L}^*+\epsilon)}(x) - \tilde{F}^{(\tilde{L}^*)}(x) = \epsilon \sum_{i=1}^{T} \theta_i L_i^{(1)}(x_i) \leq \epsilon\overline{\theta}T, \forall x \in \mathscr{X}_f \qquad (21)$$

Combining Equation (20) and (21)

$$\tilde{F}^{(\tilde{L}^*)}(x) \geq \tilde{F}^{(\tilde{L}^*+\epsilon)}(x) - \epsilon\overline{\theta}T \geq -\epsilon\overline{\theta}T \quad \square$$

LEMMA 8. *Let $\tilde{L}^*$ be the final lower bound of* PASAQ*. Let L^* be the final lower bound if we solve `CF-OPT` exactly. It follows that*

$$L^* - \tilde{L}^* \leq C_1 \frac{1}{K} + C_2\epsilon$$

Proof: According to Lemma 1, $F^{(L^*)}(\nu^{(L^*)}) \leq 0$. At the same time,

$$\begin{aligned}
&F^{(L^*)}(\nu^{(L^*)}) = F^{(\tilde{L}^*)}(\nu^{(L^*)}) + (L^* - \tilde{L}^*) \sum_{i=1}^{T} \theta_i e^{-\beta_i \nu_i^{(L^*)}} \\
&\Rightarrow (L^* - \tilde{L}^*) \sum_{i=1}^{T} \theta_i e^{-\beta_i \nu_i^{(L^*)}} \leq -F^{(\tilde{L}^*)}(\nu^{(L^*)}) \qquad (22)
\end{aligned}$$

Furthermore, Lemma 6 indicates that

$$\begin{aligned}
&- F^{(\tilde{L}^*)}(\nu^{(L^*)}) \leq -\tilde{F}^{(\tilde{L}^*)}(\nu^{(L^*)}) + (|\tilde{L}^*| + \overline{P^d})\overline{\theta} \sum_{i=1}^{T} \frac{\beta_i}{K} + +\overline{\alpha}\overline{\theta}\frac{T}{K} \\
&\leq -\tilde{F}^{(\tilde{L}^*)}(\nu^{(L^*)}) + \overline{\theta}((\overline{R^d} + \overline{P^d}) \sum_{i=1}^{T} \frac{\beta_i}{K} + \frac{T}{K}\overline{\alpha}) \\
&\leq -\tilde{F}^{(\tilde{L}^*)}(\nu^{(L^*)}) + \overline{\theta}(\frac{(\overline{R^d} + \overline{P^d})\overline{\beta}}{K} + \frac{\overline{\alpha}}{K})T \qquad (23)
\end{aligned}$$

since $|\tilde{L}^*| \leq \overline{|R_i^d|}$. Combining Equation (19),(22) and (23)

$$(L^* - \tilde{L}^*) \sum_{i=1}^{T} \theta_i e^{-\beta_i \nu_i^{(L^*)}} \leq \epsilon \overline{\theta} T + \overline{\theta}(\frac{(\overline{R^d} + \overline{P^d})\overline{\beta}}{K} + \frac{\overline{\alpha}}{K})T$$

Furthermore, $\sum_{i=1}^{T} \theta_i e^{-\beta_i \nu_i^{(L^*)}} \geq T\underline{\theta} e^{-\overline{\beta}}$, so

$$L^* - \tilde{L}^* \leq (\overline{\theta}/\underline{\theta}) e^{\overline{\beta}} \{\epsilon + \frac{1}{K}((\overline{R^d} + \overline{P^d})\overline{\beta} + \overline{\alpha})\} \quad \square$$

Finally we prove Theorem 1:

Proof: The first inequality is implied since $\tilde{x}^*$ is a feasible solution. Furthermore,

$$\begin{aligned} p^* - Obj_{P1}(\tilde{x}^*) =& (p^* - L^*) + (L^* - \tilde{L}^*) + (\tilde{L}^* - \tilde{Obj}_{P1}(\tilde{x}^*)) \\ &+ (\tilde{Obj}_{P1}(\tilde{x}^*) - Obj_{P1}(\tilde{x}^*)) \end{aligned}$$

Algorithm 1 indicates that $L^* \leq p^* \leq U^*$, hence $p^* - L^* \leq \epsilon$. Additionally, Lemma 4, 5 and 8 provide an upper bound on $\tilde{Obj}_{P1}(\tilde{x}^*) - Obj_{P1}(\tilde{x}^*)$, $\tilde{L}^* - \tilde{Obj}_{P1}(\tilde{x}^*)$ and $L^* - \tilde{L}^*$, therefore

$$p^* - Obj_{P1}(\tilde{x}^*) \leq \epsilon + C_1 \frac{1}{K} + C_2 \epsilon + C_1 \frac{1}{K} \leq 2C_1 \frac{1}{K} + (C_2 + 1)\epsilon \quad \square$$

References

Agmon, Noa, Sarit Kraus, Gal A. Kaminka, Vladimir Sadov. 2009. Adversarial uncertainty in multi-robot patrol. *Proc. of The 21st International Jont Conference on Artifical Intelligence (IJCAI)*. 1811–1817.

Agmon, Noa, Vladimir Sadov, Gal A. Kaminka, Sarit Kraus. 2008. The impact of adversarial knowledge on adversarial planning in perimeter patrol. *Proc. of The 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 55–62.

An, Bo, Manish Jain, Milind Tambe, Christopher Kiekintveld. 2011a. Mixed-initiative optimization in security games: A preliminary report. *Proc. of the AAAI Spring Symposium on Help Me Help You: Bridging the Gaps in Human-Agent Collaboration*. 8–11.

An, Bo, David Kempe, Christopher Kiekintveld, Eric Shieh, Satinder Singh, Milind Tambe, Yevgeniy Vorobeychik. 2012. Security games with limited surveillance. *Proc. of the 26th Conference on Artificial Intelligence*. 1241–1248.

An, Bo, James Pita, Eric Shieh, Milind Tambe, Christopher Kiekintveld, Janusz Marecki. 2011b. GUARDS and PROTECT: Next generation applications of security games. *SIGECOM* **10** 31–34.

An, Bo, Milind Tambe, Fernando Ordóñez, Eric Shieh, Christopher Kiekintveld. 2011c. Refinement of strong Stackelberg equilibria in security games. *Proc. of the 25th Conference on Artificial Intelligence*. 587–593.

Avenhaus, Rudolf, Bernhard von Stengel, Shmuel Zamir. 2002. *Inspection games*, vol. 3, chap. 51. North-Holland, Netherlands, 1947–1987.

Babu, L., L. Lin, R. Batta. 2006. Passenger grouping under constant threat probability in an airport security system. *European Journal of Operational Research* **168**(2) 633 – 644.

Basar, Tamer, Geert Jan Olsder. 1995. *Dynamic Noncooperative Game Theory*. Academic Press, San Diego, CA.

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

Bier, Vicki M. 2007. Choosing what to protect. *Risk Analysis* **27**(3) 607–620.

Blair, Dennis. 2010. Annual threat assessment of the us intelligence community for the senate select committee on intelligence. *http://www.isisnuclearіran.org/assets/pdf/2010_NIE.pdf* .

Breton, Michele, A. Alg, Alain Haurie. 1988. Sequential Stackelberg equilibria in two-person games. *Optimization Theory and Applications* **59**(1) 71–97.

Brown, G., M. Carlyle, J. Kline, K. Wood. 2005a. A two-sided optimization for theater ballistic missile defense. *Operations Research* **53**(5) 263–275.

Brown, G., M. Carlyle, J. Royset, K. Wood. 2005b. *The Next Wave in Computing, Optimization and Decision Technologies*, chap. On the complexity of delaying an adversary's project. Springer, NY, USA, 3–17.

Brown, Gerald, Matthew Carlyle, Javier Salmeron, Kevin Wood. 2006. Defending critical infrastructure. *Interfaces* **36**(6) 530 – 544.

Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.

Conitzer, Vincent, Tuomas Sandholm. 2006. Computing the optimal strategy to commit to. *Proc. of the 7th ACM conference on Electronic commerce*. 82–90.

Dozier, Kimberly. 2011. Bin laden trove of documents sharpen US aim. *http://www.msnbc.msn.com/id/43331634/ns/us_news-security/t/bin-laden-trove-documents-sharpen-us-aim/* .

Fudenberg, D., J. Tirole. 1991. *Game Theory*. MIT Press, Cambridge, MA.

Gilpin, Andrew. 2009. Algorithms for abstracting and solving imperfect information games. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.

Jain, Manish, Jason Tsai, James Pita, Christopher Kiekintveld, Shyamsunder Rathi, Milind Tambe, Fernando Ordóñez. 2010. Software assistants for randomized patrol planning for the LAX airport police and the federal air marshal service. *Interfaces* **40**(4) 267–290.

Kahneman, D., A. Tvesky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47**(2) 263–291.

Kiekintveld, Christopher, Janusz Marecki, Milind Tambe. 2011. Approximation methods for infinite bayesian Stackelberg games: modeling distributional uncertainty. *Proc. of The 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 1005–1012.

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

Larson, R.C. 1974. A hypercube queueing modeling for facility location and redistricting in urban emergency services. *Journal of Computers and Operations Research* **1**(1) 67–95.

Leitmann, George. 1978. On generalized Stackelberg strategies. *Optimization Theory and Applications* **26**(4) 637–643.

McKelvey, R. D., T. R. Palfrey. 1995. Quantal response equilibria for normal form games. *Games and Economic Behavior* **10**(1) 6–38.

Nie, X., R. Batta, Drury, Lin. 2007. Optimal placement of suicide bomber detectors. *Military Operations Research* **12**(2) 65–78.

Paruchuri, Praveen, Jonathan P. Pearce, Janusz Marecki, Milind Tambe, Fernando Ordóñez, Sarit Kraus. 2008. Playing games with security: An efficient exact algorithm for Bayesian Stackelberg games. *Proc. of The 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 895–902.

Paruchuri, Praveen, Milind Tambe, Fernando Ordonez, Sarit Kraus. 2006. Security in multiagent systems by policy randomization. *Proc. of The 5th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 273–280.

Pita, James, Manish Jain, Fernando Ordóñez, Milind Tambe, Sarit Kraus, Reuma Magori-Cohen. 2009. Effective solutions for real-world Stackelberg games: When agents must deal with human uncertainties. *Proc. of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 369–376.

Pita, James, Manish Jain, Milind Tambe, Fernando Ordóñez, Sarit Kraus. 2010. Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition. *Artificial Intelligence* **174**(15) 1142–1171.

Pita, James, Manish Jain, Craig Western, Christopher Portway, Milind Tambe, Fernando Ordóñez, Sarit Kraus, Praveen Parachuri. 2008. Deployed ARMOR protection: The application of a game-theoretic model for security at the Los Angeles International Airport. *Proc. of The 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 125–132.

Pita, James, Milind Tambe, Christopher Kiekintveld, Shane Cullen, Erin Steigerwald. 2011. GUARDS - game theoretic security allocation on a national scale. *Proc. of The 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 37–44.

Rogers, Brian W., Thomas R. Palfrey, Colin F. Camerer. 2009. Heterogeneous quantal response equilibrium and cognitive hierarchies. *Journal of Economic Theory* **144**(4) 1440–1467.

Ruan, S., C. Meirina, F. Yu, K. R. Pattipati, R. L. Popp. 2005. Patrolling in a stochastic environment. *Proc. the 10th International Command and Control Research and Technology Symposium*.

Sandler, Todd, Daniel Arce. 2003. Terrorism and game theory. *Simulation and Gaming* **34**(3) 319–337.

Shieh, Eric, Bo An, Rong Yang, Milind Tambe, Craig Baldwin, Joseph DiRenzo, Ben Maule, Garrett Meyer. 2012. PROTECT: A deployed game theoretic system to protect the ports of the United States. *Proc. of The 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 13–20.

Simon, H. 1956. Rational choice and the structure of the environment. *Psychological Review* **63**(2) 129–138.

Srivastava, Vivek, James Neel, Allen B. MacKenzie, Rekha Menon, Luiz A. Dasilva, James E. Hicks, Jeffrey H. Reed, Robert P. Gilles. 2005. Using game theory to analyze wireless ad hoc networks. *IEEE Communications Surveys and Tutuorials* **7**(4) 46–56.

Stahl, Dale, Paul Wilson. 1994. Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization* **25**(3) 309–327.

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

Tsai, Jason, Shyamsunder Rathi, Christopher Kiekintveld, Fernando Ordóñez, Milind Tambe. 2009. IRIS: a tool for strategic security allocation in transportation networks. *Proc. of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 37–44.

Vanek, Ondrej, Michal Jakob, Ondrej Hrstka, Michal Pechoucek. 2011. Using multi-agent simulation to improve the security of maritime transit. *Proc. of 12th International Workshop on Multi-Agent-Based Simulation (MABS)*. 1–15.

Vavasis, Stephen A. 1995. *Handbook of Global Optimization*, chap. Complexity issues in global optimization: a survey. Kluwer, 27–41.

von Stengel, Bernhard, Shmuel Zamir. 2004. Leadership with commitment to mixed strategies. Tech. Rep. LSE-CDAM-2004-01, CDAM Research Report.

wei Lye, Kong, Jeannette M. Wing. 2005. Game strategies in network security. *International Journal of Information Security* **4**(1–2) 71–86.

Wein, Lawrence M. 2008. Homeland security: From mathematical models to policy implementation. *Operations Research* **57**(4) 801–811.

Wilcox, R. R. 2003. *Applying Contemporary Statistical Techniques*. 2nd ed. Academic Press, Amsterdam, Netherlands.

Wright, J., K. Leyton-Brown. 2010. Beyond equilibrium: Predicting human behavior in normal form games. *Proc. of The 24th AAAI Conference on Artificial Intelligence (AAAI)*. 901–907.

Yang, Rong, Chris Kiekintveld, Fernando Ordóñez, Milind Tambe, Richard John. 2011. Improving resource allocation strategy against human adversaries in security games. *Proc. of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*. 458–464.

Yang, Rong, Milind Tambe, Fernando Ordonez. 2012. Computing optimal strategy against quantal response in security games. *Proc. of The 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 847–854.

Yin, Zhengyu, Manish Jain, Milind Tambe, Fernando Ordóñez. 2011. Risk-averse strategies for security games with execution and observational uncertainty. *Proc. of The 25th AAAI Conference on Artificial Intelligence (AAAI)*. 758–763.

Yin, Zhengyu, Dmytro Korzhyk, Christopher Kiekintveld, Vincent Conitzer, , Milind Tambe. 2010. Stackelberg vs. nash in security games: interchangeability, equivalence, and uniqueness. *Proc. of The 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 1139–1146.

Young, S., D Orchard. 2011. Remembering 9/11: Protecting America's ports. *http://coastguard.dodlive.mil/2011/09/remembering-911-protecting-americas-ports/* .