# A Game Theoretic Approach on Addressing Collusion among Human Adversaries

Shahrzad Gholami, Bryan Wilder, Matthew Brown, Arunesh Sinha, Nicole Sintov, Milind Tambe
University of Southern California, USA,
{sgholami,bwilder,matthew.a.brown,aruneshs,sintov,tambe}@usc.edu

## ABSTRACT

Several models have been proposed for Stackelberg security games (SSGs) and protection against perfectly rational and bounded rational adversaries; however, none of these existing models addressed the collusion mechanism between adversaries. In a large number of studies related to SSGs, there is one leader and one follower in the game such that the leader takes action and the follower responds accordingly. These studies fail to take into account the possibility of existence of group of adversaries who can collude and cause synergistic loss to the security agents (defenders). The first contribution of this paper is formulating a new type of Stackleberg security game involving a beneficial collusion mechanism among adversaries. The second contribution of this paper is to develop a parametric human behavior model which is able to capture the bounded rationality of adversaries in this type of collusive games. This model is proposed based on human subject experiments with participants on Amazon Mechanical Turk (AMT).

## Categories and Subject Descriptors

H.4 [**Security and Multi-agent Systems**]:

## General Terms

Algorithms, Experimentation, Security

## Keywords

Game Theory, Stackelberg Security Games, Human Behavior Models, Collusion

## 1. INTRODUCTION

Security agencies including the US Coast Guard (USCG), the Federal Air Marshal Service (FAMS) and the Los Angeles Airport (LAX) police are several major domains that have been deploying Stackelberg security games (SSGs) and related algorithms to protect against adversaries strategically [12]. The security games introduced in these domains, mostly, include two players: a defender and an adversary. The interaction between the defender and the attacker was modeled as a single-shot game and the attacker was defined as a perfectly rational player. A major characteristic of this class of SSGs is that, they are sequential. In other words, one player (the leader or the defender) commits to a strategy which can be observed by the other player (the follower or adversary) before choosing his own strategy.

There are different variations of the SSGs in literatures. As an example, to address the idea that the leader might be uncertain about the types of adversary that might attack (known as Bayesian Stackelberg games), an efficient exact algorithm is proposed in [11] to develop the optimal strategy for the leader. As an another example, repeated interactions of defender and the adversary is studied in [7]. This type of game is famous in the wildlife security domain and fisheries protection. In this game the defender deploys new patrolling strategies periodically and the adversary observes these strategies and responds accordingly. [5] and [15] propose models and algorithms against boundedly rational adversaries using behavioral models such as quantal response (QR) [16] and subjective utility quantal response (SUQR) [10] to model human adversaries. In protecting wildlife domain which is an active area of research in security game, preventing the poachers from hunting animals in forest area by effieviently and strategically patrol allocation is vital. In [3] and [2] Green security games are introduced, algorithms and field optimization techniques for planning effective sequential defender strategies are proposed to tackle the problem of protection of endangered animals and fish stocks.

In wildlife protection domain, international illegal trade is increasing incredibly and based on the estimations, it is worth at least $5 billion, annually. The main types of wildlife commodities that are subject to these illegal trades include elephant ivory, rhino horn, tiger parts and caviar, to name a few. These activities have the potential to introduce several threats to the national security and evironment arround the world. Biodiversity loss, potential extinctions, introduction of invasive species and desease transmission into healthy ecosystems, all can impact the environment adveresly. In addition to that, some connections have been observed among wildlife trafficking, organized crime and drug trafficking which means that poor law enforcement, poor patrol scheduling or corrupt rangers at wildlife sources, corrupt governments at transit countries and porous borders can all threaten the national security [14]. Despite the evidence of illegal exchange between different groups of criminals, the destructive synergistic effect of collusion among adversaries is unexplored in related literature in security game domain.

To combat this illegal wildlife trade, exploitation and col-

**Appears in:** *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), John Thangarajah, Karl Tuyls, Stacy Marsella, Catholijn Jonker (eds.), May 9–13, 2016, Singapore.*

laboration among criminals and adversaries, this papers addresses a new type of security game in which there are three players, one defender, who is the leader, and two adversaries who are the followers and have the option of collusion with each other. Each adversary has access to his own targets but he can team up with another adversary to share all of the pay-offs with him. Each of these adversaries can be a representative for either a poacher who is directly hunting in the field or a trader who is illegally exchanging the animals or financing other illegal commodities via animal trafficking.

# 2. COLLUSIVE SECURITY GAME

In this section, a detailed analysis of the collusive security games for rational adversaries is presented.

## 2.1 Collusive game model: Tertiary case

A generic security game problem as a normal form Stackelberg game has two players, a defender and an attacker. In the collusive form of the game which we study in this paper, there can be one defender, $\Theta$, and more than one attacker, $\Psi_1$,...,$\Psi_N$, where $N$ is the total number of attackers and similar to normal Stackelberg games, defender is the leader and attackers are the followers. In this subsection, we focus on the zero-sum games with one leader and two followers, such that followers can attack separate targets, but they have two options: i) attack their own targets individually and earn pay-offs independently or ii) attack their own targets individually but collude with each other and share all of the pay-offs equally. Attackers pay-off are not identical in the two above mentioned cases. There are some extra bonus reward, $\epsilon$, for collusive attacking that can motivate the adversaries for collusion.

To discuss this more precisely, let $T = \{t_1, ..., t_n\}$ be a set of targets that may be attacked by adversaries such that $T_1$ is a subset of targets available to the first attacker and $T_2$ is a subset of targets available to the second attacker, where $T_2 = T - T_1$. The defender has $m$ resources to cover the targets. Depending on whether a target is covered by the defender, two different cases might happen at each target. For the game with two adversaries, there two targets that are attacked by the adversaries, so four different situations might happen in total. Table 1 summarizes the players' pay-off in all possible cases when the attackers are attacking individually. $U^u_\Theta(t_1)$ and $U^u_\Theta(t_2)$ indicates the defender's pay-off at uncovered targets $t_1$ and $t_2$ attacked by attacker one, $\Psi_1$, and attacker two $\Psi_2$, respectively. Similarly, $U^c_\Theta(t_1)$ and $U^c_\Theta(t_2)$ indicates the defender's pay-off for the case of covered targets. $U^u_{\Psi_1}(t_1)$ and $U^u_{\Psi_2}(t_2)$ indicates the pay-off of attackers, $\Psi_1$ and $\Psi_2$, at uncovered targets $t_1$ and $t_2$, respectively. Likewise, $U^c_{\Psi_1}(t_1)$ and $U^c_{\Psi_2}(t_2)$ indicates the attackers' pay-off for the case of covered targets. If attackers

**Table 1: Pay-offs for individual attacks**

| Attackers: $\Psi_1, \Psi_2$ | Defender: $\Theta$ |
|---|---|
| $U^u_{\Psi_1}(t_1)$, $U^u_{\Psi_2}(t_2)$ | $U^u_\Theta(t_1)$+$U^u_\Theta(t_2)$ |
| $U^u_{\Psi_1}(t_1)$, $U^c_{\Psi_2}(t_2)$ | $U^u_\Theta(t_1)$+$U^c_\Theta(t_2)$ |
| $U^c_{\Psi_1}(t_1)$, $U^u_{\Psi_2}(t_2)$ | $U^c_\Theta(t_1)$+ $U^u_\Theta(t_2)$ |
| $U^c_{\Psi_1}(t_1)$, $U^c_{\Psi_2}(t_2)$ | $U^c_\Theta(t_1)$+$U^c_\Theta(t_2)$ |

collude with each other they will share all of their achievements fifty-fifty. Additionally, they will achieve a bonus reward, $\epsilon$, per any uncovered attack by them. As we assumed a zero-sum game, this bonus value will be deducted from the defender's pay-off. Table 2 summarize the adversaries and defender pay-offs in all possible situations when attackers are colluding. In more details, in both Tables 1 and 2, the first row indicates the pay-offs for the successful attacks by both adversaries. The second and third rows show the pay-offs for the situations that only one of the attackers succeeds and the last rows indicates the case of failure for both attackers.

**Table 2: Pay-offs for collusive attacks**

| Each attacker: $\Psi_1$ or $\Psi_2$ | Defender: $\Theta$ |
|---|---|
| $(U^u_{\Psi_1}(t_1) + U^u_{\Psi_2}(t_2) + 2\epsilon)/2$ | $U^u_\Theta(t_1)$+$U^u_\Theta(t_2) - 2\epsilon$ |
| $(U^u_{\Psi_1}(t_1) + U^c_{\Psi_2}(t_2) + \epsilon)/2$ | $U^u_\Theta(t_1)$+$U^c_\Theta(t_2) - \epsilon$ |
| $(U^c_{\Psi_1}(t_1) + U^u_{\Psi_2}(t_2) + \epsilon)/2$ | $U^c_\Theta(t_1)$+ $U^u_\Theta(t_2) - \epsilon$ |
| $(U^c_{\Psi_1}(t_1) + U^c_{\Psi_2}(t_2))/2$ | $U^c_\Theta(t_1)$+$U^c_\Theta(t_2)$ |

The coverage vector, $C$, gives the probability that each target is covered, $c_t$, and the attack vector $A$ gives the probability of attacking a target, which we restrict to attack a single target with the probability 1 (With this assumption SSE solution still exists [8].). For a given coverage and attack vector, expected utility of the defender is shown in Equation 1 and for a given coverage vector, the expected utility of the defender, at each target, is shown in Equation 2. By replacing $\Theta$ with $\Psi$, the same notation applies for expected utility of the attacker. The attack set, $\Gamma(C)$, is also defined in Equation 3 which contains all targets with the maximum expected utility for the attackers given coverage vector $C$.

$$U_\Theta(C, A) = \sum_{t \in T} a_t.(c_t.U^c_\Theta + (1 - c_t)U^u_\Theta) \quad (1)$$

$$U_\Theta(t, C) = c_t.U^c_\Theta + (1 - c_t)U^u_\Theta \quad (2)$$

$$\Gamma(C) = \{t : U_\Psi(t, C) \geq U_\Psi(t', C) \quad \forall t' \in T\} \quad (3)$$

## 2.2 ERASER based solution for generating the optimal defender strategy

The ERASER algorithm, proposed in [8], takes a security game as input and solves for an optimal defender coverage vector corresponding to a SSE strategy through a mixed integer linear program (MILP). The original formulation was developed for SSGs including one defender and one adversary. Using the same idea, we developed a new form of MILP which solves for an optimal defender coverage vector in presence of collusion between two adversaries, presented in Equation 4-20. In all of the equations, $nc$ stands for not colluding cases and $c$ stands for colluding cases. Equation 5 defines the integer variables $a^{nc}_t, a^c_t$ are, respectively, attackers' actions for when they do not collude and when they collude. $\alpha_1$ and $\alpha_2$ are decision variables that indicate each adversary's decision for collusion and $\beta$ is the decision made in the game based on $\alpha_1$ and $\alpha_2$. Meaning that, if both adversaries are inclined to collude, then $\beta$ will be equal to 1. Equation19 and 20 enforce this constraint. Equations 6 and 7 along with 5 forces that attack vector to assign a single target probability 1. Equation 8 forces that coverage vector to probabilities in range $[0, 1]$ and Equation 9 restricts the coverage by the number of the defender resources. Equations 10 and 11 indicate the defender expected utilities in colluding and not colluding cases. In equations 12 to 18, $Z$

is a large constant relative to the maximum pay-off value. Equation 12 and 13 define the defender's expected pay-off, contingent on the target attacked when attackers are not colluding and colluding, respectively. Equation 14 and 15 defines the expected utility of the attackers in colluding and non-colluding situations.

$$max \quad d \tag{4}$$

$$a_t^{nc}, a_t^c, \alpha_1, \alpha_2, \beta \in \{0,1\} \qquad \forall t \in T_1 \cup T_2 \tag{5}$$

$$\sum_{t_i \in T_i} a_{t_i}^{nc} = 1 \qquad i = 1,2 \tag{6}$$

$$\sum_{t_i \in T_i} a_{t_i}^{c} = 1 \qquad i = 1,2 \tag{7}$$

$$c_t \in [0,1] \qquad \forall t \in T_1 \cup T_2 \tag{8}$$

$$\sum_{t \in T} c_t \leq m \tag{9}$$

$$\forall t_i \in T_i, \quad i = 1,2:$$

$$U_\Theta^{nc}(t_1, t_2, C) = U_\Theta(t_1, C) + U_\Theta(t_2, C) \tag{10}$$

$$U_\Theta^{c}(t_1, t_2, C) = U_\Theta(t_1, C) + U_\Theta(t_2, C) - (1 - c_{t_1})\epsilon - (1 - c_{t_2})\epsilon \tag{11}$$

$$d - U_\Theta^{nc}(t_1, t_2, C) \leq (1 - a_{t_1}^{nc})Z + (1 - a_{t_2}^{nc})Z + \beta Z \tag{12}$$

$$d - U_\Theta^{c}(t_1, t_2, C) \leq (1 - a_{t_1}^{c})Z + (1 - a_{t_2}^{c})Z + (1 - \beta)Z \tag{13}$$

$$U_{\Psi_i}^{nc}(t_i, C) = U_{\Psi_i}(t_i, C) \tag{14}$$

$$U_{\Psi_i}^{c}(t_i, C) = U_{\Psi_i}(t_i, C) + (1 - c_{t_i})\epsilon \tag{15}$$

$$0 \leq k_i^{nc} - U_{\Psi_i}^{nc}(t_i, C) \leq (1 - a_{t_i}^{nc})Z \tag{16}$$

$$0 \leq k_i^{c} - U_{\Psi_i}^{c}(t_i, C) \leq (1 - a_{t_i}^{c})Z \tag{17}$$

$$i = 1,2:$$

$$-\alpha_i Z \leq k_i^{nc} - \frac{1}{2}(k_1^c + k_2^c) \leq (1 - \alpha_i)Z \tag{18}$$

$$\beta \leq \alpha_i \tag{19}$$

$$(\alpha_1 + \alpha_2) \leq \beta + 1 \tag{20}$$

Equation 16 and 17 constrain the attackers to select a strategy in attack set of $C$ in each situation. So the last four constraints are mutual best responses of defender and attacker in either colluding or non-colluding situations. Equation 18 forces each attacker to make his decision based on comparing the shared pay-off in collusion and his individual pay-off for non-colluding situation.

To formalize the solution concept further, the leader choose a strategy first, then given this strategy the followers play a Nash equilibrium. Ties between equlibria are broken as:

1. Equilibria with $\beta = 1$ are chosen over equlibria in which $\beta = 0$ if both followers obtain strictly higher utility in the $\beta = 1$ equilibrium.
2. In all other cases, the followers break ties in favor of the leader.

Given this, the leader's strategy is chosen to maximize his utility.

**THEOREM 1.** *Any solution to the above MILP is an equilibrium of the game.*

PROOF. We start by showing that the followers play a Nash equilibrium. Let $(a_{t_i}^*, \alpha_i^*)$ be the action of one of the followers produced by the MILP where $t_i$ is the target to attack and $\alpha_i$ is the decision of whether to collude. Let $(a_{t_i}, \alpha_i)$ be an alternative action. We need to show that the follower cannot obtain strictly higher utility by switching from $(a_{t_i}^*, \alpha_i^*)$ to $(a_{t_i}, \alpha_i)$.

If $\alpha_{t_i}^* = \alpha_{t_i}$, then Equations 16 and 17 imply that $a_{t_i}$ already maximizes the follower's utility.

If, $\alpha_{t_i}^* \neq \alpha_{t_i}$ then Equations 18 implies that $(a_{t_i}^*, \alpha_i^*)$ yields at least as much utility as $(a_{t_i}, 1 - \alpha_i^*)$, for the $a_{t_i}$ which maximizes the follower's utility given that they make the opposite decision about collusion. So, $(a_{t_i}^*, \alpha_i^*)$ yields at least as much utility as $(a_{t_i}, \alpha_i)$ as well.

Lastly, we need to verify that the two tie-breaking rules are respected. For the first, note that in Equation 18, both followers compute the utility for collusion assuming that the other will also collude. So, if follower $i$ would be best off in an equilibria with $\beta = 1$, the MILP requires that $\alpha_i = 1$. This implies that if both followers receive strictly highest utility in an equilibrium with $\beta = 1$, both will set $\alpha = 1$ as required. In all other cases, the objective is simply maximizing $d$, so ties will be broken in favor of the defender.

The following observations and propositions hold for the games with symmetric reward distribution between the two adversaries.

**OBSERVATION 1.** *The defender's main strategy is to break the collusion between them by enforcing an imbalance in resource allocation on both sides.*

In other words, the optimal solution satisfy $\theta \neq 0$ where $\theta = |x_1 - x_2|$, $x_i = \sum_{t_i \in T_i} c_{t_i}$ is the resource fraction on side of the attacker $i$ such that $x_1 + x_2 = m$ for the case of two adversaries in the game. This approach put one of the attackers in a better situation so he refuses to collude.

To analyze the effect of the imbalance in resource allocation on defender expected pay-off, we added another constraint to the MILP formulation shown in Equation 21. With this constraint, we will be able to keep the resource imbalance at an arbitrary level, $\delta$. For the case of symmetric reward distribution, WLOG, we can fix the first attacker to be the one who receives higher payoff and simply linearize the following equation; however generally, we can divide the equation into two separate linear constraints.

$$|k_1^{nc} - k_2^{nc}| = \delta \tag{21}$$

**OBSERVATION 2.** *By varying the $\delta$, one of the following cases can happen:*

1. *For $\delta < \delta^*$, $k_i^{nc} - \frac{1}{2}(k_1^c + k_2^c) < 0$ for both attackers and consequently $\alpha_i = 1$ for $i = 1,2$. In other words,*

*the defender is not able to break the collusion between the attackers and* $\beta = 1$.

2. *For* $\delta = \delta^*$, $k_1^{nc} - \frac{1}{2}(k_1^c + k_2^c) = 0$ *for one of the attackers and* $k_2^{nc} - \frac{1}{2}(k_1^c + k_2^c) < 0$ *for the other one, so consequently* $\alpha_1$ *can be either* 0 *or* 1 *and* $\alpha_2 = 1$. *In this case, the followers break ties in favor of the leader, so* $\alpha_1 = 0$ *and* $\beta = 0$.

3. *For* $\delta > \delta^*$, $k_1^{nc} - \frac{1}{2}(k_1^c + k_2^c) > 0$ *for one of the attackers and consequently* $\alpha_1 = 0$. *For the other attacker* $k_2^{nc} - \frac{1}{2}(k_1^c + k_2^c) < 0$ *and* $\alpha_2 = 1$. *In other words, the defender is able to break the collusion between the attackers and* $\beta = 0$.

**PROPOSITION 1.** *The switch-over point,* $\delta^*$, *introduced in the observation* 2 *is lower bounded by* 0 *and upper bounded by* $2\epsilon$.

PROOF. Using Equation 16, we know that at any target $t_i$, $k_i^{nc} \geq U_{\Psi_i}^{nc}(t_i, C)$. If we assume that the attacker attacks target $t_i^c$ with coverage $c_{t_i}^c$ by adding and subtracting a term as $\epsilon(1 - c_{t_i}^c)$, we can conclude that $k_i^{nc} \geq k_i^c - \epsilon(1 - c_{t_i}^c)$. Consequently, $k_1^c + k_2^c \leq k_1^{nc} + k_2^{nc} + \epsilon(1 - c_{t_1}^c) + \epsilon(1 - c_{t_2}^c)$. On the other hand, according to observation 2.2, at $\delta = \delta^*$, we have $k_1^{nc} - \frac{1}{2}(k_1^c + k_2^c) = 0$. Combining these last two equations, we will get $(k_1^{nc} - k_2^{nc}) \leq \epsilon(1 - c_{t_1}^c) + \epsilon(1 - c_{t_2}^c)$ . The LHS is equal to $\delta^*$ and the RHS can be rearranged as $2\epsilon - \epsilon(c_{t_1}^c + c_{t_2}^c)$, so we will have $\delta^* \leq 2\epsilon - \epsilon(c_{t_1}^c + c_{t_2}^c)$. Given the fact that coverage at each target is in range $[0, 1]$, the upper bound for $-(c_{t_1}^c + c_{t_2}^c)$ will be zero. Finally, by aggregating these results, we can conclude that $\delta^* \leq 2\epsilon$. Following the same analysis, the lower bound for $\delta^*$ can be found starting from $k_1^c + k_2^c \geq k_1^{nc} + k_2^{nc} + \epsilon(1 - c_{t_1}^{nc}) + \epsilon(1 - c_{t_2}^{nc})$ and as a result, $0 \leq \delta^*$.

Given the facts presented in Proposition 1, by enforcing an imbalance of maximum $2\epsilon$, the defender will be able to break the collusion. These bounds can be tighter, if we have more information about the distribution of reward at targets. For instance, if reward distribution over targets is close enough to uniform distribution, then the average coverage on each side will be $\bar{c}_{t_1} = \frac{2x_1}{n}$ and $\bar{c}_{t_2} = \frac{2x_2}{n}$, where $x_1$ and $x_2$ are fraction of resources assigned to each side and there are $\frac{n}{2}$ targets on each side. As a result, $-(c_{t_1}^c + c_{t_2}^c) \simeq -(\bar{c}_{t_1} + \bar{c}_{t_2})$. So we will be able to find an approximate upper bound of $2\epsilon(1 - \frac{m}{n})$, where $m = x_1 + x_2$. These results also implies that the larger the ratio of $\frac{m}{n}$, the less imbalance in resource allocation needed to break the collusion. In human subject experiments that will be discussed in the next section, we also observed that the wider the range of rewards over targets, the harder we can break the collusion among attackers.

# 3. HUMAN SUBJECT EXPERIMENTS

The linear program model developed in previous section assumes the rational behavior for the attackers. However, we know that human adversaries are bounded rational and taking that behavior into account will improve the attack prediction accuracy and optimal defender strategy. To that end, we simulated the game in wildlife domain and asked real human subjects to play this game. Then we analyzed the human subject decisions to derive a more accurate model to describe the human adversary behavior in security games in presence of collusion.

## 3.1 Game Interface Design

In our game, human subjects are asked to play the role of a poacher in a national park in Africa. There are different number of hippopotamus distributed over the park which indicates animal density distribution over the area. The entire park area is divided into two sections (right and left) and each human subject can only attack in one section (either right or left); however, they can explore the whole park. The other section of the park is only available to another player who is playing the same game. Each section of the park is divided into $3 \times 3$ grid, i.e. each player has 9 cells (sub-regions) accessible to him to attack. Players are able to choose different sub-regions and all of the information about success and failure likelihood, reward for the attacker (which is animal density in each sub-region) and penalty at each sub-region (either on left or side of park) will be shown to them. To avoid any bias on part of the human subjects, we assigned the sides to each player randomly and kept the other player anonymous but we used a dummy name as either Alice or Bob (chosen on a random basis) to indicate the other player's side and information. To help the human subjects to have a better view of the success/failure percentage (which is defender coverage) over all the sub-regions, we put a heat-map of that overlaid on Google Map view of the park. Also, to help the players to have a better understanding of the collusion in this game, we provided a table that summarizes all possible pay-offs for collusive attacks based on the collusion bonus considered for each game. The human subjects need to make decisions about: i)whether they are inclined to collude with the other player or not and ii)which region of the park to put their snare (trap) where there is less chance of getting caught and also a high chance of capturing a hippopotamus. So the human subjects may decide to attack "individually and independently" or attack "collusively" with the other player. In both situations, they will attack different sections separately but if both of them agree to attack collusively, they will share all of their pay-offs with each other, equally (fifty-fifty). To enhance understanding of the game, participants were asked to play one trial game to become familiar with the game interface and procedures. Then we provided a validation game to make sure that the players have read the instructions of the game and are fully aware of the rules and options of the game. For our analysis, we selected the valid players based on their performance in validation game and our validating criteria. Finally, the third game which is the main game is shown to the human subjects and their decisions are recorded and analyzed.

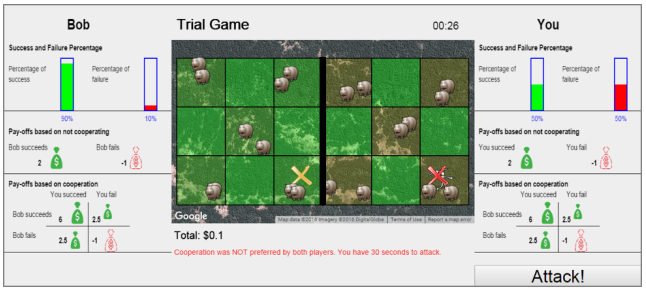



**Figure 1: Hunters vs Rangers game interface**

## 3.2 Game Pay-off Design

The "Hunters vs Poachers" game described in the previous sub-section is designed as a three-player zero-sum security game with 9 targets available to each attacker. There is one leader (defender) with $m$ resources to cover all the 18 targets (sub-regions in the park) and there are two followers (attackers) that can attack a side of the park. Reward of the adversaries at each cell for an uncovered attack is equal to the animal density at that cell and the penalty of the adversaries at each cell for a covered attack is equal to $-1$. We designed two different reward structures (animal density distributions), $RS1$ and $RS2$, shown in Figure2(a) and 2(b) and deployed on Amazon Mechanical Turk (AMT). In both of these symmetric structures, both players have identical reward distribution and we assumed a bonus of 1 for both setups.

| | | | | | |
|---|---|---|---|---|---|
| 0 | 0 | 5 | 5 | 0 | 0 |
| 0 | 6 | 3 | 3 | 6 | 0 |
| 0 | 0 | 4 | 4 | 0 | 0 |

(a) RS1

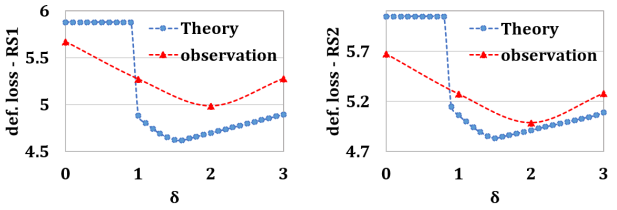
| | | | | | |
|---|---|---|---|---|---|
| 3 | 0 | 4 | 4 | 0 | 3 |
| 0 | 7 | 0 | 0 | 7 | 0 |
| 6 | 0 | 5 | 5 | 0 | 6 |

(b) RS2

**Figure 2: Reward structures deployed on AMT**

## 3.3 Experiment Results

For rational adversaries, based on the linear formulation developed in previous section, the defender can obtain the maximum expected utility by breaking the collusion between two adversaries. The main idea for breaking the collusion is to put one adversary in a better condition in terms of defender coverage and the other one in a worse condition, then collusion will not be preferred by one of adversaries and collusion breaks. The corresponding optimal strategy results in an imbalance between the maximum expected utilities on left and right side of the park. We refer to this difference as $\delta$ which indicates the level of asymmetry in allocating resources on both sides. The correlation between $\delta$ and aggregated coverage imbalance, $\theta$, is illustrated in Figure 5(b). Blue plots with circular markers in Figure3(a) and 3(b) show the changes in defender loss while $\delta$ varies for $RS1$ and $RS2$, respectively. A key point of this figure is that there is threshold $\delta$ in which we can break the collusion between rational adversaries which is equal to 0.9 for $RS1$ and 0.8 for $RS2$. Another important point is that as we increase the difference between the fraction of resources allocated on both sides, the defender loss will decrease and at $\delta$ equal to 1.5 the optimum point will be reached. To see how deviating from balanced resource allocation can affect human adversaries' decisions about collusion, we ran human subjects experiments on AMT for various $\delta$ values. Figure4(a) and 4(b) illustrate two sample cases that we have deployed on AMT for $RS2$ such that in the first case, resources are distributed symmetrically but in the second case $\delta$ was set equal to 1 and one side is covered more in comparison with the other one. For each reward structure, we tested 4 different coverage distribution such that $\delta \in \{0, 1, 2, 3\}$. The experiments showed that the level of collusion (percentage of population who decided to collude) decreased by increasing $\delta$ for both $RS1$ and $RS2$ as shown in Figure 5(a) for advantaged attacker who are in a better situation, RS1-A and RS2-A. But for the attackers that are in the disadvantaged situation, RS1-DA and RS2-DA, for both reward structures, we can see a high level of collusion at all levels of $\delta$. Average defender loss based on the observations are plotted in dashed red lines with rectangular markers in Figure3(a) and 3(b). Instead of a sharp switch-over point from colluding situation into non-colluding situation, we can see a smooth change in average defender loss along with a delayed optimum point in comparison with rational assumption situation. Based on the observations, not all of the targets are identical in terms of attractiveness to the attackers. To illustrate this fact, frequency of attack for both reward structures for the player in a better situation at different levels of $\delta$ are shown in Figure6(a) and 6(b) and the related human behavior models are discussed in the next section. These figures show that human subjects are showing more risk averse behavior in $RS1$ relative to $RS2$. In more details, in similar situations in terms of $\delta$, players in $RS2$ are not only more interested in collusion but also more interested in attacking cells with higher rewards and consequently higher coverage.




(a) Defender loss, RS1 (b) Defender loss, RS2

**Figure 3: Defender loss vs $\delta$**

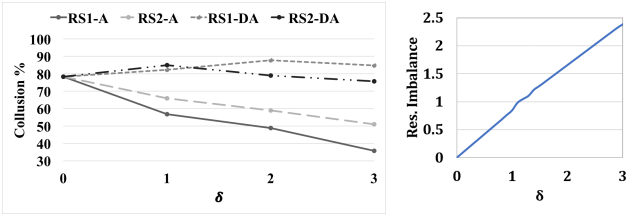
| | | | | | |
|---|---|---|---|---|---|
| 15 | 0 | 32 | 32 | 0 | 15 |
| 0 | 58 | 0 | 0 | 58 | 0 |
| 52 | 0 | 43 | 43 | 0 | 52 |

(a) $\delta = 0$, RS2

| | | | | | |
|---|---|---|---|---|---|
| 3 | 0 | 22 | 42 | 0 | 28 |
| 0 | 51 | 0 | 0 | 64 | 0 |
| 44 | 0 | 35 | 52 | 0 | 59 |

(b) $\delta = 1$, RS2

**Figure 4: Defender strategy deployed on AMT**



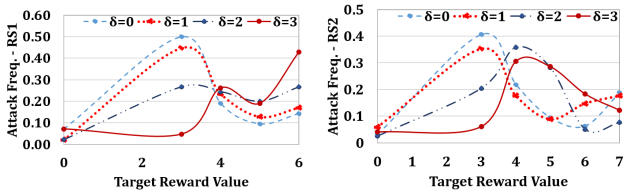

(a) Collusion level (b) Resource imbalance

**Figure 5: Collusion level and resource imbalance**

# 4. BOUNDED RATIONALITY

## 4.1 Human behavior models

**Subjective Utility Quantal Response (SUQR):** To incorporate the effect of bounded rational adversaries, we use the SUQR model, [10], to predict the probability of attack at each target $t_i$. This model is an extension to QR




(a) Attack Frequency, RS1 (b) Attack Frequency, RS2

**Figure 6: Attack frequency at targets**

model presented in [9]. The key idea behind QR model is that, there is higher probability for the adversary to attack a target with higher expected utility. In SUQR, a new utility function called Subjective Utility, is defined which is a linear combination of key features such as defender's coverage probability, adversary's reward and penalty at each target. These features are assumed to be the most important factors in adversary decision-making process.

In this paper, we assume there are two attackers in the security game, so we might see different behaviors from attackers. Since the main idea for breaking the collusion is to impose a resource imbalance between two adversaries, one adversary will be in the better position and the other one will be in the worse position. Assuming perfectly rational adversaries, we expect an inevitable inclination towards collusion from the disadvantaged attacker and an inevitable declination from the advantaged attacker. However, our observation from human subjects experiment did not support this expectation. So to model human behavior, we need to consider all of the possible cases: i) a disadvantaged attacker who is inclined to collude, DA-C , ii) a disadvantaged attacker who is not inclined to collude, D-NC, iii) an advantaged attacker who is inclined to collude, A-C, and iv) an advantaged attacker who is not inclined to collude, A-NC. Given this classification of adversaries, we define a revised version of expected utility in Equation 22 which can be adopted in security games involving collude. In this equation $i$ indicates the attacker that can attack $t_i \in T_i$ and $\beta$ indicate each adversarys' decision about collusion. The vector $\mathbf{w}_i^\beta = (\omega_{i,1}^\beta, \omega_{i,2}^\beta, \omega_{i,3}^\beta)$ contains information about each adversary type behavior and each component of $\mathbf{w}_i^\beta$ indicates the relative weights the adversary gives to each feature in the decision making process. $U_{\Psi_i}^c(t_i)$, $U_{\Psi_i}^u(t_i)$ and $\hat{c}_{t_i}$ shows the penalty, reward and modified coverage probability of the attackers, respectively. Modified coverage probability is a function of the actual coverage probability and will be discussed soon.

$$\hat{U}_{\Psi_i}(t_i, \hat{C}, \beta) = \omega_{i,1}^\beta . \hat{c}_{t_i} + \omega_{i,2}^\beta . U^u c_{\Psi_i}(t_i) + \omega_{i,3}^\beta . U_{\Psi_i}^c(t_i) \quad (22)$$

According to the SUQR model, the probability that the adversary will attack target $t_i$ for each group of adversaries that the defender might face, is given by:

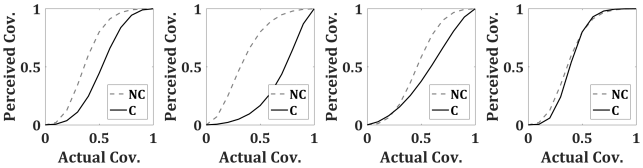
$$q_{t_i}(C \mid \beta) = \frac{e^{\hat{U}_{\Psi_i}(t_i, \hat{C}, \beta)}}{\sum_{t_i \in T_i} e^{\hat{U}_{\Psi_i}(t_i, \hat{C}, \beta)}} \quad (23)$$

**Probability weighting function:** Prospect Theory provides a descriptive model of how humans make decision among alternatives choices in presence of risk [6], [13]. According to this model, individuals overestimate low probability and underestimate high probability. Following this idea, there are literature in this domain that propose parametric models which capture the non-uniform weighting schemes including both inverse S-shaped as well as S-shaped probability curves, [1], [4]. With the notion of Prospect Theory, the modified coverage observed by the attackers is assumed to be related to the actual probability based on Equation 24, where $\gamma$ and $\eta$ determine the elevation and curvature of the function, respectively.

$$\hat{c}_{t_i} = \frac{\eta c_{t_i}^\gamma}{\eta c_{t_i}^\gamma + (1 - c_{t_i})^\gamma} \quad (24)$$

## 4.2 Results

Figures 7(a) and 7(b) show the probability weighting functions learned for the disadvantaged and advantaged adversaries for both groups who are colluding and not colluding. Figures 7(c) and 7(d) show the same results for reward structure 2.




(a) RS1: DA (b) RS1: A (c) RS2: DA (d) RS2: A

**Figure 7: Curves learned based on Prospect Theory**

The vector $\mathbf{w}_i^\beta = (\omega_{i,1}^\beta, \omega_{i,2}^\beta, \omega_{i,3}^\beta)$, $\eta_i^\beta$ and $\gamma_i^\beta$ are computed by performing Maximum Likelihood Estimation (MLE) on available attack data from human subject experiments for four classes of attackers. Table 3 and 4 show the results for both reward structures.

**Table 3: Params. learned from data for RS1**

| Class | $(i, \beta)$ | $\omega_{i,1}^\beta$ | $\omega_{i,2}^\beta$ | $\omega_{i,3}^\beta$ | $\eta_i^\beta$ | $\gamma_i^\beta$ |
|---|---|---|---|---|---|---|
| DA-NC | $(1, 0)$ | $-4.4$ | 0.8 | 0.3 | 4 | 2.4 |
| DA-C | $(1, 1)$ | $-22.8$ | 3.3 | 0.3 | 0.8 | 2.2 |
| A-NC | $(2, 0)$ | $-6.7$ | 1.5 | 0.3 | 4 | 1.8 |
| A-C | $(2, 1)$ | $-32$ | 0.8 | 0.3 | 0.2 | 1.6 |

**Table 4: Params. learned from data for RS2**

| Class | $(i, \beta)$ | $\omega_{i,1}^\beta$ | $\omega_{i,2}^\beta$ | $\omega_{i,3}^\beta$ | $\eta_i^\beta$ | $\gamma_i^\beta$ |
|---|---|---|---|---|---|---|
| DA-NC | $(1, 0)$ | $-44.5$ | 6 | 0.3 | 1.4 | 2.2 |
| DA-C | $(1, 1)$ | $-40.8$ | 4 | 0.3 | 0.6 | 1.4 |
| A-NC | $(2, 0)$ | $-14.5$ | 1.5 | 0.3 | 4 | 2.4 |
| A-C | $(2, 1)$ | $-7.6$ | 1 | 0.3 | 4 | 3 |

# 5. CONCLUSIONS

This paper provides two contributions: the first one is formulating a new type of Stackleberg security game involving a beneficial collusion mechanism among adversaries and

developing a MILP program that enables us to find the optimal defender strategy. The second contribution of this paper is to develop a parametric human behavior model which is able to capture the bounded rationality of adversaries in this type of collusive games. This model is proposed based on prospect theory, SUQR model and real data collected from conducting human subject experiments with participants on Amazon Mechanical Turk. The observation showed that the collusion between adversaries can be broken by imposing security resource imbalance among adversaries' targets. However, human adversaries are not perfectly rational and do not follow the exact patterns predicted by the MILP developed in this paper. To address this mismatch, the related human behavior models were proposed and discussed.

# REFERENCES


[1] M. Abdellaoui, O. l'Haridon, and H. Zank. Separating curvature and elevation: A parametric probability weighting function. *Journal of Risk and Uncertainty*, 41(1):39–65, 2010.

[2] F. Fang, T. H. Nguyen, R. Pickles, W. Y. Lam, G. R. Clements, B. An, A. Singh, M. Tambe, and A. Lemieux. Deploying paws: Field optimization of the protection assistant for wildlife security. 2016.

[3] F. Fang, P. Stone, and M. Tambe. When security games go green: Designing defender strategies to prevent poaching and illegal fishing. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.

[4] R. Gonzalez and G. Wu. On the shape of the probability weighting function. *Cognitive psychology*, 38(1):129–166, 1999.

[5] W. B. Haskell, D. Kar, F. Fang, M. Tambe, S. Cheung, and E. Denicola. Robust protection of fisheries with compass. In *AAAI*, pages 2978–2983, 2014.

[6] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291, 1979.

[7] D. Kar, F. Fang, F. Delle Fave, N. Sintov, and M. Tambe. 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*, pages 1381–1390. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[8] C. Kiekintveld, M. Jain, J. Tsai, J. Pita, F. Ordóñez, and M. Tambe. Computing optimal randomized resource allocations for massive security games. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 689–696. International Foundation for Autonomous Agents and Multiagent Systems, 2009.

[9] D. L. McFadden. Quantal choice analaysis: A survey. In *Annals of Economic and Social Measurement, Volume 5, number 4*, pages 363–390. NBER, 1976.

[10] T. H. Nguyen, R. Yang, A. Azaria, S. Kraus, and M. Tambe. Analyzing the effectiveness of adversary modeling in security games. In *AAAI*, 2013.

[11] P. Paruchuri, J. P. Pearce, J. Marecki, M. Tambe, F. Ordonez, and S. Kraus. Playing games for security: An efficient exact algorithm for solving bayesian stackelberg games. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, pages 895–902. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[12] M. Tambe. *Security and game theory: Algorithms, deployed systems, lessons learned*. Cambridge University Press, 2011.

[13] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323, 1992.

[14] L. S. Wyler and P. A. Sheikh. International illegal trade in wildlife: Threats and us policy. DTIC Document, 2008.

[15] R. Yang, B. Ford, M. Tambe, and A. Lemieux. Adaptive resource allocation for wildlife protection against illegal poachers. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 453–460. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[16] R. Yang, C. Kiekintveld, F. Ordonez, M. Tambe, and R. John. Improving resource allocation strategy against human adversaries in security games. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, volume 22, page 458, 2011.