Exploiting Uncertain Real-Time Information from Deep Learning in Signaling Games for Security and Sustainability

Elizabeth Bondi
Harvard University
ebondi@g.harvard.edu

Hoon Oh
Carnegie Mellon University
hooh@andrew.cmu.edu

Haifeng Xu
University of Virginia
hx4ad@virginia.edu

Fei Fang
Carnegie Mellon University
feifang@cmu.edu

Bistra Dilkina
University of Southern California
dilkina@usc.edu

Milind Tambe
Harvard University
tambe@seas.harvard.edu

Abstract

Motivated by real-world deployment of drones for conservation, this paper advances the state-of-the-art in security games with signaling. The well-known defender-attacker security games framework can help in planning for such strategic deployments of sensors and human patrollers, and warning signals to ward off adversaries. However, we show that defenders can suffer significant losses when ignoring real-world uncertainties, such as detection uncertainty resulting from imperfect deep learning models, despite carefully planned security game strategies with signaling. In fact, defenders may perform worse than forgoing drones completely in this case. We address this shortcoming by proposing a novel game model that integrates signaling and sensor uncertainty; perhaps surprisingly, we show that defenders can still perform well via a signaling strategy that exploits the uncertain real-time information primarily from deep learning models. We provide a novel algorithm, scale-up techniques, and experimental results from simulation based on our ongoing deployment of a conservation drone system in South Africa.

1 Introduction

Conservation drones are currently deployed in South Africa to prevent wildlife poaching in national parks (Fig. 1). The drones, equipped with thermal infrared cameras, fly throughout the park at night when poaching typically occurs. Should anything suspicious be observed in the videos, nearby park rangers can prevent poaching, and a warning signal (e.g., drone lights) can be deployed for deterrence. This requires a great deal of planning and coordination, as well as constant video monitoring. Rather than constant monitoring, an automatic detection system based upon deep learning has recently been deployed to locate humans and animals in these videos (Anonymous). Although helpful, its detections are uncertain. Potential false negative detections, in which the system fails to detect actual poachers, may lead to missed opportunities to deter or prevent poaching. This work is motivated by this ongoing, real-world deployment of drones for conservation.

Security challenges similar to those in conservation must be addressed around the world, from protecting large public gatherings such as marathons (Yin, An, and Jain (2014)) to protecting cities. Security game models have been shown to be effective in many of these real-world domains, e.g., Tambe (2011); Bucarey et al. (2017). Recently, these models have begun to take into account real-time information, for example by using information from footprints when tracking poachers, or images

Abstract for: Smooth Games Optimization and Machine Learning Workshop (NeurIPS 2019), Vancouver, Canada.

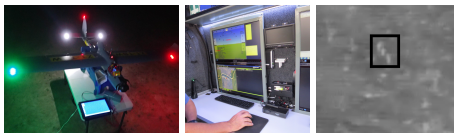

Figure 1: A drone and drone team member who are currently searching for poachers in a South African park at night.

from sensors, e.g., Wang et al. (2019); Basilico, De Nittis, and Gatti (2015). In particular, signaling based on real-time information, e.g., signaling to indicate the presence of law enforcement as in Xu et al. (2018), has been introduced and established as a fundamental area of work.

Despite the rising interest in real-time information and signaling, unfortunately, security games literature has failed to consider uncertainty in sensing real-time information and signaling, hindering real-world applicability of the game models. Previously, only some types of uncertainty have been considered, such as uncertainty in the attacker's observation of the defender's strategy, attacker's payoff values, or attacker's rationality as in Yin et al. (2011); Nguyen et al. (2014); Yang et al. (2011). However, there are fundamentally new insights when handling uncertainties w.r.t. real-time sensing and signaling, which we discuss now.

We make contributions in (i) modeling, (ii) algorithmic design and (iii) empirical evaluation. (i) We are the first to model uncertainty in sensing and signaling settings for security games. We introduce a novel reaction stage to the game model and construct a new signaling scheme, allowing the defender to mitigate the impact of uncertainty. In fact, this signaling scheme *exploits uncertain real-time information and the defender's informational advantage*. For example, both the defender and attacker may know that there is detection uncertainty; however, the defender has an informational advantage in knowing that she has or has not actually detected the attacker, which she can exploit via a signaling scheme to "mislead" the attacker who is uncertain as to whether he has been detected. (ii) To compute the defender's optimal strategy given uncertainty, we develop a novel algorithm, GUARDSS, that not only uses six states to represent the type of protection a target has in a defender's pure strategy but also uses a new matching technique in a branch-and-bound framework. (iii) We conduct extensive experiments on simulation based on our real-world deployment of a conservation drone system.

2 Related Work

Among the rich literature of Stackelberg security games (SSGs), such as Tambe (2011); Bucarey et al. (2017), SSGs with real-time information have been studied recently. Some rely on human patrollers for real-time information as in Zhang et al. (2019); Wang et al. (2019), others rely on sensors that can notify the patroller when an opponent is detected as in de Cote et al. (2013); Basilico, De Nittis, and Gatti (2015); De Nittis and Gatti (2018). Sensor placement (He et al. (2017)) and drone patrolling (Rosenfeld, Maksimov, and Kraus (2018)) have also been studied. Spatial and detection uncertainties in alarms are examined in Basilico, De Nittis, and Gatti (2016); Basilico, De Nittis, and Gatti (2017). In all of these works, the sensors are only used to collect information, and do not actively and possibly deceptively disseminate information to the attacker. One work that does consider mobile sensors with detection and signaling capability is Xu et al. (2018). However, it does not consider uncertainty in detection, which limits its capability in real-world settings. We add a new reaction stage and signaling strategy without detection, and compactly encode the different states that the defender resources can have at a target. Our model is therefore strictly more general than that in Xu et al. (2018).Our work is also related to multistage game models, e.g., defender-attacker-defender sequential games (DAD), e.g., Brown et al. (2006); Alderson et al. (2011). In DAD, the defender and attacker take turns to commit to strategies while in our game, the defender commits to a strategy of all stages at once. Extensive-form games (EFGs) also naturally model the sequential interaction between players, e.g., Kroer et al. (2017); Brown and Sandholm (2017); Moravčik et al. (2017), and recent works develop algorithms to efficiently solve the Stackelberg equilibrium in general two-player EFGs, e.g., Černý, Boỳanskỳ, and Kiekintveld (2018); Cermak et al. (2016). However, GUARDSS is more scalable than the general EFG approach.

3 Model

We consider a security game played between a defender and an attacker who seeks to attack one target. The defender has k human patrollers and l sensors to be allocated to targets in set $[N] = \{1, 2, ..., N\}$. The sensor is the same as a drone in our motivation domain, and the attacker is the same as a poacher. Let $U^{d/a}_{+/-}(i)$ be the defender/attacker (d/a) utility when the defender successfully protects/fails to protect $(+/-)$ the attacked target i. By convention, we assume $U^d_+(i) \geq 0 > U^d_-(i)$ and $U^a_+(i) \leq 0 < U^a_-(i)$ for any $i \in [N]$. The underlying geographic structure of targets is captured by an undirected graph $G = (V, E)$ (e.g., Fig. 3). A patroller can move to any neighboring target and successfully interdict an attack at the target at no cost.

Sensors cannot interdict an attack, but they can notify nearby patrollers to respond and signal to deter the attacker. If the attacker is deterred by a signal (e.g., runs away), both players get utility 0. In practice, often one signal (σ_1, e.g., illuminating the lights on the drone) is a warning that a patroller is nearby, while another signal (σ_0, e.g., turning no lights on) indicates no patroller is nearby, although these may be used deceptively. Theoretically, Kamenica and Gentzkow (2011) also showed two signals suffice (without uncertainty). We thus use two signals: σ_1 is a *strong signal* and σ_0 is a *weak signal*. When the attacker chooses one target to attack, he encounters one of four *signaling states*, based on the target either having a patroller, nothing, or a drone. The attacker may encounter: (1) a patroller and immediately get caught (state p); (2) nothing (state n); (3) a drone with signal σ_0 (state σ_0); (4) a drone with signal σ_1 (state σ_1).

3.1 Modeling Uncertainty

In this paper, we focus on two prominent uncertainties motivated directly by the use of conservation drones. The first is the *detection uncertainty*, when there is a limitation in the sensor's capability, e.g., a sensor's detection could be incorrect due to the inaccuracy of image detection techniques in the conservation domain as in Bondi et al. (2018); Olivares-Mendez et al. (2015). We consider only false negative detection in this paper because patrollers often have access to sensor videos, so the problem of false positives can be partly resolved by having a human in the loop. In contrast, verifying false negatives is harder, e.g., the attacker is easy to miss in the frame (Fig. 1) or is occluded although there is an attacker. We therefore denote the false negative rate as γ for any sensor[1].The second type of uncertainty we consider is the *observational uncertainty*, where the true signaling state of the target may differ from the attacker's observation (e.g., a poacher may not be able to detect the drone's signal). We omit details regarding this type of uncertainty, as it primarily adds notational complexity.

3.2 Handling Uncertainty

Uncertainty motivates us to (1) add an explicit reaction stage, during which the defender can respond *or* re-allocate patrollers to check on extremely uncertain sensors or previously unprotected targets, for example, and (2) to signal occasionally when we do not detect anything. The timing of the game is summarized in Fig. 2. In words, (i) the defender commits to a mixed strategy and then executes a pure strategy allocation; (ii) the attacker chooses a target to attack; (iii) the sensors detect the attacker with detection uncertainty; (iv) the sensors signal based on the signaling scheme; (v) *the defender re-allocates patrollers based on sensor detections and matching;* (vi) the attacker observes the signal with observational uncertainty; (vii) the attacker chooses to either continue the attack or run away. In (v), if a sensor detects the attacker, then nearby patroller(s) (if any) always go to that target, and the game ends; *or if no sensors or patrollers detect the attacker, the patroller moves to another target to check for the attacker.* The attacker reaction occurs after the defender reaction because the attacker reaction does not affect the defender reaction. In other words, there is no cost in reallocating the defender even if the attacker runs away, so the defender should begin moving right away.

To solve this game model, we introduce Games with Uncertainty And Response to Detection with Signaling Solver (GUARDSS), which employs the multiple LP approach for solving security games [Conitzer and Sandholm (2006)], along with the branch-and-price framework to accelerate our solver. This framework is well-known for solving large-scale optimization programs, but we modify the subroutine called the slave problem for solving each LP, and carefully design an upper bound for pruning LPs.

[1]False negative rate: $P(\text{no detection} \mid \text{poacher is present})$.

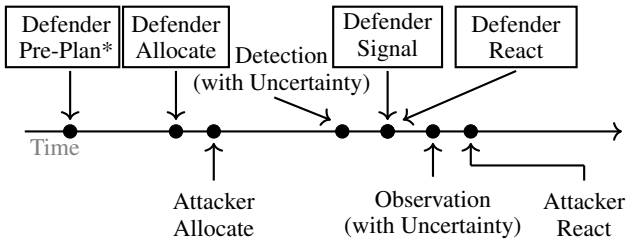

Figure 2: Game timing. Top and bottom are defender and attacker actions, respectively. *Defender fixes strategy offline.

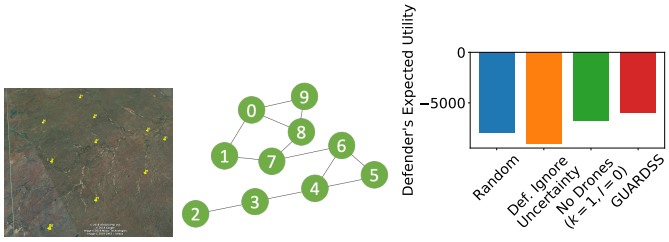

Figure 3: A park in Google Maps with potential poaching hotspots and the resulting graph (edges for < 5 km). Includes results from the case study, where GUARDSS performs best.

4 Conservation Drones

We have deployed a drone in South Africa, equipped with a thermal camera and detection system. A photo of the drone team in South Africa currently is included in Fig. 1 (center). To ease the challenges faced by these operators in coordination of drones with imperfect sensors and patrollers, we apply GUARDSS and show that it provides positive results in simulation to support future potential deployment. To facilitate the most realistic simulation possible, we utilize example poaching hotspots in a real park. We cannot provide the exact coordinates in order to protect wildlife, but we selected points based on geospatial features, and selected utilities (largest defender penalty is -16000 units) to reflect the fact that the reward and penalty of the attackers are impacted by animal presence, price, and distance to several park features used in Gholami et al. (2018). The targets are shown in Fig. 3 (left). Any targets within 5 km are connected via edges in the graph, as park rangers could cover 5km for response. The resulting graph is shown in Fig. 3 (center). For the simulation, we use 3 drones and 1 patroller. In the "no drones" scenario only, there are 0 drones and 1 patroller. We use $\gamma = 0.3$ for detection uncertainty and no observational uncertainty. These details are directly input to GUARDSS, and then a mixed strategy is determined to cover the park. Fig. 3 (right) shows the defender expected utility in this park using GUARDSS with and without uncertainty, and several baselines. A negative defender expected utility indicates that animals were lost, so a higher positive number is ideal. Therefore, we perform better with GUARDSS than using a random allocation, ignoring uncertainty, or forgoing drones. In fact, *ignoring uncertainty is worse than forgoing drones completely.* These results emphasize the importance of correctly optimizing to get value from drones even with uncertainty.

5 Conclusion

The loss due to ignoring uncertainty can be high such that sensors are no longer useful. Nevertheless, by carefully accounting for uncertainty, uncertain information can still be exploited via a novel reaction stage and signaling even upon no detection. Thriving under this uncertainty makes real-world deployment of GUARDSS promising, as shown through simulation.

6 Acknowledgements

This was supported by Microsoft AI for Earth, NSF CCF-1522054 and IIS-1850477, and MURI W911NF-17-1-0370.

References

Alderson, D. L.; Brown, G. G.; Carlyle, W. M.; and Wood, R. K. 2011. Solving defender-attacker-defender models for infrastructure defense. Technical report, Naval Postgraduate School.

Anonymous. Withheld double-blind review.

Basilico, N.; De Nittis, G.; and Gatti, N. 2015. A security game model for environment protection in the presence of an alarm system. In *GameSec*.

Basilico, N.; De Nittis, G.; and Gatti, N. 2016. A security game combining patrolling and alarm-triggered responses under spatial and detection uncertainties. In *AAAI*.

Basilico, N.; De Nittis, G.; and Gatti, N. 2017. Adversarial patrolling with spatially uncertain alarm signals. *Artificial Intelligence*.

Bondi, E.; Fang, F.; Hamilton, M.; Kar, D.; Dmello, D.; Choi, J.; Hannaford, R.; Iyer, A.; Joppa, L.; Tambe, M.; and Nevatia, R. 2018. Spot poachers in action: Augmenting conservation drones with automatic detection in near real time. In *IAAI*.

Brown, N., and Sandholm, T. 2017. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*.

Brown, G.; Carlyle, M.; Salmerón, J.; and Wood, K. 2006. Defending critical infrastructure. *Interfaces*.

Bucarey, V.; Casorrán, C.; Figueroa, Ó.; Rosas, K.; Navarrete, H.; and Ordóñez, F. 2017. Building real stackelberg security games for border patrols. In *GameSec*.

Cermak, J.; Bosansky, B.; Durkota, K.; Lisy, V.; and Kiekintveld, C. 2016. Using correlated strategies for computing stackelberg equilibria in extensive-form games. In *AAAI*.

Černỳ, J.; Boỳanskỳ, B.; and Kiekintveld, C. 2018. Incremental Strategy Generation for Stackelberg Equilibria in Extensive-Form Games. In *EC*.

Conitzer, V., and Sandholm, T. 2006. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM conference on Electronic commerce*, 82–90. ACM.

de Cote, E.; Stranders, R.; Basilico, N.; Gatti, N.; and Jennings, N. 2013. Introducing alarms in adversarial patrolling games. In *AAMAS*.

De Nittis, G., and Gatti, N. 2018. Facing Multiple Attacks in Adversarial Patrolling Games with Alarmed Targets. *arXiv preprint arXiv:1806.07111*.

Gholami, S.; Mc Carthy, S.; Dilkina, B.; Plumptre, A.; Tambe, M.; Driciru, M.; Wanyama, F.; Rwetsiba, A.; Nsubaga, M.; Mabonga, J.; et al. 2018. Adversary models account for imperfect crime data: Forecasting and planning against real-world poachers. In *AAMAS*.

He, Y.; Ma, X.; Luo, X.; Li, J.; Zhao, M.; An, B.; and Guan, X. 2017. Vehicle Traffic Driven Camera Placement for Better Metropolis Security Surveillance. *arXiv preprint arXiv:1705.08508*.

Kamenica, E., and Gentzkow, M. 2011. Bayesian persuasion. *American Economic Review*.

Kroer, C.; Waugh, K.; Kilinc-Karzan, F.; and Sandholm, T. 2017. Theoretical and practical advances on smoothing for extensive-form games. *arXiv preprint arXiv:1702.04849*.

Moravčik, M.; Schmid, M.; Burch, N.; Lisy, V.; Morrill, D.; Bard, N.; Davis, T.; Waugh, K.; Johanson, M.; and Bowling, M. 2017. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*.

Nguyen, T. H.; Yadav, A.; An, B.; Tambe, M.; and Boutilier, C. 2014. Regret-Based Optimization and Preference Elicitation for Stackelberg Security Games with Uncertainty. In *AAAI*.

Olivares-Mendez, M. A.; Fu, C.; Ludivig, P.; Bissyandé, T. F.; Kannan, S.; Zurad, M.; Annaiyan, A.; Voos, H.; and Campoy, P. 2015. Towards an autonomous vision-based unmanned aerial system against wildlife poachers. *Sensors*.

Rosenfeld, A.; Maksimov, O.; and Kraus, S. 2018. Optimal cruiser-drone traffic enforcement under energy limitation. In *IJCAI*.

Tambe, M. 2011. *Security and game theory: algorithms, deployed systems, lessons learned*. Cambridge University Press.

Wang, Y.; Shi, Z. R.; Yu, L.; Wu, Y.; Singh, R.; Joppa, L.; and Fang, F. 2019. Deep reinforcement learning for green security games with real-time information. In *AAAI*.

Xu, H.; Wang, K.; Vayanos, P.; and Tambe, M. 2018. Strategic coordination of human patrollers and mobile sensors with signaling for security games. In *AAAI*.

Yang, R.; Kiekintveld, C.; Ordonez, F.; Tambe, M.; and John, R. 2011. Improving resource allocation strategy against human adversaries in security games. In *IJCAI*.

Yin, Y.; An, B.; and Jain, M. 2014. Game-theoretic resource allocation for protecting large public events. In *AAAI*, 826–834.

Yin, Z.; Jain, M.; Tambe, M.; and Ordonez, F. 2011. Risk-averse strategies for security games with execution and observational uncertainty. In *AAAI*.

Zhang, Y.; Guo, Q.; An, B.; Tran-Thanh, L.; and Jennings, N. R. 2019. Optimal interdiction of urban criminals with the aid of real-time information. In *AAAI*.