Know Your Adversary: Insights for a Better Adversarial Behavioral Model

Yasaman D. Abbasi[1], Noam Ben-Asher[3], Cleotilde Gonzalez[2],
Debarun Kar[1], Don Morrison[2], Nicole Sintov[1], Milind Tambe[1]
[1]{ydehghan,dkar,sintov,tambe}@usc.edu; [2]{coty, dfm2}@cmu.edu; [3]nbenash@us.ibm.com
[1]941 Bloomwalk, SAL 300, University of Southern California, SAL (300), Los Angeles, CA 90089, USA
[2] Social and Decision Sciences, 5000 Forbes Avenue, BP 208, Carnegie Melon University, Pittsburg, PA 15213, USA
[3]US Army Research Labs & IBM T.J.Watson Research Center, 1101 route 134 Kitchawan Rd, Yorktown Heights, NY 10598

Abstract

Given the global challenges of security, both in physical and cyber worlds, security agencies must optimize the use of their limited resources. To that end, many security agencies have begun to use "security game" algorithms, which optimally plan defender allocations, using models of adversary behavior that have originated in behavioral game theory. To advance our understanding of adversary behavior, this paper presents results from a study involving an opportunistic crime security game (OSG), where human participants play as opportunistic adversaries against an algorithm that optimizes defender allocations. In contrast with previous work which often assumes homogeneous adversarial behavior, our work demonstrates that participants are naturally grouped into multiple distinct categories that share similar behaviors. We capture the observed adversarial behaviors in a set of diverse models from different research traditions, behavioral game theory, and Cognitive Science, illustrating the need for heterogeneity in adversarial models.

Keywords: Human Behavioral Modeling, Opportunistic Security Game, Cognitive Models, Heterogonous Adversaries

Introduction

Given the global challenges of security, optimizing the use of limited resources to protect a set of targets from an adversary has become a crucial challenge. In terms of physical security, the challenges include optimizing security resources for patrolling major airports or ports, screening passengers and cargo, scheduling police patrols to counter urban crime (Tambe 2011; Pita et al., 2008; Shieh et al., 2012). The challenge of security resource optimization carries over to cybersecurity (Gonzalez, Ben-Asher, Oltramari & Lebiere, 2015), where it is important to assist human administrators in defending networks from attacks.

In order to build effective defense strategies, we need to understand and model adversary behavior and defender-adversary interactions. For this purpose, researchers have relied on the insights from Stackelberg Security Games (SSGs) to provide ways to optimize defense strategies (Korzyk, Conitzer, & Parr, 2010; Tambe, 2011). SSGs model the interaction between a defender and an adversary as a leader-follower game (Tambe 2011). A defender plays a particular defense strategy (e.g., randomized patrolling of airport terminals) and then the adversary takes an action after having observed the defender's strategy. Past SSG research often assumed a perfectly rational adversary in computing the optimal defense (mixed or randomized) strategy. Realizing the limitation of this assumption, recent SSG work has focused on bounded rationality models from behavioral game theory, such as the Quantal Response behavior model (McFadden 1976, Camerer 2003), but typically a homogeneous adversary population is assumed, and a single adversary behavior model is prescribed (Kar et al., 2015).

In contrast to this previous work which often assumes a homogeneous adversary population with a single behavioral model, this paper focuses on the heterogeneity in adversary behavior. Our results are based on the study conducted in Opportunistic Security Games (OSGs). In that experiment, (Abbasi et al., 2015) evaluated behavioral game theory models assuming a homogeneous adversary population. However, our 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 (nearly 50%) is seen to act in a completely random fashion. We show through experiments that considering distinct groups of adversaries leads to interesting insights about their behavioral model, including the defender strategies being generated based on the learned model.

There are two strands of previous work related to this paper. First, in behavioral game theory models in security games, mostly homogenous adversary models have been studied, but some recent research has considered the heterogeneity of human adversarial behavior. 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 paper, 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 paper, 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

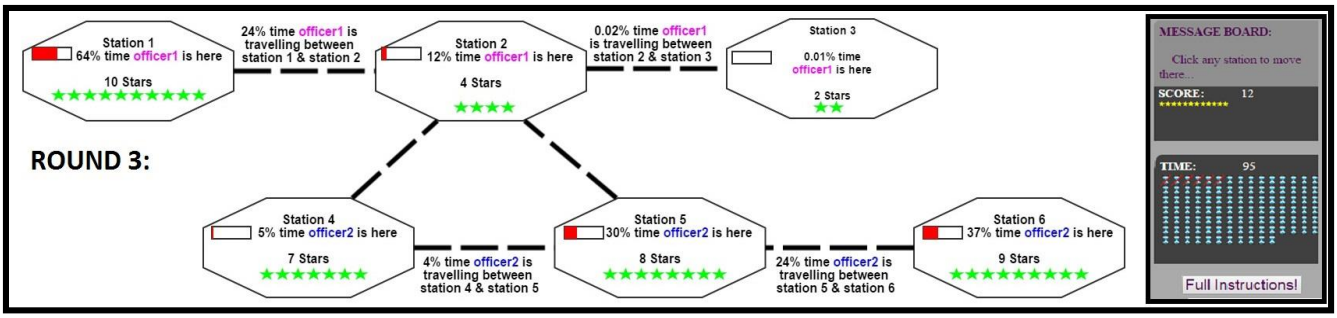

Figure 1. Game Interface

human adversaries. Our observations are also with respect to 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.

To model the different categories of human behavior, we provide a family of behavioral game theory 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). These models have been shown to successfully capture human rationality in decision making in the security games domain (Tambe, 2011). 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 paper 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) showing the impact of considering heterogeneity on the defender strategies generated.

A behavioral study in an OSG

To collect data regarding adversarial behavior from playing an OSG repeatedly, we used data collected from experiments by (Abbasi et al., 2015) using a simulation of urban crime in a metro transportation system with six stations (Figure 1).

Methods

Game Design. The players' goal is to maximize their score by collecting rewards (represented by stars in Figure 1) 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 1.

There are two moving officers, each protecting three stations. The probability of their presence at each station or route, i.e. patrolling strategy, is determined beforehand using an optimization algorithm similar to the one presented in (Zhang et al., 2014). The algorithm optimizes defender strategies given an opportunistic adversary behavior model.

The stationary coverage probabilities for each station and trains are revealed to the players. This means that players can see the percentage of the time that officers spend on average at each station and on the train, so they can determine the chance of encountering an officer at a station. However, during the game, the players cannot observe where officers are actually located unless they encounter the officer at a station.

The game can finish either if the player uses up all the 100 units of available time in each game, or the game is randomly terminated after a station visit, which may happen with a 10% probability after each station visit. The random termination encourages players to choose each action carefully, as there is a chance the game may terminate after each visit.

The player's objective is to maximize his total reward in limited time. Players must carefully choose which stations to visit, considering the available information about rewards, officers' coverage distribution on stations and time to visit the station.

Procedures. Each participant played eight games in total; starting with two practice rounds to become familiar with the game, followed by two validation rounds (two simple versions of the main games), in which the participants were presented with obvious choices to confirm they understood the rules and game's objective, and finally, four main games from which we collect and use the data for our analyses. To ensure that the collected data is independent of the graph structure, the four main games were played on four different graphs, presented in a random order to the participants. Each graph had six stations with a different route structure and patrolling strategy.

Participants. The participants were recruited from Amazon Mechanical Turk. They were eligible if they had previously played more than 500 games and had an acceptance rate of minimum 95%. To motivate the subjects to participate in the study, they were rewarded based on their total score ($0.01 for each gained point) in addition to a base compensation ($1). In total, 70 participants took part in the

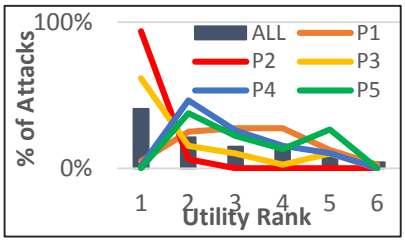

Figure 2: % of Attacks on utility rank

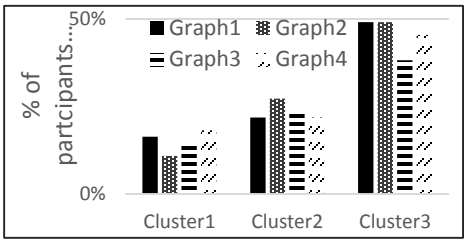

Figure 3: Clustering Distribution

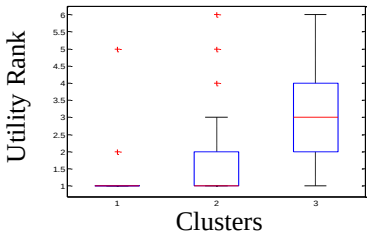

Figure 4: Utility Rank by Cluster

game and went through a validation test. Data from 15 participants who did not pass validation were excluded.

Human Adversarial Behavior

Using data from all the main games, Figure 2 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 graph illustrates significant heterogeneity behavior among individuals (line charts), and comparison to the average behavior (bar chart).

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 5: % of participants based on the Mobility Score

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.

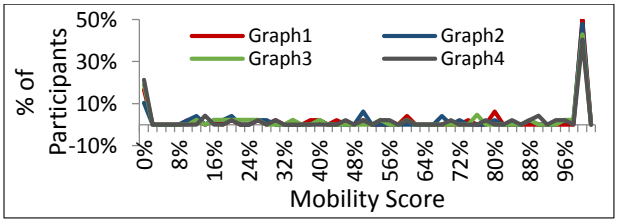

Figure 5: % of participants based on the Mobility Score

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

Figure 3 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 4, 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.

Models of Adversarial Behavior in OSG

In what follows, we present a series of models that have been proposed recently to represent adversarial behavior.

Quantal Response Model (QR)

Quantal Response models the bounded rationality of a human player by capturing the uncertainty in the decisions made by the player (McKelvey & Palfrey 1995; McFadden 1976). Instead of maximizing the expected utility, QR posits that the decision-making agent chooses an action that gives high expected utility, with probability higher than another action which gives a lower expected utility. In the context of OSG, given the defender's strategy s (e.g., stationary coverage probability at station i (s_i) shown in Figure 1), the probability of the adversary choosing to attack target i when he is in target j and when the defender's coverage is s, $q_{i,j}(s)$, 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 of the adversary as 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 player is in station j

Subjective Utility Quantal Response (SUQR)

The SUQR model combines two key notions of decision making: Subjective Expected Utility, SEU, (Fischhoff et al., 1981) and Quantal Response; it essentially replaces the expected utility function in QR with the SEU function (Nguyen et al., 2013). In this model, the probability that the adversary chooses station i when he is at station j, when the defender's coverage is s, is given by $q_{i,j}(s)$. $SEU_{i,j}$ (s) is a linear combination of three key factors. The key factors are (a) r_i, (b) s_i, and (c) $time_{i,j}$, $w = <w_r, w_{sta}, w_{time}>$ denotes the weights for each decision making feature:

$$q_{i,j}(s) = \frac{e^{SEU_{i,j}(s)}}{\sum_{t' \in T} e^{SEU_{i,t}(S)}} \text{ where}$$

$$SEU_{i,j}(s) = w_r . r_i + w_{sta} . s_i + w_{time} . time_{i,j}$$

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 draw form 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:

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

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 σ is a free parameter that reflects noisy memory retrieval.

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. RMSE represents the deviation between model's predicted probability of adversary's attack ($\hat{p}$) and the actual proportion of attacks of participants from each station to others (p).

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

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 a number of estimation parameters).

$$AIC = 2 * \#\, model's\ pramters - 2 * \ln(likelihood)$$

Table 1 shows the results on the full data set. The model parameters obtained from the training data set 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 parameter values and performance of the models for each cluster (Table 2).

Table 1: 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

[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 >

ranks of targets chosen by adversaries in each cluster, as shown in Figure 5. 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, thus modeling perfectly rational and completely random adversaries respectively.

Moreover, considering the fact that 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 three where participants do not have a clear movement pattern. Therefore, identifying different groups of adversaries gives us valuable insight into the types of behavioral models that can be used in different scenarios to generate accurate future predictions

Table 2: Metrics and Parameters on each Cluster

	Model	*Parameters*	*RMSE*	*AIC*
Cluster 1	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
Cluster 3	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

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 done 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. We show on data collected from their domain that 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). The bar charts in Figure 6 shows the percentage of change in defender strategy, for example, for target 16, the change in coverage probability from defender strategy generated against a homogeneous to that against a heterogeneous model is 110%.

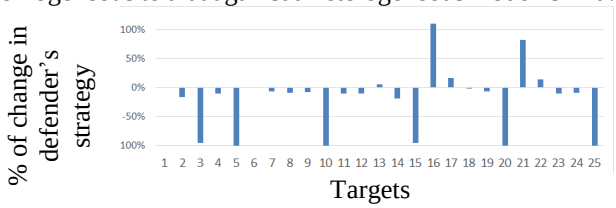

Figure 6: Strategy against homogenous & heterogeneous

Conclusions

Significant research has been conducted towards understanding 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 (Shieh et al. 2012) and ARMOR for scheduling of police patrols at major airports such as LAX (Pita et al. 2008). Although researchers 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 paper, we showed the heterogeneity in adversary behavior by clustering adversaries into distinct groups based on their exploration patterns. Three clusters emerged based on the adversaries' exploration patterns, two of which fall at two extreme ends of the parameter spectrum, capturing perfectly rational and completely random behavior. 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 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) QR captures the behavior of utility maximizing adversaries much better than SUQR or IBL based models; (ii) the behavioral and cognitive models have similar prediction performance for adversaries who do not act in a perfectly rational fashion. 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.

Acknowledgments

This research was partly supported by the Army Research Laboratory under Cooperative Agreement Number W911NF-13-2-0045 (ARL Cyber Security CRA) to Cleotilde Gonzalez. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. This research is also supported by MURI grant W911NF-11-1-0332, and award no. 004525-00001 by US-Naval Research laboratory.

References

Abbasi, Y. D., Short, M., Sinha, A., Sintov, N., Zhang, Ch., Tambe, M. (2015). Human Adversaries in Opportunistic Crime Security Games: Evaluating Competing Bounded Rationality Models. *Advances in Cognitive Systems.*

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

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

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

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

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*.

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

Kar, D., Fang, F., Delle Fave, F., Sintov, N., Tambe, M. (2015) "A Game of Thrones": When Human Behavior Models Compete in Repeated Stackelberg Security Games. (AAMAS).

Korzhyk, D., Conitzer, V., & Parr, R. (2010, July). Complexity of Computing Optimal Stackelberg Strategies in Security Resource Allocation Games. InAAAI.

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

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

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

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

Shieh, E., An, B., Yang, R., Tambe, M., Baldwin, C., ... & Meyer, G. (2012). Protect: A deployed game theoretic system to protect the ports of the United States. *AAMAS.*

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

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

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

Zhang, C., Jiang, A.X., Short, M.B., Brantingham, J.P. and Tambe, M. (2014). Defending Against Opportunistic Criminals: New Game-Theoretic Frameworks and Algorithms *Gamesec*.