# Adversaries Wising Up: Modeling Heterogeneity and Dynamics of Behavior

**Yasaman Dehghani Abbasi[1], Noam Ben-Asher[3], Cleotilde Gonzalez[2], Don Morrison[2], Nicole Sintov[1], Milind Tambe[1]**

**[1]{ydehghan,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, 1101 Route 134 Kitchawan Rd, Yorktown Heights, NY 10598

**Abstract**

Security is an important concern worldwide. Stackelberg Security Games have been used successfully in a variety of security applications, to optimally schedule limited defense resources by modeling the interaction between attackers and defenders. Prior research has suggested that it is possible to classify adversary behavior into distinct groups of adversaries based on the ways humans explore their decision alternatives. However, 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 paper, we advance this work by showing how adversaries' behavior changes as they learn the defenders' behavior over time. Furthermore, we show how behavioral game theory models can be modified to capture 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.

**Keywords:** Cognitive Models, Decision Making, Artificial Intelligence, Game Theory

## Introduction

Building effective defense strategies requires a profound understanding of adversary goals and behaviors. This can be achieved by constructing models that predict the adversary's attack patterns. For example, to have an optimized patrolling strategy for the defender, it is crucial to understand and model adversary behavior and defender-adversary interactions. Given that defense resources are limited, computational models also provide a method for optimizing resource allocation to maximize defense efficiency using the minimum quantity of resources.

To this end, researchers have used insights from Stackelberg Security Games (SSGs) to offer solutions that optimize defense strategies (Korzyk, Conitzer, & Parr, 2010; Tambe, 2011). Generally, SSGs model the interaction between a defender and an adversary as a leader-follower game (Tambe 2011), in which a defender plays a particular defense strategy (e.g., randomized patrolling of an airport's terminals) and then, having observed the defender's strategy, the adversary takes an action. Traditionally SSG research assumes a perfectly rational model of the adversary's behavior, but recent advances have shown this is not a valid assumption. To overcome the limitations of this assumption, bounded rationality models from behavioral game theory have been adopted in recent SSG work, such as the Quantal Response behavior model (McFadden 1976, Camerer 2003). These models, however, typically assume a homogeneous adversary population, creating a single adversary behavior model (Kar et al., 2015).

Again, this assumption has been challenged, and recently some researchers modeled heterogeneous behavior by either assuming a smooth distribution of the model parameters for the entire adversary population (Yang et al., 2014), or by using a single behavioral model for each adversary (Haskell et al., 2014; Yang et al., 2014). In a recent study, using data collected in an Opportunistic Security Games (OSGs), Abbasi et al. (2016) demonstrated that a population of human attackers can be naturally divided into clusters, according to their exploration of the choice options. Furthermore, exploration is negatively correlated with utility maximization, leading to different attack strategies.

The current paper addresses possible limitations in Abbasi et al. (2016) study. 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 paper presents a new experimental study using the same OSG used in Abbasi et al. (2016), 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.

With new experimental data, we demonstrate that the population of human attackers found in Abbassi et al. (2016) is robust to the number of decisions that participants make, and not determined by chance. We replicate the clusters of adversarial behavior. Furthermore, given that all participants had a fixed number of decisions in the game, we are able to investigate the adversary behavior dynamics. We show that the categories of adversarial behavior change over the course of the game, and adversary behavior shifts among these categories as adversaries learn the defender's behavior over time.

To account for the change in adversary behavior, we modified traditional economic models of bounded rationality models used in Abbassi et al. (2016), 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.

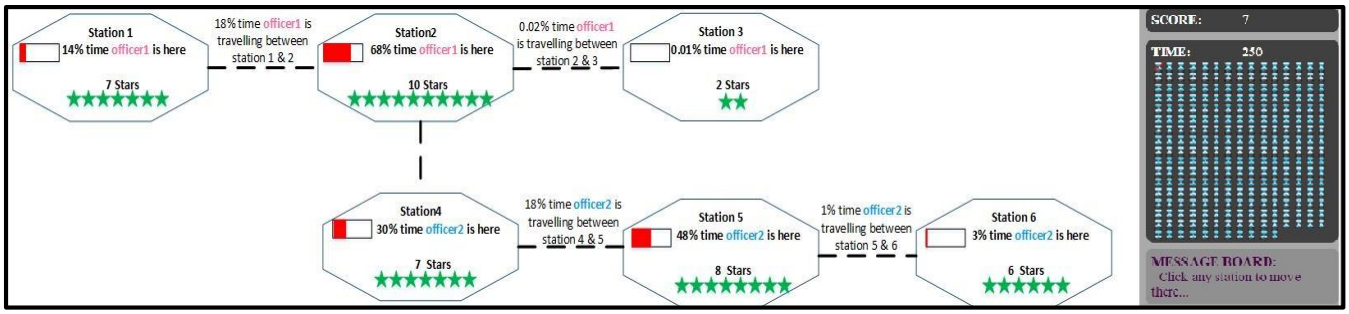


Figure 1: Experiment Game Interface

## Experimental Study of Adversarial Behavior in an Opportunistic Security Game

### Methods.

**Game Design**. To collect data on adversarial behavior in the OSG domain, we adopted the experimental design used in Abbasi et al. (2015) using a simulation of urban crime in a metro transportation system with six stations (Figure 1). The players' goal is to maximize their score by collecting rewards (represented by stars in Figure 1) in limited time while avoiding officers on patrol. Each player can visit any station, including the current one. The player can travel to any station by train as represented by the dashed lines in Figure 1. 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 (i.e. for the transportation system represented in Figure 1, if the player is currently at station 1, revisiting station 1 will take one unit of time, and visiting station station 4 will take three units of time). By traveling to a station (or visiting the current 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. The randomized terminator is shown as a fortune wheel with two parts. One part has 10% of the area and colored red, the other part is 90% of the area and colored green. If the arrow stops at the red area, the game is over, and if it lands on the green area, the participant will continue playing the game.

To encourage participants to make each decision responsibly, we showed the randomized terminator to players from the first attack, but for the first 50 trials, we forced the arrow to stop at the green area and start using random number generator after the 50$^{th}$ trial.

In the end, a player's 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.

**Procedures.** 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 was randomly assigned to play two practice rounds and the main game on a single graph.

**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%.

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 took part in the game and went through a validation test (correctly answered all the questions about the game at the end of the instructions). Data from 24 participants who did not pass validation were excluded from further analyses.

### Results: Adversarial Attack Patterns

Abbasi et al. (2016) showed 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 little or no

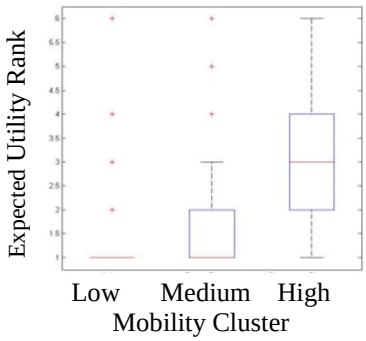


Figure 2: Utility Rank by Cluster

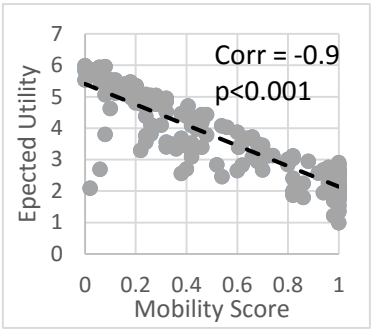


Figure 3: Clustering Distribution

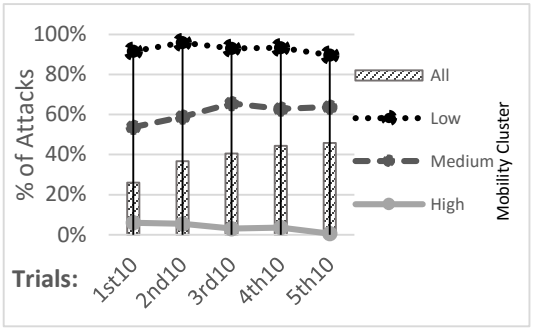


Figure 4: % of attacks on the highest EU station

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.

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 from Mobility.

Following Abbasi et al. (2016), we applied hierarchical clustering approach 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. Low), participants whose mobility score is greater than or equal to 70%: High Mobility (i.e. High), and participants whose mobility score is greater than 10% and less than 70%: Medium Mobility (i.e. Medium).

Figure 2 shows the three clusters and their corresponding distribution of the Expected Utility Rank of the choices they made (EU-rank distribution). Expected Utility is defined as:

$$EU = \frac{(1 - stationary\ coverage) * reward}{time}$$

Note that the higher expected utility, the better performance. 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).

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=1.7, SD=1.15). These results replicated those in Abbasi et al., (2016). 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 3 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 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 research 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 4 demonstrates the percentage of attacks on the stations with the highest Expected Utility (EU) in each of 5 consecutive blocks of 10 trials each (each participant was re-assigned to different clusters based on his mobility scores in each of the five blocks).

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 percentage of Low Mobility participants increases over time while the percentage of High Mobility participants decreases (Figure 5).

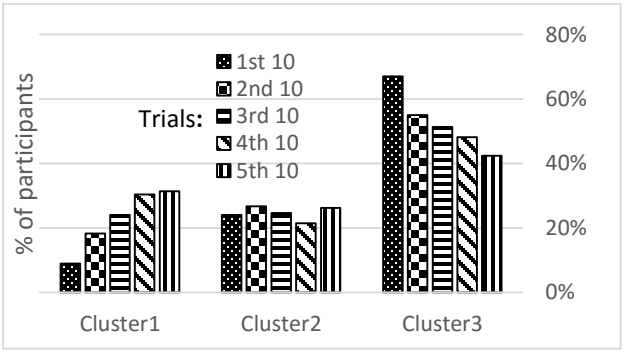


Figure 5: Cluster distribution over time

In other words, participants' shifts toward clusters with lower mobility score and higher rationality level 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.

## 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.

### Quantal Response Model (QR)

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, QR posits that the decision maker chooses an action that gives a high expected utility, with a 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 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 $\lambda$ is his degree of rationality (higher value of λ corresponds to higher rationality level) and $EU_{i,j}(s)$ is the expected utility (EU) 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 a 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 at station j, when the defender's coverage is $s$, is given by $q_{i,j}(s)$. $SEU_{i,j}$ 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:

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

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

### Bayesian Update of Human Behavior Models

In the previous study (Abbasi et al., 2016), the human behavior models did not have the power to predict how human behavior changed over time. On the other hand, our results show participants 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 Abbasi et al. (2016) 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 important 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 predictions more similar to those that the IBL model can make, 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-QR and Bayesian-SUQR, after each 10 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.

### 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 the expected 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

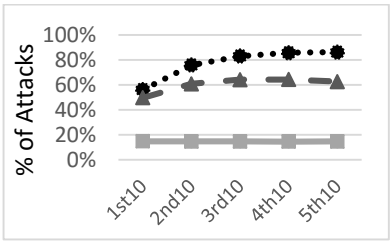


Figure 6: % of attack on the highest EU station predicted by IBL

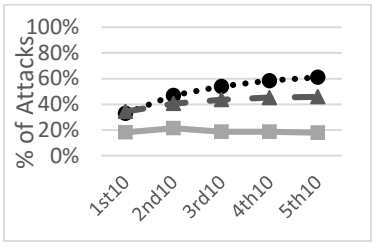


Figure 7: % of attack on the highest EU station predicted by B-QR

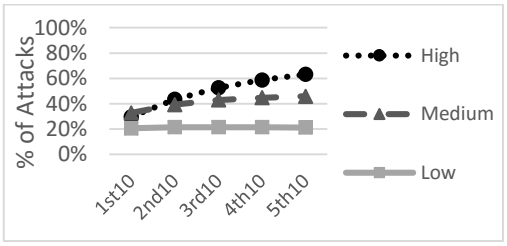


Figure 8: % of attack on the highest EU station predicted by B-SUQR

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 $\tau$ is a noise value defined as $\sigma \cdot \sqrt{2}$. The $\sigma$ variable is a free noise parameter. The activation of instance *i* represents how readily available the information is in memory:

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

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 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 σ 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), Abbasi et al. (2016) 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 the current work, 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.

## 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 σ 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) representing the deviation between a model's predicted probability of an adversary's attack ($\hat{p}$) and the actual proportion of attacks from each station to others in the human data ($p$).

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

Table 1 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[1] update the transition probabilities of attacker after each ten trials. Table 2 shows the performance of different models in different clusters.

Table 1: Metrics and Parameter on the full data set

| *Model* | *Parameters* | *RMSE* |
|---|---|---|
| ***QR*** | 0.41 | 0.25 |
| ***SUQR*** | <2.9,-2.1,-2.7>[2] | 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>[3] | 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.

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

[2] $< w_r, w_{sta}, w_{time} >$

[3] $<noise, decay>$

Table 2: 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>[2] | 0.35 |
| | ***B-QR*** | 0.69 | 0.20 |
| | ***B-SUQR*** | <2.8,-2.2,-5.1> | 0.17 |
| | IBL | <0.46, 0.01>[3] | 0.50 |
| ***Medium Mobility Cluster*** | ***QR*** | 0.69 | 0.27 |
| | ***SUQR*** | <4.5,-2.3,-5.1>[2] | 0.28 |
| | ***B-QR*** | 0.49 | 0.17 |
| | ***B-SUQR*** | <3.0,-1.7,-2.5> | 0.24 |
| | IBL | <3.64, 1.82>[3] | 0.35 |
| ***High Mobility Cluster*** | ***QR*** | 0.07 | 0.25 |
| | ***SUQR*** | <2.1,-1.7,-0.5>[2] | 0.25 |
| | ***B-QR*** | 0.14 | 0.27 |
| | ***B-SUQR*** | <1.9,-1.5,-1.5> | 0.28 |
| | IBL | <0.1, 2.71>[3] | 0.27 |

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 9 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.

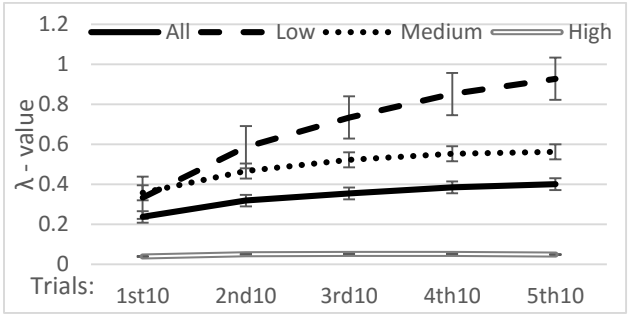


Figure 9: lambda value over time

This observation is consistent with Figure 4, extracted from participants' data, where the bar chart shows the percentage of attacks on the highest expected utility stations which increase over time.

Figure 6, Figure 7 and Figure 8 all focus on the percentage 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.

## Conclusions

In security game researches, understanding human adversary behavior has led to several deployed real-world applications (Tambe 2011), for example, PROTECT for the protection of major ports in the US by the US Coast Guard (Shieh et al. 2012). 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 paper, 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. 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 process model, a cognitive model known as Instance-Based Learning, to predict learning patterns. This study 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.

## 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.*

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

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

Bernard, J. M. (2005). An introduction to the imprecise Dirichlet model for multinomial data. *International Journal of Approximate Reasoning*, 39(2), 123-150.

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

Fischhoff, B., Goitein, B., & Shapira, Z. (1981). Subjective expected utility: A model of decision-making. *Journal of*

*the American Society for Information Science*, 32(5), 391-399.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological review*, 118(4), 523.

Gonzalez, C., Ben-Asher, N., Martin, J. & Dutt, V. (2015). A cognitive model of dynamic cooperation with varied interdependency information. *Cognitive Science,* 39(3), 457-495.

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*, 27(4), 591-635.

Haskell, W., Kar, D., Fang, F., Tambe, M., Cheung, S., & Denicola, L. E. (2014). Robust protection of fisheries with compass. *IAAI* (pp. 2978-2983).

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. *In Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (pp. 1381-1390). International Foundation for Autonomous Agents and Multiagent Systems.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*, 25(2), 143-153.

McFadden, D. L. (1976). Quantal choice analysis: A survey. In *Annals of Economic and Social Measurement, Volume 5, number 4* (pp. 363-390). NBER.

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1), 6-38.

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

Shieh, E. A., An, B., Yang, R., Tambe, M., Baldwin, C., DiRenzo, J., ... & Meyer, G. (2012). PROTECT: An Application of Computational Game Theory for the Security of the Ports of the United States. In *AAAI*.

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. *In Proceedings of the 2014 International Conference on Autonomous agents and multi-agent systems* (pp. 453-460). International Foundation for Autonomous Agents and Multiagent Systems.

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* (pp. 3-22). Springer International Publishing.