On Repeated Stackelberg Security Game with the Cooperative Human Behavior Model for Wildlife Protection

(Extended Abstract)

Binru Wang
State Key Laboratory for Novel Software Technology
Nanjing University
Nanjing, China 210023
wbinru@gmail.com

Yuan Zhang
State Key Laboratory for Novel Software Technology
Nanjing University
Nanjing, China 210023
zhangyuan05@gmail.com

Sheng Zhong
State Key Laboratory for Novel Software Technology
Nanjing University
Nanjing, China 210023
zhongsheng@nju.edu.cn

ABSTRACT

Inspired by successful deployments of Stackelberg Security Game in real life, researchers are working hard to optimize the game models to make them more practical. Recent security game work on wildlife protection makes a step forward by taking the possible cooperation among attackers into consideration. However, it models attackers to have complete rationality, which is not always possible in practice given they are human beings. We aim to tackle attackers' bounded rationality in the complicated, cooperation-enabled and multi-round security game for wildlife protection. Specifically, we construct a repeated Stackelberg game, and propose a novel adaptive human behavior model for attackers based on it. Despite generating defender's optimal strategy requires to solve a non-linear and non-convex optimization problem, we are able to propose an efficient algorithm that approximately solve this problem. We perform extensive real-life experiments, and results show our solution effectively helps the defender to deal with attackers who might cooperate.

Keywords

Game Theory, Cooperation Mechanism, Repeated Stackelberg Games, Human behavior, Wildlife Protection

1. INTRODUCTION

Stackelberg security games (SSGs) are receiving an increasing amount of attention after their successful deployments in various real-life occasions. Inspired by recent outstanding game theoretical efforts on wildlife protection [4, 3], we revisit the security game design problem, taking into account the poachers' possible collaboration mechanisms and their bounded rationality in this paper.

Nowadays, many endangered species, such as tiger and rhino, are facing danger of extinction from poaching and illegal animal parts trading. The current, major method for security agencies (e.g. park rangers) to prevent poaching is patrolling. However, catching a poacher in this way is never an easy task. First, rangers only have limited security resources in general, so they cannot cover every corner of the large protected area at the same time. Moreover, the poachers can observe rangers' patrolling routine, and design their own routine correspondingly to avoid being caught. To address these difficulties and generate the optimal strategy, a few excellent game theoretical models [6, 2, 1] have been proposed. In contrast to the work above which studies individual attacks, a recent work [3] demonstrates the destructive synergistic effect of collaboration among poachers, and presents a novel cooperative game the solution of which could effectively help the defender reduce its loss due to attackers' cooperation.

Unlike the attackers in the counter-terrorism domain, the attackers in the wildlife protection domain are not generally perfectly rational when making their decisions [6]. Hence, we extend Subjective Utility Quantal Response (SUQR) [5] to model attackers' behavior in a cooperative game in this paper.

To reveal attackers' adaptive nature in decision-making and further improve the performance of our human behavior model, especially in the domain of wildlife crime, we model the strategic interactions between the defender and the attackers as a repeated SSG. In a traditional one-shot SSG, the defender (leader) deploy the patrolling strategy first and the attacker (follower) picks a target to attack accordingly after observing the strategy. While in a repeated SSG, where repeated interactions between players are involved, the defender is able to change the strategy periodically in different rounds of a game based on the attacker's previous behavior.

In this paper, we address attackers' adaptivity in decision-making by defining the inclination of an attacker, and propose an adaptive human behavior model to optimize the usage of limited resources in repeated SSG where two attckers who might cooperate to cause more damage to the defender. Furthermore, we design an efficient algorithm which approximately solve underlying non-linear, non-convex optimization problem derived from the game model. We also conduct extensive real-life experiments to test the effectiveness of our model, and results show that our model has excellent performance.

2. ADAPTIVE HUMAN BEHAVIOR MODEL IN COOPERATIVE GAME

In this section we study the cooperation mechanism in a

Appears in: *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

zero-sum repeated SSG involving one defender and two attackers. Based on this cooperative game, we present a new adaptive human behavior model and form the corresponding optimization problem to solve for defender strategy. The game consists of several rounds so that the defender may change his/her patrolling strategy based on the data collected in the previous round. An attacker will get his payoff based on the success or failure of his own attack if he chooses to attack independently. On the contrary, if both attackers choose to cooperate, they will evenly share all of the payoffs they earn. There is also an extra bonus reward, ϵ, for motivating attacker cooperation. Briefly speaking, the attackers have two major decisions to make in a game: 1) the targets they choose to attack and 2) whether to cooperate with the other attacker.

To address bounded rationality and adaptivity, we extend SUQR [5] and propose a new model. In our game, both attackers can choose to cooperate or not. In each situation, they are modeled to have different attacking probability at target i. Without loss of generality, the probability that a non-cooperative attacker attacks target i is given by:

$$q_i^{nc}(x) = \frac{e^{ASU_1^{nc}(x_i)}}{\sum_{t_i \in T_i} e^{ASU_1^{nc}(x_{t_i})}}$$

where $ASU_1^{nc}(x_i)$ denotes ***adaptive subjective utility***, which is a utility function we define. $ASU_1^{nc}(x_i)$ is determined by the defender's strategy, the attacker's reward and penalty, and ***inclination***, which we introduce to evaluate attackers' adaptivity in choosing a target based on the historical attacking data.

Given the model and the parameters learned from available data, we may form the non-linear optimization problem with non-linear constraints.

3. GENERATING OPTIMAL STRATEGIES

In order to obtain the optimal defender strategy, we need to solve a non-linear optimization problem. First we simplify the problem by decomposing it into four subproblems. After that, we make some variable substitutions and relax the constraints by introducing piecewise linear functions. Then we transform the subproblem into a mixed integer quadratically constrained quadratic program (MIQCQP) problem and the global optimal solution can be numerically computed using any MIQCQP solver.

4. EXPERIMENTS

We conduct multiple experiments and analyse the results derived from them. In our experiments, players are asked to play the role of poachers looking to hunt the wild animals in a protected area. The interface of the game is shown in Figure 1. The protected area is divided into two regions, each consists of 9 small cells representing 9 targets. Players are allowed to observe all 18 targets in the area to get information of them before they attack, but only choose an attacking target form one region. In addition to choosing a target, players also have to decide whether they agree to cooperate with the other player.

There are five rounds of the game in total. In the first round, the defender strategy is generated using the approach presented in [3]. This strategy, which we refer to as "Maximin" strategy in this paper, is used because of unavailability to data in the first round. To evaluate the effectiveness of our model, we deploy "Maximin" strategy in all the five rounds and the defender strategies generated based on our model respectively. There are two sets of games with payoff structure $S1$ and $S2$ respectively in our experiments, and for each five-round game, we recruit a new group of players to eliminate the impact of bias.

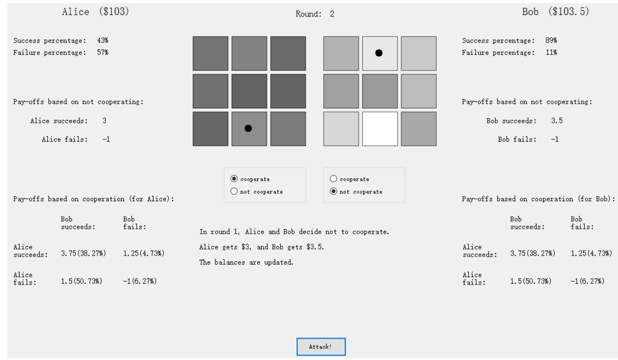
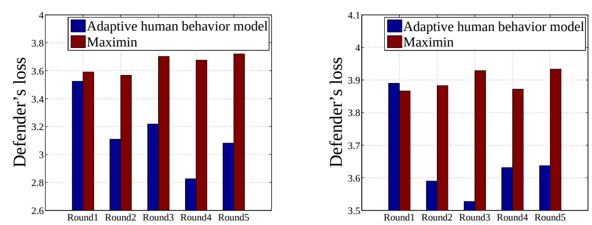


Figure 1: Game interface

To illustrate the effectiveness of our model in terms of reducing defender's average loss, the comparison diagrams are given in Figure 2.



(a) Results on $S1$ (b) Results on $S2$

Figure 2: Defender loss

Except for the first round, where two models use the same strategy, our model provides significantly lower defender losses than the "Maximin" model does on both structures.

5. CONCLUSIONS

In this paper we address the security issues in wildlife protection. First, we propose a new model in a repeated SSG, which takes into account the bounded rationality, the historical behaviors and cooperation of human attackers. Moreover, we approximately solve the nonlinear and nonconvex optimization problem generating the defender strategy against the model we propose. Finally, through extensive experiments we show that our model outperforms the existing model assuming perfectly rational attackers in terms of reducing defender loss.

Acknowledgments

Yuan Zhang was supported in part by NSFC-61402223. This work was supported in part by NSFC-61425024, NSFC-61300235, the Jiangsu Province Double Innovation Talent Program, and NSFC-61321491.

REFERENCES

[1] 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. In *Proceedings of the Twenty-Eighth Innovative Applications of Artificial Intelligence Conference*, pages 3966–3973, 2016.

[2] F. Fang, P. Stone, and M. Tambe. When security games go green: Designing defender strategies to prevent poaching and illegal fishing. In *Proceedings of the 24th International Conference on Artificial Intelligence*, pages 2589–2595, 2015.

[3] S. Gholami, B. Wilder, M. Brown, A. Sinha, N. Sintov, and M. Tambe. A game theoretic approach on addressing cooperation among human adversaries. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, 2016.

[4] 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.

[5] T. H. Nguyen, R. Yang, A. Azaria, S. Kraus, and M. Tambe. Analyzing the effectiveness of adversary modeling in security games. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, pages 718–724, 2013.

[6] 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, 2014.