Improved Computational Models of Human Behavior in Security Games

(Extended Abstract)

Rong Yang, Christopher Kiekintveld*, Fernando Ordonez, Milind Tambe, Richard John
University of Southern California, Los Angeles, CA, 90089
* University of Texas El Paso, El Paso, TX, 79968
{yangrong,tambe,fordon,richardj}@usc.edu
ckiekint@gmail.com

Categories and Subject Descriptors

H.4 [**Computing Methodology**]: Game Theory

General Terms

Algorithms, Security

Keywords

Human Behavior, Stackelberg Games, Decision-making

1. INTRODUCTION

Security games refer to a special class of attacker-defender Stackelberg games. In these non zero-sum games, the attacker's utility of attacking a target decreases as the defender allocates more resources to protect it (and vice versa for the defender). The defender (leader) first commits to a mixed strategy, assuming the attacker (follower) decides on a pure strategy after observing the defender's strategy. This models the situation where an attacker conducts surveillance to learn the defender's mixed strategy and then launches an attack on a single target. Given that the defender has limited resources, she must design her mixed-strategy optimally against the adversaries' response to maximize effectiveness.

One leading family of algorithms to compute such mixed strategies are DOBSS and its successors [3, 5], which are used in the deployed ARMOR [5] and IRIS [8] applications. One key set of assumptions these systems make is about how attackers choose strategies based on their knowledge of the security strategy. Typically, such systems apply the standard game-theoretic assumption that attackers are perfectly rational. This is a reasonable proxy for the worst case of a highly intelligent attacker, but it can lead to a defense strategy that is not robust against attackers using different decision procedures, and it fails to exploit known weaknesses in human decision-making. Indeed, it is widely accepted that standard game-theoretic assumptions of perfect rationality are not ideal for predicting the behavior of humans in multi-agent decision problems [1]. Thus, integrating more realistic models of human decision-making has become necessary in solving real-world security problems.

Cite as: Improved Computational Models of Human Behavior in Security Games (Extended Abstract), Rong Yang, Christopher Kiekintveld, Fernando Ordonez, Milind Tambe, Richard John, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems – Innovative Applications Track (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. XXX-XXX.

The current leading contender that accounts for human behavior in security games is COBRA [6], which assumes that adversaries can deviate to $\epsilon-$optimal strategies and that they have an anchoring bias when interpreting a probability distribution. It remains an open question whether other models yield better solutions than COBRA against human adversaries. The literature has introduced a multitude of candidate models, but there is an important empirical question of which model best represents the salient features of human behavior in applied security contexts.

We address these open questions by developing three new algorithms to generate defender strategies in security games, based on using two fundamental theories of human behavior to predict an attacker's decision: Prospect Theory (PT) [2] and Quantal Response Equilibrium (QRE) [4]. PT is a noble-prize-winning theory, which describes human decision making as a process of maximizing 'prospect'. 'Prospect' is defined as the weighted sum of the benefit of all possible outcomes for each action. QRE suggests that instead of strictly maximizing utility, individuals respond stochastically in games: the chance of selecting a non-optimal strategy increases as the cost of such an error decreases.

2. METHODOLOGY

Methods for computing PT: Best Response to Prospect Theory (**BRPT**) is a a mixed integer programming formulation for the optimal leader strategy against players whose response follows a PT model. Only the adversary is modeled using PT in this case, since the defender's actions are recommended by the decision aid. The defender has a limited number of resources to protect the set of targets. BRPT maximizes the defender's expected utility by selecting the optimal mixed strategy, which describes the probability that each target will be protected by a resource. The attacker chooses a target to attack after observing such mixed strategy. PT comes into the algorithm by adjusting the weighting and value functions that are used by adversary to decide the benefit ('prospect') of attacking each target. We use a piecewise linear function to approximate the non-linear weighting function. BRPT enforces the adversary to select the target which yields the highest prospect.

Robust-PT (**RPT**) modifies the base BRPT method to account for uncertainty about the adversaries choice, caused (for example) by imprecise computations [7]. RPT assumes

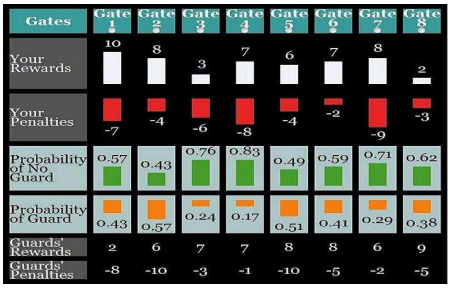

Figure 1: Game Interface

that the adversary may choose any strategy within ϵ of the best choice (i.e. attacking the target with the highest prospect). It optimizes the worst-case outcome for the defender among this $\epsilon-$optimal set of strategies, so the minimum expected utility of the defender against the $\epsilon-$optimal strategies of the adversary is maximized.

Methods for computing QRE: In applying the QRE model to our domain, we only add noise to the response function for the adversary, so the defender computes an optimal strategy assuming the attacker responses with a noisy best-response. The parameter λ represents the amount of noise in the attacker's response. We estimate λ using the standard Maximum Likelihood Estimation method based on the data collected by Pita et al. [6]. Given λ and the defender's mixed-strategy x, the adversary's quantal response q_i (i.e. probability of i) can be represented by a logit function [4]. The goal is to maximize the defender's expected utility given q_i, i.e. $\sum_i q_i U_i^d(x)$, where $U_i^d(x)$ is the expected defender's utility if she plays mixed strategy x and the subject selects target i. Essentially, we need to solve a non-linear optimization problem to find the optimal mixed strategy for the defender. However, the objective function is non-linear and non-convex in its most general form, so finding the global optimum is extremely difficult. Therefore, we focus on methods to find local optima. We develop the Best Response to Quantal Response (**BRQR**) heuristic to compute an approximately optimal QRE strategy efficiently.

3. EVALUATION

We conducted empirical tests with human subjects playing an web-based game to evaluate the performances of leader strategies generated using five candidate algorithms: BRPT, RPT, BRQR, DOBSS and COBRA. The game was designed to simulate a security scenario similar to the one analyzed by ARMOR [5] for the LAX airport. Fig. 1 shows the interface of the game. Players were introduced to the game through a series of explanatory screens describing how the game is played. In each game instance, the subjects played as the attackers and were asked to choose one of the eight gates to open (attack). They were rewarded based on the reward/penalty shown for each gate and the probability of winning/losing on each choice. To motivate the subjects, they would earn or lose money based on whether or not they succeed in attacking a gate.

We tested seven different payoff structures (four new, three from Pita et al. [6]). For each payoff structure, we generated the mixed strategies for the defender using the five algorithms. There are a total of 35 payoff structure/strategy combinations and each subject played all 35 combinations. The order of the 35 game instances played by each subjects was randomized to mitigate the order effect on their response. Besides, no feedback on success or failure was given to the subjects until the end of the experiment to mitigate learning. A total of 40 human subjects played the game. The experiment results will be available on `http://teamcore.usc.edu/yangrong/experiment.htm`.

4. CONCLUSIONS

The unrealistic assumptions of perfect rationality made by existing algorithms applying game-theoretic techniques to real-world security games need to be addressed due to their limitation in facing human adversaries. This paper successfully integrates two important human behavior theories, PT and QRE, into building more realistic decision-support tool. To that end, the main contributions of this paper are, (i) Developing efficient new algorithms based on PT and QRE models of human behavior; (ii) Conducting the most comprehensive experiments to date with human subjects for security games (40 subjects, 5 strategies, 7 game structures).

5. ACKNOWLEDGEMENT

This research was supported by Army Research Office under the grand # W911NF-10-1-0185. We also thank Mohit Goenka and James Pita for their help on developing the web-based game.

6. REFERENCES

[1] C. F. Camerer, T. Ho, and J. Chongn. A congnitive hierarchy model of games. *QJE*, 119(3):861–898, 2004.

[2] D. Kahneman and A. Tvesky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.

[3] C. Kiekintveld, M. Jain, J. Tsai, J. Pita, F. Ordonez, and M. Tambe. Computing optimal randomized resource allocations for massive security games. *In AAMAS*, 2009.

[4] R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 2:6–38, 1995.

[5] J. Pita, M. Jain, F. Ordonez, C. Portway, M. Tambe, C. Western, P. Paruchuri, and S. Kraus. Deployed armor protection: The application of a game theoretic model for security at the los angeles international airport. *In AAMAS*, 2008.

[6] J. Pita, M. Jain, F. Ordonez, M. Tambe, and S. Kraus. Solving stackelberg games in the real-world: Addressing bounded rationality and limited observations in human preference models. *Artificial Intelligence Journal*, 174(15):1142–1171, 2010.

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

[8] J. Tsai, S. Rathi, C. Kiekintveld, F. Ordonez, and M. Tambe. Iris - a tool for strategic security allocation in transportation networks. *In AAMAS*, 2009.