Strategic Play By Resource-Bounded Agents in Security Games

Extended Abstract

Xinming Liu
Massachusetts Institute of Technology
Cambridge, United States of America
xinming@mit.edu

Joseph Y. Halpern
Cornell University
Ithaca, United States of America
halpern@cs.cornell.edu

ABSTRACT

Many studies have shown that humans are “predictably irrational”: they do not act in a fully rational way, but their deviations from rational behavior are quite systematic. Our goal is to see the extent to which we can explain and justify these deviations as the outcome of rational but resource-bounded agents doing as well as they can, given their limitations. We focus on the well-studied ranger-poacher game, where rangers are trying to protect a number of sites from poaching. We capture the computational limitations by modeling the poacher and the ranger as probabilistic finite automata (PFAs). We show that, with sufficiently large memory, PFAs learn to play the Nash equilibrium (NE) strategies of the game and achieve the NE utility. However, if we restrict the memory, we get more “human-like” behaviors, such as *probability matching*, and avoiding sites where there was a bad outcome, that we also observed in experiments conducted on Amazon Mechanical Turk. Interestingly, we find that adding human-like behaviors such as probability matching and overweighting significant events actually improves performance, showing that this seemingly irrational behavior can be quite rational.

KEYWORDS

Bounded rationality; Security games; Probabilistic Finite Automata

ACM Reference Format:

Xinming Liu and Joseph Y. Halpern. 2023. Strategic Play By Resource-Bounded Agents in Security Games: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

While standard economic theory assumes that people are rational, many studies (see, e.g., [1]) have shown that humans are irrational in a systematic way. Our goal in this paper is to see the extent to which computational limitations can explain and justify human behaviors in *security games* [9]. Specifically, we consider a (finitely) repeated two-player ranger-poacher game (based on [5]). At each stage of the repeated game, the poacher tries to catch a rhino at one of n sites, and the ranger tries to prevent the poacher from doing so. We assume that there is a commonly-known probability of a rhino being at any particular site, which does not change over time. We can formulate the stage game (i.e., the game played at each step of the repeated game) as a zero-sum normal-form game, which we show has a unique Nash equilibrium (NE).

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

We show through simulations that if the PFA has sufficiently many states, players eventually converge to playing the NE strategies of the stage game and achieve the NE utilities. As we limit the number of states of the PFA, the PFA acts more human-like; the strategy it uses is somewhere between the NE strategy and *probability matching* (i.e., visiting sites in proportion to the probability of a rhino being there). In addition, it seems reasonable to think that a poacher views getting caught by the ranger as particularly significant, because it gives him negative utility. Our simulations show that taking significance into account, even in this naive way, can lead to higher utility. We also show that the greater the weight of significant events, the greater the improvement in the utility, although the effect of the increasing the weight has diminishing returns. To understand the effects of bounded memory and taking significance into account, for various choices of ranger strategy, we compared the performance (in terms of utility) of various parameter settings of our PFA to each other and to other poacher strategies. Our results showed that probability matching and overweighting significance can often lead to higher utility. This supports one of our key hypotheses: It can be quite rational to be (somewhat) “irrational”, at least in the ranger-poacher game!

To see how humans actually play the ranger-poacher game, we ran experiments on Amazon Mechanical Turk (MTurk), using a number of different rhino distributions, with people playing the role of poacher for 100 rounds. Our experimental data show that most human poachers tend to probability match. As we show, this is also the case for poachers that use a PFA with small memory size M. As M increases, our PFA will play a combination of probability matching and NE strategy. These and other observations suggest that modeling people as PFAs does capture important aspects of human behavior.

2 THE RANGER-POACHER GAME

As in the wildlife poaching game introduced by Kar et al. [5], there are two players, a ranger and a poacher, and a fixed number n of sites that rhinos might go to. We assume that the rhino distribution $d = (d_1, \ldots, d_n)$ is commonly known, where $d_i \in [0, 1]$ is the probability that there is a rhino at site i (we do not assume that $\sum_{i=1}^{n} d_i = 1$; there could be more than one rhino!). We denote by $\Gamma^K(d)$ the ranger-poacher game with rhino distribution d and K stages, whose stage game is denoted $\Gamma(d)$. (Note that the distribution also implicitly encodes the number of sites.)

We take the poacher's utility in $\Gamma^K(d)$ to be his average utility in each of the K stage games, and similarly for the ranger. Although, in general, zero-sum games can have multiple equilibria, as we show in the full paper, the ranger-poacher game has a unique NE.

Proposition 2.1. *For all d, $\Gamma(d)$ has a unique NE.*

3 PFAS PLAY THE RANGER-POACHER GAME

To understand the effects of memory and taking significance into account, we compared the performance (in terms of utility) of various parameter settings of our PFA to each other and to other poacher strategies for various choices of ranger strategies. We considered eight poacher strategies: (1) the NE strategy, which can be viewed as a baseline; (2) *fictitiuos play* (FP), where a player keeps track of what the other player has done, and best responds to it, with unbounded memory; (3) *multiplicative weight updating* (MWU) [2], a strategy that has been shown to lead to NE quickly; (4) *utility matching* (UM) (instead of best responding, a site is chosen with probability proportional to its expected utility); (5) PFA1: a PFA with limited memory and no overweighting of significant events; (6) PFA2: a PFA with limited memory that overweights significant events; (7) PFA3: a PFA with very limited memory and no overweighting; (8) PFA4: a PFA with very limited memory that overweights significant events. We want to see how each of these eight poacher strategies plays against the various ranger strategies. We consider four ranger strategies: (a) the NE strategy; (b) probability matching (PM) based on the rhino distribution; (c) FP with unbounded memory; (d) a PFA with small memory and no overweighting. Notice that ranger strategies (a) and (b) are nonadaptive; the ranger visits a site according to a predetermined distribution at each step. In contrast, strategies (c) and (d) are adaptive; The ranger decides which site to visit next based on what the poacher has done in previous rounds. For each ranger-poacher pair, we simulated the game for 1000 rounds, using various rhino distributions, and repeated each game 100 times.

The results show that if the ranger uses a nonadaptive strategy (NE or PM), then all the poacher's strategies do equally well. However, if the ranger's strategy is adaptive, using a PFA with limited memory can significantly *improve* performance, especially when significance is taken into account. Thus, far from being irrational, overweighting significant events and probability matching are completely rational if the ranger is using FP. If the ranger is also using a PFA with a relatively small memory size, the ranger will also switch more often, so using FP or a PFA with a larger memory will make the poacher stickier and lead to lower utility for the poacher. However, overweighting significance and probability matching still help improve the poacher's utility significantly.

4 EXPERIMENTS

We wanted to understand the extent to which our PFAs capture human behavior in the ranger-poacher game. We conducted experiments on MTurk with 94 participants. In these experiments, human subjects play the role of the poacher; they must decide which site to visit in each round. We used the PFA to play the role of the ranger, with $M = 100$ and $s = 0$ (i.e., it does not take significance into account). Each game lasts for 100 rounds. Subjects are given the rhino distribution and are told that the ranger knows it as well. They are also told that, in each round, they and the ranger will simultaneously choose a site to visit. After these choices are made, the subjects discover where the rhinos were, so they can see whether they caught a rhino or were caught. They get 1 point if they catch a rhino without being caught, -1 point if they are caught, and 0 points otherwise. Subjects get $1 for completing the task plus a bonus of $0.10 for each point they obtain. Since the game lasts for 100 rounds, the bonus is usually significantly more than the fixed payment. Thus, they are (somewhat) incentivized to maximize their payoff by playing strategically. We submitted an IRB consent form and qualified for exemption from IRB review.

Applying k-means clustering, we clustered players on MTurk into three groups: (1) level-0 poachers, who visit all sites with equal probability or simply stick to one site; (2) level-1 poachers, who visit each site with probability roughly proportional to the rhino distribution; and (3) level-2 poachers, who seem to visit sites with probability proportional to their utility under the assumption that the rangers are playing a level-1 strategy. We suspect that level-0 players are often ones who simply want to get the game done as quickly as possible, so that they can collect the fixed payment. Therefore, we focus on level-1 and level-2 players. We can best approximate level-1 poacher behavior using a PFA with $M = 2$ and $s = 1$, where M denotes the memory size and s denotes whether or not we take significance into account. As explained earlier, a PFA that has a small memory will do more probability matching. We can best approximate level-2 poacher with a PFA with $M = 10$ and $s = 0$.

5 DISCUSSION AND RELATED WORK

There has been a great deal of recent interest in modeling human behavior in the ranger-poacher game. Perhaps most relevant to us is a sequence of papers by Tambe and his collaborators. Nguyen et al. [8] proposed the *Subjective Utility Quantal Response* (SUQR) model, which allows poachers to not best respond. Yang et al. [14] refined the model by allowing different poachers to be characterized by different weight vectors. Kar et al. [5] further refined the model by considering successes and failures of the poacher's past actions. Feng et al. [3] considered poachers with a fixed memory Γ, which is the number of rounds of past observations they respond to, similar to our use of M in this paper. Kar et al. [6] presented a behavior model based on an ensemble of decision trees. The goal of all this work was essentially to learn and predict human behavior. This is part of a more general thrust of trying to model human play in games (see, e.g., [11–13]). By way of contrast, our goal is to see the extent to which we can *explain* and *justify* apparently irrational human behavior (like probability matching and overweighting of significant events) as the outcome of computational limitations. To that end, we model resource-bounded poachers and rangers as PFAs. We showed that quite rational behavior (i.e., best responding) can lead to behaviors that have been viewed as irrational, namely, probability-matching and overweighting, as we limit the memory size. However, our results show that this so-called irrational behavior actually leads to *better* outcomes.

The fact that computational limitations lead both to more human-like behavior and (often) to better outcomes in the ranger-poacher game reinforces similar results obtained in other contexts [4, 7, 10]. This suggests a rather rich line of future research. As a first step, it would be of interest to see if behavior in other games, such as coordination games, can be explained and justified by computational limitations. We look forward to investigating and reporting on these issues in the future.

ACKNOWLEDGMENTS

We thank the reviewers for their insightful comments. This work was supported in part by NSF grants IIS-178108 and IIS-1703846 and MURI grant W911NF-19-1-0217.

REFERENCES

[1] D. Ariely. 2010. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. Harper Perennial, New York, NY.

[2] S. Arora, E.Hazan, and S. Kale. 2012. The multiplicative weights update method: A meta-algorithm and applications. *Theory of Computing* 8 (2012), 121–164.

[3] F. Feng, P. Stone, and M. Tambe. 2015. When security games go green: Designing defender strategies to prevent Poaching and illegal fishing. In *Proc. 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*. 2585–2595.

[4] J. Y. Halpern, R. Pass, and L. Seeman. 2012. I'm doing as well as I can: modeling people as rational finite automata. In *Proc. Twenty-Sixth National Conference on Artificial Intelligence (AAAI '12)*. 1917–1923.

[5] D. Kar, F. Fang, F. Delle Fave, N. Sintov, and M. Tambe. 2015. "Game of Thrones": when human behavior models compete in repeated Stackelberg security games. In *Proc. 2015 International Conference on Autonomous Agents and Multiagent Systems (AAAMAS '15)*. 1381–1390.

[6] D. Kar, B. Ford, S. Gholami, F. Fang, A. Plumptre, M. Tambe, M. Driciru, Wanyama F., A. Rwetsiba, M. Nsubaga, and J. Mabonga. 2017. Cloudy with a chance of poaching: Adversary behavior modeling and forecasting with real-world poaching data. In *Proc. 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '16)*. 159–167.

[7] X. Liu and J. Y. Halpern. 2020. Bounded rationality in Las Vegas: Probabilistic finite automata play multi-armed bandits. In *Proc. 36th Conference on Uncertainty in Artificial Intelligence (UAI 2020)*. 1298–1307. The proceedings are published as *Proceedings of Machine Learning Research*, Vol. 124.

[8] T. H. Nguyen, R. Yang, A. Azaria, S. Kraus, and M. Tambe. 2013. Analyzing the effectiveness of adversary modeling in security games. In *Proc. Twenty-Seventh National Conference on Artificial Intelligence (AAAI '13)*. 718–724.

[9] M. Tambe. 2012. *Security and Game Theory*. Cambridge University Press, Cambridge/New York.

[10] A. Wilson. 2015. Bounded memory and biases in information processing. *Econometrica* 82, 6 (2015), 2257–2294.

[11] J. R. Wright and K. Leyton-Brown. 2010. Beyond equilibrium: Predicting human behavior in normal-form games. In *Proc. Twenty-Fourth National Conference on Artificial Intelligence (AAAI '10)*. 901–907.

[12] J. R. Wright and K. Leyton-Brown. 2017. Predicting human behavior in unrepeated, simultaneous-move games. *Games and Economic Behavior* 106 (2017), 16–37.

[13] J. R. Wright and K. Leyton-Brown. 2019. Level-0 models for predicting human behavior in normal-form games. *Journal of A.I. Research* 64 (2019), 357–383.

[14] R. Yang, B. Ford, M. Tambe, and A. Lemieux. 2014. Adaptive resource allocation for wildlife protection against illegal poachers. In *Proc. Thirteenth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014)*. 453–460.