ECAI 2012
Luc De Raedt et al. (Eds.)

doi:10.3233/978-1-61499-098-7-660

A Robust Approach to Addressing Human Adversaries in Security Games

James Pita and **Richard John** and **Rajiv Maheswaran** and **Milind Tambe** and **Sarit Kraus*** [1]

Abstract. Game-theoretic approaches have been proposed for addressing the complex problem of assigning limited security resources to protect a critical set of targets. However, many of the standard assumptions fail to address human adversaries who security forces will likely face. To address this challenge, previous research has attempted to integrate models of human decision-making into the game-theoretic algorithms for security settings. The current leading approach, based on experimental evaluation, is derived from a well-founded solution concept known as quantal response and is known as BRQR. One critical difficulty with opponent modeling in general is that, in security domains, information about potential adversaries is often sparse or noisy and furthermore, the games themselves are highly complex and large in scale. Thus, we chose to examine a completely new approach to addressing human adversaries that avoids the complex task of modeling human decision-making. We leverage and modify robust optimization techniques to create a new type of optimization where the defender's loss for a potential deviation by the attacker is bounded by the distance of that deviation from the expected-value-maximizing strategy. To demonstrate the advantages of our approach, we introduce a systematic way to generate meaningful reward structures and compare our approach with BRQR in the most comprehensive investigation to date involving 104 security settings where previous work has tested only up to 10 security settings. Our experimental analysis reveals our approach performing as well as or outperforming BRQR in over 90% of the security settings tested and we demonstrate significant runtime benefits. These results are in favor of utilizing an approach based on robust optimization in these complex domains to avoid the difficulties of opponent modeling.

1 Introduction

Game theory has gained attention in security resource allocation decisions in important settings [2, 6]. Traditionally, Stackelberg games have been used to model these problems because they encapsulate the commitment a defender must make in allocating her resources before an attacker surveys their defensive strategy and chooses an attack method. In fact, algorithms utilizing a Stackelberg framework have been featured in real-world resource allocation decision aids [6]. However, a key assumption underlying the technique in these systems is that the attacker is a perfectly-rational player. Thus, these systems optimize their strategy against an expected-value-maximizing opponent and are not robust to deviations from this strategy.

It is well known that standard game-theoretic assumptions of expected-value-maximizing rationality are not ideal for addressing human behavior in game-theoretic settings [3]. In addressing this bounded rationality of human adversaries, different models have been proposed [11, 13]; however, within the security game setting, BRQR [13], has emerged as the leading approach. BRQR is based on the quantal response (QR) model [9] of human decision-making. This QR model is a well-founded solution concept in game theory derived from Nobel-prize-winning work in choice modeling theory [8], and there has been significant support for this QR model elsewhere in the literature [12]. A major difficulty of this modeling approach however is that it requires the estimation of a parameter that determines the level of noise in the human adversary's response function. In real-world scenarios that are often complex and large in scope, creating an accurate model of human-decision making can be a difficult task. This difficulty is exacerbated in security settings where information about potential adversaries is often sparse and noisy.

We take an alternative approach that has not been explored before based on robust optimization [1] where the defender strategy is robust to certain worst-case deviations from the attacker. However, we significantly modify the standard worst-case assumption of robust optimization and, instead, bound the defender's loss for a potential deviation by the human attacker based on the degree of the deviation from the expected-value-maximizing strategy. Our new algorithm, MATCH, provides three key benefits: (i) it provides significant runtime benefits over BRQR; (ii) it strongly couples the attacker's and defender's performance, robustly guarding against potential deviations by human adversaries and avoiding situations where minor deviations (i.e., deviations that result in minor losses in expected value) by the attacker may result in large losses for the defender; (iii) it avoids the dilemma of creating an accurate opponent model. We will refer to this new type of optimization as graduated optimization and show in Section 3 that it lies within a space between MAXIMIN and the standard game-theoretic optimal solution.

To evaluate the advantages of our new approach, we make the most comprehensive investigation to date. Whereas previous work has examined few security settings ($\leq$ 10 settings) [11, 13], we examine 104 security settings. We examine the four recommended security settings from Yang et al. [13] and we also intelligently select 100 additional payoff structures, which we will describe in detail in Section 4. Furthermore, in Section 4 we defend why we believe this experimental setup is superior to previous setups [11, 13]. We test our 104 security game settings against 363 human subjects playing 8823 games in total to compare the performance of MATCH against BRQR. Our results reveal that MATCH performs as well as or better than BRQR against human adversaries in over 90% of the settings tested and in Section 5 we give an analysis of these results.

[1] University of Southern California, Los Angeles, CA 90089 and *Bar-Ilan University, Ramat-Gan 52900, Israel and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742

2 Background and Related Work

Security games refer to a special class of Stackelberg games where there are two agents – the defender and an attacker – who act as the leader and the follower respectively [14]. There exists a number of game-theoretic optimal solvers for security games [4, 6]. These algorithms compute a strong Stackelberg equilibrium (SSE) [4]. A SSE assumes an attacker will both choose a strategy that maximizes his expected value and (with a technical justification [6]) break ties in the defender's favor. However, in real-world settings, security forces often face human adversaries who may not perceive minor differences in expected value and may not break ties in the defender's favor.

To that end, COBRA was introduced [11], which assumes the opponent plays an ϵ-optimal response strategy. That is, COBRA assumes the attacker is willing to choose any strategy with an expected value within ϵ of the maximum expected value strategy. COBRA then attempts to maximize the defender's utility for the worst-case outcome of any ϵ-optimal response strategy, avoiding the issue of tie breaking by the attacker.

One critical issue with COBRA is that it has a hard cutoff point (i.e., ϵ) and if the attacker deviates to any strategy beyond an ϵ-optimal response the result can once again be arbitrarily bad for the defender. To address this dilemma, Yang et al. [13] introduced Best Response to Quantal Response (BRQR) as an efficient model for computing a resource allocation strategy based on quantal response equilibrium (QRE) [9], a well-founded solution concept within game theory based on Nobel prize winning work in Choice Modeling theory [8]. QRE suggests that instead of strictly maximizing expected value, individuals respond stochastically in games: the chance of selecting non-optimal strategies increases as the cost of such an error decreases. In BRQR, noise is only added to the response function for the adversary, so the defender computes an optimal strategy assuming the attacker responds with a noisy best-response. BRQR thus allows for a more gradual approach to defending against deviations as opposed to a hard-cutoff point and has been experimentally shown to be the current leading approach for addressing human adversaries.

Two issues with BRQR are that it critically depends on the appropriate estimation of λ, which represents the amount of error or noise in the attacker's response function, and that its runtime is slow. In security domains where data can be sparse, noisy, and the games are highly complex; determining an appropriate setting for λ can be difficult. Furthermore, if human adversaries deviate from the predicted response distribution (opponent model), the result can once again be grossly negative for the defender. We will present an algorithm, MATCH, that addresses the issues in both COBRA and BRQR.

3 MATCH Algorithm

The key concept behind the MATCH algorithm is the new idea of graduated robust optimization. Whereas standard robust optimization robustly guards against a worst-case outcome within some error bound, MATCH assumes a utility maximizing outcome on behalf of the attacker, but constrains the impact of deviations depending on the magnitude of the deviation. Specifically, the defender's loss for a potential deviation by the attacker is bounded based on the distance of that deviation from the expected-value-maximizing strategy. We present the Mixed Integer Linear Program (MILP) for MATCH in Equation 1, however, we first formally define our problem space.

In a security game [14], the defender has $K \in \mathbb{N}$ resources to protect a set of targets $T = \{t_1, \ldots, t_n\}$ from the attacker. Each player has a set of pure strategies: the defender can *cover* a set of K targets, and the attacker can *attack* one target. The payoffs for each player depend on the target attacked, and whether or not that target was *covered*. If a target $t_i \in T$ is *covered* the defender will receive $R_i^d \in (0, \infty)$ and the attacker will receive $P_i^a \in (-\infty, 0)$ otherwise the defender will receive $P_i^d \in (-\infty, 0)$ and the attacker will receive $R_i^a \in (0, \infty)$. The defender's strategy is denoted by x where $x_i \in [0, 1]$ represents the probability with which the defender covers target $t_i \in T$. This coverage vector is equivalent to obtaining a probability distribution over pure strategies of selecting K targets (i.e., a mixed strategy) [7]. The attacker's strategy is denoted by $q \in \{1, \ldots, n\}$ and represents the target the attacker chooses to attack. Let $U^a(q, x) = x_q \cdot P_q^a + (1 - x_q) \cdot R_q^a$ and $U^d(q, x) = x_q \cdot R_q^d + (1 - x_q) \cdot P_q^d$. We define $\beta \in [0, \infty)$. In the following MILP the defender's goal is to maximize her expected utility which we represent as γ:

$$
\begin{aligned}
\max \quad & \gamma \\
\text{s.t.} \quad & \sum_{i \in T} x_i = K && (1) \\
& 0 \leq x_i \leq 1 && (2) \\
& q = \arg \max_{\hat{q} \in \{1, \ldots, n\}} U^a(\hat{q}, x) && (3) \\
& \gamma \leq U^d(q, x) && (4) \\
& \beta \cdot (U^a(q, x) - U^a(\hat{q}, x)) \geq \\
& \gamma - U^d(\hat{q}, x) \; \forall \, x_{\hat{q}} < 1 && (5)
\end{aligned}
$$

Constraints (1) and (2) ensure that the defender utilizes all her resources and that no target has more than 1 resource assigned to it. Constraint (3) ensures that the attacker chooses the target that maximizes his expected value. Constraint (4) ensures that the defender obtains the corresponding expected value (γ) to the attacker's optimal strategy. Constraint (5) is the most crucial portion of the formulation. The left portion calculates the attacker loss in expected value for a deviation from the optimal strategy. The right hand side constrains the defender loss in expected value for this deviation by the attacker to be no more than a factor of β times the loss the attacker receives. For example, if the defender does not want to lose any more than twice what the attacker loses for a potential deviation, then we can set $\beta = 2$. If the defender does not want to lose any more than half what the attacker loses, then we can set $\beta = .5$. This provides a direct trade-off between the defender's maximum utility for the attacker's optimal strategy and additional protection on potential weaknesses.

MATCH addresses the issues in both COBRA and BRQR. *A fundamental property of MATCH is that it does not rely on some complex non-linear non-convex optimization problem (e.g., as in BRQR or other approaches such as RPT* [13]*).* Indeed, its power is in its perceived simplicity, which not only means its simple to implement, but its orders of magnitude faster than its competitors. In addition, similar to BRQR, it allows for a more gradual defense against deviations as opposed to a hard cutoff point. However, MATCH avoids the challenge of creating an accurate opponent model of human decision-making by relying instead on a form of robust optimization. Nonetheless, MATCH still faces one crucial consideration, which is a trade-off between robustness and defender utility. The key difference between MATCH and BRQR is that BRQR attempts to model human decision-making; but if this model is inaccurate, the defender's performance suffers. MATCH in contrast bypasses modeling of the human decision-making process; it instead directly focuses on how much maximum utility a defender is willing to trade off to protect against the human attacker's potential deviations from the rational

strategy. While the β-parameter can be adjusted, in our experimental sections we will consistently keep $\beta = 1$ and show that even with this flat setting without any tuning, MATCH outperforms BRQR with careful tuning of λ. The performance of MATCH might be enhanced with alternative β-settings, however, finding an appropriate procedure for estimating the β-parameter is left for future work.

Proposition 1 *If β is sufficiently large, then MATCH solves for a strong Stackelberg equilibrium (SSE).*

Intuitive Justification: If β is sufficiently large then Constraint (5) is trivially satisfied when $U^d(q, x) > U^a(\hat{q}, x)$, effectively removing it. In order for x, q and γ to be a SSE they must meet three defined criteria [7]: (i) the leader plays a best response, (ii) the follower plays a best response, (iii) the follower breaks ties optimally for the leader. Constraint (3) ensures the follower plays a best response to the leader strategy x. The objective, γ, and Constraint (4) ensure the leader plays a best response. If there were an alternative x such that γ could be increased it would be selected. Finally, Constraint (5) actually enforces tie breaking in favor of the leader. If $U^a(q, x) = U^a(\hat{q}, x)$ for any $\hat{q} \neq q$ then Constraint (5) becomes $0 \geq \gamma - U^d(\hat{q}, x)$. There are three possible outcomes. First, $\gamma = U^d(\hat{q}, x)$, which means that either choice is favorable for the leader. Second, $\gamma > U^d(\hat{q}, x)$, which means that Constraint (5) is no longer satisfied and the attacker has an alternative optimal strategy that is not in favor of the leader. Here, in order to induce the favorable outcome, the MILP will enforce $U^a(q, x) > U^a(\hat{q}, x) + \epsilon$, where ϵ will be the smallest possible increment given that β is sufficiently large. By definition of a security game, this can be done by removing the smallest increment of probability from x_q and placing it on $x_{\hat{q}}$ to enforce the favorable outcome (breaking the tie in favor of the leader). Finally, $\gamma < U^d(\hat{q}, x)$, making this deviation favorable for the leader, and by the same logic can be induced[2].

Proposition 2 *If $\beta = 0$, then MATCH provides an equivalent worst-case bound to MAXIMIN.*

Intuitive Justification: If $\beta = 0$ it follows that Constraint (5) becomes $\gamma \leq U^d(\hat{q}, x) \ \ \forall \ x_{\hat{q}} < 1$. Assuming that $x_i < 1 \ \ \forall \ i \in \{1, \ldots, n\}$, by definition Constraints (4) and (5) maximize the leader's minimum utility since $\gamma \leq U^d(i, x) \ \ \forall \ i \in \{1, \ldots, n\}$. If $\exists \ x_j = 1$ then we will show that $U^d(j, x) \leq U^d(i, x) \ \ \forall \ x_i < 1, i \in \{1, \ldots, n\}$, guaranteeing that $U^d(j, x)$ is the best worst-case bound by definition of a security game since x_j cannot be increased further. Consider (x, q) an optimal solution for MATCH with $\beta = 0$. Let $x_j = 1$ and assume $\exists \ i \in \{1, \ldots, n\} : U^d(i, x) < U^d(j, x), x_i < 1$. It follows from Constraint (5) that $\gamma \leq U^d(i, x)$. By definition of a security game, the defender's expected utility could be improved by increasing the value of x_i and this could be accomplished by directly trading probability from x_j to x_i at least until $U^d(j, x) = U^d(i, x)$, a contradiction since (x, q) is an optimal solution.

4 Experiments

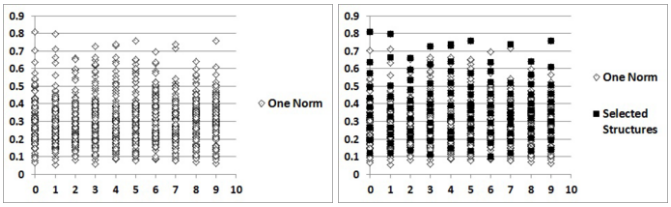
The bulk of this paper is now devoted to carefully constructed extensive experiments comparing MATCH with BRQR. Our experimental methodology and the scope of the experiments is a key contribution. In these experiments, we utilize the same eight-target scenario used by Yang et al. [13] where three guards – jointly acting as the defender – guard eight gates, and each human subject acts as a single attacker who will choose one of the eight possible gates. Before beginning the game, subjects were given a brief tutorial about how to play and a short test to ensure that they understood the general game play. In order to simulate the Stackelberg setting, we presented subjects with the following information before they chose a gate: (i) the subject's reward and penalty for each gate; (ii) the defender strategy (i.e., the probability distribution of the guards over the 8 gates); (iii) the guard's reward and penalty for each gate. We use this Stackelberg framework because in real-world scenarios an attacker can conduct extensive surveillance of his potential target and the corresponding defensive strategy before choosing to attack, which would allow them to learn this information. In each game instance, the guards would choose 3 gates based on their strategy and the subject's goal was to choose the gate that would maximize expected value given the defender's strategy. However, in a particular game instance the subject would fail or succeed based on where the guards were actually stationed. Subjects were given unlimited time to make a decision.

Our experiments were run in Amazon Mechanical Turk and participants were paid a base amount of US \$1.50 for participating. To ensure that subjects were not choosing gates arbitrarily we introduced two obvious games where a gate with the highest reward and lowest penalty possible had the lowest probability (5%) of being covered. If subjects did not choose this gate in these two games their results were removed. To further motivate the subjects, we allowed them to earn additional money based on their performance in the game. For each reward point earned or lost, a subject would receive or lose an additional US \$0.15 from their net money. Before playing, subjects were informed that only a small sample of their games would be selected from all games played to determine the actual bonus payment. Since subjects were not aware which games would be chosen, they would have incentive to perform the best they could in each game and they were given immediate feedback at the end of each game. Subjects were paid their final earnings, but were not required to pay money back if their net money earned was less than zero.

Given this experimental setup, we ran experiments in two sets of reward structures. We first explored the four reward structures proposed by Yang et al. [13], which were chosen to be the most representative of the entire payoff structure space for security games based on metrics proposed by Yang et al. [13]. However, in these particular reward structures the strategies produced by BRQR and MATCH were highly similar. Thus, to more fully compare the performance of BRQR and MATCH, for the second set of experiments we systematically chose 100 new reward structures based on covariant games in GAMUT [10]. We chose covariant games because they are naturally able to capture the adversarial nature of security games. In covariant games, we can adjust a parameter $r \in [-1, 1]$, which determines a correlation between player rewards. Specifically, when $r = -1$ the game is zero-sum between players and when $r = 1$ the game is perfectly cooperative. We slightly modified the code in GAMUT to restrict the resulting payoffs to meet the criteria of a security game.

To select the 100 new reward structures we first generated 1000 reward structures with r ranging from -.9 to 0 by .1 increments (i.e., 100 games for each setting of r). We chose 0 as the boundary because, in an adversarial setting, it does not make sense that the payoffs would be positively correlated. To examine how different MATCH and BRQR are overall, for each structure we computed the 1-norm distance between MATCH and BRQR where we set $\lambda = .75$ and $\beta = 1$. While we discuss estimating appropriate λ-settings in Section 6, our choice of $\lambda = .75$ was inspired by the original estimate made by Yang et al. [13]. We chose $\beta = 1$ to constrain the defender losses to be no worse than the attacker losses for deviations.

[2] For proofs of both propositions visit http://matchproof.webs.com.



(a) 1-Norm (b) Selected Structures

Figure 1. 1-Norm Scatter Plots

In Figure 1(a) we present the scatter plot for these 1000 reward structures. For readability, on the x-axis we display the setting of r from 0 to $-.9$ as 0 to 9 (i.e., on the x-axis 3 represents -.3). On the y-axis we display the 1-norm value. We see from these scatter plots that there is a wide range of possibilities for the difference between MATCH and BRQR strategies. While extreme differences don't occur as frequently, they do not appear to be completely rare. Given these 1000 structures, we attempted to select 100 reward structures that would best cover the possible space based on the 1-norm values. We present the 100 structures selected in Figure 1(b).

We believe this experimental setup is superior to previous setups [11, 13] for three critical reasons: i) by examining the 1-norm distance we can explore a spectrum of reward structures where the strategies produced are most different (high 1-norm) to where they are most similar (low 1-norm); ii) by utilizing covariant games, we can control the correlation between player rewards to ensure rewards and penalties are not positively correlated where previous experiments have ignored this crucial issue; and iii) previous results may give a distorted view of the overall performance of an algorithm compared to other algorithms since they look at such a narrow portion of the entire security game space (i.e., 4 to 10 potential settings versus over 100). In the following, we will first present the results for the reward structures proposed by Yang et al. [13], then the results for the newly selected structures, and finally we will give an analysis of these results. We evaluate the statistical significance of our results using the bootstrap-t method used by Yang et al. [13] previously.

4.1 Results for Original Structures

In these experiments, we tested the mixed strategies generated by BRQR and MATCH using the original λ estimate made by Yang et al. [13] (i.e., $\lambda = .76$) and $\beta = 1$ to constrain the defender's losses to be no worse than the attacker's losses. To avoid boredom in the subjects, we limited the number of games they would have to play by separating the reward structures into two groups. Subjects either played all strategies against reward structures 1 and 2 or all strategies against reward structures 3 and 4. We present the results of these experiments in Figure 2. In total, 36 subjects played against reward structures 1 and 2 while 33 played against reward structures 3 and 4. In Figure 2, the y-axis represents the average defender expected value over all the choices made by each individual subject against a particular strategy. For example, examining Structure 1 in Figure 2, we see that the defender received -0.29 on average against human subjects if using the strategy generated by BRQR.

In reward structures 1 and 4 we find that MATCH is statistically significantly better than BRQR ($p = .028$ and $p = .004$ respectively). In reward structures 2 and 3 neither strategy was statistically significantly better than the other ($p = .15$ and $p = .392$ respectively). However, in general across all 4 payoff structures, MATCH and BRQR create highly similar strategies (i.e., the probability difference on any particular gate is relatively low [$\leq 12.6\%$]). Regardless,

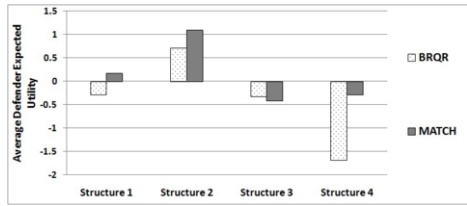


Figure 2. Original Reward Structures

we still see that MATCH outperforms BRQR with statistical significance in two of the four reward structures and does at least as well in the other two. Based on these results we can conclude that there exists conditions where MATCH is the superior algorithm to BRQR.

4.2 Results for New Reward Structures

While the results from our first experiment were promising, we wanted to take the most extensive look to date at a large space of potential security game settings, examining 100 potential reward structures compared to 10 or less in previous experiments [13, 11]. As stated previously, in these experiments we tested the mixed strategies generated by BRQR and MATCH using $\lambda = .75$ and $\beta = 1$. Once again to avoid boredom in the subjects, we limited the number of games they would have to play by separating the reward structures into the following groups: (i) Structures 5-9 (we start at 5 to account for the previous 4 reward structures) [25 participants], (ii) Structures 10-21 [33 participants], (iii) Structures 22-36 [37 participants], (iv) Structures 37-53 [40 participants], (v) Structures 54-70 [37 participants], (vi) Structures 71-87 [42 participants], and (vii) Structures 88-104 [39 participants]. Subjects would play against all the strategies for a given group of reward structures.

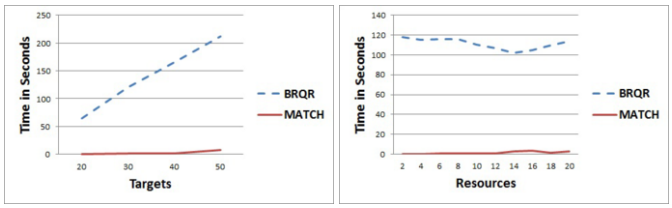
We present an overview of the results from these experiments in Table 1. Here, we show the number of settings where MATCH won with statistical significance, both strategies were approximately equivalent (i.e., neither strategy won with statistical significance), and BRQR won with statistical significance. What we find is that in 42 of the 100 reward structures MATCH outperformed BRQR with statistical significance and in an additional 52 of the 100 reward structures MATCH did at least as well as BRQR given that neither strategy won with statistical significance. These results combined with our previous experiment show MATCH performing at least as well as or outperforming BRQR in 98 out of 104 potential security settings. In Section 5 we will give further analysis of these results.

	MATCH	Draw	BRQR
$\alpha = .05$	42	52	6

Table 1. Overview of Results

Runtime Results: In Figure 3 we present runtime results for BRQR versus MATCH. In Figure 3(a), the number of resources are fixed at 10 and on the x-axis we vary the number of targets from 10 to 50. On the y-axis we present the runtime in seconds averaged over 20 randomly generated payoff structures. In Figure 3(b), the number of targets are fixed at 30 and on the x-axis we vary the number of resources from 2 to 20. Once again on the y-axis we present the runtime in seconds averaged over 20 randomly generated payoff structures. Based on these results, we see that MATCH provides orders of magnitude speedup over BRQR further demonstrating the benefits of such an approach. The reason for this runtime improvement is that BRQR requires the solution to a non-linear and non-convex objective

function in its most general form. In fact, because of the complexity of the objective function, BRQR is only a heuristic solution for solving the objective. MATCH on the other hand is a mixed-integer linear program, which can be solved with standard packages.



(a) Increasing Targets (b) Increasing Resources

Figure 3. Runtime results

5 Analysis

The QR model is a well-established solution concept and so an important question to address is whether BRQR was actually an accurate model of human decision-making in these security settings. To determine whether BRQR is accurate, in each reward structure we run a Pearson's chi-squared goodness of fit test [5] on the predicted distribution of attacker choices against the observed attacker choices for our subjects. We present the results in Table 2. In the first three columns we denote reward structures where MATCH won with significance, neither strategy won with significance, and BRQR won with significance. In the last column we give the overall result for all 100 reward structures. In the rows we denote whether Pearson's chi-squared goodness of fit test rejected the null hypothesis that the observed distribution of choices could have been drawn from the expected distribution of choices ($\alpha = .05$).

	MATCH	Draw	BRQR	TOTAL
Rejected	40	40	3	83
Not Rejected	2	12	3	17

Table 2. Pearson Chi-squared Results

Our first observation is that in 83% of all reward structures tested the model proposed by BRQR did not fit the data observed. This is a significant number and suggests that perhaps BRQR is not a good model of human decision-making in security games. However, it is possible that this result is due to a poor estimation of the λ-parameter for these particular security settings and so in Section 6 we will re-estimate the λ-parameter based on the observed data and run additional experiments for a key subset of the reward structures. Even so, in real-world security settings it may be even more difficult to appropriately estimate λ since data can often be sparse or noisy, and the problem instances can be much larger and more complex.

The fact that, for this set of results, BRQR does not provide a good fit for the data observed in general is one potential explanation for why MATCH is outperforming BRQR in the majority of the security settings. BRQR attempts to exploit an assumed model of the human attacker and if the attacker deviates from that model in a significant way it can severely impact the performance of the defender. For instance, if BRQR assumes that an attacker is not likely to attack a certain target it will provide minimal coverage for that target. If, however, a large number of attackers choose this target, it can have severe effects on the defender's average expected utility. An inaccurate model of human decision-making can lead to severe consequences in security domains. This is one of the key advantages of MATCH since it does not assume any decision-making model and specifically bounds the impact of such potential deviations.

Our second observation is that of the 17 structures where BRQR was potentially a good fit of human decision-making, it only outperformed MATCH in 3. Thus, of the six cases where BRQR won with statistical significance, only three of the cases can potentially be justified by accurate opponent modeling. These results show that even a decent model of human decision-making may not be sufficient enough and an approach based on robust optimization is a potentially strong alternative to modeling human decision-making processes. It may be necessary to have a highly accurate model of human decision-making before it becomes sufficient in security settings.

A second important question is, given the space of possible reward structures, can we determine when MATCH or BRQR will likely perform better. In Figure 4, we present where the results for our 100 structures appear in the scatter plot of 1000 structures. In the cases where neither MATCH nor BRQR won with statistical significance we present which strategy had the higher average defender expected utility. While this does not imply that the strategy is better in these cases, this was done to see if it would provide some insight into what portion of the payoff structure space MATCH or BRQR performed better. Here, it is evident that the results are diverse within the potential space and so without further investigation we cannot determine where BRQR will likely be better than MATCH overall. However, these results support our earlier argument that when doing experimental analysis in security settings, examining few potential settings can give misleading results. For example, given a small sample size from the potential space, it is possible we could have chosen only structures where BRQR outperformed MATCH, which would be contrary to what our results have demonstrated. Thus, we have rightly raised the standard for future experimental investigations.

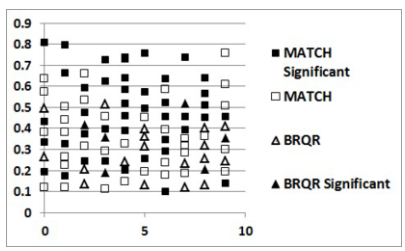


Figure 4. Scatter Plot of Results

6 λ-Re-estimation

As suggested previously, to confirm whether BRQR is actually a poor predictor of human decision-making we are required to examine it with appropriately estimated λ-values. To focus our analysis we selected three groups of five reward structures from the 42 reward structures where MATCH outperformed BRQR with statistical significance as follows: (i) the five structures where BRQR and MATCH had the most significant strategy difference averaged over the 1-norm, 2-norm, infinite-norm, and KL distances; (ii) the five structures where BRQR and MATCH had the least difference in average expected utility; and (iii) the five structures where BRQR and MATCH had the highest difference in average expected utility. For these experiments we will refer to these as structures 1 through 15. To re-estimate the λ-parameter we used the Maximum Likelihood Estimation procedure proposed by Yang et al. [13] using the data from the previous experiments in each of the 15 reward structures yielding 15 new λ-values. In Table 3 we present the new λ-estimates along with the 1-norm distance between the strategy produced by the original λ-setting ($\lambda = .75$) and the re-estimated λ-setting.

Our first observation is that λ is largely dependent on the reward structure implying that for each potential security domain the defender would be required to make a new estimate. As previously stated, estimating λ can be difficult in real-world settings where data may be sparse or noisy. This problem is further exacerbated since we have shown that data cannot likely be pooled from different settings. This is an additional advantage of our approach since the level of robustness is not dependent on the reward structure. That is, upon deciding a β-setting it is consistent across all reward structures where a λ-setting is not equivalently accurate for all reward structures.

Structure:	1	2	3	4	5	6	7	8
λ:	.18	.71	.25	1.14	.01	1.39	1.09	.67
1-norm:	.376	.012	.384	.177	3.10	.221	.110	.050
Structure:	9	10	11	12	13	14	15	
λ:	.84	.48	.15	.43	.55	.23	.42	
1-norm	.018	.069	.424	.396	.127	.244	.356	

Table 3. New λ-estimates

Our second observation is that the actual impact of altering the λ-parameter varies significantly depending on the reward structure. For example, in structure 10 we vary λ from .75 to .48 and see only a 1-norm difference of .069 while in structure 12 we reduce λ to .43 and see a 1-norm difference of .396 (i.e., there is a 40% difference of probability across targets in one case and only a 7% difference across targets in another case). This once again demonstrates the difficulty in appropriately estimating λ since minor changes can lead to significant differences in some reward structures.

We had all 41 subjects play against both MATCH ($\beta = 1$) and BRQR with newly estimated λ-values in all 15 reward structures. We present the results in Figure 5. In these results, MATCH remained statistically significantly better in 8 of the 15 reward structures (structures 1-5, 11, 13, and 15) and neither strategy was statistically significantly better in the remaining 7. Additionally, we ran Pearson's chi-squared goodness of fit test and found that, even after re-estimation, the null hypothesis that the observed choice distribution could have been drawn from the predicted choice distribution was rejected in all 15 cases. Thus, even if we obtain tailored λ estimates, MATCH continues to perform as well as or outperform BRQR.

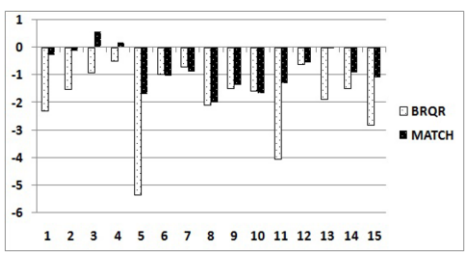


Figure 5. Re-estimated Reward Structures

7 Summary

Game theory continues to be a useful tool for motivating resource allocation decisions in important security settings. However, one critical assumption of standard game-theoretic approaches is that the attacker maximizes his expected value. Such an approach is not robust to deviations by a potential attacker and in the real-world deviations are likely since defender's face a boundedly rational human adversary. A number of models have been proposed to try and address these potential deviations including BRQR, which was formerly the best known approach for addressing humans. Our work provides five fundamental contributions to this research: (i) we develop an approach, MATCH, to addressing human adversaries based on robust optimization rather than relying on finding appropriate models of human decision-making; (ii) we introduce a systematic way to generate meaningful reward structures based on covariant games where previous work has simply generated completely random reward structures; (iii) we make the most comprehensive evaluation to date involving 363 human subjects playing 8823 games in 104 security game settings, whereas previous work has examined at most 10 security game settings; (iv) we demonstrate that MATCH performs as well as or better than BRQR in over 94% of the security settings tested (42 of 104 settings with statistical significance); and (v) we demonstrate the significant runtime benefits of MATCH over BRQR. These results demonstrate the potential benefits of using an approach based on robust optimization over previous algorithms that rely on creating more efficient models of human-decision making.

8 Acknowledgments

This research was supported by the United States Department of Homeland Security through the Center for Risk and Economic Analysis of Terrorism Events (CREATE) under grant number 2007-ST-061-000001. However, any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the United States Department of Homeland Security.

REFERENCES

[1] Michele Aghassi and Dimitris Bertsimas, 'Robust game theory', *Math. Program.*, **107**(1-2), 231–273, (2006).
[2] Noa Agmon, Vladimir Sadov, Sarit Kraus, and Gal Kaminka, 'The impact of adversarial knowledge on adversarial planning in perimeter patrol', in *AAMAS*, (2008).
[3] Colin Camerer, in *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, (2003).
[4] Vincent Conitzer and Tuomas Sandholm, 'Computing the optimal strategy to commit to', in *EC*, (2006).
[5] Priscilla E Greenwood and Michael S Nikulin, *A Guard to Chi-squared Testing*, John Wiley & Sons, Inc., 1996.
[6] Manish Jain, Jason Tsai, James Pita, Christopher Kiekintveld, Shyamsunder Rathi, Fenando Ordóñez, and Milind Tambe, 'Software assistants for randomized patrol planning for the LAX airport police and the Federal Air Marshals Service', *Interfaces*, **40**(4), 267–290, (2010).
[7] Chris Kiekintveld, Manish Jain, Jason Tsai, James Pita, Milind Tambe, and Fernando Ordóñez, 'Computing Optimal Randomized Resource Allocations for Massive Security Games', in *AAMAS*, (2009).
[8] Daniel L McFadden, 'The sveriges riksbank prize in economic sciences in memory of alfred nobel', (2000).
[9] Richard McKelvey and Thomas Palfrey, 'Quantal response equilibria for normal form games', *Games and Economic Behavior*, **10**, 6–38, (1995).
[10] Eugene Nudelman, Jennifer Wortman, Yoav Shoham, and Kevin Leyton-Brown, 'Run the GAMUT: A comprehensive approach to evaluating game-theoretic algorithms', in *AAMAS*, (2004).
[11] James Pita, Manish Jain, Milind Tambe, Fernando Ordóñez, and Sarit Kraus, 'Robust solutions to stackelberg games: Addressing bounded rationality and limited observations in human cognition', *Artificial Intelligence Journal*, **174**(15), 1142–1171, (2010).
[12] James R Wright and Kevin Leyton-Brown, 'Beyond equilibrium: Predicting human behavior in normal-form games', in *AAAI*, (2010).
[13] Rong Yang, Christopher Kiekintveld, Fernando Ordóñez, Milind Tambe, and Richard John, 'Improving resource allocation strategy against human adversaries in security games', in *IJCAI*, (2011).
[14] Zhengyu Yin, Dmytro Korzhyk, Christopher Kiekintveld, Vincent Conitzer, and Milind Tambe, 'Stackelberg vs. Nash in security games: Interchangeability, equivalence, and uniqueness', in *AAMAS*, (2010).