Toward Addressing Collusion among Human Adversaries in Security Games

Shahrzad Gholami and **Bryan Wilder** and **Matthew Brown** and **Dana Thomas** and **Nicole Sintov** and **Milind Tambe**[1]

Abstract. Security agencies including the US Coast Guard, the Federal Air Marshal Service and the Los Angeles Airport police are several major domains that have been deploying Stackelberg security games and related algorithms to protect against a single adversary or multiple, independent adversaries strategically. However, there are a variety of real-world security domains where adversaries may benefit from colluding in their actions against the defender. Given the potential negative effect of these collusive actions, the defender has an incentive to break up collusion by playing off the self-interest of individual adversaries. This paper deals with problem of collusive security games for rational and bounded rational adversaries. The theoretical results verified with human subject experiments showed that behavior model which optimizes against bounded rational adversaries provides demonstrably better performing defender strategies against human subjects.

1 Introduction

Models and algorithms based on Stackelberg security games have been deployed by many security agencies including the US Coast Guard, the Federal Air Marshal Service, and Los Angeles International Airport [12] in order to protect against attacks by strategic adversaries in counter-terrorism settings. More recently, security games research has explored new domain such as wildlife protection, where planning of effective strategies is needed to tackle sustainability problems such as illegal poaching and illegal fishing [3]. Most of these previous works on security games assumes that different adversaries can be modeled independently [8]. However, there are many real-world security domains in which adversaries may collude in order to more effectively evade the defender. Three example domains are:

i) *Wildlife Protection Domain*: International trade of illicit wildlife products is growing rapidly and the most common types of illicitly traded wildlife products include elephant ivory, rhino horn, tiger parts, and caviar. Biodiversity loss, species extinction, invasive species introduction, and disease transmission resulting from illicit wildlife trade can all have disastrous impacts on the environment. Additionally, connections have been observed between illicit wildlife trade and organized crime as well as terrorist organizations, and thus activities such as poaching can serve to indirectly threaten national security [16]. Different forms of collusion have been observed among different groups of poachers. For example, these groups may coordinate to reduce the cost for storage, handling, and transportation of goods as well as to gain access to trade markets. This coordination can result in overall higher levels of poaching and damage to the environment [15].

ii) *Illegal Drug Trade:* Due to an ever growing demand for drugs, international organized crime syndicates have increased cooperation in order to facilitate drug trafficking, expand to distant markets, and evade local law enforcement [1]. In some cases, drug traders must cooperate with terrorist organizations to send drugs through particular areas. More broadly, expansion of global transportation networks and free trade has motivated cooperation between criminal organizations across different countries [11].

iii) *Finance - "rent-a-tribe" model:* Authorities in the US attempt to regulate payday lenders. These are lenders which offer extremely high interest rates to low-income borrowers, who cannot obtain loans from traditional banks. Recently, payday lenders have begun to operate in partnership with Native American tribes, which are exempt from state regulations. In this domain, the defender (a regulator) seeks a policy which prevents collusion between the adversaries (payday lenders and Native American tribes) [6].

Despite mounting evidence of the destructive influence of collusive behavior, strategies for preventing collusion have not been explored in the security games literature.

2 Background and Related Work

To better understand how humans make decisions regarding collusion, the following frameworks and theories are helpful in analysing the problem of collusive security game for rational and bounded rational adversaries.

Stackelberg Security Game model: The Stackelberg Security Game model, introduced almost a decade ago, has led to a large number of applications and has been discussed widely in the literature [12]. All of these works consider adversaries as independent entities and the goal is for a defender (leader) to protect a set of targets with a limited set of resources from a set of adversaries (followers)[2]. The defender commits to a strategy and the adversaries observe this strategy and each select a target to attack. The solution concept for security games involves computing a strong Stackelberg equilibrium which assumes that the adversaries maximize their own expected utility and break ties in favor of the defender. Security game models where an adversary is capable of attacking multiple targets simultaneously have been explored in [17]. To address cooperation between adversaries, [5] introduced a communication network based approach for adversaries to share their skills and form coalitions in

[1] University of Southern California, Los Angeles, CA 90089

[2] We use the convention in the security game literature where the defender is referred as "she" and an adversary is referred to as "he".

order to execute more attacks. However, no previous work on security games has conducted behavioral analysis or considered the bounded rationality of human adversaries in deciding whether to collude in the first place.

We now introduce key behavioral models and concepts that are useful for modeling and analyzing adversary behaviors in collusive security games.

Quantal Response and Subjective Utility Quantal Response: In real-world settings, human adversaries do not strictly maximize their expected utility, rather, they choose strategies stochastically [9]. Quantal Response (QR) model is a solution concept based on the assumption of bounded rationality. SUQR [10] has been proposed as an extension to QR and is the model used in this paper to predict the probability of attack at each target. In SUQR, subjective utility replaces expected utility and is defined as a linear combination of key domain features including the defender's coverage probability and the adversary's reward and penalty at each target. These features are assumed to be the most salient factors in the adversary's decision-making process.

Prospect Theory and Probability Weighting Functions: Another aspect of bounded rationality is the misperception by the adversary of key factors that influence decision making. Prospect Theory provides a descriptive model of how humans make decisions among alternative choices in the presence of probabilistic risk [7, 14]. According to this model, individuals overestimate low probability and underestimate high probability. Following this idea, literature in this domain proposes parametric models which capture different non-uniform weighting schemes including both inverse S-shaped as well as S-shaped probability curves. Based on these curves, the adversaries perceive a modified coverage probability and this fact can be exploited to the benefit of the defender. Human subject experiments have been conducted for security games to test both bounded rationality and probability perception [8], but such work never considered the type of collusive actions of concern in this paper.

Inequity Aversion Theory: Decisions regarding the interaction between humans in strategic settings can be influenced by the relative advantage of participants. According to Inequity Aversion theory humans are sensitive to inequity of outcome regardless of whether they are in the advantaged or disadvantaged situation and they make decisions in a way that minimizes inequity [4]. Inequity aversion has been widely studied in economics and psychology and is consistent with observations of human behavior in standard economic experiments such as the dictator game and ultimatum game in which the most common choice of people is to split the reward 50-50 [2]. Along these lines and contrary to the theoretical predictions, the IA theory also supports our experiments and analyses in security game domain.

Individualism Collectivism Analysis: Similarly, the personal attitudes and attributes of participants can also influence their interactions in strategic settings. A key characteristic is the well-established individualism-collectivism paradigm, which describes cultural differences in the likelihood of people to prioritize themselves versus their in-group. Specifically, those who identify as part of collectivistic cultures, compared to people in individualistic cultures, tend to identify as part of their in-groups, prioritize group-level goals, define most relationships with in-group members as communal, and are more self-effacing. Individualism-collectivism can be reliably measured using psychometrically-validated survey instruments [13].

3 Conclusion

We introduced a type of security games involving potential collusion among adversaries. Also we discussed the underlying frameworks and theories that are helpful in understanding how human makes decision regarding collusion in security games. Theoretical results verified by real human subject experiments showed that human adversaries are far from rational when deciding whether or not to collude and human behavioral model that incorporates bounded rationality of adversaries outperforms models assuming rational human adversaries.

REFERENCES

[1] Horace A Bartilow and Kihong Eom, 'Free traders and drug smugglers: The effects of trade openness on states' ability to combat drug trafficking', *Latin American Politics and Society*, **51**(2), 117–145, (2009).

[2] Nathan Berg, 'Behavioral economics', *21st century economics: A reference handbook*, **2**, 861–872, (2010).

[3] Fei Fang, Thanh H Nguyen, Rob Pickles, Wai Y Lam, Gopalasamy R Clements, Bo An, Amandeep Singh, Milind Tambe, and Andrew Lemieux, 'Deploying paws: Field optimization of the protection assistant for wildlife security', (2016).

[4] Ernst Fehr and Klaus M Schmidt, 'A theory of fairness, competition, and cooperation', *Quarterly journal of Economics*, 817–868, (1999).

[5] Qingyu Guo, Bo An, Yevgeniy Vorobeychik, Long Tran-Thanh, Jiarui Gan, and Chunyan Miao, 'Coalitional security games', (2016).

[6] Creola Johnson, 'America's first consumer financial watchdog is on a leash: Can the cfpb use its authority to declare payday-loan practices unfair, abusive, and deceptive', *Cath. UL Rev.*, **61**, 381, (2011).

[7] Daniel Kahneman and Amos Tversky, 'Prospect theory: An analysis of decision under risk', *Econometrica: Journal of the Econometric Society*, 263–291, (1979).

[8] Debarun Kar, Fei Fang, Francesco Delle Fave, Nicole Sintov, and Milind Tambe, 'a game of thrones: When human behavior models compete in repeated stackelberg security games', in *International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, (2015).

[9] Daniel L McFadden, 'Quantal choice analaysis: A survey', in *Annals of Economic and Social Measurement, Volume 5, number 4*, 363–390, NBER, (1976).

[10] Thanh Hong Nguyen, Rong Yang, Amos Azaria, Sarit Kraus, and Milind Tambe, 'Analyzing the effectiveness of adversary modeling in security games.', in *AAAI*, (2013).

[11] Andres Lopez Restrepo and Álvaro Camacho Guizado, 'From smugglers to warlords: twentieth century colombian drug traffickers', *Canadian Journal of Latin American and Caribbean Studies*, **28**(55-56), 249–275, (2003).

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

[13] Harry C Triandis and Michele J Gelfand, 'Converging measurement of horizontal and vertical individualism and collectivism.', *Journal of personality and social psychology*, **74**(1), 118, (1998).

[14] Amos Tversky and Daniel Kahneman, 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and uncertainty*, **5**(4), 297–323, (1992).

[15] Greg L Warchol, Linda L Zupan, and Willie Clack, 'Transnational criminality: An analysis of the illegal wildlife market in southern africa', *International Criminal Justice Review*, **13**(1), 1–27, (2003).

[16] Liana S Wyler and Pervaze A Sheikh, 'International illegal trade in wildlife: Threats and us policy'. DTIC Document, (2008).

[17] Zhengyu Yin, Dmytro Korzhyk, Christopher Kiekintveld, Vincent Conitzer, , and Milind Tambe, 'Stackelberg vs. nash in security games: Interchangeability, equivalence, and uniqueness', in *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, (2010).