Analyzing the Effectiveness of Adversary Modeling in Security Games

Thanh H. Nguyen[1], Rong Yang[1], Amos Azaria[2], Sarit Kraus[2,3], Milind Tambe[1]

[1]University of Southern California, Los Angeles, CA 90089
{thanhhng, yangrong, tambe}@usc.edu

[2]Bar-Ilan University, Ramat Gan 52900, Israel
[3]University of Maryland, College Park MD 20742
{azariaa1, sarit}@cs.biu.ac.il

Abstract

Recent deployments of Stackelberg security games (SSG) have led to two competing approaches to handle boundedly rational human adversaries: (1) integrating models of human (adversary) decision-making into the game-theoretic algorithms, and (2) applying robust optimization techniques that avoid adversary modeling. A recent algorithm (MATCH) based on the second approach was shown to outperform the leading modeling-based algorithm even in the presence of significant amount of data. Is there then any value in using human behavior models in solving SSGs? Through extensive experiments with 547 human subjects playing 11102 games in total, we emphatically answer the question in the affirmative, while providing the following key contributions: (i) we show that our algorithm, SU-BRQR, based on a novel integration of human behavior model with the subjective utility function, significantly outperforms both MATCH and its improvements; (ii) we are the first to present experimental results with security intelligence experts, and find that even though the experts are more rational than the Amazon Turk workers, SU-BRQR still outperforms an approach assuming perfect rationality (and to a more limited extent MATCH); (iii) we show the advantage of SU-BRQR in a new, large game setting and demonstrate that sufficient data enables it to improve its performance over MATCH.

Introduction

The recent multiple deployments of Stackelberg Security Games (SSG) assist security agencies ("defenders") to optimally allocate their limited resources against human adversaries (Tambe 2011; Basilico, Gatti, and Amigoni 2009; Letchford and Vorobeychik 2011). While these deployments have often assumed that the adversary is a perfectly rational player, who maximizes expected value, it is well understood that such an assumption is not ideal for addressing human adversaries (Camerer 2011). As a result, researchers have been pursuing alternative approaches to handle adversary's bounded rationality in SSGs (Pita et al. 2010; Yang et al. 2011; Pita et al. 2012).

Two competing approaches have emerged to address human bounded rationality in SSGs. One approach integrates models of human decision-making into algorithms for computing defender strategies; the other adopts robust optimization techniques to intentionally avoid adversary modeling. The BRQR algorithm (Yang et al. 2011), based on modeling adversary decision-making with the Quantal Response (QR) (McKelvey and Palfrey 1995) model, leads to significantly better defender strategies than any previous leading contenders. However, the more recent robust algorithm MATCH (Pita et al. 2012) outperforms BRQR. It is indeed surprising that despite the long history of modeling success of QR, MATCH still performs better, even when significant amount of data were used to tune the key parameter in QR and no tuning was done to MATCH's key parameter.

Thus, there is now an important open question of whether there is any value in adversary modeling in SSGs. Our first contribution in answering this question builds on the significant support for QR (Haile, Hortacsu, and Kosenok 2008; Choi, Gale, and Kariv 2012): we hypothesize that QR's stochastic response is crucial in building a human decision-making model. Where we part company with the original QR model however is in its assumption that human stochastic response is based on expected value. Instead, we propose a new model based on integration of a novel *subjective utility function* (SU) into QR, called the SUQR model. We show that the SUQR model, given learned parameters (from limited data), has superior predictive power compared to the QR model. We then derive the SU-BRQR algorithm, similar to BRQR, to compute the defender strategy assuming the adversary response follows the SUQR model. We evaluate SU-BRQR's performance by conducting two sets of experiments using an online game with Amazon Mechanical Turk (AMT) workers and show that: (i) SU-BRQR significantly outperforms MATCH in previously used settings; (ii) SU-BRQR usually outperforms (and always performs at least as well as) improved versions of MATCH such as ones offering it the same SU functions or tuning its key parameter.

SU-BRQR's parameters were learned from previously available (albeit limited) game data; we now test SU-BRQR in domains without the benefit of such *a-priori* data. Indeed, while some domains of SSG application, e.g., deterring fare evasion (Yin et al. 2012) or forest protection (Johnson, Fang, and Tambe 2012), could provide significant amounts of data to tune SU-BRQR, would we be better off with MATCH or other algorithms in applications that do not? Our second contribution answers this question by conducting experiments with security intelligence experts, where we do not

have any previous modeling data. These experts, who serve as proxies for real-world adversaries, serve in the best Israeli Intelligence Corps unit or are alumna of that unit, and are found to be more rational than the AMT workers. Against these experts, SU-BRQR with its earlier learned parameters, significantly outperforms both an algorithm assuming perfect adversary rationality (Paruchuri et al. 2008) and (to a more limited extent) MATCH. Finally, our third contribution tests SU-BRQR in a new large game with AMT workers. We show that SU-BRQR with previously learned parameters still outperforms MATCH; and learning from more data, SU-BRQR performance can be further improved.

Background and Related Work

SSGs are defender-attacker games where the defender attempts to allocate her ("she" by convention) limited resources to protect a set of targets, and the adversary plans to attack one such target (Conitzer and Sandholm 2006; Tambe 2011). In SSGs, the defender first commits to a mixed strategy assuming that the adversary can observe that strategy. Then, the adversary takes his action.

Let T be the number of targets and K be the number of defender resources. The payoffs of both players depend on the attacked target and whether that target is covered by the defender. When the adversary attacks a target t, he will receive a reward R_t^a if the target is not covered by the defender; otherwise, he will receive a penalty P_t^a. In contrast, the defender will get a penalty P_t^d in the former case and a reward R_t^d in the latter case. We assume, as usual, $R_t^a, R_t^d > 0$ and $P_t^a, P_t^d < 0$. Let x_t be the coverage probability of the defender on target t. The defender's expected value at target t can be calculated as:

$$U_t^d = x_t R_t^d + (1 - x_t) P_t^d$$

Similarly, the expected value for the attacker is given by:

$$U_t^a = x_t P_t^a + (1 - x_t) R_t^a$$

Traditionally, the algorithms to compute the defender strategy in SSGs have assumed a perfectly rational adversary, who tries to maximize his expected value given the defender's strategy(Conitzer and Sandholm 2006; Paruchuri et al. 2008; Korzhyk, Conitzer, and Parr 2010). However, in real-world problems, the adversary's decision may be governed by his bounded rationality (March 1978; Conlisk 1996) due to effects such as task complexity and the interplay between emotion and cognition, which may cause him to deviate from the optimal action.

Recent research has therefore focused on developing algorithms to address the adversary's bounded rationality. In particular, BRQR(Yang et al. 2011) and MATCH(Pita et al. 2012) are the two leading contenders for handling adversary bounded rationality in SSGs. BRQR subscribes to modeling human decision making; it computes an optimal strategy for the defender assuming that the adversary's response follows the QR model. The QR model predicts a stochastic distribution of the adversary response: the greater the expected value of a target the more likely the adversary will attack that target. QR's key parameter λ represents the level of rationality in adversary's response: as λ increases, the predicted response by the QR model converges to the optimal action of the adversary. In contrast, instead of using a human behavior model, MATCH computes a robust defender strategy by guaranteeing a bound on the defender's loss in her expected value if the adversary deviates from his optimal choice. More specifically, the defender's loss is constrained to be no more than a factor of β times the adversary's loss in his expected value. The key parameter β describes how much the defender is willing to sacrifice when the adversary deviates from the optimal action.

A comparison of these two algorithms by (Pita et al. 2012), using over 100 payoff structures, showed that MATCH significantly outperforms BRQR. They also showed that even with sufficient data, with carefully re-estimated λ of the QR model, and no effort to estimate MATCH's β parameter, MATCH still outperformed BRQR.

A Simulated Security Game: A simulated online SSG, called "The guards and treasures" has previously been used as the platform for human subject experiments (Yang et al. 2011; Pita et al. 2012). We will also use it in our experiments. The game is designed to simulate the security scenario at the LAX airport, which has eight terminals that can be targeted in an attack. Figure 1 shows the interface of the game. Before playing the game, all the subjects are given detailed instructions about how to play. In each game, the subjects are asked to select one target to attack, given the following information: subject's reward and penalty at each target, the probability that a target will be covered by the guard, and the reward and penalty of the defender at each target (more details in (Yang et al. 2011)).

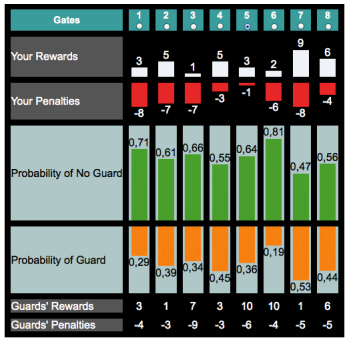


Figure 1: Game Interface

The SUQR Model

The key idea in subjective expected utility (SEU) as proposed in behavioral decision-making (Savage 1972; Fischhoff, Goitein, and Shapira 1981) is that individuals have their own evaluations of each alternative during decision-making. Recall that in an SSG, the information presented to the human subject for each choice includes: the marginal coverage on target t (x_t); the subject's reward and penalty (R_t^a,P_t^a); the defender's reward and penalty (R_t^d, P_t^d). Inspired by the idea of SEU, we propose a subjective utility function of the adversary for SSG as the following:

$$\hat{U}_t^a = w_1 x_t + w_2 R_t^a + w_3 P_t^a \quad (1)$$

The novelty of our subjective utility function is the linear combination of the values (rewards/penalty) and *probabilities*. (Note that we are modeling the decision-making of the general population not of each individual as we do not

have sufficient data for each specific subject). While unconventional at first glance, as shown later, this model actually leads to higher prediction accuracy than the classic expected value function. A possible explanation for that is that humans might be driven by simple heuristics in their decision making. Indeed, several studies in other research domains have demonstrated the prediction power of simple combination of features (Meehl 1963; Dawes 1979) while complex models could possibly lead to over-fitting issues (Meehl 1963). Other alternatives to this subjective utility function are feasible, e.g., including all the information presented to the subjects ($\hat{U}_t^a = w_1 x_t + w_2 R_t^a + w_3 P_t^a + w_4 R_t^d + w_5 P_t^d$), which we discuss later.

We modify the QR model by replacing the classic expected value function with the SU function, leading to the SUQR model. In the SUQR model, the probability that the adversary chooses target t, q_t, is given by:

$$q_t = \frac{e^{\lambda \hat{U}_t^a}}{\sum_{t'} e^{\lambda \hat{U}_{t'}^a}} = \frac{e^{\lambda(w_1 x_t + w_2 R_t^a + w_3 P_t^a)}}{\sum_{t'} e^{\lambda(w_1 x_{t'} + w_2 R_{t'}^a + w_3 P_{t'}^a)}} \quad (2)$$

The problem of finding the optimal strategy for the defender can therefore be formulated as:

$$\max_x \sum_{t=1}^{T} \frac{e^{\lambda(w_1 x_t + w_2 R_t^a + w_3 P_t^a)}}{\sum_{t'} e^{\lambda(w_1 x_{t'} + w_2 R_{t'}^a + w_3 P_{t'}^a)}} (x_t R_t^d + (1 - x_t) P_t^d)$$

$$\text{s.t.} \sum_{t=1}^{T} x_t \leq K, 0 \leq x_t \leq 1 \quad (3)$$

Here, the objective is to maximize the defender's expected value given that the adversary chooses to attack each target with a probability according to the SUQR model. Constraint (3) ensures that the coverage probabilities on all the targets satisfy the resource constraint. Given that this optimization problem is similar to BRQR we use the same approach as BRQR to solve it (Yang et al. 2011). We refer to the resulting algorithm as SU-BRQR.

Learning SUQR Parameters Without loss of generality, we set $\lambda = 1$. We employ Maximum Likelihood Estimation (MLE) (Hastie, Tibshirani, and Friedman 2009) to learn the parameters (w_1, w_2, w_3). Given the defender strategy $\mathbf{x}$ and N samples of the players' choices, the log-likelihood of (w_1, w_2, w_3) is given by:

$$logL(w_1, w_2, w_3|\mathbf{x}) = \sum_{j=1}^{N} log[q_{t_j}(w_1, w_2, w_3|\mathbf{x})]$$

where t_j is the target that is chosen in sample j and q_{t_j} $(w_1, w_2, w_3|\mathbf{x})$ is the probability that the adversary chooses the target t_j. Let N_t be the number of subjects attacking target t. Then we have:

$$logL(w_1, w_2, w_3|\mathbf{x}) = \sum_{t=1}^{T} N_t log[q_t(w_1, w_2, w_3|\mathbf{x})]$$

Combining with equation (2),

$$logL(w_1, w_2, w_3|\mathbf{x}) = w_1(\sum_{t=1}^{T} N_t x_t) + w_2(\sum_{t=1}^{T} N_t R_t^a) + w_3(\sum_{t=1}^{T} N_t P_t^a) - Nlog(\sum_{t=1}^{T} e^{w_1 x_t + w_2 R_t^a + w3 P_t^a})$$

$logL(w_1, w_2, w_3|\mathbf{x})$ can be shown to be a concave function: we can show that the Hessian matrix of $logL(w_1, w_2, w_3|\mathbf{x})$ is negative semi-definite. Thus, this function has an unique local maximum point and we can hence use a convex optimization solver to compute the optimal weights (w_1, w_2, w_3), e.g., *fmincon* in Matlab.

Prediction Accuracy of SUQR model As in some real-world security environments, we would want to learn parameters of our SUQR model based on limited data. To that end, we used a training set of 5 payoff structures and 2 algorithms MATCH and BRQR (10 games in total) from (Pita et al. 2012) to learn the parameters of the new SU function and the alternatives. In total, 33 human subjects played these 10 games using the setting of 8-targets and 3-guards from our on-line game. The parameters that we learnt are: (w_1, w_2, w_3)=$(-9.85, .37, .15)$ for the 3-parameter SU function; and $(w_1, w_2, w_3, w_4, w_5)$ = $(-8.23, .28, .12, .07, .09)$ for the 5-parameter function.

Table 1: Prediction Accuracy

QR	3-parameter SUQR	5-parameter SUQR
8%	51%	44%

We ran a Pearson's chi-squared goodness of fit test (Greenwood and Nikulin 1996) in a separate test set which includes 100 payoff structures in (Pita et al. 2012) to evaluate the prediction accuracy of the two proposed models as well as the classic QR model. The test examines whether the predicted distribution of the players' choices fits the observation. We set $\lambda = .76$ for QR model, the same as what was learned in (Yang et al. 2011). The percentages of the payoff structures that fit the predictions of the three models (with statistical significance level of $\alpha = 0.05$) are displayed in Table 1. The table clearly shows that the new SUQR model (with the SU function in Equation (1)) predicts the human behavior more accurately than the classic QR model. In addition, even with more parameters, the prediction accuracy of the 5-parameter SUQR model does not improve. Given this result, and our 3-parameter models demonstrated superiority (as we will show in the Experiments section), we leave efforts to further improve the SUQR model for future work.

Improving MATCH

Since SUQR better predicts the distribution of the subject's choices than the classic QR, and as shown later, SU-BRQR outperforms MATCH, it is natural to investigate the integration of the subjective utility function into MATCH. In particular, we replace the expected value of the adversary with subjective utility function. Therefore, the adversary's loss caused by his deviation from the optimal solution is measured with regard to the subjective utility function.

$$\max_{x,h,\eta,\gamma} \quad \gamma \quad (4)$$

$$\text{s.t.} \sum_{t \in T} x_t \leq K, 0 \leq x_t \leq 1, \qquad \forall t \quad (5)$$

$$\sum_{t \in T} h_t = 1, h_t \in \{0, 1\}, \qquad \forall t \quad (6)$$

$$0 \leq \eta - (w_1 x_t + w_2 R_t^a + w_3 P_t^a) \leq M(1 - h_t) \quad (7)$$

$$\gamma - (x_t R_t^d + (1 - x_t) P_t^d) \leq M(1 - h_t) \quad (8)$$

$$\gamma - (x_t R_t^d + (1 - x_t) P_t^d) \leq \beta \cdot (\eta - (w_1 x_t + w_2 R_t^a + w_3 P_t^a)), \qquad \forall t \quad (9)$$

We refer to this modified version as SU-MATCH, which is shown in Equation (4)-(9) where h_t represents the adversary's target choice, η represents the maximum subjective utility for the adversary, γ represents the expected value for the defender if the adversary responds optimally and M is a large constant.

Constraint (7) finds the optimal strategy (target) for the adversary. In constraint (8), the defender's expected value is computed when the attacker chooses his optimal strategy. The key idea of SU-MATCH is in constraint (9). It guarantees that the loss of the defender's expected value caused by adversary's deviation is no more than a factor of β times the loss of the adversary's subjective utility.

Selecting β for MATCH: In MATCH, the parameter β is the key that decides how much the defender is willing to lose if the adversary deviates from his optimal strategy. Pita et al. set β to 1.0, leaving its optimization for future work. In this section, we propose a method to estimate β based on the SUQR model.

Initialize $\gamma^* \leftarrow -\infty$;
for *i = 1 to N* **do**
  $\beta \leftarrow$ *Sample*$([0, \texttt{MaxBeta}], i)$, $x \leftarrow$ MATCH(β);
  $\gamma \leftarrow \sum_t q_t U_t^d$;
  if $\gamma \geq \gamma^*$ **then**
    $\gamma^* \leftarrow \gamma, \beta^* \leftarrow \beta$;
return (β^*, γ^*);

In this method, N values of β are uniformly sampled within the range (0, MaxBeta). For each sampled value of β, the optimal strategy x for the defender is computed using MATCH. Given this mixed strategy x, the defender's expected value, γ, is computed assuming that the adversary will respond stochastically according to the SUQR model. The β leading to the highest defender expected value is chosen. In practice, we set MaxBeta to 5, to provide an effective bound on the defender loss, given that penalties/rewards of both players range from -10 to 10; and N to 100, which gives a grid size of 0.05 for β for the range of $(0, 5)$. We refer to the algorithm with *carefully selected β as MATCHBeta*.

Experimental Results

The tested algorithms in our experiments include: SU-BRQR, MATCH, SU-MATCH, MATCHBeta, SU-MATCHBeta, i.e., MATCH embedded with both SU and selecting β, and DOBSS, i.e., a robust algorithm again perfectly rational opponents.

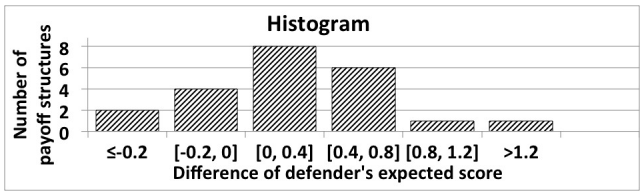
Results with AMT Workers, 8-target Games

Our first experiment compares SU-BRQR against MATCH and its improvements, in the setting where we learned the parameters of the SUQR model, i.e., the 8-target and 3-guard game with the AMT workers. In this 8-target game setting, for each game, our reported average is over at least 45 human subjects. The experiments were conducted on the AMT system. When two algorithms are compared, we ensured that identical human subjects played both on the same payoff structures. Participants were paid a base amount of US $1.00. In addition, each participant was given a bonus based on their performance in the games to motivate them. Similar to (Pita et al. 2012)'s work, we ensured that players were not choosing targets arbitrarily by having each participant play two extra trivial games (i.e., games in which there is a target with the highest adversary reward and lowest adversary penalty and lowest defender coverage probability). Players' results were removed if they did not choose that target.

We generated the payoff structures based on covariance games in GAMUT (Nudelman et al. 2004). In covariance games, we can adjust the covariance value $r \in [-1, 1]$ to control the correlation between rewards of players. We first generate 1000 payoff structures with r ranging from -1 to 0 by 0.1 increments (100 payoff structures per value of r). Then, for each of the 11 r values, we select 2 payoff structures ensuring that the strategies generated by each candidate algorithm (e.g., SU-BRQR and versions of MATCH) are not similar to each. One of these two has the maximum and the other has the median sum of 1-norm distances between defender strategies generated by each pair of the algorithms. This leads to a total of 22 payoff structures. By selecting the payoffs in this way, we explore payoff structures with different levels of the 1-norm distance between generated strategies so as to obtain accurate evaluations with regard to performance of the tested algorithms. We evaluate the statistical significance of our results using the bootstrap-t method (Wilcox 2002).

SU-BRQR vs MATCH This section evaluates the impact of the new subjective utility function via a head-to-head comparison between SU-BRQR and MATCH. In this initial test, the β parameter of MATCH was set to 1.0 as in (Pita et al. 2012). Figure 2a first shows all available comparison results for completeness (without regard to statistical significance). More specifically, we show the histogram of the difference between SU-BRQR and MATCH in the average defender expected reward over all the choices of the participants. The x-axis shows the range of this difference in each bin and the y-axis displays the number of payoff structures (out of 22) that belong to each bin. For example, in the third bin from the left, the average defender expected value achieved by SU-BRQR is higher than that achieved by MATCH, and the difference ranges from 0 to 0.4. There are 8 payoffs that fall into this category. Overall, SU-BRQR achieves a higher average expected defender reward than MATCH in the 16 out of the 22 payoff structures.

In Figure 2b, the second column shows the number of payoffs where SU-BRQR outperforms MATCH with statistical significance ($\alpha = .05$). The number of payoff structures where MATCH is better than SU-BRQR with statistical significance is shown in the fourth column. In the 22 payoff structures, SU-BRQR outperforms MATCH 13 times with statistical significance while MATCH defeats SU-BRQR only once; in the remaining 8 cases, no statistical significance is obtained either way. This result stands in stark contrast to (Pita et al. 2012)'s result and directly answers the question we posed at the beginning of this paper: there is indeed value to integrating models of human decision making in computing defender strategies in SSGs, but

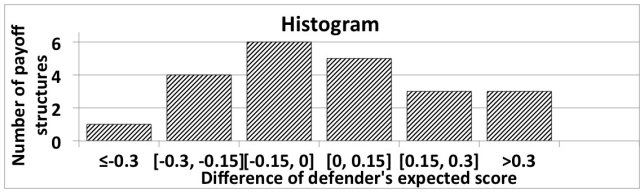


(a) All comparison data

	SU-BRQR	Draw	MATCH
$\alpha = .05$	13	8	1

(b) Results with statistical significance

Figure 2: SU-BRQR vs MATCH, AMT workers, 8 targets



(a) All comparison data

	SU-BRQR	Draw	MATCH
$\alpha = .05$	6	13	3

(b) Results with statistical significance

Figure 3: SU-BRQR vs MATCH, security experts

use of SUQR rather than traditional QR models is crucial.

Table 2: Performance comparison, $\alpha = .05$

	SU-MATCH	MATCHBeta	SU-MATCHBeta
MATCH	3, 11	1, 6	1, 8
SU-BRQR	8, 2	8, 2	5, 3

SU-BRQR vs Improved MATCH In Table 2, we compare MATCH and SU-BRQR against the three improved versions of MATCH: SU-MATCH, MATCHBeta, and SU-MATCHBeta (i.e., MATCH with both the subjective utility function and the selected β) when playing our 22 selected payoff structures. Here, we only report results that hold with statistical significance ($\alpha = .05$). The first number in each cell in Table 2 shows the number of payoffs (out of 22) where the row algorithm obtains a higher average defender expected reward than the column algorithm; the second number shows where the column algorithm outperforms the row algorithm. For example, the second row and second column shows that MATCH outperforms SU-MATCH in 3 payoff structures with statistical significance while SU-MATCH defeats MATCH in 11.

Table 2 shows that the newer versions of MATCH achieve a significant improvement over MATCH. Additionally, SU-BRQR retains a significant advantage over both SU-MATCH and MATCHBeta. For example, SU-BRQR defeats SU-MATCH in 8 out of the 22 payoff structures with statistical significance, as shown in Table 2; in contrast, SU-MATCH is better than SU-BRQR only twice.

Although SU-BRQR in this case does not outperform SU-MATCHBeta to the extent it does against MATCH (i.e., SU-BRQR performs better than SU-MATCHBeta only 5 times with statistical significance while SU-MATCHBeta is better than SU-BRQR thrice (Table 2)), SU-BRQR remains the algorithm of choice for the following reasons: (a) SU-BRQR does perform better than SU-MATCHBeta in more cases with statistical significance; (b) selecting the β parameters in SU-MATCHBeta can be a significant computational overhead for large games given that it requires testing many values of β. Thus, we could just prefer SU-BRQR .

Results with New Experimental Scenarios

All previous experiments are based on the 8-target and 3-guards game, which were motivated by the LAX security scenario (Tambe 2011). In addition, the games have been played by AMT workers or college students. To evaluate the performance of the SUQR model in new scenarios, we introduce two new experimental settings: in one the experiments are conducted against a new type of human adversary, i.e., security intelligence experts; and in the other, we change the game to 24 targets and 9 guards.

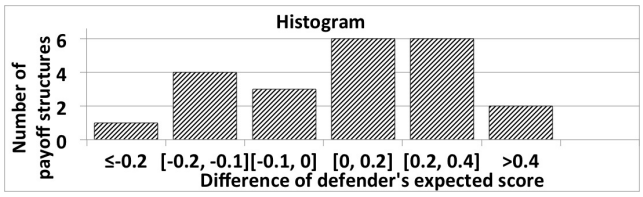
Security Intelligence Experts, 8-target games

In this section, we evaluate our algorithm with security intelligence experts who serve in the best Israeli Intelligence Corps unit or are alumna of that unit. Our purpose is to examine whether SU-BRQR will work when we so radically change the subject population to security experts. We use the same 22 payoff structures and the same subjective utility function as in the previous experiment with AMT workers. Each result below is averaged over decisions of 27 experts.

SU-BRQR vs DOBSS DOBSS (Paruchuri et al. 2008) is an algorithm for optimal defender strategies against perfectly rational opponents. DOBSS performed poorly in 8-target games against AMT workers(Pita et al. 2010; Yang et al. 2011). However, would DOBSS perform better in comparison to SU-BRQR against security experts? Our results show that SU-BRQR is better than DOBSS in all 22 tested payoff structures; 19 times with statistical significance. Thus, even these experts did not respond optimally (as anticipated by DOBSS) against the defender's strategies.

SU-BRQR vs MATCH Figure 3a shows that SU-BRQR obtains a higher expected defender reward than MATCH in 11 payoff structures against our experts. Furthermore, SU-BRQR performs better than MATCH in 6 payoff structures with statistical significance while MATCH is better than SU-BRQR only in 3 payoff structures with statistical significance (Figure 3b). These results still favor SU-BRQR over MATCH, although not as much as when playing against AMT workers (as in Figure 2).

Nonetheless, what is crucially shown in this section is that changing the subject population to security experts does not undermine SU-BRQR completely; in fact, despite using parameters from AMT workers, SU-BRQR is still able to perform better than MATCH. We re-estimate the parameters (w_1, w_2, w_3) of the SU function using the data of experts. The result is: $w_1 = -11.0, w_2 = 0.54$, and $w_3 = 0.35$. This result shows that while the experts evaluated all the criteria differently from the AMT workers they gave the same



(a) All comparison data

	SU-BRQR	Draw	MATCH
$\alpha = .05$	8	11	3

(b) Results with statistical significance

Figure 4: SU-BRQR vs MATCH, 24 targets, original

importance level to the three parameters. Because of limited access to experts, we could not conduct experiments with these re-estimated parameters; we will show the impact of such re-estimation in our next experimental setting.

Bounded Rationality of Human Adversaries We now compare the AMT workers and security experts using the traditional metric of "rationality level" of the QR model. To that end, we revert to the QR-model with the expected value function to measure how close these players are to perfect rationality. In particular, we use QR's λ parameter as a criterion to measure their rationality. We use all the data from AMT workers as well as experts on the chosen 22 games in previous experiments to learn the λ parameter. We get $\lambda = 0.77$ with AMT workers and $\lambda = 0.91$ with experts. This result implies that security intelligence experts tend to be more rational than AMT workers (the higher the λ, the closer the players are to perfect rationality). Indeed, in 34 of 44 games, experts obtains a higher expected value than AMT workers. Out of these, their expected value is higher than AMT workers 9 times with statistical significance while AMT workers is higher only once ($\alpha = .05$). Nonetheless, the lambda for experts of 0.91 suggests that the experts do not play with perfect rationality (perfect rational $\lambda = \infty$).

AMT Workers, 24-target Games

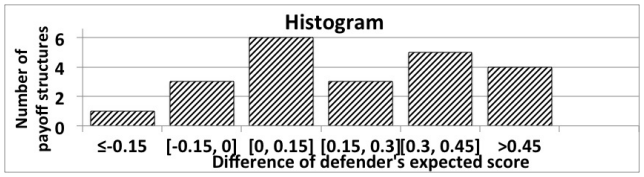
In this section, we focus on examining the performance of the algorithms in large games, i.e., 24 targets and 9 defender resources. We expect that the human adversaries may change their behaviors because of tedious evaluation of risk and benefit for each target. Two algorithms were tested: SU-BRQR, MATCH. We first run experiments with the new subjective utility function learned previously using the data of the 8-target game.

SU-BRQR vs MATCH with Parameters Learned from the 8-target Games Figure 4a shows that SU-BRQR obtains a higher average defender expected value than MATCH in 14 out of 22 payoff structures while MATCH is better than SU-BRQR in 8 payoff structures. These averages are reported over 45 subjects. In addition, as can be seen in Figure 4b, SU-BRQR performs better than MATCH with statistical significance 8 times while MATCH outperforms SU-BRQR 3 times. While SU-BRQR does perform better than MATCH, its superiority over MATCH is not as much as it was in previous 8-target games.

We can hypothesize based on these results that the learned parameters of the 8-target games do not predict human behaviors as well in the 24-target games. Therefore, we re-estimate the values of the parameters of the subjective utility function using the data of the previous experiment in the 24-target games. The training data contains 388 data points. This re-estimating results in $w_1 = -15.29, w_2 = .53$, $w_3 = .34$. Similar to the experts case, the weights in 24-target games are different from the ones in 8-target games but their order of importance is the same.

SU-BRQR vs MATCH with Re-estimated Parameters In this experiment, we evaluate the impact of the new subjective utility function with the re-estimated parameters on the performance of SU-BRQR in comparison with MATCH.

Figure 5a shows that SU-BRQR outperforms MATCH in 18 payoff structures while MATCH defeats SU-BRQR in only 4 cases. Moreover, it can be seen in Figure 5b that SU-BRQR defeats MATCH with statistical significance 11 times while MATCH defeats SU-BRQR only once with statistical significance. In other words, the new weights of the subjective utility function indeed help improve the performance of SU-BRQR . This result demonstrates that a more accurate SU function can help improve SU-BRQR's performance.



(a) All comparison data

	SU-BRQR	Draw	MATCH
$\alpha = .05$	11	10	1

(b) Results with statistical significance

Figure 5: SU-BRQR vs MATCH, 24 targets, re-estimated

Summary

This paper demonstrates the importance of integrating models of human decision making in computing defender strategies in SSGs using a novel *subjective utility function*. Through extensive experiments, the paper provides the following contributions: (i) we show that our SU-BRQR algorithm, which involves a novel integration of QR with SU function, significantly outperforms both MATCH and its improved versions; (ii) we are the first to present experimental results with security intelligence experts, and find that even though the experts are more rational than the AMT workers, SU-BRQR performs better than its competition against these experts; (iii) we show the advantage of SU-BRQR in a new, larger game setting and demonstrate that additional data can further boost the performance of SU-BRQR over MATCH.

Acknowledgement

This research was supported by MURI under the grant # W911NF-11-1-0332 and by the Google Inter-university center for Electronic Markets and Auctions, ARO under the grants # W911NF0910206 and # W911NF1110344.

References

[Basilico, Gatti, and Amigoni 2009] Basilico, N.; Gatti, N.; and Amigoni, F. 2009. Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In *AAMAS*, 57–64.

[Camerer 2011] Camerer, C. 2011. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

[Choi, Gale, and Kariv 2012] Choi, S.; Gale, D.; and Kariv, S. 2012. Social learning in networks: a quantal response equilibrium analysis of experimental data. *Review of Economic Design* 1–23.

[Conitzer and Sandholm 2006] Conitzer, V., and Sandholm, T. 2006. Computing the optimal strategy to commit to. In *EC*, 82–90.

[Conlisk 1996] Conlisk, J. 1996. Why bounded rationality? *Journal of economic literature* 669–700.

[Dawes 1979] Dawes, R. M. 1979. The robust beauty of improper linear models in decision making. *American psychologist* 34(7):571–582.

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

[Greenwood and Nikulin 1996] Greenwood, P., and Nikulin, M. 1996. A guard to chi-squared testing.

[Haile, Hortacsu, and Kosenok 2008] Haile, P. A.; Hortacsu, A.; and Kosenok, G. 2008. On the empirical content of quantal response equilibrium. *The American Economic Review* 98(1):180–200.

[Hastie, Tibshirani, and Friedman 2009] Hastie, T.; Tibshirani, R.; and Friedman, J. 2009. The elements of statistical learning 2nd edition.

[Johnson, Fang, and Tambe 2012] Johnson, M.; Fang, F.; and Tambe, M. 2012. Patrol strategies to maximize pristine forest area. In *AAAI*.

[Korzhyk, Conitzer, and Parr 2010] Korzhyk, D.; Conitzer, V.; and Parr, R. 2010. Complexity of computing optimal stackelberg strategies in security resource allocation games. In *AAAI*, 805–810.

[Letchford and Vorobeychik 2011] Letchford, J., and Vorobeychik, Y. 2011. Computing randomized security strategies in networked domains. In *AARM Workshop In AAAI*.

[March 1978] March, J. 1978. Bounded rationality, ambiguity, and the engineering of choice. *The Bell Journal of Economics* 587–608.

[McKelvey and Palfrey 1995] McKelvey, R., and Palfrey, T. 1995. Quantal response equilibria for normal form games. *Games and economic behavior* 10(1):6–38.

[Meehl 1963] Meehl, P. E. 1963. *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press.

[Nudelman et al. 2004] Nudelman, E.; Wortman, J.; Shoham, Y.; and Leyton-Brown, K. 2004. Run the gamut: A comprehensive approach to evaluating game-theoretic algorithms. In *AAMAS*, 880–887.

[Paruchuri et al. 2008] Paruchuri, P.; Pearce, J.; Marecki, J.; Tambe, M.; Ordonez, F.; and Kraus, S. 2008. Playing games for security: an efficient exact algorithm for solving bayesian stackelberg games. In *AAMAS*, 895–902.

[Pita et al. 2010] Pita, J.; Jain, M.; Tambe, M.; Ordóñez, F.; and Kraus, S. 2010. Robust solutions to stackelberg games: Addressing bounded rationality and limited observations in human cognition. *Artificial Intelligence* 174(15):1142–1171.

[Pita et al. 2012] Pita, J.; John, R.; Maheswaran, R.; Tambe, M.; and Kraus, S. 2012. A robust approach to addressing human adversaries in security games. In *ECAI*, 660–665.

[Savage 1972] Savage, L. J. 1972. *The Foundations of Statistics*. Dover Publications.

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

[Wilcox 2002] Wilcox, R. 2002. *Applying contemporary statistical techniques*. Academic Press.

[Yang et al. 2011] Yang, R.; Kiekintveld, C.; Ordonez, F.; Tambe, M.; and John, R. 2011. Improving resource allocation strategy against human adversaries in security games. In *IJCAI*, 458–464.

[Yin et al. 2012] Yin, Z.; Jiang, A.; Johnson, M.; Tambe, M.; Kiekintveld, C.; Leyton-Brown, K.; Sandholm, T.; and Sullivan, J. 2012. Trusts: Scheduling randomized patrols for fare inspection in transit systems. In *IAAI*.