Modeling Human Bounded Rationality in Opportunistic Security Games

By

Yasaman Dehghani Abbasi

A Dissertation Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the Requirements for the Degree
DOCTOR OF PHILOSOPHY
(Industrial and System Engineering Department)
May 2016

Ph.D. Dissertations Committee

Professor Milind Tambe, Committee Chair
Professor Detlof von Winterfeldt, Committee Member
Professor Najmedin Meshkati, Committee Member
Professor Richard John, Committee Member
Professor Nicole Sintov, Committee Member

Acknowledgments

It would have been impossible to accomplish this dissertation without the help and support of many people who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, I am most grateful to my Ph.D. advisor Professor Milind Tambe, without whom I would not be anywhere near where I am now. He has been an extraordinary mentor to me, and I want to thank him sincerely for all I have achieved. His enthusiasm for pursuing problems at the highest levels of scientific integrity and rigor has been a constant source of inspiration, excitement, advice, and guidance throughout my study. I genuinely thank him for fostering my capacity critically as an independent researcher. His advice and encouragement were always important guiding lights towards my professional development. I consider myself extremely lucky to have known such a technical pioneer.

I would also like to thank my other dissertation committee members, Dr. Winterfeldt, Dr. Meshkati, Dr. John and Dr. Sintov for their insightful suggestions and for their valuable time.

Mainly thanks to Dr. Meshkati and Dr. Sintov. Thank you Dr. Meshkati for believing in me and supporting me in every step of my graduate studies at USC; words cannot describe how grateful I am to have your guidance and support during my time at USC. You are always the source of help and encouragement for all Iranian students. I would also want to thank Nicole for always supporting me technically and emotionally. It was an honor to collaborate with you.

I would also like to thank the many excellent researchers that I have had the privilege to work with including Dr. Zhang, Dr. Gonzalez, Dr. Ben-Ashef, Don Morrison, Dr. Short, Dr. Sinha, and Debarun Kar.

I also thank all those who supported me in any aspect during my Ph.D. Thank you Dr. Sima Parisay for always being a great mentor and encouraging me to be a successful woman. Thank you all my best and dearest friends and my lab mates who fill my Ph.D. with fun memories.

Last, but certainly not least, I would like to thank my lovely family and amazing husband, Mohammad, for their unconditional love and support and encouragement. Maman, Baba, I cannot begin to imagine how I could have made this journey if it weren't for your blessing. Thanks for always supporting me and being the greatest parents. Mamanhoma, Pedarjoon, Khale Anna, thank you so much for always believing in me and supporting me with your love. Ghazal, thank you and your beloved family for your kindness and encouragement, thank you for bringing Alborz to our life. No matter how far away you may be physically, you are always my source of energy. Thank you Khale Azi, for always supporting me and making me feel like home. Thank you Emranis and Ghasemazars for supporting me with your love. And finally, I would have definitely not been here if I didn't have you, Mohammad. Thank you so much for always being there for me, during the happy times with laughter and love, during the deadlines with patience and support, and during the sad times with encouragement and hope. Thank you for keeping me sane!

Abstract

Security has been an important, world-wild concern over the past decades. Security agencies have been established to prevent different types of crimes in various domains, such as illegal poaching, human trafficking, terrorist attacks to ports and airports, and urban crimes. Unfortunately, in all these domains, security agencies have limited resources and cannot protect all potential targets at all time. Therefore, it is critical for the security agencies to allocate their limited resources optimally to protect potential targets from the adversary.

Recently, game-theoretic decision support systems have been applied to assist defenders (e.g. security agencies) in allocating and scheduling their limited resources. Stackelberg Security Game (denoted as SSG), is an example of a game-theoretic model that has been deployed to assign the security resources to the potential targets. Indeed, decision-support systems based on SSG models have been successfully implemented to assist real-world security agencies in protecting critical infrastructure such as airports, ports, or suppressing crime in urban areas. SSG provides an approach for generating randomized protection strategies for the defender using a mathematical representation of the interaction between the defender and the attacker. Therefore, one of the key steps in applying the SSG algorithm to real-world security problems is to model adversary decision-making process.

Building upon the success of SSGs applications, game theory is now being applied to adjacent domains such as Opportunistic Security. In this domain, the defender is faced with adversaries with special characteristics. Opportunistic criminals carry out repeated, and frequent illegal activities (attacks), and they generally do not conduct extensive surveillance before performing an attack and spend less time and effort in planning each attack.

To that end, in my thesis, I focus on modeling the opportunistic criminals' behavior in which modeling adversary decision-making process is particularly crucial to develop efficient patrolling strategies for the defenders. I provide an empirical investigation of adversary behavior in opportunistic crime settings by conducting extensive human subject experiments and analyzing how participants are making their decisions to create adversary behavior prediction models to be deployed in many opportunistic crime domains. More specifically, this thesis provides (i) a comprehensive answer to the question that "which of the proposed human bounded rationality models best predicts adversaries' behavior in the Opportunistic Crime domain?", (ii) enhanced human behavior models which outperform existing state-of-the-art models (iii) a detailed comparison between human behavior models and well-known Cognitive Science model: Instance-Based Learning model (iv) an extensive study on the heterogeneity of adversarial behavior, and (v) a thorough study of human behavior changing over time, (vi) as well as how to improve human behavior models to account for the adversaries' behavior evolve over time.

Contents

List of Figures

List of Tables

Chapter 1. Introduction

Security is a critical worldwide concern over the years, and there have been enormous efforts to prevent different types of crimes. For instance, tremendous efforts have been dedicated to protecting critical infrastructures such as ports and airports, or to suppress crimes in urban areas and securing cyberspace.

In all these domains, there is a common challenge: the adversary might attack any target while the number of resources is limited and not all the potential targets can be protected at all time. Therefore, it is critical to have an algorithm which optimally allocates the limited defender resources to given targets, and reduces the number of crimes or attacks as much as possible. Meanwhile, the adversaries are observing and investigating about defenders' patrolling strategy, and hence, any deterministic allocation of defense resource may be exploited by these adversaries. Therefore, it is crucial for the security agencies to allocate their resources in a randomized fashion.

In recent years, Game Theory has become a well-established modeling and optimization tool for complex resource allocation problems in security and sustainability domains. Game-theoretic decision support systems assist defenders in allocating and scheduling their limited resources to protect targets from adversaries. One particular game-theoretic model is the Stackelberg Security Game (SSG) model which has received significant attention for its success in modeling security problems and application in real-world settings (Figure 1), such as scheduling patrols conducted by the US Coast Guard 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 2008).

Figure 1: Stackelberg Security Games applied to protect ports and airports

Stackelberg Security Game provides a Game Theory based representation (Tambe 2008) of the interaction between a leader (the defender) and a follower (the attacker). SSGs offers a sophisticated approach for generating unpredictable, randomized strategies that alleviate attackers surveillance ability and provides computational tools to optimize the defender's action based on possible attacker's moves (Tambe 2011, Korzyk, Conitzer and Parr 2010, Gatti 08).

In an SSG, the defender commits to a mixed strategy, which is a randomized allocation of her security resources. Then, considering the defender strategy, the attacker chooses the best response - the one which gives him the highest expected utility. In result, the defender and the attacker receive payoffs that depend on the target that was attacked and whether or not it was defended; this payoff is considered the utility of attacker or defender. The Stackelberg equilibrium computation considers the adversary's response and provides defender with a mixed strategy which maximizes her expected utility.

Following the success of applying SSG-based approaches in protecting infrastructures such as airports and ports, researchers are now studying the application of SSG to Urban Crimes (Figure

2). In this domain, it is known that adversaries are more flexible in planning and executing their attacks. The adversaries generally do not conduct extensive surveillance before performing an attack and spend less time and effort on each attack. Considering this opportunistic behavior of criminal in this domain, Opportunistic Security Game (OSG) is used to design optimal patrolling strategy for the defender.

Figure 2: Examples of Opportunistic Crimes

1.1. Problem Addressed

As mentioned earlier in this chapter, the Stackelberg equilibrium computation involves finding the mixed strategy that maximizes the utility for the defender, while taking into consideration the adversary's response. One of the key steps in SSGs is how attackers choose strategies based on their knowledge of the defender strategy. Traditionally, Stackelberg Security Game assumes a perfectly rational adversary and adopts a perfect rationality model. This model is a reasonable starting point and is justifiable when considering domains such as counter-terrorism, where the adversary has sufficient time and motivation to launch a single carefully planned attack; whereas in other domains, such as urban crime, this assumption appears weak. Indeed, it is known that adversaries in OSG domain are boundedly rational and may deviate from the optimal choice (Zhang *et al*. 2014). 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 OSG and hence significant

losses for the defender. Therefore, constructing a reliable model of the adversary behavior is vital for enhancing security against urban crime.

To address bounded rationality of human subjects (as opposed to perfect rationality assumptions), we need to answer multiple open questions. First, among various bounded rationality models presented in behavioral game theory and cognitive psychology (Camerer et al., 2004; Costa-Gomes *et al.*, 2001), which model can best represent the fundamental features of human behavior in the opportunistic crime domain? Next, how these models can be improved by considering the particular characteristic of opportunistic criminals?

In addition, in most behavioral game theory models in security games, adversary assumed to be homogenous while some recent researches have focused on the heterogeneity of human adversarial behavior. These researches have considered the heterogeneous behavior of adversary by either assuming a smooth distribution of the model parameters for the entire adversary population (Yang *et al.*, 2014) or by utilizing a single behavioral model for each adversary (Haskell et al., 2014; Yang *et al.*, 2014). However, it is unclear if, in OSG domain, defenders are dealing with heterogeneous opportunistic criminals; and if so, which approach can capture this heterogeneity the best.

Lastly, this thesis is the first to study opportunistic adversary behavior dynamics systematically. This is a significant advancement as previously, despite the widespread use of SSGs, there has been no systematic study of how adversary behavior changes over time. This new study is the first to provide such a systematic study in Opportunistic Security Game Setting.

1.2. Dissertation Contributions

This thesis focuses on an empirical investigation of human adversary behavior modeling in opportunistic crime setting. In the Opportunistic Security Game domain, to model the adversary's decision-making process in opportunistic crimes (Zhang et al. 2014), Quantal Biased Random Movement (QBRM) was previously proposed, but not evaluated. In this thesis, I compare QBRM with variations of two commonly used bounded rationality models: Quantal Response (QR), and Subjective Utility Quantal Response (SUQR) combined with Prospect Theory (PT) functions. Quantal Response (McKelvey and Palfrey 1995) models the bounded rationality of human subjects by introducing uncertainty into their decision-making process. The SUQR model is mathematically equivalent to the conditional logit model in discrete choice theory in which Subjective Utility (SU) (Fischhoff, Goitein, and Shapira 1981) and Quantal Response (McKelvey and Palfrey 1995) are combined. Prospect Theory (Kahneman, Tversky 1979) models decision making under uncertainty and captures the bounded rationality by mapping the real probability values to a person's interpretation of probabilities through a non-linear probability weighting function, while also accounting for the adversary's risk preference.

In this thesis, I have conducted extensive human subject experiments, compare various bounded rationality models, and show that: (1) although previous the research proposed the use of well-known stochastic choice Quantal Response (QR) model for 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 the original prospect theory in terms of how humans are seen to weigh 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 (4) improvement over the SUQR model by introducing two additional indicator features that refer to an attacker's preference to attack the same target (different target) when he was successful (failed) in the previous round.

Moreover, Instance-Based Learning (IBL) model have been evaluated in the Opportunistic Security Domain. The results showed the similar performance of the IBL model as the Quantal Response behavior models when the models applied to the aggregated data.

In addition, this thesis indicates that participants have heterogeneous behavior and can be categorized into multiple distinct groups, and each such group can be represented by the distinct degree of rationality. Interestingly, for the group with the highest rationality level, behavioral game theory models provide a significantly better fit compared to the IBL cognitive model.

Finally, for the first time in the opportunistic security domain, I systematically study adversary behavior dynamics. This is a significant advancement; previously, despite the widespread use of SSGs, there has been no systematic study of how adversary behavior changes over time. Furthermore, in order to account for adversary behavior change, I modified the traditional human bounded rationality models with a Bayesian update method so these models would also be able to predict behavior change over time.

1.3. Overview of this Dissertation

The rest of this thesis is organized as follows: Chapter 2 provides a background and related work on the fundamentals of Stackelberg Security Games, and the particular case of SSG: Opportunistic Security Game. The chapter also focuses on the literature review about optimal defender strategy used in designing the experiment, behavioral game theory and instance based learning models, as well as, related work on the heterogeneous adversary, followed up with some

relevant information about the human subject experiment. Chapter 3 describes the methodologies to model behavior of human adversaries with bounded rationality, as well as the experiment setup to collect data on human behavior, followed by its results. Chapter 4 describes the performance of instance-based model compared to the discrete choice model, as well as, studying the heterogeneity of the adversaries. Chapter 5 focuses on adversary behavior dynamics and modified versions of human behavior models. Finally, Chapter 6 presents the concluding points of this thesis and presents possible future directions.

Chapter 2. Background and Related Work

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). The defender protects a set of targets with a 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 represented as a vector of size $|T|$ in which each element of the 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 which 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 , x_t refers to the marginal probability that the defender protects target t, and the adversary's expected utility in attacking target *t* is given by the following equation:

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

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, the above equation assumes a perfectly rational adversary, which may be more appropriate in domains such as counter-terrorism. On the other hand, in domains such as opportunistic crime settings, the adversary's behavior may be governed by models of bounded rationality (Zhang et al. 2014).

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 later on. However, in domains such as urban crime (i.e. theft on trains), the attackers (adversary) are opportunistic, i.e., they are flexible in executing their plan and seek opportunities for crime rather than strategically planning their attacks. For example, a thief may decide not to steal if he observes a police officer, and may move to another area to seek another opportunity for committing a crime. In particular, the creation of opportunistic crime can be thought of as an interaction between three essential groups: attackers (opportunistic criminal) victim/target and defenders.

These types of crime are really important to be considered; based on the Census report, there were almost eight million such crimes in 2009 which resulted in around eleven billion dollars loss. Recent work by Zhang et al. (2014) explores a model of opportunistic attackers, the authors describe three characteristics of an opportunistic attacker: (i) opportunistically and repeatedly seeks to commit crimes; (ii) reacts to real-time information at execution time rather than planning attacks in advance; and (iii) has limited observation of defender strategy and selects the crime location with limited information and computations i.e. bounded rationality.

Figure 3: Examples of defender's patrolling strategy against opportunistic criminals

2.3. Bounded Rationality Models

Bounded rationality, first introduced by Herbert Simon (Simon, 1956) as an alternative for the mathematical modeling of decision theory. Bounded rationality implies that people have a limited ability to compute the expected utility of every single decision alternative and thus might not be able to choose the most optimum alternative; instead, they use a heuristic approach that may lead to suboptimal decision making. Simon believes this sub-optimality is due to the human limitation on the ability to processes information, as well as, formulate and solve the decision problem.

While Herbert Simon focuses on an imperial challenge of the actual decision-making process, other economists such as Ariel Rubinstein (Rubinstein 1998) focuses more on exploring the formal models and relaxing some less realistic standard assumptions in decision-making theory. He proposes bounded rationality models by emphasizing on the decision-making procedures involved, and using examples when the assumption of perfect rationality can and should be relaxed. In this section, we describe details of some of these models that have been explored in the literature, including Quantal Response, Subjective Utility Quantal Response, and Prospect Theory.

2.3.1. Quantal Response Model

Quantal Response model captures the bounded rationality of a human player through the uncertainty in the decisions making process (McKelvey & Palfrey 1995; McFadden 1976). Instead of maximizing the expected utility, Quantal Response posits that the decision maker chooses an action that gives a high expected utility, with a probability greater than another action which gives a lower expected utility. In the context of OSG, given the defender's strategy, the probability of the adversary choosing to attack target i when in target j ,$q_{i,j}$ (s), is given by the following equation:

$$q_{i,j}(s) = \frac{e^{\lambda * EU_{i,j}(s)}}{\sum_k e^{\lambda * EU_{(k,j)}(s)}}$$

where λ is his degree of rationality and $EU_{i,j}(s)$ is the expected utility (EU) of the attacker in this specific situation.

Previously, in the Opportunistic Security Game domain, Quantal Biased Random Movement (QBRM) has been proposed, but not evaluated, to model opportunistic attacker's behavior in a defender-attacker interaction (Zhang *et al.* 2014). QBRM is categorized as a Quantal Response model, and its performance is evaluated in the next chapter.

2.3.2. Subjective Utility Quantal Response

Subjective Utility Quantal Response or LensQR (Karelaia *et al*. 2008, Kaufmann *et al*. 2009, Grove *et al*. 1996) model is an integration of the Lens utility function and Quantal Response. More specifically, SUQR combined two key notions of decision making: Subjective Utility (SU) (Fischhoff, Goitein, and Shapira 1981) and Quantal Response (McKelvey and Palfrey 1995), and proposed the SUQR model. The SUQR model is mathematically equivalent to the conditional logit model in discrete choice theory.

In Subjective Utility (SU) - as suggested in behavioral decision-making (Savage 1972; Fischhoff *et al.* 1981) - the key idea is that individuals have their own evaluations of different factors during the decision-making process. In 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 for the attacker (R_t^a, P_t^a). Inspired by the idea of SU, a subjective utility function for the adversary in an SSG setting is as follows:

$$SU = w_1 x_t + w_2 R_t^a + w_3 P_t^a$$

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 SU function. In the SUQR model, the probability that the adversary chooses target t while the defender strategy is x is given by:

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

2.3.3. 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).

There are two functions in prospect theory: the probability weighting function and the value function. The probability weighting function models human interpretation of probability and suggests that people weight probability non-uniformly. More specifically, the original PT suggests that people tend to overweight 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 that capture a broad range of probability weighting functions (Gonzalez and Wu 1999) given by the following formula:

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

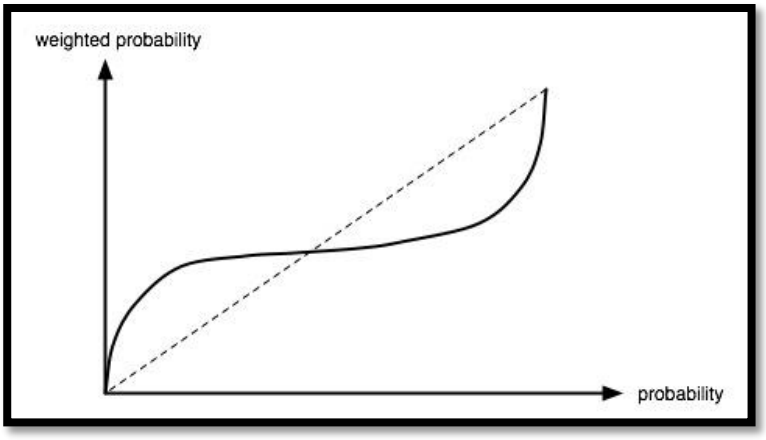

Figure 4: Prospect Theory Weighting Function

Another aspect of prospect theory is the value function, which passes through a reference point, *c*, and is S-shaped and asymmetrical and given by the following formula:

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

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.

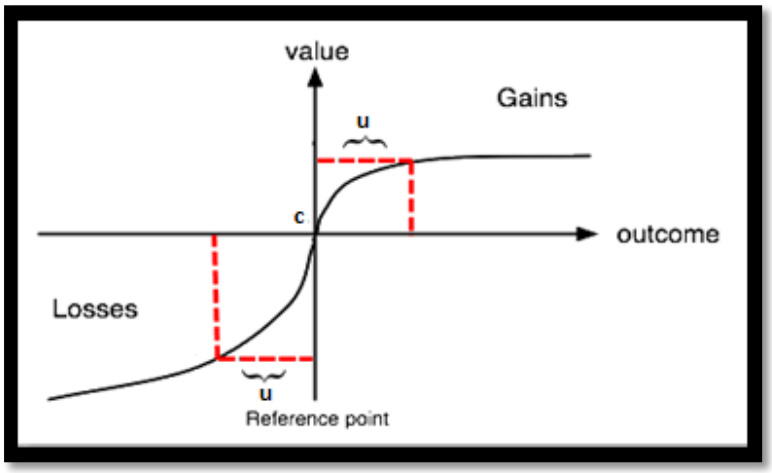

Figure 5: Prospect Theory Value function

Finally, the model assumes that humans evaluate each target by the following equation:

$$U(t,x) = f(x_t) * V(U_a^c(t)) + f(1 - x_t)V(U_a^u(t))$$

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

2.4. Instance-Based Learning Theory

Instance-Based Learning Theory (IBLT) has been a prominent approach to explain learning, and decisions from experience in a wide diversity of tasks where decisions are made from experience (Gonzalez, Lerch & Lebiiere, 2003; Gonzalez & Dutt, 2011; Hertwig, 2015). This model has demonstrated to be a general process and globally account for multiple variations of the dual choice paradigms commonly used to study decisions from experience (e.g., Gonzalez & Dutt, 2011; Lejarraga *et al*., 2012).

IBLT was developed to explain human decision-making behavior in dynamic tasks (Gonzalez *et al*., 2003), where individuals make repeated decisions attempting to maximize gains over the long run (Edwards, 1961; 1962; Rapoport, 1975). IBLT proposed that decisions in dynamic tasks were made by referencing past similar experiences and applying the decisions that worked in the past. IBLT's most important development was the description of the learning process and mechanisms by which experiences may be built, retrieved, evaluated, and reinforced during the interaction with a dynamic environment.

The model makes a choice by selecting the choice option with the highest *blended* value. An option's blended value is a weighted average of all observed payoffs on that option in previous trials. Formally, the model works as follows:

Select an option with the highest blended value. The blended value *V* of option *j* is:

$$V_j = \sum_{i=1}^{n} p_{ij} x_{ij}$$

where x_{ij} is the observed payoff (utility) in instance *i* for the option *j*, and p_{ij} is the probability of retrieving that instance for blending from memory (Gonzalez & Dutt, 2011; Lejarraga et al., 2012). The blended value of an option j is the sum of all x_{ij} stored in instances in memory, weighted by their probability of retrieval p_{ij}. The *n* value is the number of different instances containing observed payoffs on option j up to the last trial. For example, if by trial *t = 2*, option *j* revealed two different payoffs stored in two instances, then *n* = 2 for option j. If the two observed payoffs on option j are the same in the previous two trials, then only one instance is created in memory and *n* = 1.

In any trial, the probability of retrieving from memory an instance i containing a payoff observed for option *j* is a function of that instance's activation relative to the activation of all other instances that contain observed payoffs l occurring within the same option. This probability is given by:

$$p_{ij} = \frac{e^{\frac{A_i}{\tau}}}{\sum_l e^{\frac{A_l}{\tau}}}$$

where l refers to the total number of payoffs observed for option j up to the last trial, and τ is a noise value defined as $\sigma \cdot \sqrt{2}$. The σ variable is a free noise parameter expected to capture the imprecision of recalling instances from memory from one trial to the next.

The activation of each instance in memory depends on upon the activation mechanism originally proposed in the ACT-R architecture (Anderson & Lebiere, 1998). The IBL model uses a simplified version of that activation mechanism. In each trial *t*, activation *A* of an instance i is

$$A_i = \ln\left[\sum_{t_i \in \{1,\ldots,t-1\}} (t - t_i)^{-d}\right] + \sigma \cdot \ln\left(\frac{1-\gamma_i}{\gamma_i}\right)$$

where d is a free decay parameter, and t_i refers to previous trials when payoff contained in the instance i was observed (if a payoff occurs for the first time in a trial, a new instance containing this payoff is created in memory). The summation will include a number of terms that coincides with the number of times that a payoff has been observed after it was created (the time of creation of instance itself is the first timestamp). Therefore, an instance's activation containing a payoff increases with the frequency of observing that payoff (i.e., by increasing the number of terms in the summation) and with the recency of observing that payoff (i.e., by small differences in $t - t_i$). The decay parameter d affects the activation of the instances directly, as it captures the rate of forgetting. The higher the value of the d parameter, the faster the decay of instances' activations in memory is.

The γ_i term is a random draw from a uniform distribution defined between 0 and 1, and $\sigma \cdot \ln\left(\frac{1-\gamma_i}{\gamma_i}\right)$ represents the Gaussian noise that is important for capturing variability in behavior from one trial to the next. The σ variable is the same noise parameter defined in equation 3 above. A high σ implies a high noise in activation.

2.5. Heterogeneous Adversary

In behavioral game theory models, adversaries are often considered as one homogenous population, and therefore, one model is used to explain all adversaries' decision-making process. Recently, some researches have considered the heterogeneity of human adversarial behavior. These studies have achieved it by either assuming a smooth distribution of the model parameters for the entire adversary population (Yang et al., 2014), such as a normal distribution, or by utilizing

a single behavioral model for each adversary (Haskell et al., 2014; Yang et al., 2014).). However, these researches have not categorized the adversaries into distinct groups based on their attack patterns. In Chapter 4, we show that adversaries can be categorized into multiple distinct groups, and each such group can be represented by distinct degrees of rationality.

2.6. Human Subject Experiment

In my thesis, the primary focus is modeling adversaries' behavior via conducting experiments with human subjects, so it is crucial to evaluate the effectiveness of the proposed approach. For evaluating different human behavioral models, I designed an online experiment with human subjects and launched it on Amazon Mechanical Turk (AMT). Many behavioral researchers have been using AMT as a tool to collect data (Mason and Suri, 2012). Using AMT for conducting experiment provides many advantages including subject pool access, subject pool diversity and low cost (Reips, 2002; Mason and Suri, 2012). Although there are advantages of conducting experiments with a real criminal, in reality, it is often infeasible and expensive, and experimental analysis with the general population still points to the right direction.

In fact, as criminals are so different, using human subject experiment deployed on AMT and accessing pool diversity can provide us great insight about criminal's behavior. Specifically, in urban crime domain such as ticket-less travelers or thefts in the metro train system, the criminals are close to the general population and the data gathered from Amazon Mechanical Turk can be a great source to model adversarial behavior.

Chapter 3. Modeling human adversary

3.1. Opportunistic Security Experiment

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 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. 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). 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

After the game, a brief survey was used to gather data about the players' perception of the game, and risk seeking tendencies.

Main Games Description: In the main games, human subjects collect rewards by visiting any of the six stations (see an example in Figure 6), while avoiding officers on patrol. Each station has a known reward, indicated by the number of stars (e.g., as shown in Figure 6, Station 2 has four

stars). 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, nor get arrested. The player's objective is to maximize the total reward. Players must carefully choose which stations to visit, consider 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 6).

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 definite 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 the time on Station 1 in Figure 6), and determine their probability of encountering an officer at a station. During the game, players cannot observe where officers are located unless they encounter the officer at a station.

Each player starts each main game at a random station and is given a limited amount of game time (100 units) to play each game. 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 path 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).

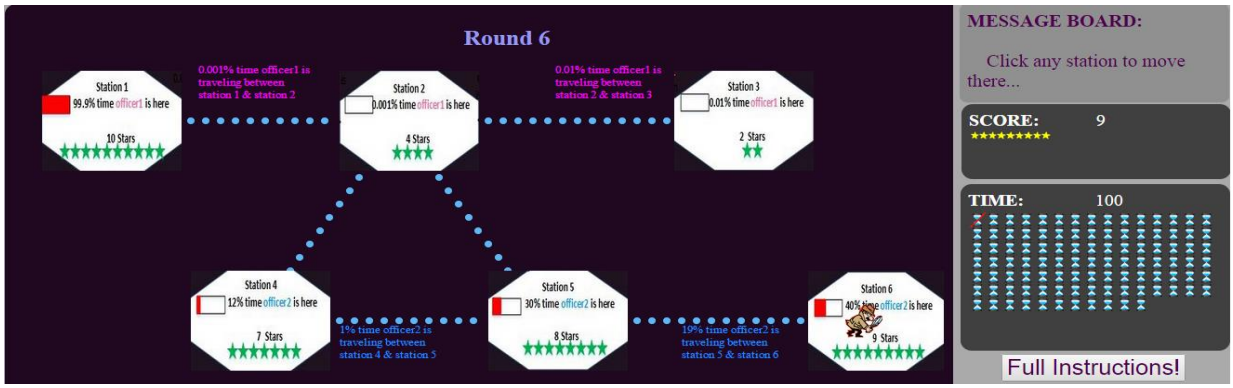

Figure 6. Game interface

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 me to determine how well they understood the game. I 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.

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 7, Figure 8, Figure 9 and Figure 10 show the layout for the four main games; two are based on real-world transportation systems, and the other two are random topologies. In the experiments, these four games were shown in random order.

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 6 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

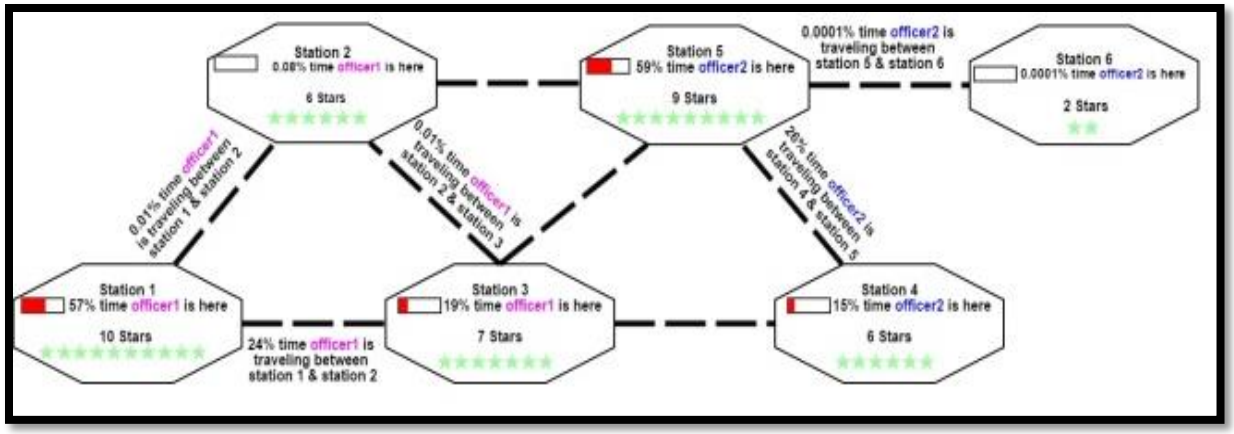

Figure 7: Main Game- Graph 1

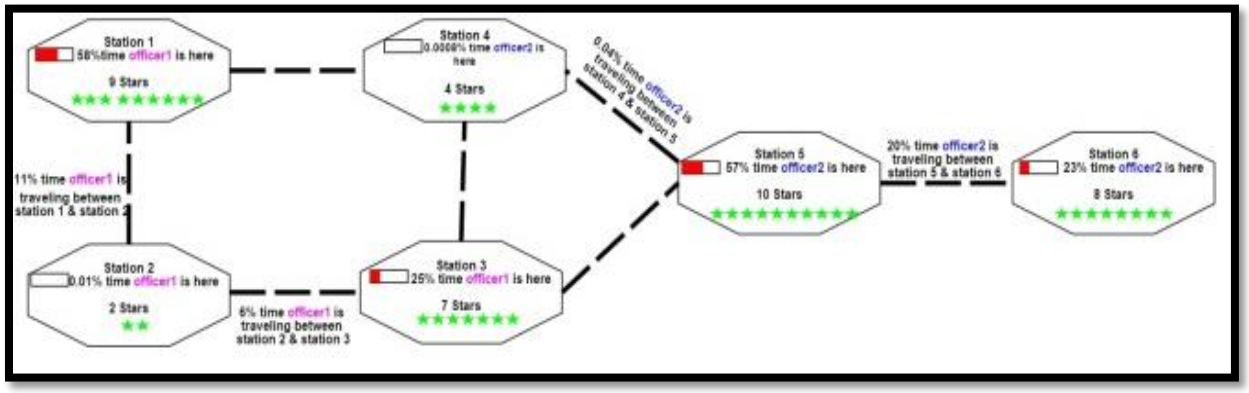

Figure 8: Main Game- Graph 2

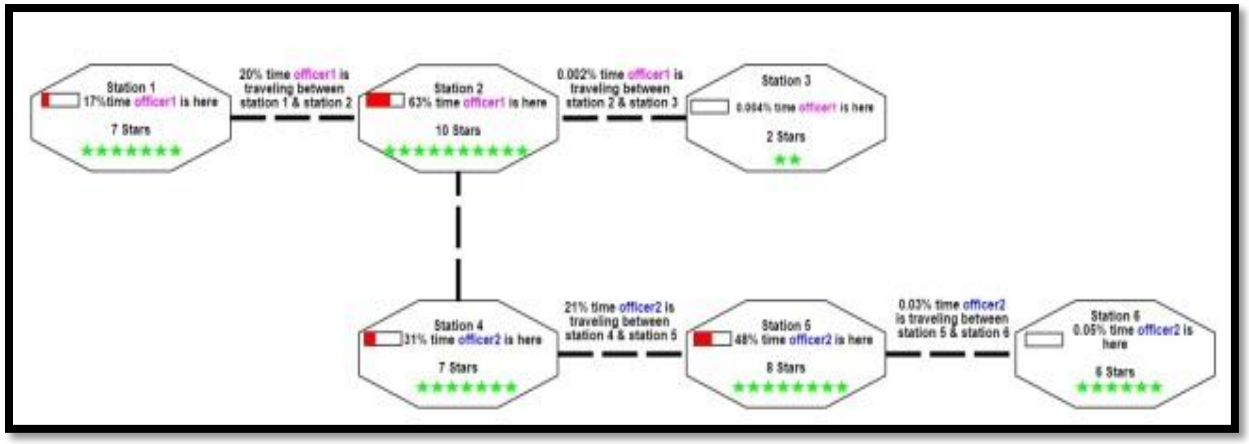

Figure 9: Main Game- Graph 3

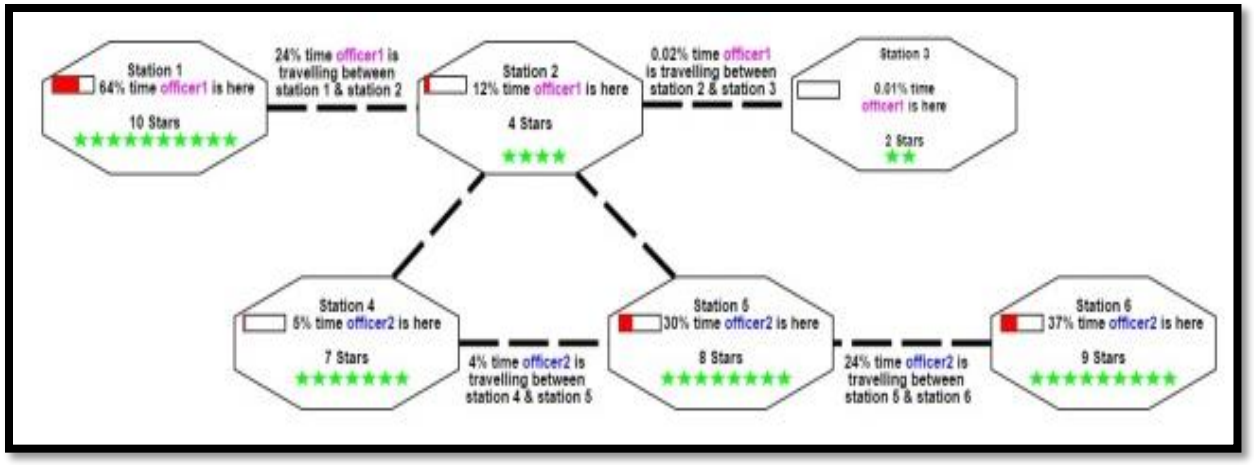

Figure 10: Main Game- Graph4

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. 218 AMT participants played the game with the average age of 35 years old where 46% were female, and 54% were male, and the majority of participants were White and Asian (65% and 22% respectively). In total, 167 unique human subjects successfully passed the validation games, and their data were used for evaluation.

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).

3.2. Optimal Defender Strategy Against Opportunistic Crime

As mentioned in the previous section, the defenders' patrolling strategies were extracted from (Zhang *et al.*, 2014). In that work, given the defender and criminal models and specific characteristics of opportunistic criminals, the Exact Opportunistic Security Game (EOSG) algorithm were used to compute the optimal defender strategy by modeling the OSG as a finite

state Markov chain. Each state of the EOSG Markov chain is a combination of the criminal's station and the defender's location. To model opportunistic criminals' behavior, Zhang et al. (2014) have used Quantal Based Random Movement, which can be considered as Power Quantal Response.

3.3. Models for Predicting Attacker Behaviors

It has been shown that addressing adversaries' bounded rationality is an essential factor in the successful performance of SSG algorithms. In previous work on opportunistic crime, QBRM was used to describe human behavior. However, to the best of my knowledge, no research has been done to descriptively study how well human behavior models capture opportunistic attacker behavior and their choices. Here, I 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. These models are listed in Table 1.

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

Category	**Model**	**Abbrev.**	**Mathematical Formulation**
Quantal Response	Logit Quantal Response Equilibrium	QR	$q_t(i) = \frac{e^{\lambda EU\,(i)}}{\sum_j e^{\lambda EU\,(j)}}$
	Quantal Biased Random Movement	QBRM	$q_t(i) = \frac{EU(i)^{\lambda}}{\sum_j EU(j)^{\lambda}}$
Subjective Utility Quantal Response	SUQR with Stationary Probability (SP)	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 SUQR-SP, 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
	QBRM with PT weighting function	PT-QBRM	Use $f(p)$ instead of p in QBRM.

Prospect Theory	SUQR-SP-C with weighting function for stationary probability	PT-SUQR-C	Use $f(p)$ instead of p in SUQR-SP.
	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.
	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

3.3.1. Quantal Response

In this category, the Quantal Response model refers to the logit quantal response formulation presented in section 2.3.1. The utility is computed as the expected reward of the adversary (expectation over the randomness in defender's strategy) in choosing a particular move (details are in (Zhang et al. 2014). In the context of OSG, given the defender's strategy s (e.g., stationary coverage probability at station i is s_i), the probability of the adversary choosing to attack target i when in target$j$,$q_{i,j}(s)$, is given by the following equation:

$$q_{i,j}(s) = \frac{e^{\lambda * EU_{i,j}(s)}}{\sum_{1 \le k \le 6} e^{\lambda * EU_{(k,j)}(s)}}$$

where λ is his degree of rationality and $EU_{i,j}(s)$ is the expected utility (EU) of the adversary when the defender strategy is s and is given by:

$$EU_{i,j}(s) = \frac{r_i}{time\ (i,j)} * (1 - s_i)$$

Where r_i is the number of stars at station $i, time\ (i,j)$ refers to time taken to attack station i when a player is in station j.

Quantal Biased Random Movement is the other model in this category which uses the power function form of quantal response. Here $q_{i,j}(s)$, is given by the following equation:

$$q_{i,j}(s) = \frac{EU_{i,j}(s)^{\lambda}}{\sum_j EU_{i,j}(s)^{\lambda}}$$

3.3.2. Subjective Quantal Response

In this category, instead of expected utility, the subjective utility is used. Subjective utility (SU) is the combination of the key factor. In the OSG experiment, the key factors I found to be the

most important are number of stars at destination station or stations i (also called the reward, or *r*), stationary coverage probability (referred to as SP or *s*) or projected coverage probability (referred to PP or *proj*) of destination station, and time taken to visit station i from station j ($time_{i,j}$), and the connectivity degree of the destination station (*con*). Thus, for examples in the SU in SUQR-SP is defined as

$$SU_{i,j}(s) = \sum_{k} w_k f_k(i,j) = w_{att} att_i + w_{sta} sta_i + w_{dis} time_{i,j} + w_{con} con_i$$

Moreover, the probability of attacking/visiting target i when the the attacker/player is at station j and the defender patrolling strategy is s is formulated as:

$$q_{i,j}(s) = \frac{e^{SU_{i,j}(s)}}{\sum_k e^{SU_{(k,j)}(s)}}$$

In this thesis, I considered three variations of non-Bayesian SUQR models. The first uses a single set of weights whether the attacker currently observes the officer or not. The second SUQR, SUQR-C, uses two sets (conditional or C) of weights, one when the officer is observed and the other when not.

The third SUQR (SUQR-S&F) model has two additional indicator features which refer to an attacker's preference to stay at (leave) the current station when he successfully (unsuccessfully) attacked the station. This is to test the hypothesis that if the player successfully attacked a station, he gives additional positive weight to that station and boosts his probability of attack at the same station in the next round. On the other hand, if the player failed to attack his current station, he gives negative weight to the current station which boosts his probability of attacking other stations. This phenomenon, often referred to as "repeat-victimization", is well-known in criminology (Short *et al.,* 2009), making it of interest in the OSG domain.

3.3.3. Prospect Theory

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 weighting function in section 2.3.3. In PT-SUQR-C-VA, in addition to the weighting function for coverage probability, the value function was also used for the number of stars at each station (rewards). In PT-SUQR-C-VRA, instead of using the absolute value of the rewards in the value function, I 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.

3.4. Model Prediction Accuracy

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

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.

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

Weighted Absolute Percentage Error (WAPE): Although RMSE is often used in statistical modeling, it is more sensitive toward outliers, and WAPE can provide a more accurate measure of model fit in these situations

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

Akaike information criterion (AIC): AIC provides a measure of the relative quality of statistical models; the lower the value, the better the model. The metric rewards goodness of fit (as assessed by the likelihood function), and penalizes overfitting (based on the number of estimation parameters).

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

Student's t-test: I used Student's t-test to evaluate the prediction accuracy of the proposed models. Student's t-test was used to study if one model is significantly better than another model.

3.5. Experimental Results

3.5.1. Extraction of Model Parameters

We divided the participants into training (70%) and testing (30%) groups. To extract various models' parameters, data belongs to the participants in training group was used. On the other hand, the data belongs to the participants in test group data kept aside for evaluation of different models.

To learn the Prospect Theory parameters, the training data set was randomly divided (80/20% split) 10 times to obtain ten 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. The models' parameter are presented in Table 2.

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	$< w_{att}, w_{sta}, w_{dis}, w_{con} >$ = <0.3853,-4.6033,-0.7031,0.145>
SUQR with PP	$< w_{att}, w_{proj}, w_{dis}, w_{con} >$ = <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 , δ=1.8, γ=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>

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 the absolute reward does not provide any improvement over *PT-SUQR-C*. Hence, we did not consider *PT-SUQR-C-VA* in our evaluation further on.

After normalizing the weights by their maximum value (multiplying w_r by 10, w_{time} with 5 and w_{con} by 4), it can be seen that attractiveness, coverage probabilities, and time are the most important factors. In the following chapters, these three factors are included for different variations of SUQR model.

3.5.2. Results

After determining the parameters of the models, we used the resulting models and test data set to predict human decisions and compared models to actual decisions on the test data set. The following are our most important 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. The 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 chapter, 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 superior performance to QR.

- SUQR outperforms QR model. Previously, (Cui *et al*. 2014) performed a comparison between QR and SUQR models in the SSG domain. My work provides a comprehensive comparison between QR and SUQR in OSG domain; in particular, it compares QR with four different versions of SUQR examining over 167 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 their decision rationally; since QR is based on expected utility (an entirely 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	0.07	0.04	0.005	0.000

Prospect Theory is a landmark theory, and the following four observations relate to the 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. 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 shown in the table, 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

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 before, addressing the human perception of *absolute* attractiveness provides no improvement over PT-SUQR-C model.

- 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 mean they do not overfit 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 three 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 11 demonstrates the weighing probability functions which indicates that human subjects' perception of small probabilities (less than 0.13) is below the actual values while their perception of high probabilities is above the real 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 12 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.

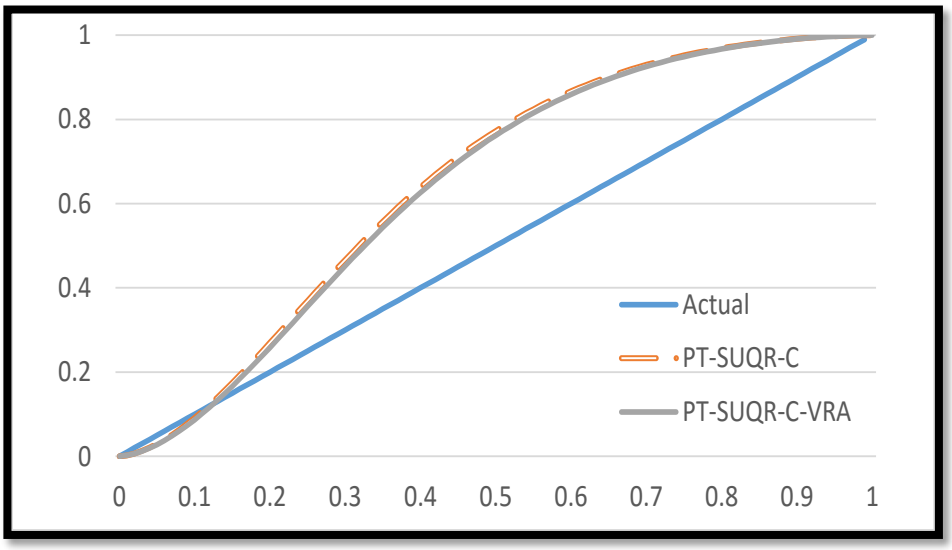

Figure 11: Extracted Prospect Theory Weighting Function

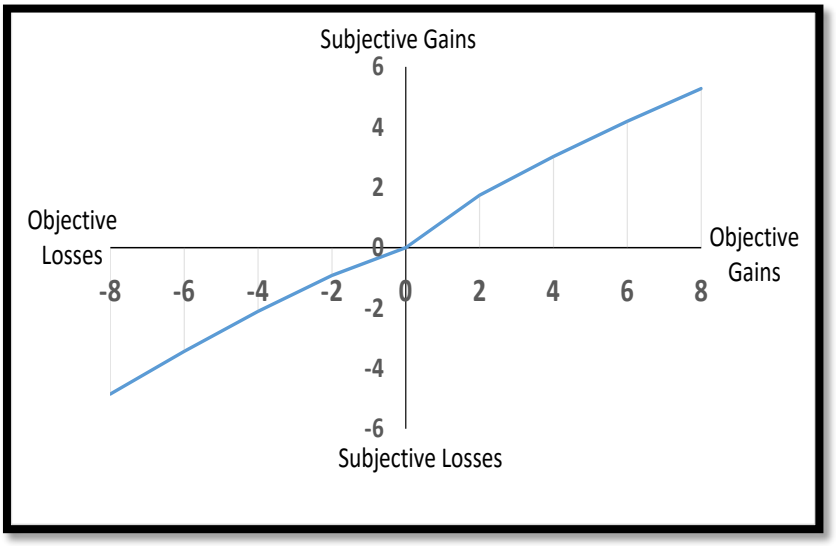

Figure 12: Extracted Prospect Theory Value function

- Adding additional weights for player success/failure results in further improvement of SUQR model. In (Kar et al., 2015), the authors investigate a human behavior model in the green security domain and reveal that the attacker adapts his future action based on past success and failure. They show that adversaries who have succeeded in attacking a target in one round tend to attack the target with "similar" features in the next round. Figure 13, Figure 14, and Figure 15 reveal similar results in our urban crime experiment. We used the data for three experiments that each include four graphs. The y-axis in the graph denotes the percentage of (i) attacks on the same station out of the total successful attacks in the previous step and (ii) attacks on the same station out of the total failed attacks in the last step. As shown in Figure 13, when the data for all the stations is aggregated, we get similar results as in (Kar et al., 2015); the players who have attacked a station successfully tend to stay at that station more than those who failed in their attack. Moreover, when the aggregated data is used, differences between the percent of successful people who stayed at the current stations and the percent of participants who failed and remained at the

current station is statistically significant (using Student' t-test with 95% CI). This difference motivates the addition of two further weights to SUQR-C (the winner of six previously used model) and results in SUQR-S&F: w_fail and w_succ. w_fail (w_succ) refers to the additional weight to staying at the same station when the player failed (succeeded). As expected, w_fail is negative (-0.9666) and w_succ is positive (0.1452). Table 3 also shows that the SUQR-S&F results in better prediction accuracy in all four categories.

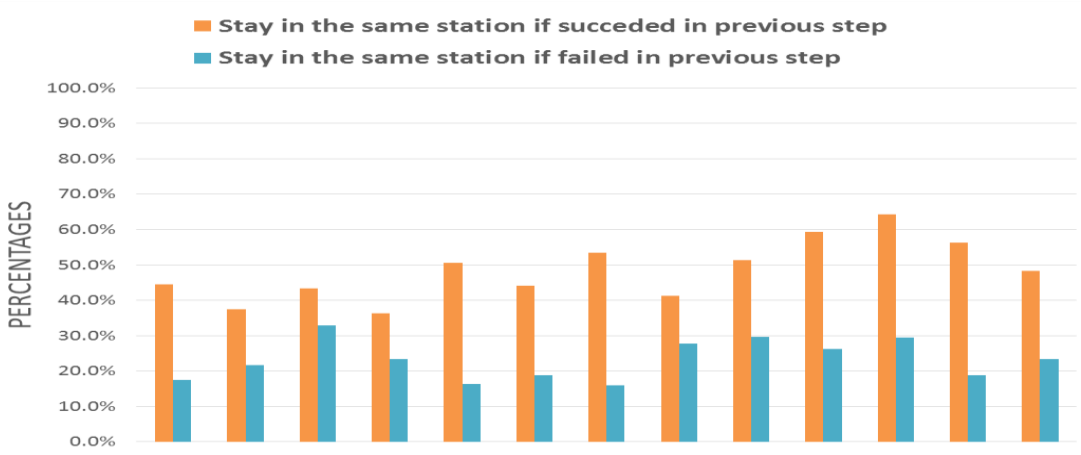

Figure 13: Attacking at the same stations

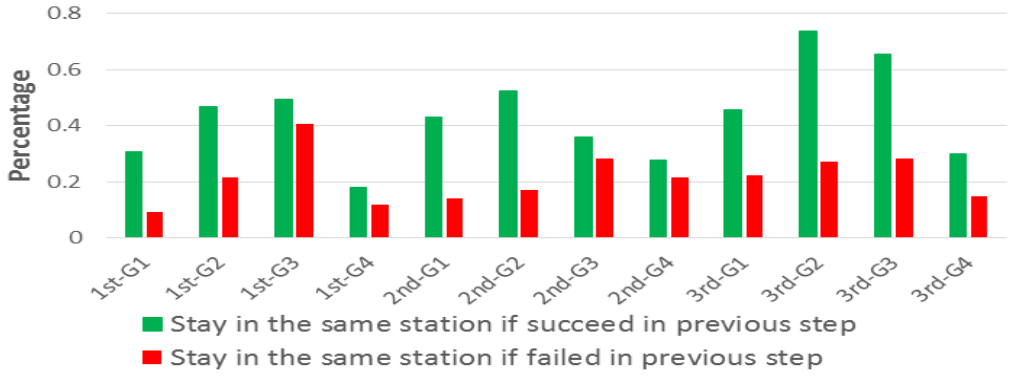

Figure 14: Attacking the same station- high reward stations

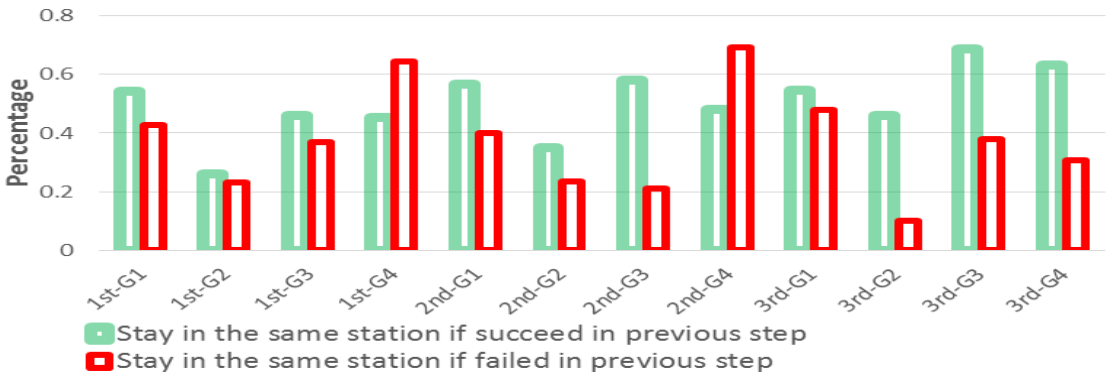

Figure 15: Attacking the same stations - Stations with Low Rewards

3.6. Summary

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. In this chapter, for the first time, I provide 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 ten 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 (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. (e) SUQR-S&F model significantly outperformed quantal

response in predicting player behavior, thus further indicating that human decision making is not based on maximizing expected utility. These and other findings outlined in this chapter provide valuable 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.

Chapter 4. Heterogeneity in Adversary

In the previous chapter, I focused on bounded rationality models from behavioral game theory (such as Quantal Response and Subjective Utility Quantal Response) while a homogeneous adversary population is assumed, and a single adversary behavior model is prescribed. Actually, in security games, mostly homogenous adversary models have been studied, but some recent research has considered the heterogeneity of human adversarial behavior. So in contrast to the previous chapter, this chapter focuses on the heterogeneity in adversary behavior.

There are two strands of previous work related to this study. First, in behavioral game theory models, they have achieved it by either assuming a smooth distribution of the model parameters for the entire adversary population (Yang et al., 2014), such as a normal distribution or by utilizing a single behavioral model for each adversary (Haskell et al., 2014; Yang et al., 2014). However, they have not categorized the adversaries into distinct groups based on their attack patterns. In this chapter, we show that adversaries can be categorized into multiple distinct groups, and each such group can be represented by distinct degrees of rationality.

The second strand of related work is with respect to the exploration of available options, which is an important aspect of decision making in many naturalistic situations (Pirolli & Card, 1999; Todd, Penke, Fasolo, & Lenton, 2007; Gonzalez & Dutt, 2011). In line with previous work (Hills & Hertwig, 2010; Gonzalez & Dutt, 2012), in this chapter, we show that there is a negative relationship between exploration behavior and maximization of rewards. However, in their work, they did not contrast behavioral models with cognitive models and did not provide insights for behavioral game theory models which we provide. In particular, we study the relationship between exploration and human reward maximization behavior by parameters of bounded rationality

models of human adversaries. Our observations are also on the security games domain where this kind of relationship between exploration behavior and maximization of rewards has not been studied before. Furthermore, in our work participants were shown all relevant information, such as rewards about all the alternative choices, while in earlier work participants had to explore and collect information about various alternatives.

In this chapter, I focus on the heterogeneity in adversary behavior. The results are based on the OSG experiment explained in Chapter 3. The results show that adversaries can be naturally categorized into distinct groups based on their attack patterns. For instance, while one group of participants (about 20% of the population) is seen to be highly rational and taking reward maximizing action, another group (almost 50%) is seen to act in an entirely random fashion. In this chapter, I show that considering distinct groups of adversaries lead to interesting insights about their behavioral model, including the defender strategies being generated based on the learned model.

To model the different categories of human behavior, I used previously presented behavioral game theory model and cognitive models. In behavioral game theory models, we have explored models such as the popular Quantal Response (McKelvey & Palfrey 1995) and the Subjective Utility Quantal Response models (Nguyen et al., 2013). In addition, based on the tradition of Cognitive Science, we use a model derived from a well-known cognitive theory, the Instance-Based Learning Theory (IBLT) (Gonzalez, Lerch, & Lebiere, 2003), developed to explain human decision making behavior in dynamic tasks and used to detect adversarial behaviors (Ben-Asher, Oltramari, Erbacher & Gonzalez, 2015). This is the first such use of cognitive models in security games.

In summary, in this chapter we build on the existent literature of security games and adversary behavior modeling by: (i) investigating the heterogeneity of adversarial behavior in an experimental study designed for OSGs, by categorizing adversaries into groups based on their exploration patterns; (ii) comparing computational models and showing the impact of heterogeneity on future behavior prediction; and (iii) indicating the effect of considering heterogeneity on the defender strategies generated.

4.1. Human Adversarial Behavior

Using data from all the main games, Figure 16 illustrates the distribution of attacks (i.e., moves) from all participants (black bars) on stations ranked by the participants' expected utility (average earning per time)[1], as well as attacks of five randomly selected individuals (P1 to P5). To normalize the utility scores among graphs, we have used the ranking of stations' utility (utility rank) instead of its absolute value (the highest utility in a graph is ranked 1). The figure illustrates significant heterogeneity behavior among individuals (line charts), and comparison to the average behavior (bar chart).

[1] $EU = (1 - stationary\ coverage) * reward/\ time$

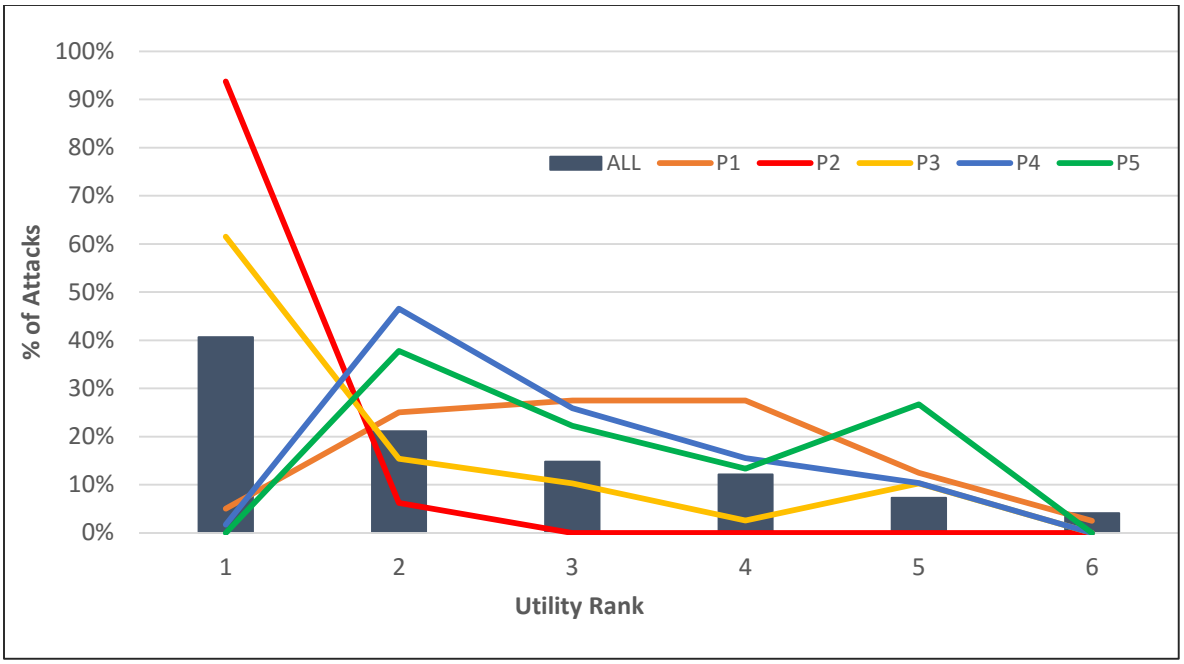

Figure 16: Percentage of Attacks on utility rank

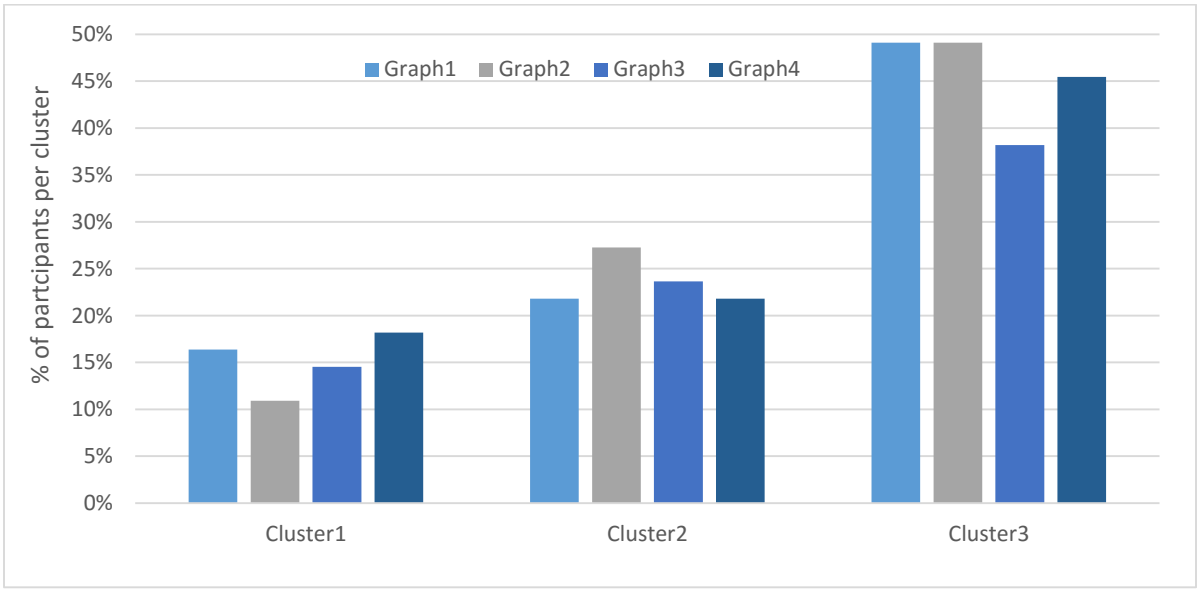

Figure 17: Clustering Distribution

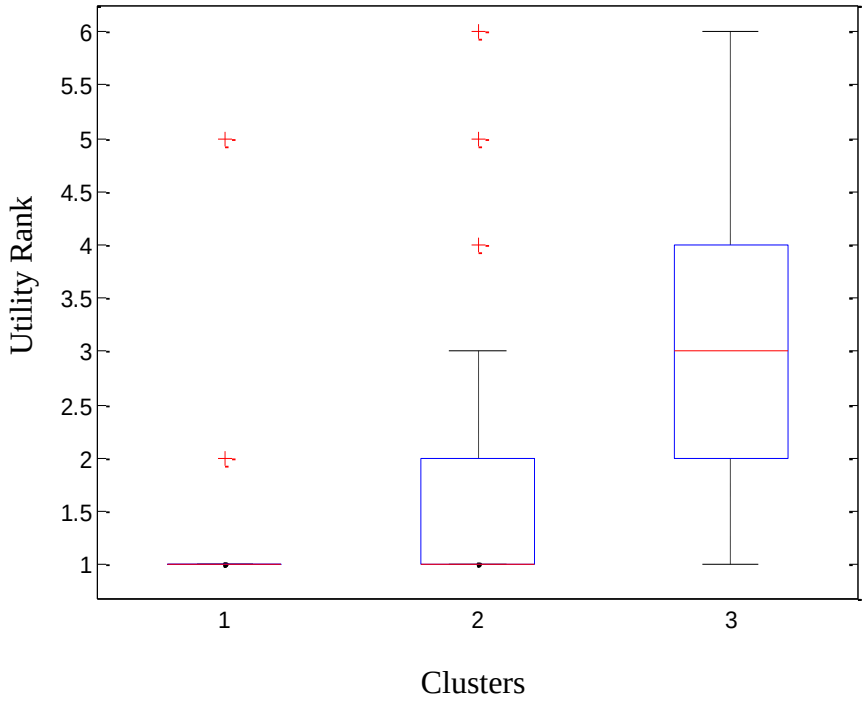

Figure 18: Utility Rank by Cluster

Given this heterogeneous behavior, we have applied the Hierarchical Clustering algorithm (Johnson, 1967) on different features related to an individuals' exploration behavior and found that *mobility score* was the best feature to cluster the participants. The *mobility score* is a measure of exploration: it is a ratio of the number of movements between stations over the number of trials (total number of movements) by a participant in the game. Figure 19 shows the distribution of participants based on their mobility score for each graph. The mobility score varied widely (0% to 100%) with a significant proportion of participants at the two extremes. Informally, the exploration behavior seems to fall into three categories: (i) those who did no exploration; (ii) those who always explored and (iii) those who engaged in a middling level of exploration. Indeed, the clustering algorithm resulted in three groups of participants: participants whose mobility score is less than

10% belong to Cluster1, participants with 10% to 80% mobility score belong to Cluster2, and participants whose mobility score is greater than 80% belong to Cluster3

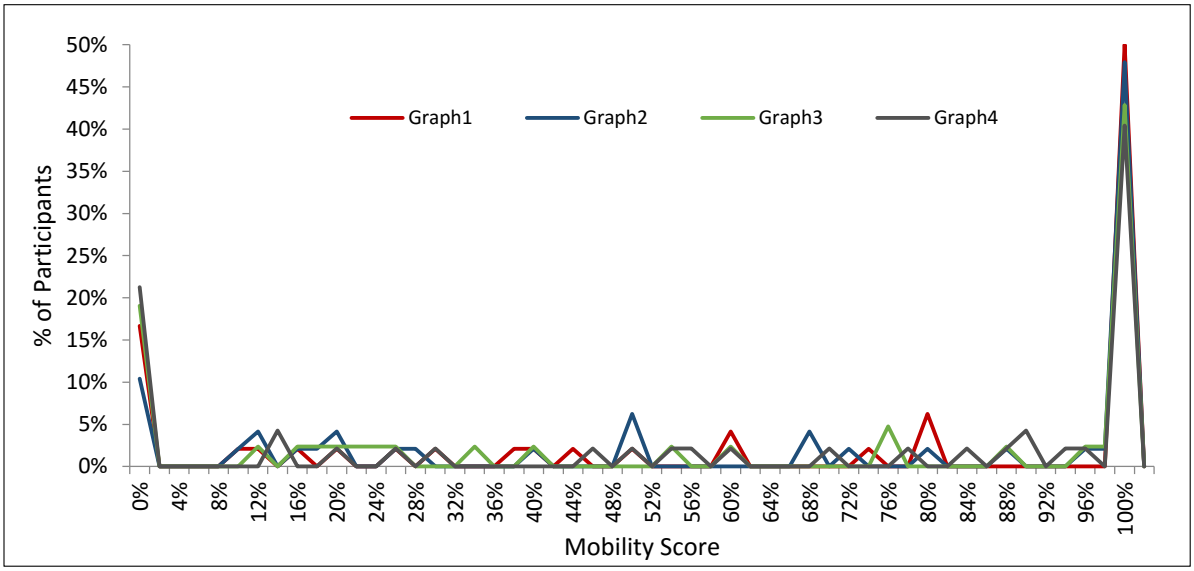

Figure 19: % of participants based on the Mobility Score

Figure 17 shows the percentage of participants belonging to each cluster for four different graphs (Graph 1 to Graph 4). The percentage of participants belonging to each cluster is nearly consistent across all graphs: approximately, 20% in Cluster1, 30% in Cluster2 and 50% in Cluster3.

In Figure 18, using the data from all the graphs per cluster, we show the distribution of utility ranks for each of the three clusters. Interestingly, mobility scores were highly correlated with the utility ranks of the attacked stations ($R^2 = .85$ & $p < .01$). We observe that participants in Cluster1 (the lowest mobility scores), attacked stations with the highest utility (average utility rank of 1.04). In contrast, participants in Cluster3 (the highest mobility score), attacked stations that varied more widely in the utility rank (average utility rank of 3.3). Participants in Cluster2 also attacked a variety of stations but were leaning (on average) towards higher utility rank stations (average

utility rank of 1.7). These observations provide interesting insights for building defender strategies, as illustrated in Section Model Results.

Following, I will provide a brief description of Instance-Based Learning Model and comparison between human bounded rationality models and IBL model in each category and overall.

4.2. Instance-Based Learning Model

In what follows, we present a well-known cognitive science named Instance-Based Learning Model. The IBL model of an adversary in the OSG makes a choice about the station to go to, by first applying a randomization rule at each time step:

If drawn from U(0,1) >= Satisficing threshold

Make a random choice

Else;

Make a choice with the highest Blended value.

This rule aims at separating highly exploratory choices from those made by the satisficing mechanism of the IBL, the Blended Value. *Satisficing* is a parameter of this model. The Blended value *V* represents value of attacking each station (option *j*):

$$V_j = \sum_{i=1}^{n} p_{ij} x_{ij}$$

where x_{ij} refers to the value (payoff) of each station (the number of stars divided by time taken) stored in memory as instance *i* for the station *j*, and p_{ij} is the probability of retrieving that instance for blending from memory (Gonzalez & Dutt, 2011; Lejarraga et al., 2012) defined as:

$$p_{ij} = e^{\frac{A_i}{\tau}} \Big/ \sum_l e^{\frac{A_l}{\tau}}$$

where l refers to the total number of payoffs observed for station j up to the last trial, and τ is a noise value defined as $\sigma \cdot \sqrt{2}$. The σ variable is a free noise parameter. The activation of instance i represents how readily available the information is in memory:

$$A_i = ln \sum_{\substack{t_p \\ \epsilon\ observed}} (t - t_p)^{-d} + \sum_{\substack{Attribute \\ \epsilon\ Situation}} P(M_{Attribute} - 1) + \sigma \ln(\frac{1 - \gamma_{i,t}}{\gamma_{i,t}})$$

Please refer to (Anderson & Lebiere, 1998) for a detailed explanation of the different components of this equation. The Activation is higher when instances are observed frequently and more recently. For example, if an unguarded nearby station with many starts (reward) is observed many times, the activation of this instance will increase, and the probability of selecting that station in the next round will be higher. However, if this instance is not observed often, the memory of such station will decay with the passage of time (the parameter *d*, the decay, is a non-negative free parameter that defines the rate of forgetting). The noise component is a free parameter that reflects noisy memory retrieval.

4.3. Model Results

We aggregated the human data and divided the data set into two groups: training and test datasets. The data from the first three graphs played by the participants were used for training and the last graph played was used for testing the models. This resulted in 1322 instances in the training set and 500 instances in the test data set. For comparison of different models, we use Root Mean Squared Error (RMSE) and Akaike Information Criterion (AIC) metrics.

Table 6 shows the results on the full data set. The model parameters obtained from the training dataset were used to make predictions on the test dataset. The prediction errors from all the models are relatively similar, even though they provide different perspectives. QR and SUQR predict the stable state transition probabilities of the attacker while the IBL is a process model that captures learning and decision dynamics over time. We also examine the performance of different models for each cluster (Table 7).

Table 6: Metrics and Parameter on the full data set

Model	*Parameters*	*RMSE*[2]	*AIC*
QR	0.4188	0.25	3962
SUQR	<3.97,-2.51,-2.55>[3]	0.23	3685
IBL	<1.4, 3.2,0.3>[4]	0.24	4359

The value of λ (higher value of λ corresponds to higher rationality level) in the QR model decreases significantly from Cluster1 (high value of λ=1.81) to Cluster3 (λ=0). These findings are consistent with our observation of the utility ranks of targets chosen by adversaries in each cluster, as shown in Figure 18.

This is significant because past research has assumed that all participants either behave based on an average value of λ or that each individual's value of λ can be sampled from a smooth distribution. In this study, however, we show that a significant number of participants (70%: 20% in Cluster1 plus 50% in Cluster3) have values of λ which fall at two extreme ends of the spectrum.

[2] the average is over 288 entries, representing moves from any of 6 stations to any other station, in four graphs, and for two cases where the player observes the officer or not

[3] $w = < w_{re}, w_{sta}, w_{time} >$

[4] <noise, decay, Satisficing threshold >

Thus their behavior should be modeled either as perfectly rational or completely random adversaries.

Table 7: Metrics and Parameters on each Cluster

	Model	*Parameters*	*RMSE*	*AIC*
Cluster1	QR	1.81	0.01	52
	SUQR	<7.16,-4.53, -13.43>[3]	0.06	67
	IBL	<2.3, 0.9, 0.9>[4]	0.27	238
Cluster2	QR	0.6582	0.28	1023
	SUQR	<5.63,-3.14 ,-4.16>[3]	0.27	927
	IBL	<0.9, 1.4, 0.8>[4]	0.30	1821
Cluster3	QR	0	0.26	2188
	SUQR	< 1.9,-1.1,0.13>[3]	0.23	2007
	IBL	<0.01, 1.8, 0.1>[4]	0.27	2529

Moreover, because SUQR weights indicate the importance of each attribute to the decision maker, the results of SUQR parameter extraction for different clusters reveal some interesting points. First, the fact that Cluster1 has the largest weights for all attributes (in the absolute terms) implies that Cluster1 participants are very attracted to the stations with high rewards and highly repelled by high defender coverage; which conforms with the observed behavior of Cluster1 participants in maximizing the expected utility.

Second, although SUQR outperforms QR overall and in Cluster2 and 3, QR has lower prediction error (statistically significant for paired t-test at $t(288) = 02.34$, $p<0.01$) on data for

Cluster1. This is intuitive if participants are utility maximizers, this would be captured better when in the QR model. On the other hand, a model like SUQR which reasons based on different features of the game capture better the propensity of the participants to switch between stations, and hence perform better on Clusters 2 and Cluster 3 where participants do not have a clear movement pattern.

Therefore, identifying different groups of adversaries (with various degree of rationality) gives us valuable insight into the types of behavioral models that can be used in various scenarios to generate accurate future predictions.

The results from the IBL model suggest that the categories of adversaries found in this study do not emerge naturally from the learning process. Indeed, in this study participants had little opportunities to learn. Instead, it appears that participants either use the information readily available to them in the OSG and attempt to maximize their gains, or they explore the choices randomly which may lead them to less optimal decisions.

These two modes of behavior were captured in the IBL model by a meta-rule with a Satisficing parameter. This meta-rule is not part of the IBL model, but it helps to overpass the natural choice by Blending (similar to the Inertia meta-rule used in Gonzalez & Dutt, 2011). This meta-rule was added to explicitly account for random exploratory behavior observed in the OSG. Therefore, the Satisficing parameter helps in selecting between the two modes of behavior to form the different clusters. The Satisficing parameter is highest in Cluster1, lower in Cluster2, and lowest in Cluster3. Cluster1 results from most choices being made by the IBL's Blending while Cluster3 results from a random choice. However, this parameter interacts with the IBL model's decay and noise parameters. For example, in Cluster1, most decisions are made for the station with highest Blended value, and there is a need for a high noise value to introduce the variability found in human

behavior. In contrast, choices in Cluster3 are mostly made randomly, but in the rare occasions when the model makes choices based on the highest Blended value, it attempts to benefit from recent past experiences (i.e., low decay) and with low noise to the decision processes. Therefore, identifying such meta-rules for accounting for explicit descriptive information in addition to the IBL model's learning mechanisms is an important aspect of capturing adversary behavior in security games.

It is interesting to observe that the behavioral game theory models provide a significantly better fit in Cluster1, compared to the IBL cognitive model, while the values of behavioral game theory models are comparable to those of the IBL model in Clusters 2 and 3. The IBL model, being a learning model, is poor at making highly accurate decisions with little or no experience as in the OSG study.

Finally, to demonstrate the impact of considering distinct heterogeneous groups of adversaries, we consider one of the most recent works (Kar et al., 2015) which advocated the use of a homogeneous adversary model. Using their data, I show there is a significant difference between the defender strategies generated by a homogeneous (SUQR) and a heterogeneous model which considers three distinct clusters (Bayesian SUQR). For example, for target 16 in Figure 20, the change in coverage probability from defender strategy generated against a homogeneous to that against a heterogeneous model is 110%.

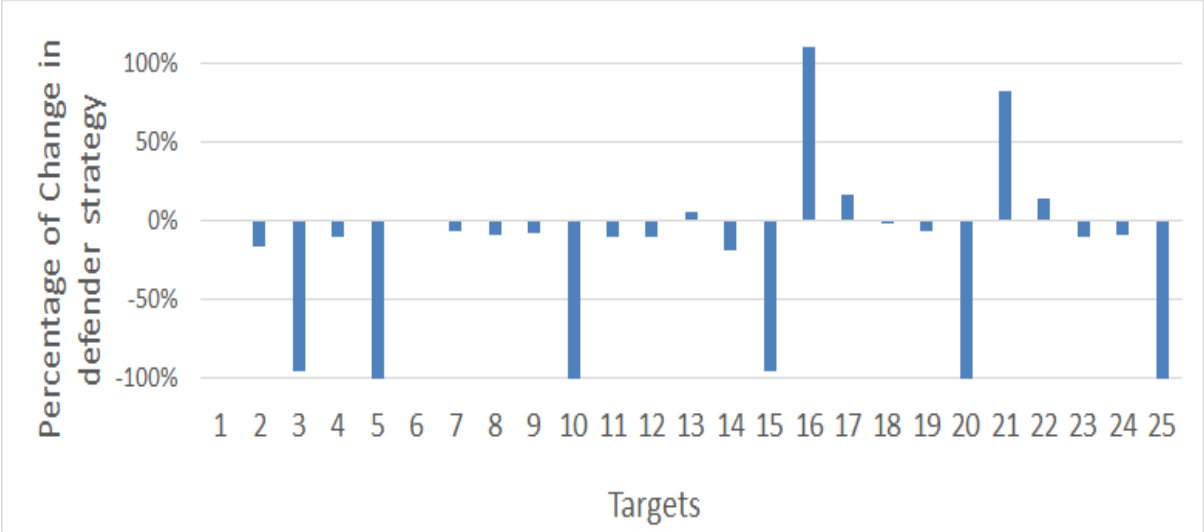

Figure 20: Strategy against homogeneous & heterogeneous

4.4. Summary

Significant research has been conducted towards understanding human adversary behavior in security games, which has led to several deployed real-world applications (Tambe 2011), such as PROTECT for the protection of major ports in the US by the US Coast Guard (Shieh et al. 2012). Moreover, ARMOR for scheduling of police patrols at major airports such as LAX (Pita et al. 2008).

Although these researches in security games have relied on modeling adversaries via a single homogeneous model or a heterogeneous model with a smooth distribution over model parameters, in this chapter, I showed the heterogeneity in adversary behavior by clustering adversaries into distinct groups based on their exploration patterns.

We found that three distinct clusters emerged based on the adversaries' exploration patterns, two of which fall at two extreme ends of the spectrum, capturing perfectly rational and completely random adversarial behavior while people in the third cluster exhibit some behavior in between.

We also observed that in our OSG domain, exploration is negatively correlated with utility maximization.

We demonstrate that accounting for the diversity of adversary behavior leads to very different model parameters and can provide more accurate predictions of future behavior. Specifically, we show on data collected based on an Opportunistic Security Game that (i) the QR model captures the behavior of utility maximizing adversaries better than the SUQR or IBL based models; (ii) the behavioral (QR, SUQR) and cognitive models (IBL) have similar prediction performance for adversaries who do not act in a perfectly rational fashion (iii) furthermore, we show that considering the heterogeneity in adversary behavior leads to different defender strategies being generated. The effectiveness of such strategies is an important area of future work.

Chapter 5. Adversary Dynamics of Behavior

Despite the widespread use of Stackelberg Security Games, there has been little research on how adversaries adapt to defense strategies over time (i.e., dynamics of behavior). In this chapter, I advance the study of human behavior by showing how adversaries' behavior changes as they learn the defenders' behavior over time. Furthermore, I show how behavioral game theory models can be modified to capture learning dynamics using a Bayesian Updating modeling approach.

More specifically, this chapter starts by addressing possible limitations in the previous chapter. In the previous experiment, participants' exploration in their experiment was, to some extent, determined by a random termination rule in the game. That is, the number of decisions that participants could make was, at least in part, determined by chance. This particular effect may have contributed to the variability in exploration processes observed. This chapter presents a new experimental study using the same OSG setting employed in previous chapters but enforced a fixed number of decisions for all participants. Furthermore, the larger number of trials in the present study enables the study of human behavior over-time. This new study is the first to provide a systematic study of how adversary behavior changes over time in Security Game Setting. Moreover, building on the previous chapter and using distinct adversary categories, in this chapter, I show how adversary behavior shifts among these categories as adversaries learn over time.

Furthermore, to account for adversary behavior change, I modified the traditional human bounded rationality models with a Bayesian update method, so that these models would also be able to predict behavior over time. These models are compared to an Instance-Based Learning (IBL) model that provides a cognitively-plausible account for overtime behavior in the OSG.

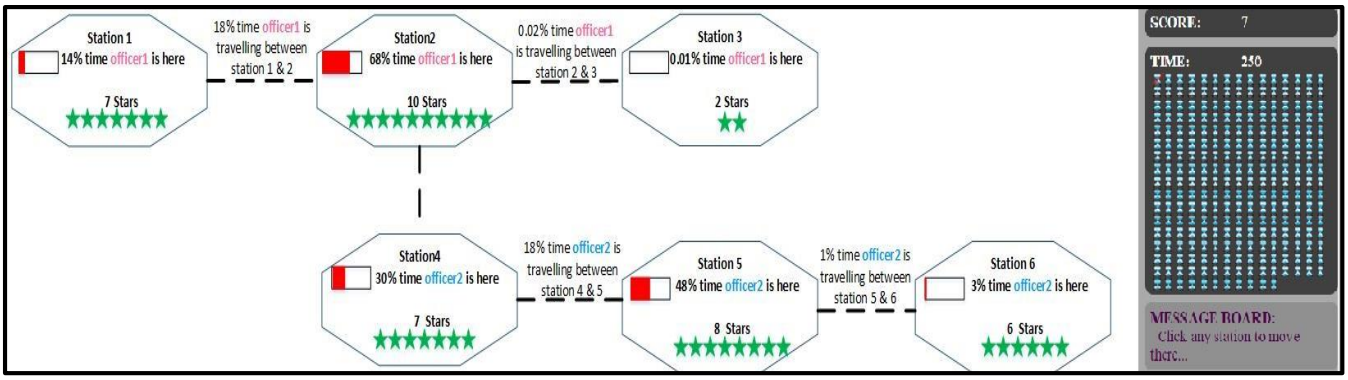

Figure 21: Experiment Game Interface

5.1. Modified OSG Experiment

To collect data on adversarial behavior in the OSG domain, we adopted the experimental design used in previous (Figure 21).

5.1.1. Experimental Procedure

In the experiment, the players' goal is to maximize their score by collecting rewards (represented by stars in Figure 21) in limited time while avoiding officers on patrol. Each player can travel to any station, including the current one, by train as represented by the dashed lines in Figure 21. By going to a station, the player attacks that station and collects the rewards.

Two officers are protecting these six stations. Each officer protects three stations where his patrolling strategy (i.e. probability of officers' presence at each station or route,) is determined by an optimization algorithm similar to the one presented in (Zhang et al. 2014). This algorithm utilizes opportunistic adversary behavior model to provide optimized defender strategies.

The stationary coverage probabilities for each station and trains are provided to the players, so players can determine the chance of encountering an officer at a station by considering the percentage of the time that officers spend on average at each station and on a train. However,

during the game, the players cannot observe the officers unless they encounter the officer at a station.

The game can finish either if the player uses up all the 250 units of available time in each game, or the game is randomly terminated after the 50^{th} attack with a 10% probability shown by the randomized terminator. To encourage the participants to make each decision responsibly, we have employed deception and the randomized terminator is shown to players from the 1^{st} attack to encourage them to choose each action carefully, as they believe there is a chance the game may terminate after each attack.

Thus, a players' objective is to maximize his total reward in limited time. The player must carefully choose which stations to attack considering the available information about available time, rewards, and officers' coverage distribution on stations and time spent to attack the station. If there is no officer at the station the player has attacked, his score will be increased by the number of stars at the station. If there is an officer at the station, his score remains the same.

Each participant began by playing two practice rounds to become familiar with the game. Next, participants played [50+] trials on the main game from which data were used in analyses. We constructed four different graphs (i.e., layouts), each of which had six stations with a different route structure and patrolling strategy. Each participant played two practice rounds and was randomly assigned to the main game on one of the four graphs.

5.1.2. Participants

Participants were recruited from Amazon Mechanical Turk. They were eligible, if they were living in the United States, had previously played more than 500 games and had an acceptance rate

of a minimum of 95%. In total 246 participants played the game with the average age of 33.6 years old where 45% were female, and 54% were male, and 1% preferred not to identify their sex.

To motivate the subjects to play games, they were compensated based on their total score ($0.01 for each gained point) in addition to a base compensation ($1). In total, 215 participants passed the validation test (correctly answered all the questions about the game at the end of the instructions). Data from participants who did not pass validation were excluded from further analyses.

5.2. Adversarial Attack pattern

The previous chapter shows that attackers can be divided into distinct groups based on their exploration behavior (i.e., Mobility Score: the ratio of the number of movements between stations over the number of trials (total number of possible movements)) by a participant in the game. Therefore, attacking the same station in consecutive trials resulted in a low mobility score while attacking a different station in each trial resulted in a high mobility score. We define three attack patterns: (i) Low Mobility for attackers who did limited or no exploration; (ii) High Mobility for attackers who tended to explore and frequently move between stations and (iii) Medium Mobility for attackers who engaged in a middling level of exploration.

In previous work where participants had different numbers of trials, it was possible that the different mobility patterns were due to variability in the number of decisions and subsequent feedback available. Alternatively, it was also possible that the participants' mobility pattern represents their intentional different attack patterns that differ in the amount of exploration and their tendency to exploit.

So in this study, in contrast with the previous study, participants had the same number of trials, allowing us to factor out the "variability in the number of decisions" and test whether clusters still emerge.

Following the previous chapter, we applied K-means (k=3) clustering analysis to the data. Participants naturally divided into three groups: participants whose mobility score is less than or equal to 10% or Low Mobility (i.e. Cluster 1), participants whose mobility score is greater than or equal to 70%: High Mobility (i.e. Cluster 3), and participants whose mobility score is greater than 10% and less than 70%: Medium Mobility (i.e. Cluster 2).

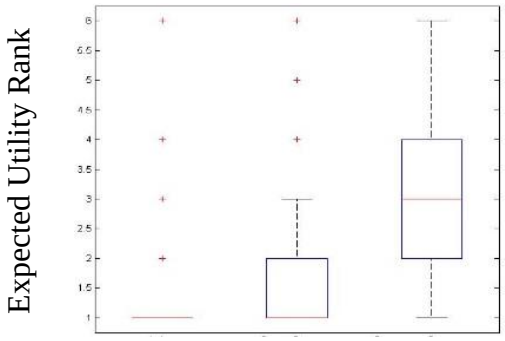

Figure 22: Expected Utility Rack for each Cluster

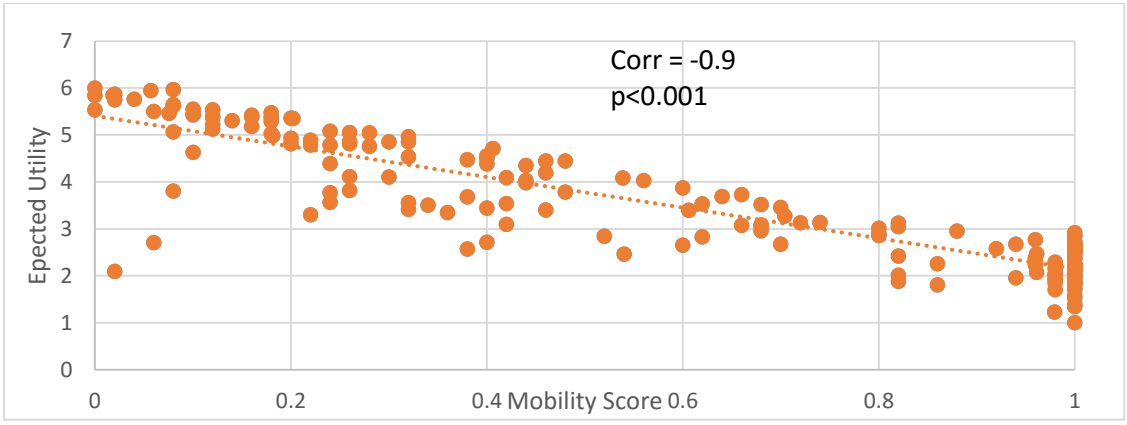

Figure 23: Utility Rank by Cluster

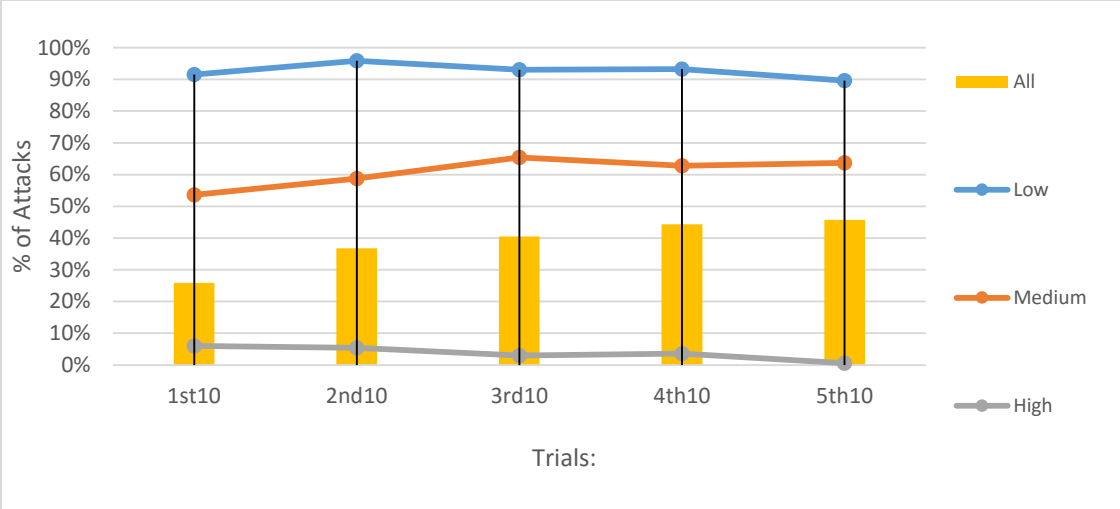

Figure 24: Percentage of attacks on the highest EU station

Figure 22 shows the three clusters and their corresponding distribution of the utility[5] of the choices they made (EU-rank[6] distribution). Participants who belong to the Low Mobility Cluster focused on the stations with highest expected utility (mean=1.2, SD=0.5). On the other hand, participants who tended to move frequently between different stations (High Mobility) attacked average stations with lower utility (mean=3.0, SD=1.3). Participants in Medium Mobility Cluster also attacked a variety of stations but were leaning (on average) towards higher utility rank stations (mean=0.7, SD=1.15). These results replicated those in the previous chapter. The robustness of these observations is very important in designing defenders' strategies as they show that attackers that belong to different clusters make decisions differently.

The cluster results are reinforced by Figure 23 that illustrates the negative correlation between mobility scores and the average utility of the attacked stations by the participant (r= -0.9, p<.001). Similarly, there is a significant negative correlation (r = -0.36, p <0.001) between the final score

[5] $EU = (1 - stationary\ coverage) * reward/\ time$
[6] To normalize the utility score among graphs, we have used the ranking of stations' utility instead of its absolute value (the highest utility in the graph is ranked 1)

of the participant and his mobility score. In other words, the more participants moved between stations, the lower their score was, and the less money they earned in the experiment. The main question of interest in this chapter is the change of adversarial behavior over the course of the 50+ trials. We expect that as attackers play the game they will learn to discover the station with higher EU and the patterns of defense behavior, and therefore learn to concentrate in the most profitable stations.

To test this hypothesis we analyzed participants' behavior over the course of the 50 trials. Figure 24 demonstrates the percentage of attacks on the stations with the highest Expected Utility (EU) in each of 5 consecutive blocks of 10 trials each. For the participants who belong to Low Mobility and High Mobility Clusters, the percentage of attacks has small fluctuation over time. On the other hand, this percentage increase for the participants in the Medium Mobility Cluster. Moreover, the bar charts show the percentage of attacks on the highest EU stations by all the participants combined. Over the course of the 50 trials, the percentage increases significantly as the proportion of Low Mobility participants increases over time while the percentage of High Mobility participants decreases. In other words, participants shifts toward clusters with lower mobility score and higher rationality level over time.

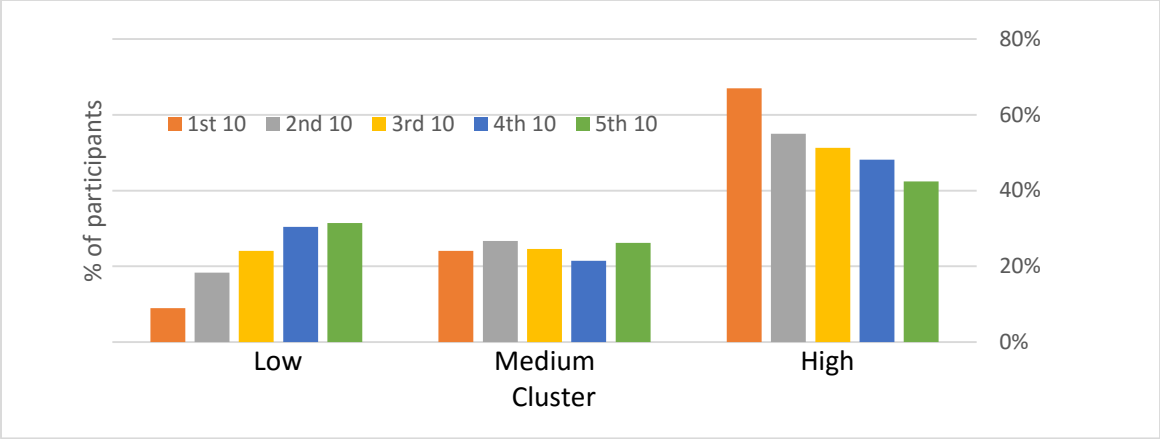

Figure 25: Clustering Change over time

Figure 25 demonstrates the percentage of participants belonging to each cluster, in each of 5 consecutive blocks of 10 trials each. The percentage of Low Mobility participants increases from the first ten trials to the last ten trials, while the percentage of High Mobility participants decreases over the course of the 50 trials. In other words, participants became more "rational" over time. These observations provide us with further insights for designing defenders' strategies, as we show that in addition to classifying attackers by their mobility behavior we also need to consider how adversaries become smarter and learn the defend strategy with more attack attempts.

5.3. Models of Adversarial Behavior in OSGs

In the following, we present competing models that represent adversarial behavior and focus on a modification to these models that capture changes in human behavior over time.

5.3.1. Bayesian Update of Human Behavior Models

In the previous chapter, the human behavior models did not have the power to predict how human behavior changed over time. On the other hand, our results in this chapter show participants (regardless of their classification based on mobility score) learn to take advantage of the defense algorithm, and they attack stations with higher utility over the course of the trials (becoming wiser).

Thus, it is important that models of adversarial behavior account for the change in participants' behavior. Furthermore, in the previous chapters, the QR, and SUQR models were compared to a process model, a cognitive model that predicts individual choices over time (the IBL model). In some way, these comparisons did not demonstrate the most significant advantages of a cognitive model of learning: to predict individual choices at each point in time. For this reason and given our experimental results, we modified the traditional QR and SUQR models described above with a Bayesian update method (B-QR and B-SUQR), so that these models would make more similar predictions to the IBL model, with individual choices over time.

To use the Bayesian update method, we focus on participants' decision at each trial: each participant made a decision of selecting one out of the six stations to attack. We modeled this problem with a Multinomial distribution over six options with a probability vector $< p_1, \dots, p_k >$ where p_i refers to probability of choosing option i in each trial.

At first, before having any data, we assumed that participants can attack any of the six stations with uniform probability. Then, after each ten trials, we gathered data on the actual number of attacks at each station, yielding data on the actual probability of attacking each station. So $< p_1, \dots, p_k >$ can be updated in the Multinomial Distribution. Luckily, Dirichlet distribution (Bernard, J. M., 2005) is a conjugate distribution for the Multinomial distribution which leads to generating a distribution for each of the probabilities in Multinomial distribution. So in Bayesian-Quantal Response and Bayesian-SUQR, after each ten trials, the distribution over probabilities of attacks get updated and then 100 random samples generated out these probability distributions and 100 human behavior models' parameters were extracted using these samples of probability of attaching each target.

5.3.2. Instance-Based Learning Model

The IBL model (Gonzalez & Dutt, 2011; Lejarraga, Dutt & Gonzalez, 2013) of an adversary makes a choice about the station to go to each trial by using the Blended Value. The Blended value V represents value of attacking each station (option j) in a particular trial:

$$V_j = \sum_{i=1}^{n} p_{ij} x_{ij}$$

where x_{ij} refers to the value (payoff) of each station (the number of stars divided by time taken) stored in memory as instance *i* for the station *j*, and p_{ij} is the probability of retrieving that instance for blending from memory (Gonzalez & Dutt, 2011; Lejarraga et al., 2012) defined as:

$$p_{ij} = e^{\frac{A_i}{\tau}} \Big/ \sum_l e^{\frac{A_l}{\tau}}$$

Where l refers to the total number of payoffs observed for station *j* up to the last trial, and τ is a noise value defined as $\sigma \cdot \sqrt{2}$. The σ variable is a free noise parameter. The activation of instance *i* represents how readily available the information is in memory:

$$A_i = ln \sum_{\substack{t_p \\ \epsilon\ observed}} (t - t_p)^{-d} + \sum_{\substack{Attribute \\ \epsilon\ Situation}} P(M_{Attribute} - 1) + \sigma \ln(\frac{1 - \gamma_{i,t}}{\gamma_{i,t}})$$

The Activation is higher when instances are observed frequently and more recently. For example, if an unguarded, nearby station with many stars (high reward) is observed many times, the activation of this instance will increase, and the probability of selecting that station in the next round will be higher. However, if this instance is not observed often, the memory of this station will decay with the passage of time (the parameter *d*, the decay, is a non-negative free parameter

that defines the rate of forgetting). The noise component is a free parameter that reflects noisy memory retrieval.

Importantly, in addition to the kernel mechanisms of the IBL model described above, and used in a multitude of studies (see Gonzalez, 2013 for a summary), the previous chapter proposed a mechanism that would allow the IBL model to account for the various mobility clusters. This mechanism was a randomization rule applied at each time step, which resulted in making a random selection of a station instead of selecting the station with the highest *Blended* value. This randomization rule served the purpose of generating the clusters of participants with diverse mobility scores. In this chapter, this rule was removed given that each participant made exactly 50 choices, and the process of learning over those should be captured by the kernel mechanisms of the IBL model without the additional randomization rule.

5.4. Modeling Results

To test the models, we divided the human data set into two groups: training and test datasets. For each cluster or block of trials, 70% of the participants were randomly selected, and their data were used to train the QR, SUQR, and their Bayesian versions (B-QR and B-SUQR) and to fit the *d* and the in the IBL model. The remaining 30% of the participants were used for testing the models.

For comparison of different models, we use Root Mean Squared Error (RMSE). *Table 8* shows the results on the full data set. Although models provide different perspectives, their prediction errors are similar. The IBL model captures learning and decision dynamics over time while QR and SUQR predict the stable state transition probabilities of the attacker while B-QR and B-SUQR[7]

[7] For Bayesian-QR (B-QR) and Bayesian-SUQR (B-SUQR), the average values over 100 data have reported in the tables

update the transition probabilities of attacker after each ten trials. Table 9 shows the performance of different models in different clusters.

Table 8: Metrics and Parameter on the full data set

Model	*Parameters*	*RMSE*
QR	0.41	0.25
SUQR	<2.9,-2.1,-2.7>[8]	0.24
Bayesian-QR (B-QR)	0.34	0.24
Bayesian-SUQR (B-SUQR)	<2.5.-1.9,-2.1>	0.23
IBL	<0.01,0.01>[9]	0.23

In Low Mobility Cluster, human behavior models outperform IBL model, and the Bayesian update on these models results in a significant improvement over their counterparts' models. For the Medium Mobility Cluster, the improvement is not significant, and all models' prediction errors are similar for the High Mobility Cluster.

Table 9: Metrics and Parameters on each Cluster

Clusters	Model	Parameters	RMSE
Low Mobility Cluster	QR	1.28	0. 33
	SUQR	<5.6,-4.4,-8.9>[8]	0.35
	B-QR	0.69	0.20
	B-SUQR	<2.8,-2.2,-5.1>	0.17
	IBL	<0.46, 0.01>[9]	0.50

[8] $< w_r, w_{sta}, w_{time} >$
[9] $<noise, decay>$

Medium Mobility Cluster	*QR*	0.69	0.27
	SUQR	<4.5,-2.3,-5.1>[8]	0.28
	B-QR	0.49	0.17
	B-SUQR	<3.0,-1.7,-2.5>	0.24
	IBL	<3.64, 1.82>[9]	0.35
High Mobility Cluster	*QR*	0.07	0.25
	SUQR	<2.1,-1.7,-0.5>[8]	0.25
	B-QR	0.14	0.27
	B-SUQR	<1.9,-1.5,-1.5>	0.28
	IBL	<0.1, 2.71>[9]	0.27

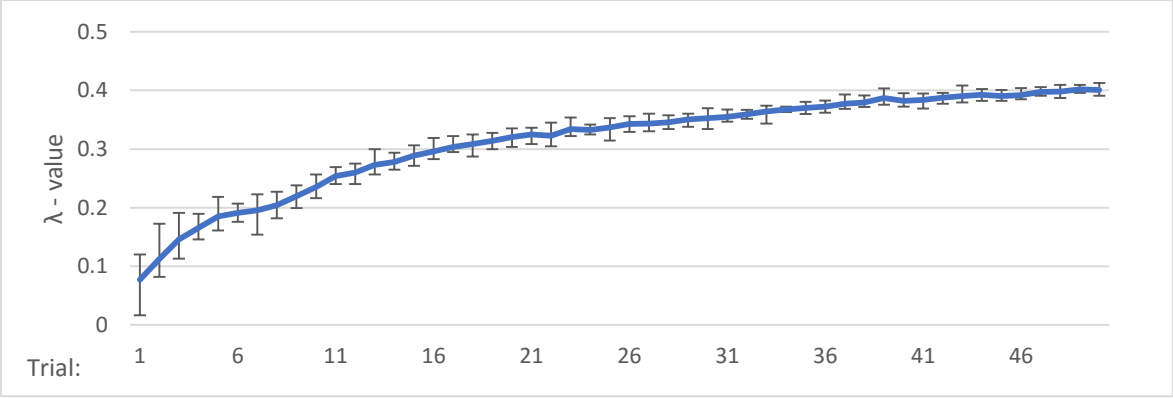

Figure 26: lambda value over time

For the Bayesian models, the reported parameters in the tables were averaged over extracted parameters from the samples. Further analyses over these parameters are shown in Figure 26 which shows the distribution of Quantal Response λ-value over time. As shown in the graph, the λ-value increases, which means the participants are becoming more rational.

This observation is consistent with Figure 24, extracted from participants' data, where the bar chart shows the percentage of attacks on the highest expected utility stations which increase over time.

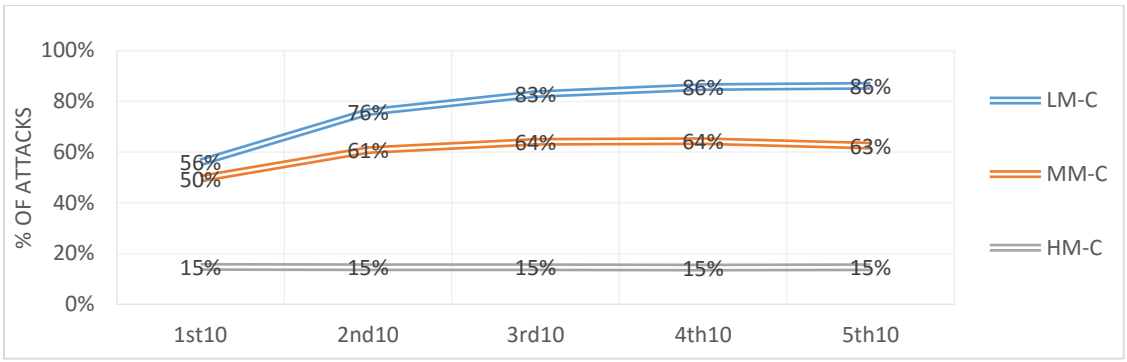

Figure 27: Percentage of attack on the highest EU station predicted by IBL

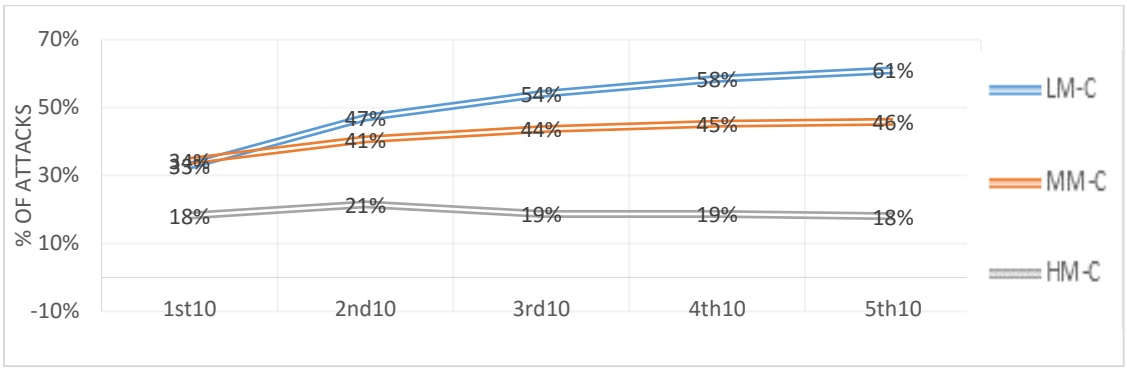

Figure 28: Percentage of attack on the highest EU station predicted by B-QR

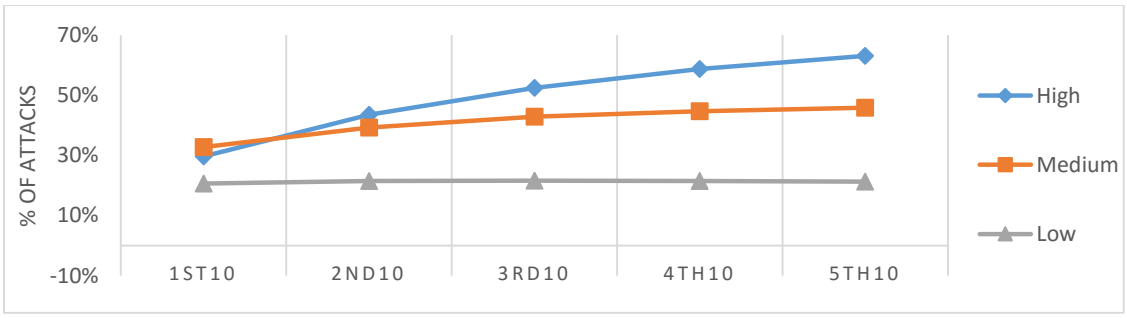

Figure 29: Percentage of Attack on the highest EU stations predicted by B-SUQR

Figure 27, Figure 28, and Figure 29 all focus on the proportion of attacks on the highest EU stations over time, predicted by the IBL, B-QR, and B-SUQR models, respectively. As shown in the graphs, all models also predict the increasing rationality of the participants over time specifically for Low Mobility Cluster and Medium Mobility Cluster.

5.5. Summary

In security game researches, understanding human adversary behavior has led to several deployed real-world applications (Tambe 2011). Although there are a significant amount of such researchers, there has been little research of heterogeneous adversary and how adversaries adapt to defense strategies over time. In this chapter, we focus on opportunistic adversaries and advance the prior research which suggested classifying adversary into distinct groups based on the ways humans explore their choice options.

More specifically, we advance this work by showing how adversaries shift among the categories as they learn the defenders' behavior over time; toward the end of the game, more participants belong to Cluster 1 (cluster with lower mobility score and higher score) and fewer participants belong to Cluster 3 (participants with higher mobility score). This might be because of the fact that toward the end of the game, participants learn more about the game, the consequences of their action and make a more rational decision and so more participants join the cluster with the higher rationality level. On the other hand, some participants did not change their behavior over the course of 50 trials and these may refer to criminals who do almost no planning before committing their crime.

Furthermore, we show how behavioral game theory models can be modified to capture the learning dynamics using a Bayesian Updating modeling approach. These models perform similarly to a cognitive model known as Instance-Based Learning, to predict learning patterns.

This chapter also provides interesting insights into building defense strategies. For example, current sophisticated defense algorithms often assume a homogeneous adversary population who behave the same over time. Given the significant impact of modeling adversarial behavior to

designing optimum patrolling strategies for the defenders, it is critical to account for this heterogeneity in behavior also we need to have defenders' strategy which adapts to change in human behavior over time.

Chapter 6. Conclusion

6.1. Contributions

With the growing number of automated decision aids based on a game-theoretic algorithm, investigations of human bounded rationality models become crucial to ensure effective security resource allocation and scheduling. The findings outlined in this thesis provides valuable advice for modeling human behavior in Opportunistic Security domain

This work 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 collected by running extensive human subject experiments and comparing different bounded rationality models, I have illustrated 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; (b) combinations of the well-known prospect theory model and 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; and (d) models based on relative weighting of values, i.e., gain and loss from current state, provide better modeling accuracy than absolute weighting. These are some of the findings that provide valuable advice for practical implementations of decision-aids.

In the second part of this thesis, I investigated the heterogeneity in adversary behavior by clustering adversaries into distinct groups based on their exploration patterns. We found that three distinct clusters emerged based on the adversaries' exploration patterns, two of which fall at two

extreme ends of the spectrum, capturing perfectly rational and completely random adversarial behavior while people in the third cluster exhibit some behavior in between. It has also been observed that in our OSG domain, exploration is negatively correlated with utility maximization.

I demonstrate that accounting for the diversity of adversary behavior leads to very different model parameters and can provide more accurate predictions of future behavior. Specifically, on the data collected from our OSG experiment, I show that (a) the QR model captures the behavior of utility maximizing adversaries better than SUQR or IBL based models; (b) the behavioral models (QR and SUQR) and cognitive models (IBL) have similar prediction performance for adversaries who do not act in a perfectly rational fashion, and (c) considering the heterogeneity in adversary behavior leads to different defender strategies. The effectiveness of strategies generated by adopting heterogeneous behavioral models is an important area of future work.

And finally, in the third part of this thesis, I extend the previous chapter on classifying human adversaries into distinct groups and show (a) how adversaries shift among the categories as they learn the defenders' behavior over time; toward the end of the 50 trials, more participants belong to Cluster 1(cluster with the lower mobility) and less to Cluster 3 (high mobility cluster). Or in other word, participants become more rational over time, and (b) how behavioral game theory models can be modified to capture the learning dynamics using a Bayesian Updat approach.

6.2. Future Work

In this thesis, I have shown how human behavior can be best predicted by human bounded rationality models as well as cognitively plausible models. I show how detailed analysis of our data can be used to categorize different types of the adversary and also show how human behavior changes over time. To that end, these results provide insight to design optimal patrolling strategy

for the defenders, and there are still many research opportunities, challenges and questions to be answered.

One possible direction for future work is more detailed analyses of the data. First, extensive study of participants' behavior can be done in regards to the complexity of the graph; initial results show participants' rationality level highly correlates with the graph structure. Second, we can change the incentives (constant vs. bonus payments) in the AMT experiment and study how individual risk attitude change with respect to different incentives. One other possible future work would be an extensive study of human behavior over time. The results show that some participants changed their behavior over time while some perform invariably and remained in the same category. It would be insightful to study if there is any correlation between participant' behavioral change pattern and the specific game the participant played.

Additionally, it would be very helpful to perform a sensitivity analysis by running modified versions of the experiment. For example, by hiding the information on stationary coverage from the players, we can study the impact of partial information on different models' performance. Designing the experiment with partial information would also provide the opportunity to expand the study from Opportunistic Security Game to Cyber Security.

Moreover, the results show that defenders are facing heterogeneous adversaries whose behavior change over time. Therefore, designing defending strategies which can adopt with heterogeneous adversaries over time would be interesting.

The mentioned directions are only some ideas out of many interesting research opportunities that could advance this field in many ways.

Bibliography

Alarie, Y., Dionne G. (2001) Lottery decisions and probability weighting function. *Journal of Risk and Uncertainty*.

Allison, P. (2012). Handling missing data by maximum likelihood. Haverford, PA: Sage. *Presented at the SAS Global Forum*.

Anderson, J. R., & Lebiere, C. (1998). The atomic components of thought. Lawrence Erlbaum Associates.

Auria, L., & Moro, R. A. (2008). Support vector machines (SVM) as a technique for solvency analysis.

Ben-Asher, N., Oltramari, A, Erbacher, R.F., and Gonzalez, C. (2015). Ontology-based Adaptive Systems of Cyber Defense. (STIDS).

Bernard, J. M. (2005). An introduction to the imprecise Dirichlet model for multinomial data. *International Journal of Approximate Reasoning*.

Bishop, C. M. (2006). Pattern Recognition. *Machine Learning*.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

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

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of artificial intelligence research*.

Chen, H., Chung, W., Xu, J. J., Wang, G., Qin, Y., & Chau, M. (2004). Crime data mining: a general framework and some examples. *Computer*.

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

Cristianini, N., & Shawe-Taylor, J. (2000*). An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press.

Critchlow, R., Plumptre, A. J., Driciru, M., Rwetsiba, A., Stokes, E. J., Tumwesigye, C., ... & Beale, C. M. (2015). Spatiotemporal trends of illegal activities from ranger-collected data in a Ugandan national park. *Conservation Biology*.

Cui, J., & John, R. S. (2014). Empirical comparisons of descriptive multi-objective adversary models in Stackelberg security games. *In Decision and Game Theory for Security*. Springer International Publishing.

Das, S., Zook, A., & Riedl, M. O. (2015, April). Examining Game World Topology Personalization. *In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM.

Dehghani Abbasi, Y., Gonzalez, C., Ben-Asher, N., Morrison, D., Kar, D., Sintov, N., Tambe, M. (2016) Know Your Adversary: Insights for a Better Adversarial Behavioral Model. *Cognitive Science Society*.

Abbasi, Y. D., Short, M., Sinha, A., Sintov, N., Zhang, C., & Tambe, M. (2015). Human adversaries in opportunistic crime security games: evaluating competing bounded rationality

models. *In Proceedings of the Third Annual Conference on Advances in Cognitive Systems ACS*.

Dehghani Abbasi, Y., Short, M., Sinha, A., Sintov, N., Zhang, Ch., Tambe, M., (2015) Human Adversaries in Opportunistic Crime Security Games: How Past success (or failure) affects future behavior. *In Workshop on Behavioral, Economic and Computational Intelligence for Security held at the 24th International Joint Conference on Artificial Intelligence (IJCAI)*.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*.

Edwards, W. (1962). Dynamic decision theory and probabilistic information processings. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.

Fischhoff, B., Goitein, B., & Shapira, Z. (1981). Subjective expected utility: A model of decision-making. *Journal of the American Society for Information Science*.

Gatti, N. (2008, June). Game Theoretical Insights in Strategic Patrolling: Model and Algorithm in Normal-Form. In *ECAI*.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating decisions from experience in sampling and repeated choice paradigms. *Psychological Review*.

Gonzalez, C., Ben-Asher, N., Martin, J. & Dutt, V. (2015). A cognitive model of dynamic cooperation with varied interdependency information. *Cognitive Science*.

Gonzalez, C., Ben-Asher, N., Oltramari, A., & Lebiere, C. (2015). Cognition and Technology. Cyber defense and situational awareness.

Gonzalez, C., Lerch, F. J., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*.

Haskell, W., Kar, D., Fang, F., Tambe, M., Cheung, S., & Denicola, L. E. (2014). Robust protection of fisheries with a compass. In *IAAI*.

Hills, T. T., & Hertwig, R. (2010). Information Search in Decisions From Experience Do Our Patterns of Sampling Foreshadow Our Decisions, *Psychological Science.*

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*.

Kar, D., Fang, F., Delle Fave, F., Sintov, N., & Tambe, M. (2015, May). A game of thrones: when human behavior models compete in repeated Stackelberg security games. *In Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems.

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: a meta-analysis of lens model studies. *Psychological Bulletin*

Kaufmann, E., & Athanasou, J. A. (2009). A meta-analysis of judgment achievement as defined by the lens model equation. *Swiss Journal of Psychology*.

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

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

Laurikkala, J. (2001). *Improving identification of difficult small classes by balancing class distribution*. Springer Berlin Heidelberg.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*.

Lemieux, A. M. (2014). *Situational prevention of poaching. Routledge*.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods.*

McFadden, D. L. (1976). Quantal choice analysis: A survey. *In Annals of Economic and Social Measurement*.

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

Nath, S. V. (2006, December). Crime pattern detection using data mining. In Web Intelligence and Intelligent Agent Technology Workshops, 2006. *WI-IAT 2006 Workshops. 2006 IEEE/WIC/ACM International Conference on.* IEEE.

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

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.*

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*.

Rapoport, A. (1975). Research paradigms for studying dynamic decision behavior. *In Utility, probability, and human decision making*. Springer Netherlands.

Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental psychology*.

Rubinstein, A. (1998). *Modeling bounded rationality*. MIT Press.

Savani, R., & Von Stengel, B. (2004, October). Exponentially many steps for finding a Nash equilibrium in a bimatrix game. *In Foundations of Computer Science, 2004. Proceedings. 45th Annual IEEE Symposium*. IEEE.

Shieh, E., An, B., Yang, R., Tambe, M., Baldwin, C., DiRenzo, J., ... & Meyer, G. (2012, June). 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*. International Foundation for Autonomous Agents and Multiagent Systems.

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*.

Short, M. B., D'Orsogna, M. R., Brantingham, P. J., & Tita, G. E. (2009). Measuring and modeling repeat and near-repeat burglary effects. *Journal of Quantitative Criminology*.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*.

Soley-Bori, M. (2013*). Dealing with missing data: Key assumptions and methods for applied analysis (No. 4)*. Technical Report.

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*.

Volzhanin, I., Hahn, U., Jönsson, M. L., & Olsson, E. J. (2015). Individual Belief Revision Dynamics in a Group Context. In CogSci 2015.

Ficici, S., 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.*

Yang, R., Ford, B., Tambe, M., & Lemieux, A. (2014). Adaptive resource allocation for wildlife protection against illegal poachers. AAMAS.

Yang, R., Ford, B., Tambe, M., & Lemieux, A. (2014). Adaptive resource allocation for wildlife protection against illegal poachers. *In Proceedings of the 7th international joint conference on Autonomous Agents and Multiagent*. AAMAS.

Yang, R., Kiekintvled, 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.

Zhang, C., Jiang, A. X., Short, M. B., Brantingham, P. J., & Tambe, M. (2014). Defending against opportunistic criminals: New game-theoretic frameworks and algorithms. *In Decision and Game Theory for Security*. Springer International Publishing.