Human Adversaries in Opportunistic Crime Security Games: Evaluating Competing Bounded Rationality Models

Yasaman Dehghani Abbasi* YDEHGHAN@USC.EDU
Martin Short** MBSHORT@MATH.GATECH.EDU
Arunesh Sinha* ARUNESHS@USC.EDU
Nicole Sintov* SINTOV@USC.EDU
Chao Zhang* ZHAN661@USC.EDU
Milind Tambe* TAMBE@USC.EDU
*University of Southern California, Los Angeles, CA 90089, USA
**Georgia Institute of Technology, Atlanta, GA 30332, USA

Abstract

There are a growing number of automated decision aids based on game-theoretic algorithms in daily use by security agencies to assist in allocating or scheduling their limited security resources. These applications of game theory, based on the "security games" paradigm, are leading to fundamental research challenges: one major challenge is modeling human bounded rationality. More specifically, the security agency, assisted with an automated decision aid, is assumed to act with perfect rationality against a human adversary; it is important to investigate the bounded rationality of these human adversaries to improve effectiveness of security resource allocation. This paper for the first time provides an empirical investigation of adversary bounded rationality in *opportunistic crime settings*, where modeling bounded rationality is particularly crucial. We conduct extensive human subject experiments, comparing ten different bounded rationality models, and illustrate that: (a) while previous research proposed the use of the stochastic choice "quantal response" model of human adversary, this model is significantly outperformed by more advanced models of "subjective utility quantal response"; (b) Combinations of the well-known prospect theory model with these advanced models lead to an even better performance in modeling human adversary behavior; (c) while it is important to model the non-linear human weighing of probability, as advocated by prospect theory, our findings are the exact opposite of prospect theory in terms of how humans are seen to weigh this non-linear probability.

1. Introduction

In recent years, the Stackelberg Security Games (SSG) model has received significant attention for its success in modeling physical security problems and application to real world settings, such as scheduling patrols conducted by the US Coast Guards at multiple major US ports (Shieh *et al*. 2012), scheduling police patrols at major airports such as LAX (Pita *et al*. 2008), allocating federal air marshals on flights of US Air Carriers and several other applications (Tambe 2011). SSG provides a game theory-based representation of the interaction between an attacker and defender,

and provides computational tools to optimize the defender's action based on possible attacker moves (Tambe 2011, Korzyk, Conitzer and Parr 2010, Gatti 08).

In SSG, the defender (leader) moves first by choosing to play a particular defense strategy. The adversary (follower) observes this strategy and then chooses a best response strategy. In order to prevent the adversary from predicting the defenders actions, the defender must play a distribution over strategies, known as a mixed strategy, rather than a single fixed one. The Stackelberg equilibrium computation involves finding the utility maximizing mixed strategy, taking into consideration the adversary's response. Traditionally, SSG assumes a model of a perfectly rational adversary. This model is justified when considering domains such as counter-terrorism, where the adversary has sufficient time and motivation to launch a single carefully planned attacks (e.g., see (Southers 2011)); in other domains such as urban crime, this assumption appears weak. It is known that adversaries in these domains are boundedly rational (Zhang *et al*. 2014) and moreover, human subjects do not generally demonstrate perfect rationality in their decisions (Camerer and Chongn 2004; Costa-Gomes *et al.* 2011). Failure to account for this bounded rationality can lead to non-optimal defender strategies in SSG and hence significant losses for the defender. Therefore, constructing a reliable model of the adversary behavior is vital for security against urban crime.

Models of bounded rationality have received lot of attention recently (Gal, Pfeffer 2007; Ficici, Pfeffer 2008; Nguyen *et al.* 2013). Two commonly used models are Prospect Theory (PT) and Quantal Response. PT (Kahneman, Tversky 1979) models decision making under uncertainty and captures bounded rationality by mapping the real probability values to a person's interpretation of probabilities through a non-linear probability weighing function, while also accounting for his risk preference. Quantal Response (QR) (McKelvey and Palfrey 1995) models bounded rationality of human subjects by introducing uncertainty into their decision making process.

In SSG literature, variations and combinations of PT and QR models have been investigated (Yang *et al. 2013*; Nguyen *et al.* 2013; Cui *et al. 2014*). However, only recently a different category of adversaries have been investigated: opportunistic adversaries (Zhang *et al.* 2014). A significant portion of urban crime is opportunistic in nature (Zhang *et al.* 2014); and only recently a different category of adversaries have been investigated: opportunistic adversaries which is significant portion of urban crime is opportunistic in nature (Zhang *et al.* 2014); these adversaries, in addition to not being completely rational, are flexible in execution of crime and seeking opportunities for crime.

This paper for the first time focuses on empirical investigation of human adversary modeling in opportunistic crime setting. Quantal Biased Random Movement (QBRM) is one of few models proposed to model adversary decision making in opportunistic crimes (Zhang *et al.* 2014). We conduct extensive human subject experiments, comparing ten different bounded rationality models, and show that: (1) although previous research proposed the use of the well-known stochastic choice QR model of human adversary, this model is significantly outperformed by more advanced models of Subjective Utility Quantal Response (SUQR) in the context of opportunistic crime; (2) while it is important to model the non-linear human weighing of probability, proposed by prospect theory, our findings contradict with prospect theory in terms of how humans are seen to weigh this non-linear probability; and (3) combinations of the well-known prospect theory model with SUQR models lead to an even better performance in modeling human adversary behavior.

2. Background

2.1. Stackelberg Security Game

A Stackelberg Security Game (SSG) is a game model that captures the interaction between a single defender (leader) and one adversary (follower) (Tambe 2011). The defender protects a set of targets T with limited number of resources from attack by the adversary. A pure strategy of the defender is an assignment of the security resources to the targets. A mixed strategy is a probability distribution over the set of all possible pure strategies, which is succinctly represented as a vector x of size |T| in which each element of vector represents the probability of covering a target (Korzhyk, Conitzer, and Parr 2010). SSG assumes strategic adversaries who learn the defender's strategy by conducting long-term surveillance; the adversary's pure strategy best response is then to choose a target to attack that maximizes the adversary's expected utility. The utility of the adversary is given by $U_a^c(t)$ and $U_a^u(t)$ when the adversary attacks the target t and it is covered or uncovered, respectively (the utility of the defender is given by $U_d^c(t)$ and $U_d^u(t)$). Given the defender mixed strategy x, the adversary's expected utility in attacking target t is given by the following equation

$$U_a(t, x) = x_t U_a^c(t) + (1 - x_t) U_a^u(t). \quad (1)$$

The equilibrium in this game corresponds to the optimal strategy x that maximizes the defender's utility assuming the adversary provides his best response. However, Equation 1 assumes a perfectly rational adversary, which may be appropriate in domains such as counter-terrorism. However, in domains such as opportunistic crime settings discussed next, the adversary's behavior may be governed by models of bounded rationality (Nguyen *et al*. 2013). We review human behavior model accounting for adversary bounded rationality in SSG in Section 2.3.

2.2. Opportunistic Security Game

SSG assumes strategic adversaries who learn the defender's strategy and then decide an attack plan which will not change. However, in domains such as urban crime and theft on trains, the attackers (adversary) are opportunistic, i.e., they are flexible about their plan and seek opportunities for crime rather than strategically planning attacks. For example, a thief may decide not to steal if he observes a police officer, and may move to another area to seek opportunities for committing a crime. Recent work (Zhang *et al*. 2014) explores a model (Quantal Biased Random Movement) of opportunistic attackers within a game interaction between the defender and attackers. Specifically, the authors describe three characteristics of an opportunistic attacker: (i) opportunistically and repeatedly seeks to commit crimes, using a boundedly rational process to select the next crime location; (ii) reacts to real-time information at execution time rather than planning attacks in advance; and (iii) has limited observation of defender strategy. Section 4 of this paper evaluates QBRM model along with other models for bounded rationality in OSG domain.

2.3. Human Behavior Models

In this section, we describe details of some human behavior models that have been explored in the literature, including SUQR, PT and QBRM.

2.3.1. Subjective Utility Quantal Response (SUQR) (Conditional Logit)

In a recent work on SSG (Nguyen *et al.* 2013), the authors combined two key notions of decision making: Subjective Expected Utility (SEU) (Fischhoff, Goitein, and Shapira 1981) and Quantal Response (QR) (McKelvey and Palfrey 1995), and proposed the SUQR model. The SUQR model is mathematically equivalent to the conditional logit model in discrete choice theory.
QR models the uncertainty in the decisions made by any agent. Traditionally, the utility maximizing rational agent chooses the action a_i that provides highest utility u_i. In the logit QR model, the rationality assumption is relaxed by positing that the decision making agent chooses an action a_i with probability proportional to e^{u_i}. In the context of SSG, given the defender's mixed strategy x, the probability of the adversary choosing to attack target t is given by

$$q_t(x) = \frac{e^{\lambda U_a(t,x)}}{\sum_{t' \in T} e^{\lambda U_a(t',x)}} \tag{2}$$

Other models of QR have assumed a power function for formulation of q_t, which is given by

$$q_t(x) = \frac{U_a(t,\ x)^{\lambda}}{\sum_{t' \in T} U_a(t',\ x)^{\lambda}} \tag{3}$$

In Subjective Expected Utility (SEU) - as proposed in behavioral decision-making (Savage 1972; Fischhoff *et al.* 1981) - the key idea is that individuals have their own evaluations of different factors during decision making process. In an SSG, the factors considered by an adversary in choosing the target to attack include the marginal coverage on target t (x_t) and the subject's reward and penalty (R_t^a, P_t^a). Inspired by the idea of SEU, a subjective utility function for the adversary in an SSG setting is as follows: $w_1 x_t + w_2 R_t^a + w_3 P_t^a$, where the weights, w_i, denote the relative importance given to these factors by the adversary. While unconventional at first glance, this model leads to higher prediction accuracy than the classic expected value function (Nguyen *et al.* 2013). This might be due to the fact that human decision making process may be based on simple heuristics.

The SUQR model replaces the expected value function in logit QR model with the SEU function. In the SUQR model, the probability that the adversary chooses target t is given by:

$$q_t(x) = \frac{e^{w_1 x_t + w_2 R_t^a + w_3 P_t^a}}{\sum_{t' \in T} e^{w_1 x_{t'} + w_2 R_{t'}^a + w_3 P_{t'}^a}} \tag{4}$$

2.3.2. Prospect Theory

Prospect theory (PT) is an alternative model to expected utility theory that models decision making under risk. It is one of the most successful theories of decision making under risk and has been applied in a wide variety of contexts (Tversky and Kahneman 1992). This model is descriptive and tries to model real-life choices rather than optimal decisions.

There are two functions in prospect theory: the probability weighing function (equation 5) and the value function (equation 6). The probability weighing function models human interpretation of probability and suggests that people weigh probability non-uniformly. More specifically, the original PT suggests that people tend to overweigh low probabilities and underweight high probabilities, capturing the idea that people tend to overreact to small probability events, but

underreact to large probabilities. Some works in this domain propose and experiment with parametric models (equation 5) that capture a wide range of probability weighing functions (Gonzalez and Wu 1999)

$$f(p) = \frac{\delta p^{\gamma}}{(\delta p^{\gamma} + (1-p)^{\gamma})} \tag{5}$$

Another aspect of prospect theory is the value function (equation 6), which passes through a reference point, *c*, and is S-shaped and asymmetrical.

$$V(U) = \begin{cases} U^{\alpha} & U \geq c \\ -\theta(-U)^{\beta} & U < c \end{cases} \tag{6}$$

The interpretation of this equation is that, contrasting a rational agent, who uses expected utility, and cares about the absolute value, an agent with bounded rationality cares about relative value to c, while unit losses hurt more than unit gains feel good. Finally, the model assumes that humans evaluate each target with the following equation:

$$U(t,x) = f(x_t) * V(U_a^c(t)) + f(1-x_t)V(U_a^u(t)) \tag{7}$$

where $U(t,x)$ is the overall or expected utility of the outcomes to the individual making the decision, with respect to potential outcomes and their respective probabilities.

2.3.3. Quantal Biased Random Movement

Quantal Biased Random Movement (QBRM) is a model proposed to describe an opportunistic attacker's behavior in a defender-attacker interaction on graphs (Zhang *et al.* 2014). The defender moves from node to node on the graph according to some strategic transition matrix, in the hope of dissuading the attacker from committing crimes (not in an effort to catch the attacker). The adversary likewise moves from node to node seeking opportunities for crime at each node. The adversary may commit a crime if no defender is present at a node, but does not commit a crime if a defender is present, which models the flexibility of his plans. The adversary also has a belief about the defender's position, which is based on the defender's coverage probability over different stations and trains. The belief of the attacker, c_b, about the defender's position at any given time is a probability distribution over possible defender positions, and it also depends on real-time observations. That is, if a defender is currently observed, then the attacker's belief about the position of the observed defender matches the actual position, otherwise it is a probability distribution based on the stationary distribution. Given belief c_b^t at time t and current position i of the adversary, the adversary has an expected utility $U(j|i, c_b^t)$ of moving to position j.

The adversary is also assumed to have bounded rationality, hence the QR model is used to model his choice of actions. More specifically, the probability of the adversary moving from position i to position j is given by the following equation:

$$q_j(i, c_b^t) = \frac{U(j|\, i, c_b^t)^{\lambda}}{\sum_k U(k|i, c_b^t)^{\lambda}} \tag{8}$$

The authors provide algorithms to compute the optimal strategy of the defender, given parameters of the adversary model. In this paper we use a setting similar to OSG, but explore a variety of different models of the adversary. To the best of our knowledge, our work is the first paper to perform human subject experiments in the context of OSG.

3. Experiments

3.1. Experimental Procedure

We conducted online experiments with human subjects to evaluate the performance of various models of human behavior in OSG settings. To simulate urban crimes, we deployed an online treasure hunting game, set in a metro transportation system, in which human subjects, recruited from Amazon Mechanical Turk (AMT), played the role of a treasure hunter. To reduce potential bias that could arise from asking participants to engage in illegal behavior, we created a familiar gaming scenario of treasure hunting to simulate opportunistic theft. These players attempt to maximize the rewards they receive by accumulating stars from metro stations in a limited time. Each participant played eight games in total: two practice games, two validation games, and four main games. To ensure that players understood instructions, each player first played two practice games. After each player action in these practice games, we provided them with feedback on their choices. Players then played two simple validation games, but they were not informed that these were validation games. The results of players who did not score a set threshold in the validation rounds were discarded, in order to eliminate invalid data

Before playing the game, players were provided with detailed instructions explaining the game mechanics (which were also available for review at any time during the game). After the game, a brief survey was used to gather data about the players' perception of the game, demographics, and risk seeking tendencies.

3.2. Main Games Description

In the main games, human subjects collect rewards by visiting any of the six stations (see an example in Figure 1), while avoiding officers on patrol. Each station has a known reward, indicated by the number of stars (e.g., Station 2 has 4 stars in Figure 1). These stars are guarded by two officers, and each officer patrols three stations. If a player (human) arrives at a station when there is no officer present, his total reward increases by the number of stars of that station; if the officer is present at the station, he does not gain any reward, but does not pay any penalty either. The player's objective is to maximize the total reward. Players must carefully choose which stations to visit, considering the available information about rewards and officers' coverage distribution on stations. Players can travel to any station (including the current one) from their current station by train (the dotted lines in Figure 1). Sub-windows contain additional information including total reward, remaining game time, link to full instructions, and the message board. The message board provides information about future available actions, warning messages in case of illegal moves, and also descriptions of the current situation.

The officers patrol (move around) stations according to a pre-determined strategy which is calculated offline using an optimization algorithm similar to the one presented in (Zhang *et al.* 2014). Given the topology of the metro system and the stars at each station, a randomized patrolling strategy is generated, which can be used to determine the stationary coverage. The stationary coverage probabilities of each station and trains are revealed to the players, but the exact transition matrix is hidden. This means that players can see the percentage of the time that officers spend on average at each station and on the trains (e.g., 64% of time on Station 1 in Figure 1), and determine

their probability of encountering an officer at a station. During the game, players cannot observe where officers are actually located, unless they encounter the officer at a station.

Each player starts the game at a random station, and is given a limited amount of game time (100 units). For both the player and the officer, visiting a station takes one unit of time, and traveling to a new station takes a number of time units equal to the minimum distance between source and destination station along train routes. A connected line between two stations in the graph (called an edge) illustrates a route between the two stations with unit distance.

The game can finish in one of three ways: (1) the player exceeds 45 minutes limit to read the instruction and play all the games or (2) uses up all 100 units of time for each game, and finally (3) each game is randomly terminated after a station visit, which happens with a 10% probability after each such visit. The random termination encourages the players to choose each action carefully, as there is a chance the game may terminate after each visit. The termination randomizer is also used to model attackers exiting the metro system (Zhang *et al.* 2014).

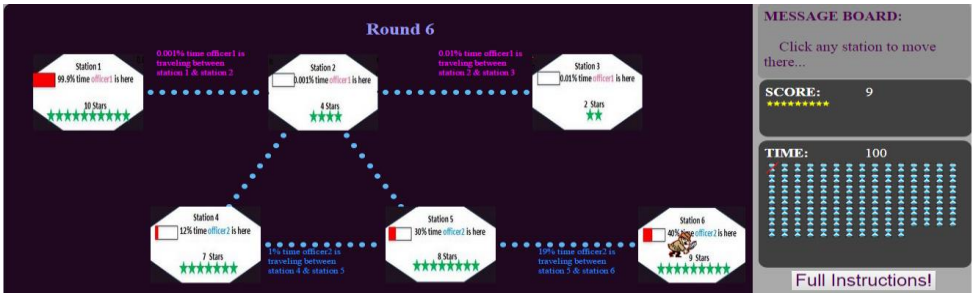

Figure 1. Game interface

3.2.1. Pilot Study

We ran pilot studies to help design the game and prepare the instructions in order to avoid future confusion for AMT players. We interviewed subject matter experts to create an initial prototype. Then, we recruited 21 undergraduate students from the University of Southern California's subject pool to serve as research participants in a set of pilot experiments. Participants played the game, and were then interviewed by study staff to determine how well they understood the game. We asked participants a set of questions about the instructions, their decisions in playing the game (such as staying at any station or moving to other stations), as well as their understanding of the coverage probability of a selected station. Based on the feedback, we improved the game interface in multiple ways, such as adding animated cartoons, graphic representation of coverage probabilities with red bars, etc.

3.2.2. Main Games Design

Recall that our study has practice games, validation games and main games. In all main games, there were six stations, but each game had different layouts, different distributions of rewards at each station, and different stationary coverage probabilities. Figure 2 shows the layout for the four main games; two are based on real world transportation systems and the other two are random topologies (Expanding the graph structure is part of our future work). In the experiments, these four games were shown in random order.

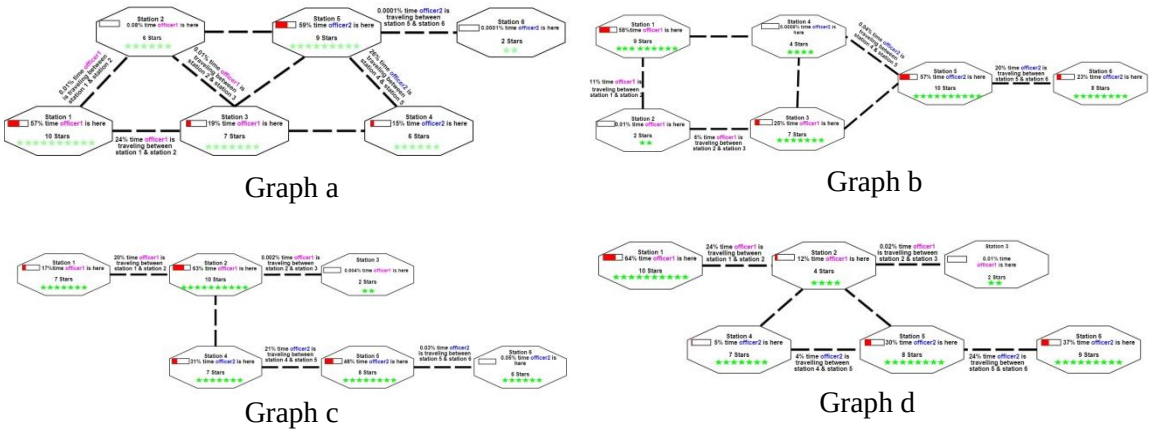

Figure 2. Four main game layouts

To factor out the influence of range and summation of stars, the number of stars at each station is a random integer between 2 and 10; additionally, the summation of stars present at all six stations is kept constant and equal to 40 for all games.

The text along the edges in Figure 2 shows the partitioning of the stations covered by two officers. This 2-way partitioning is determined offline based on the following requirements: (i) each officer must patrol half of the total stations, (ii) each station is patrolled by one and only one officer, (iii) the nodes patrolled by an officer must form a connected graph, and (iv) an officer can get to any of his covered stations without passing any stations covered by the other officer.

3.2.3. Participants

To be eligible to participate, AMT participants must have played more than 500 games on AMT with at least 95% acceptance rate. The games were played in three sets of experiments. In each set, out of about 70 participants, at least 55 unique human subjects successfully completed the games; i.e. successfully passed the validation games, and played a set of four games. In total, 167 unique human subjects successfully passed the validation games and their data were used for evaluation.

3.2.4. Compensation

To motivate the subjects to participate in the study, they were compensated based on their total reward on top of a base compensation. We paid AMT users per the following policy: $1 for participating in the study, plus $0.01 per reward unit (star).

4. Opportunistic Adversary Behavior Model

It has been shown that addressing adversaries' bounded rationality is an essential factor in successful performance of SSG algorithms. In previous work on opportunistic crime, QBRM was used to describe human behavior. However, to the best of our knowledge, no research has been done to descriptively study how well human behavior models capture opportunistic attacker behavior and their choices. Here, we propose ten different models and compare them using different metrics computed on the experimental data. These models are based on variations of Prospect

Theory, Logit Quantal Response and Subjective Utility Quantal Response. We list these models in Table1 below.

Table 1- Summary of models used for human bounded rationality in OSG

Category	Model	Abbrv.	Generic Mathematical Formulation
Quantal Response	Logit Quantal Response Equilibrium	QR	$q_t(i) = \frac{e^{\lambda U(i)}}{\sum_j e^{\lambda U(j)}}$
	Quantal Biased Random Movement	QBRM	$q_t(i) = \frac{U(i)^{\lambda}}{\sum_j U(j)^{\lambda}}$
SUQR (see the factors $\boldsymbol{f_k}$ in text)	SUQR with Stationary Probability (SP) as one factor	SUQR-SP	$q_t(i) = \frac{e^{\sum_k w_k f_k(i)}}{\sum_j e^{\sum_k w_k f_k(j)}}$
	SUQR with Projected Probability (PP)	SUQR-PP	Same as above except using PP instead of SP
	SUQR with Stationary Probability with conditional weights for observed (*o*) and not observed (*n*) case	SUQR-SP-C	Same as above, except two different set of weights
	SUQR with Projected Probability with conditional weights	SUQR-PP-C	Same as above except using PP instead of SP
Prospect Theory	QBRM with PT weighting function on projected probability	PT-QBRM	Use $f(p)$ instead of p in QBRM. See equation 5.
	SUQR-SP-C with weighting function for stationary probability	PT-SUQR-C	Use $f(p)$ instead of p in SUQR-SP. See equation 5
	SUQR-SP-C with weighting function & value function on Absolute Attractiveness	PT-SUQR-C-VA	Use $f(p)$ instead of p, $V(att)$ instead of att in SUQR-SP. See eqn. 5,6
	SUQR-SP-C with weighting function for stationary coverage and value function on relative attractiveness	PT-SUQR-C-VRA	Same as above except using V on relative att

4.1. Models Description

4.1.1. First Category: Quantal Response

In this category, the Quantal Response model refers to the logit quantal response formulation presented in equation 2. The utility is computed as the expected reward of the adversary (expectation over the randomness in defender's strategy) in choosing a certain move (details are in [Zhang *et al.* 2014]).

Quantal Biased Random Movement is the other model in this category which uses the power function form of quantal response as in equation 3.

4.1.2. Second Category: Subjective Utility Quantal Response

In this category, instead of expected reward, subjective utility is used. The key factors we found to be the most important are number of stars at destination station (also called the attractiveness, or att), stationary coverage probability (referred to as SP or *sta*) or projected coverage probability

(referred to PP or *proj*) of destination station, the distance between the current station and the destination station (*dist*), and the connectivity degree of the destination station (*con*). Thus, for examples in SUQR-SP we get

$$\sum_k w_k f_k(i) = w_{att} att_i + w_{sta} sta_i + w_{dis} dis_i + w_{con} con_i$$

We considered two variation of SUQR models. The first uses a single set of weights whether the attacker currently observes the officer or not. The second uses two sets (conditional or C) of weights, one when the officer is observed and the other when not.

4.1.3. Third Category: Using Prospect Theory Weighting Function and (or) Value Function

In this category, we developed models based on combinations of SUQR/QR with Prospect Theory. In PT-QBRM, PT-SUQR-C, and PT-SUQR-C-VA, instead of using the actual projected/stationary coverage probability, we used the weighted probability obtained through equation 5. In PT-SUQR-C-VA, in addition to the weighting function for coverage probability, the value function (equation 6) was also used for the number of stars at each station (attractiveness). In PT-SUQR-C-VRA, instead of using the absolute value of the attractiveness in the value function, we used its relative value given the current and destination station, i.e. if the player moves from a station with 10 stars to a station with 2 stars, he loses 8 stars and if he moves from a station with 6 stars to a station with 8 stars, he gains 2 stars. The idea of using the relative attractiveness and relative coverage probability is based on the anchoring bias theory, and the fact that people make decisions based on comparison between available options.

4.2. Model Prediction Accuracy

We used four metrics to evaluate how well different models predict human decision making compared to the actual responses of human participants in our experiments.

4.2.1. Root-Mean-Square Error (RMSE)

RMSE represents the deviation between model's prediction of attacker's movement ($\hat{p}$) and the actual proportion movements of AMT players from each station to others (p). The prediction probability ($\hat{p}$) and proportion movements (p) both distinguish between the situations that the attacker observes and not observer the officer.

$$RMSE(\hat{p}) = \sqrt{MSE}(\hat{p}) \text{ where MSE } (\hat{p}) = \frac{1}{n}\sum(\hat{p} - p)^2$$

4.2.2. Weighted Absolute Percentage Error (WAPE)

Although RMSE is used often in statistical modeling, it is more sensitive toward outliers, and WAPE can provide a more accurate measure of model fit in these situations

$$WAPE = \frac{\sum |\hat{p} - p|}{\sum P}$$

4.2.3. Akaike information criterion (AIC)

AIC is a measure of relative quality of a model and of the information lost when a given model is used. The model with the smallest AIC is the best fitted model among other models, but absolute AIC cannot provide any information about a particular model's goodness of fit.

$AIC = 2k - 2\ln(Likelihood)$ where k is the number of parameters in the model

4.2.4. Student's t-test

We used Student's t-test to evaluate the prediction accuracy of the proposed models. Model's Prediction were used to study if one model is significantly better than another model.

5. Experimental Results

5.1. Extraction of Model Parameters

We divided the data into training (70%) and testing (30%) datasets. Test data was not used in learning the parameters, and kept aside for evaluation of different models.

To learn the Prospect Theory parameters, the training data set was randomly divided into 10 training and validation sets, and for each combination of PT parameters, the training data set was randomly divided (80/20% split) 10 times to obtain 10 training and validation sets. For each combination of PT parameters, the average error on the validation sets was calculated. Then the PT parameters combination which results in the least average error were selected, and used further to estimate the QBRM or SUQR parameters. To estimate the λ in QBRM, QR, and weights (w's) in SUQR, Maximum Likelihood Estimation Method (MLE) was used.

Table 2- Model parameters and their values

Model	***Parameters***
Logit Quantal Response Equilibrium	λ =0.3645
Quantal Biased Random Movement	λ =1.1955
SUQR with SP as one factor	$< w_{att}, w_{sta}, w_{dis}, w_{con} >$ = <0.3853,-4.6033,-0.7031,0.145>
SUQR with PP as one factor	$< w_{att}, w_{proj}, w_{dis}, w_{con} >$ = <0.2136,-2.5495,-0.6937,0.0327>
SUQR with Stationary Probability with conditional weights	$< w^o_{att}, w^o_{sta}, w^o_{dis}, w^o_{con}, w^n_{att}, w^n_{sta}, w^n_{dis}, w^n_{con} >$ = <0.4206,-4.2065,-0.4281,0.2451,0.4106,-4.9489,-0.7634,0.0427>
SUQR with Projected Probability with conditional weights	$< w^o_{att}, w^o_{proj}, w^o_{dis}, w^o_{con}, w^n_{att}, w^n_{proj}, w^n_{dis}, w^n_{con} >$ = <0.1915,-1.8435,-0.7485,0.0834,0.2584,-3.3138,-0.6021,0.0418>
QBRM with PT weighting function on PP	λ =1.2351 , $\delta = 1.8$, $\gamma = 0.6$
SUQR-SP-C with weighting function for stationary probability	$< w^o_{att}, w^o_{proj}, w^o_{dis}, w^o_{con}, w^n_{att}, w^n_{proj}, w^n_{dis}, w^n_{con} >$ = <0.4159,-2.8093,-0.4286,0.2505,0.4085,-3.1953,-0.7479,0.0903 $< \delta, \gamma >$ = <3.4, 1.6>
SUQR-SP-C with weighting function & value function on Absolute Attractiveness	$< w^o_{att}, w^o_{proj}, w^o_{dis}, w^o_{con}, w^n_{att}, w^n_{proj}, w^n_{dis}, w^n_{con} >$ = <0.4159,-2.8093,-0.4286,0.2505,0.4085,-3.1953,-0.7479,0.0903> $< \delta, \gamma, \alpha >$ = <3.4, 1.6,1>
SUQR-SP-C with weighting function for stationary coverage and value function on relative attractiveness	$< w^o_{att}, w^o_{proj}, w^o_{dis}, w^o_{con}, w^n_{att}, w^n_{proj}, w^n_{dis}, w^n_{con} >$ = <0.6676,-2.8474,-0.5193,0.221,0.6114,-3.2045,-0.8618,0.0648> $< \delta, \gamma, \alpha, \theta, \beta >$ = <3.2, 1.6, 0.8, 0.4, 1.2>

In our analysis, the best α parameter for model PT-SUQR-C-VA was equal to 1, which makes the model *PT-SUQR-C-VA* equivalent to model *PT-SUQR-C*; i.e. introduction of the value function on absolute attractiveness does not provide any improvement over *PT-SUQR-C*. Hence, we did not consider *PT-SUQR-C-VA* in our evaluation.

After normalizing the weights (multiplying w_{att} by 10, w_{dis} with 5 and w_{con} by 4), it can be seen that attractiveness and coverage probabilities are more important than the other factors. The weights for these two factors are very similar across models without prospect theory's functions.

5.2. Results

After determining the parameters of the models, we used the resulting models to predict human decisions and compared them to actual decisions on the test data set. The following are our main observations from the experiment results, starting with significant deviation from perfectly rational play by human subjects.

- **Human decision-making does not conform to the traditional game theoretic assumption of perfect rationality**. Traditional game theory assumes perfect rationality on the part of all players. While an enormous body of work in behavioral game theory has questioned this assumption, research in adversary behavior models in Stackelberg security games (Yang *et al.* 2013) have further illustrated that human behavior is far from perfectly rational. In this paper, we confirm these findings in the context of opportunistic crime settings. Table 2 shows that the rationality factor (λ) for QR model is 0.3 which is extremely low considering the fact that $\lambda = 0$ corresponds to complete irrationality level (uniform decision making for all targets) and $\lambda = \infty$ corresponds to perfect rationality level.

QR is a well-established model of human decision making tracing its origins to (McFadden 1973), and with significant support over the past several decades (McKelvey and Palfrey 1995). The following observation conveys an important finding that an alternative model provides a superior performance to QR.

- **SUQR performs better than QR model**. Previously, (Nguyen *et al*. 2013) performed a brief comparison between QR and SUQR models. Indeed, our work provides the first comprehensive comparison between QR and SUQR; in particular, it compares QR with four different versions of SUQR examining over 176 human decision instances. Table 3 shows the p-value of student's t-test for absolute error on the test data set between QR and SUQR models. As seen in the table, all P-values are smaller than 0.1, which supports the claim that SUQR models result in a better fit compared to QR at 90% confidence level.

Indeed, the fact that SUQR outperforms QR provides support to the claim that human subjects are not making decision rationally; since QR is based on expected utility (a completely rational decision maker seeks to maximize the expected utility) while SUQR is a linear combination of different factors.

Table 3- Student's t-test comparing QR with SUQR models

p-value	**SUQR-SP**	**SUQR-SP-C**	**PT-SUQR-C**	**PT-SUQR-C-VRA**
QR	0.0749	0.045685	0.00571	0.000

PT is a landmark theory and the following four observations relate to incorporation of elements of this theory to improve our model prediction accuracy.

- **Considering human non-linear weighting of probability improves model prediction accuracy.** Past research in human behavior models in the context of Stackelberg security games, including research focused on SUQR (Nguyen *et al* 2013), has not comprehensively addressed human (non-linear) weighting of probability. Yet this weighting of probability is an important aspect of PT. Table 4 presents a summary of model performance on test data illustrating the importance of probability weighting. As can be seen, the models that use the PT weighting function perform better than their counterparts without the PT weighting function: PT-QBRM vs QBRM and PT-SUQR-C vs SUQR-SP-C (these models only have the probability weighting function and not the value function). More specifically, in PT-QBRM and PT-SUQR-C models, instead of using the actual coverage probability, the players' weighting of coverage probability were considered. Furthermore, PT-SUQR-C performs better than PT-QBRM, which further supports the claim that SUQR provides better performance compared to the Quantal Response models.

Table 4- Model prediction accuracy[1]

Model	AIC	RMSE	WAPE
Perfectly Rational	-	0.1964	0.8562
QR	8801.0	0.1868	0.7944
QBRM	8748.2	0.18	0.7494
SUQR-SP	8461.0	0.1748	0.7211
SUQR-PP	8572.3	0.17669	0.7385
SUQR-SP-C	8442.6	0.17019	0.7138
SUQR-PP-C	8554.7	0.17490	0.7324
PT-QBRM	8623.4	0.17831	0.7372
PT-SUQR-C	8363.9	0.16422	0.6843
PT-SUQR-C-VRA	**8361.8**	**0.16388**	**0.6841**

- **Considering human perception of relative attractiveness results in further improvement of the SUQR model.** Table 4 shows that PT-SUQR-C-VRA model has the best values among all the models across all prediction accuracy metrics. This model deploys the perception of the *relative* attractiveness as well as the perception of coverage probability. As stated in section 5.1, addressing human perception of *absolute* attractiveness provides no improvement over PT-SUQR-C model.

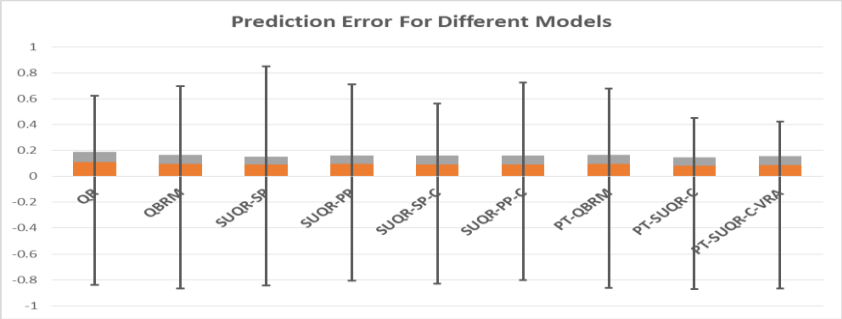

Figure 3. Box plot of actual prediction errors

[1] (Cui *et al. 2014) has also compared human behavior models using AIC metric but in SSG domain and without the test data set.*

Figure 3 shows the box-plots of the errors for all nine models. The box-plot is a common method to show the distribution of the data based on minimum, 1st quartile, median, 3rd quartile and maximum statistics. The two end whiskers represent the minimum and maximum values of the prediction error; the orange and blue boxes are the 1st and 3rd quartiles, respectively. Besides the fact that PT-SUQR-C-VRA model has the lowest average error as shown in Table 4, Figure 3 illustrates that it also has the least variability among all models, and hence, it provides the best fit for human bounded rationality in OSG domain.

- **PT-SUQR-C and PT-SUQR-C-VRA perform better than others.** PT-SUQR-C and PT-SUQR-C-VRA have the least RMSE and WAPE errors among all models. Moreover, these two models have the smallest AIC values which means they do not over fit the data, since AIC compensates for the greater number of parameters in these models compared to others. Between these two models, PT-SUQR-C-VRA has the lowest values with respect to all four error metrics; however the difference is not statistically significant.

In order to achieve high accuracy, one of these two models can be deployed depending on complexity requirements; PT-SUQR-C provides a relatively less complex model, but it still performs better than other models. Table 5 also provides the p-value of Students' t-test for comparing PT-SUQR-C-VRA with other models.

Table 5. P-value for Student's t-test for comparing models

p-value	QR	QBRM	SUQR-SP	SUQR-PP	SUQR-SP-C	SUQR-PP-C	PT-QBRM	PT-SUQR-C
PT-SUQR-C-VRA	0.005	0.097	0.3411	0.1613	0.4348	0.2111	0.1764	0.9948

- **Players' probability weighting function is S-shaped, exactly opposite of the PT standard model of human weighting of probability:** Figure 4a demonstrates the weighing probability functions which indicates that human subjects' perception of low probabilities (less than 0.13) is below the actual values while their perception of high probabilities is above the actual value. For both of these models, the weighting function is S-shaped which is opposite of the inverse S-shaped function proposed by Prospect Theory. Previous works such as (Alarie, Dionne 2001) also found S-shaped weighing functions for probability. Figure 4b illustrates the value function for players' perception of relative attractiveness. It appears that the human subjects value gains more than they dislike losses, which is also the opposite of what is usually found in PT.

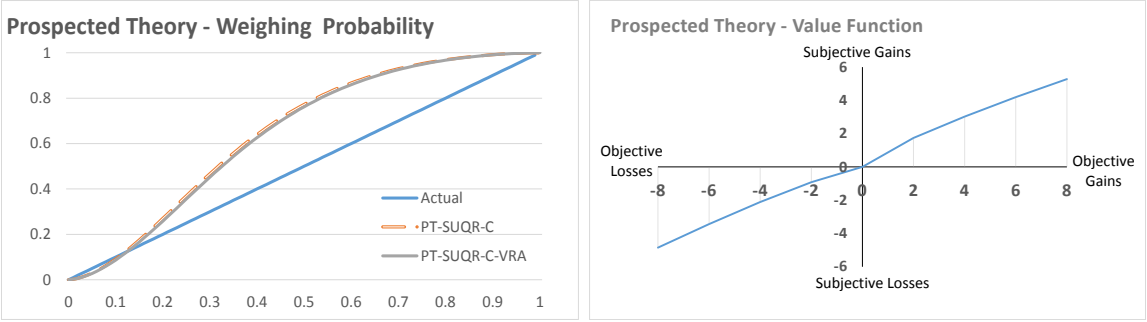

Figure 4. (a) coverage probability weighting function (b) value function

6. Conclusions

With the growing number of automated decision aids based on game-theoretic algorithms in daily use by security agencies, investigations of bounded rationality models of human adversary decision making are now critical, in order to ensure effective security resource allocation and scheduling. This paper for the first time provides an empirical investigation of adversary bounded rationality in opportunistic crime settings, where modeling bounded rationality is particularly crucial. Based on data from extensive human subject experiments, we compare nine different bounded rationality models, and illustrate that: (a) while previous research proposed the use of the quantal response model of human adversary, this model is significantly outperformed by the SUQR model which uses linear combination of target features - thus further indicating that human decision making is not based on maximizing expected utility; (b) combinations of the well-known prospect theory model SUQR leads to an even better performance in modeling human adversary behavior; (c) while it is important to model the non-linear human weighting of probability, as advocated by prospect theory, our data suggests that human weighting of probability is "S-shaped" as opposed to the "inverse S-shape" advocated in prospect theory; (d) models based on relative weighting of values, i.e., gain and loss from current state, provide better modeling accuracy than absolute weighting. These and other findings outlined in this paper provides important advice for practical implementations of decision-aids. Indeed, as police departments begin to adopt these decision aids, modeling and testing these findings in practice in the real-world provides an important next step for future work.

7. Acknowledgment

This research was supported by MURI grant W911NF-11-1-0332, and award no. 004525-00001 by US-Naval Research laboratory.

References

Alarie, Y., Dionne G. (2001) Lottery decisions and probability weighting function. Journal of Risk and Uncertainty, 22(1), 21-33.

Camerer, C.F., Ho, T., Chongn, J. (2004) A congnitive hierarchy model of games. *The Quarterly Journal of Economics, 861-898.*

Costa-Gomes, M., Crawford, V.P., Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study, *Econometrica* 69, 1193-1235.

Cui, J. and John, R.S. (2014). Empirical Comparisons of Descriptive Multi-objective Adversary Models in Stackelberg Security Games. *Decision and Game Theory for Security. Springer International Publishing*. 309-318.

Das, S., Zook, A., and Riedl, M.O. (2015) Examining Game World Topology Personalization. Proceedings of the ACM SIGCHI Conference on *Human Factors in Computing Systems*, Korea.

Ficici, S. and Pfeffer, A. (2008). Simultaneously modeling humans' preferences and their beliefs about others' preferences. *In Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1* (pp. 323-330).

Fischhoff, B., Goitein, B., and Shapira, Z. (1981). Subjective utility function: A model of decision-making. *American Society of Information Science*. 32(5), 391-399.

Gatti, N. (2008). Game theoretical insights in strategic patrolling: Model and algorithm in normal-form, in ECAI-08 403-407.

Gal, Y., Pfeffer A. (2007). Modeling reciprocal behavior in human bilateral negotiation, In *Proceedings of the National Conference on Artificial Intelligence* -Vol. 22, No. 1, p. 815.

Gonzalez, R., and Wu, G. (1999). On the shape of the probability weighting function. Cognitive psychology - Vol 38 129–166.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47 263-292.

Kifer, D., Ben-David, Sh., Gehrke, J. (2004) Detecting change in data streams. *Proceedings of the Thirtieth international conference on Very large data bases*-Volume 30, 180-191. VLDB Endowment.

Korzhyk, D., Conitzer, V., and Parr, R. (2010). Complexity of computing optimal stackelberg strategies in security resource allocation games. In AAAI.

Nguyen, T.M., Yang R., Azaria A., Kraus S., Tambe M. (2013). Analyzing the Effectiveness of Adversary Modeling in Security Games, In *AAAI*.

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

Pita, J., Jain, M., Ordónez, F., Portway, C., Tambe, M., Western, C., and Kraus, S. (2008). ARMOR Security for Los Angeles International Airport. In *AAAI*. pp 1884-1885.

Pita, J., John, R., Maheswaran, R., Tambe, M., Yang, R., Kraus, S. (2012): A robust approach to addressing human adversaries in security games, In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3,* pp. 1297-1298.

Shieh, E., Yang, R., An, B., Tambe, M., Baldwin, C., DiRenzo, J., Maule, B., Meyer, G. (2012). PROTECT: A Deployed Game Theoretic System to Protect the Ports of the United States In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1, pp. 13-20.*

Short, M.B., D'Orsogna, M.R., Pasour, V.B., Tita, G.E., Brantingham, P.J., Bertozzi, A.L., Chayes, L.B. (2008) A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences* 18, 1249–1267.

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

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

Tversky, A. and Kahneman, D. (1992) Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Yang, R., Kiekintveld, C., Ordonez, F., Tambe, M., John, R. (2013) Improving Resource Allocation Strategies Against Human Adversaries in Security Games: An Extended Study. *Artificial Intelligence Journal (AIJ)*. 195:440-469.

Zhang, C., Jiang, A.X., Short, M.B., Brantingham, J.P. and Tambe, M. (2014). Defending Against Opportunistic Criminals: New Game-Theoretic Frameworks and Algorithms, In Conference on *Decision and Game Theory for Security pp. 3-22 (Gamesec).*