Wildlife GUARDSS: Using Uncertain Real-Time Information in Signaling Games for Sustainability

Elizabeth Bondi [1] **Hoon Oh** [2] **Haifeng Xu** [3] **Fei Fang** [2] **Bistra Dilkina** [1] **Milind Tambe** [1]

Abstract

Mobile sensors, e.g., unmanned aerial vehicles (UAVs), are becoming increasingly important in security domains and can be used for tasks such as searching for poachers in conservation areas. Such mobile sensors augment human patrollers by assisting in surveillance and in signaling potentially deceptive information to adversaries, and their coordinated deployment could be modeled via the well-known security games framework. Unfortunately, real-world uncertainty in the sensor's detection of adversaries presents major challenges in the sensors' use. This leads to significant detriments in security performance. We first discuss the current shortcomings in more detail, and then propose a novel game model that incorporates uncertainty with sensors. We then briefly introduce GUARDSS, the algorithm to solve these games, and show results from a simulation based on a real-world deployment of a conservation system in South Africa.

1. Introduction

In many real-world situations, there are not enough security resources, such as human patrollers, to protect all possible targets from attackers and prevent illegal activities. Security games have been used to model and solve strategic security resource allocation in these situations in the past decade for problems such as protecting airports, traffic enforcement, protecting elections, and protecting borders (Tambe, 2011; Rosenfeld & Kraus, 2017; Bucarey et al., 2017). Concurrently, mobile sensors such as unmanned aerial vehicles (UAVs or drones) have been introduced for security purposes with an increasing importance in domains such as traffic enforcement (Rosenfeld et al., 2018) and wildlife poaching prevention (Mulero-Pázmány et al., 2014). The security game framework has been augmented and applied to the coordinated deployment of human patrollers and mobile sensors as well as strategic signaling (Xu et al., 2018).

Unfortunately, real-world circumstances inevitably involve uncertainty in the sensors' detection of adversaries, leading to challenges in successfully using sensors in security domains. Our motivation comes directly from the real-world domain of wildlife conservation, and in particular, preventing poaching. UAVs equipped with thermal infrared (heat-detecting) cameras are used to locate poachers at night when poaching typically occurs (Air Shepherd, 2018) and sometimes send warning signals to poachers through onboard lights for deterrence. In Fig. 1, a deployed conservation drone (left) equipped with a thermal infrared camera is used to locate a poacher in the rectangle (center) in order to prevent poaching in a national park (right). Although useful, detectors such as those in (Bondi et al., 2018; Olivares-Mendez et al., 2015; van Gemert et al., 2014) suffer from imperfect detection, and poachers may not even see signals due to occlusions by trees. Ignoring such uncertainties would result in significant detriments in security performance. Consider a sensor with a high false negative rate as an example. In this case, it could be beneficial for the human patroller to go and check a nearby location even if the sensor in the location does not detect any adversary. This would be done to confirm that there is no adversary there, rather than fully trusting the sensor. Fully trusting the sensors' capability of detecting adversaries leads to a wrong belief of the location of the adversary, and the efficiency of patrol can be even worse than not having any sensors. We aim to address this limitation and provide an efficient patrol plan that works in an environment with uncertainty.

We make contributions in (i) modeling, (ii) algorithmic design and (iii) empirical evaluation: (i) We are the first to model uncertainty in signaling settings for security games. We introduce the novel reaction stage to the game model, allowing the defender to mitigate the impact of such uncertainties. A fundamental novelty of our model compared to (Xu et al., 2018) is finding and compactly encoding six different states that the defender resources can have at a target. This model is more general than (Xu et al., 2018),

[1]University of Southern California, Los Angeles, California, USA [2]Carnegie Mellon University, Pittsburgh, Pennsylvania, USA [3]Harvard University, Cambridge, Massachusetts, USA. Correspondence to: Elizabeth Bondi <bondi@usc.edu>.

Appearing at the International Conference on Machine Learning AI for Social Good Workshop, Long Beach, United States, 2019.

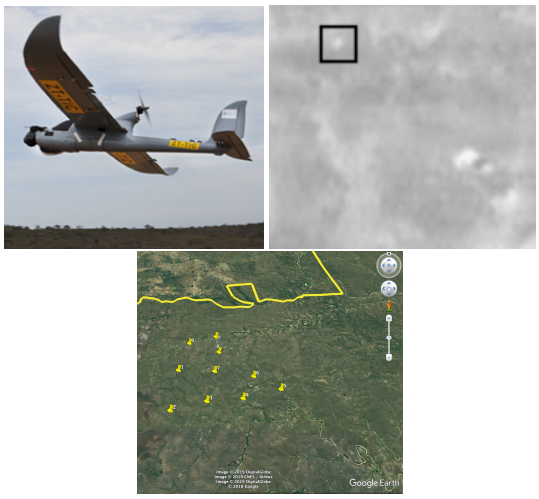

Figure 1. Drone that captures thermal images of people in Africa.

and has an augmented space of the defender strategy which includes action in the reaction stage. (ii) To compute the defender's optimal strategy given uncertainty, we develop a novel algorithm that not only extends to the six different states (compared to four states in (Xu et al., 2018)) but also uses a new matching technique. Specifically, nodes with a patroller must be matched to nodes using a directed graph and linear constraints. (iii) We provide experimental results for random instances and simulation based on our real-world deployment of a conservation system in South Africa.

2. Related Work

Among the rich literature of Stackelberg security games (SSGs) (Tambe, 2011; Bucarey et al., 2017), SSGs with uncertainty have been studied. Several 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 (Yin et al., 2011; Nguyen et al., 2014; Yang et al., 2011), but these do not focus on detection uncertainty. Spatial and detection uncertainties in alarms are examined in (Basilico et al., 2016; 2017), but the sensors are only used to collect information, and do not actively and possibly deceptively disseminate information to the attacker.

Our work is also related to multistage game models. Defender-attacker-defender sequential games (DAD) have been studied (Brown et al., 2006; Alderson et al., 2011). While our game has multiple stages, the defender commits to a strategy of all stages at once and the attacker best responds while in DAD, the defender and attacker take turns to commit to strategies. Extensive-form games (EFGs) also naturally model the sequential interaction between players [(Kroer et al., 2017; Brown & Sandholm, 2017; Moravčik et al., 2017)], and recent works develop algorithms to efficiently solve the Stackelberg equilibrium in general two-player EFGs (Černỳ et al., 2018; Cermak et al., 2016). However, GUARDSS is more scalable than the general-purpose EFG approach because the EFG approach solves exponentially many more linear programs (LPs) depending on the number of drones.

3. Model

We consider a security game played between a defender and an attacker, who seeks to attack one target. The defender possesses k human patrollers and l sensors, and aims to protect N targets. Let $[N] = \{1, 2, ..., N\}$ denote the set of all targets. 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 $G = (V, E)$. Mobile sensors cannot interdict an attack, though they can notify nearby patrollers to respond. If a target i is attacked, then we assume that a patroller at any neighboring target of i can move to i and successfully interdict the attack. Mobile sensors will send one of two signals – the quiet and warning signals to the attacker. The warning signal (lights on aboard the UAV) is used to warn the attacker off. The quiet signal (lights off aboard the UAV) means that nobody is responding. We would like a model in which the adversary would stop the attack and run away upon seeing a warning signal. We will first discuss types of uncertainty, then the multistage game model to incorporate uncertainty, the additional considerations for adding uncertainty, and finally the solution method for this game model with uncertainty.

3.1. Types of Uncertainty

Uncertainty is a crucial factor in automated applications of mobile sensors, yet has not been considered in previous work (Xu et al., 2018). With *detection uncertainty*, the sensor could fail to detect a real attacker (false negative), or it could incorrectly classify something as an attacker (false positive) due to the inaccuracy of image recognition techniques (Bondi et al., 2018; Olivares-Mendez et al., 2015; van Gemert et al., 2014). We only consider false negatives in this work because the patrollers often have access to sensor videos, and the problem of false positives can be partly resolved by having a human in the loop.

3.2. Multistage Game Model

To facilitate incorporating uncertainty, we start with a novel three-stage game model: (1) *allocation stage* where (i) the defender places security resources (defender allocation stage), and (ii) the attacker chooses a target to attack based on the defender mixed strategy (attacker allocation stage);

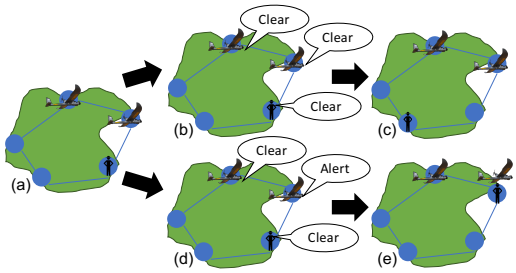

Figure 2. The reaction stage. (a) shows an initial allocation. In (b), no attacker is detected, so the patroller moves to the matched target (c). In (d), an attacker is detected, so the patroller responds (e).

(2) *signaling stage* where the mobile sensors send signals based on detection (defender signaling stage); (3) *reaction stage* where (i) defender reacts to the sensor detection and relocates a human patroller to a nearby target (defender reaction stage), and (ii) the attacker chooses to deploy the attack or run away after the observation (attacker reaction stage). In practice, the defender signaling and reaction stages can happen simultaneously. In stage (3), the human defender moves from the original assigned location to a new location. If the attacker is detected by a sensor, nearby patroller(s) react by moving to the attacker's location to interdict. Unlike (Xu et al., 2018), if no sensors or patrollers detect the attacker, the defender still reacts by moving to another target. An example is provided in Fig. 2.

As a result of this rich structure, a pure strategy in the model induces 6 possible allocation states for each target. Let $\Theta = \{\text{p}, \text{n}+, \text{n}-, \bar{\text{s}}, \text{s}+, \text{s}-\}$ denote the set of all possible allocation states of an individual target. The target is assigned a patroller (p), nothing (n), or a sensor (s). If there is no patroller near a sensor ($\bar{\text{s}}$), then no one can respond to the sensor's detection. If there is a nearby patroller, the target is either matched (n+,s+) or not matched (n−,s−). Therefore, each target is in one of the allocation states in Table 3.2. For example, n+ is the state of a target which was not allocated a patroller or sensor, but in the reaction stage has a patroller from a neighboring target ("patroller matched").

	Covered By:	Near Patroller?	Patroller Matched?	Protected Overall?
p	Patroller	N/A	N/A	Yes
n+	Nothing	Yes	Yes	Yes
n-	Nothing	N/A	No	No
$\bar{\text{s}}$	Sensor	No	N/A	No
s-	Sensor	Yes	No	Yes*
s+	Sensor	Yes	Yes	Yes

Table 1. Allocation State, *protected if sensor detects

Given Θ, a defender pure strategy can be compactly represented with an allocation state vector $\mathbf{e} \in \Theta^N$, in which $e_i \in \Theta$ denote the allocation state of a target $i \in [N]$. Let $\mathcal{E} \subseteq \Theta^N$ be the set of feasible allocation state vectors that corresponds to defender pure strategies. A *defender mixed strategy* is thus a distribution over $\mathcal{E}$ and can be described by $\{q_e\}_{\mathbf{e}\in\mathcal{E}}$ where $q_\mathbf{e}$ is the probability of playing pure strategy $\mathbf{e} \in \mathcal{E}$. Similarly, a defender mixed strategy can also be compactly represented by a marginal probability vector x, where x_i^θ represents the marginal probability that target i is in the allocation state $\theta \in \Theta$.

3.3. Adding Uncertainty

Uncertainty affects many aspects of the game model, such as the utilities, attacker behavior, and signaling and reaction strategy. For instance, in previous work (Xu et al., 2018; Ma et al., 2018), if the sensor does not detect an attacker, the patroller does not do anything. Now, patrollers should relocate to locations where sensors are placed rather than just empty locations in order to check for false negative detections. The false negative rate is denoted by γ. Similarly, we may need to signal when there is no detection. The probability of sending a quiet signal on no detection is denoted by $\varphi_i^\theta \in [0, x_i^\theta]$, where θ represents one of the six states in Θ. The probability of sending a quiet signal on a detection is denoted by $\psi_i^{\text{s}-} \in [0, x_i^\theta]$.

Each player's utility function is broken into three parts: 1) when no sensor is allocated ($U_{\text{-s}}^{d/a}$); 2) when sensor is allocated and signals nothing (σ_0) ($U_{\sigma_0}^{d/a}$); and 3) when sensor is allocated and sends the warning signal (σ_1) ($U_{\sigma_1}^{d/a}$). In words, 2) and 3) are the sum of signaling on a detection and the sum of signaling on no detection, with γ associated with no detection. We omit the full formulas due to space. The (exponentially-large) linear program (LP) formulation for computing the optimal defender strategy assuming best attacker response t, as follows:

$$\max_{x,\psi,\varphi} \quad U_{\text{-s}}^d(t) + U_{\sigma_0}^d(t) \tag{1}$$

$$\text{s.t.} \quad \sum_{\mathbf{e}\in\mathcal{E}:e_i=\theta} q_\mathbf{e} = x_i^\theta \qquad \forall\,\theta\in\Theta, \forall\, i\in[N] \tag{2}$$

$$\sum_{\mathbf{e}\in\mathcal{E}} q_\mathbf{e} = 1 \tag{3}$$

$$q_\mathbf{e} \geq 0 \qquad \forall\,\mathbf{e}\in\mathcal{E} \tag{4}$$

$$U_{\sigma_0}^a(i) \geq 0 \qquad \forall i \neq t \tag{5}$$

$$U_{\sigma_1}^a(i) \leq 0 \qquad \forall i \neq t \tag{6}$$

$$U_{\text{-s}}^a(t) + U_{\sigma_0}^a(t) \geq U_{\text{-s}}^a(i) + U_{\sigma_0}^a(i) \quad \forall i \neq t \tag{7}$$

$$0 \leq \psi_i^\theta \leq x_i^\theta \qquad \forall\,\theta\in\Theta_s, \forall\, i\in[N] \tag{8}$$

$$0 \leq \varphi_i^\theta \leq x_i^\theta \qquad \forall\,\theta\in\Theta_s, \forall\, i\in[N] \tag{9}$$

The objective function maximizes defender expected utility. The first three constraints (2)-(4) enforce that the randomized resource allocation is feasible. The next two con-

straints (5)-(6) guarantee that σ_1, σ_0 result in the attacker best responses of running away and attacking[1]. The next constraint (7) ensures the attacker expected utility at target t is bigger than the attacker expected utility at any other target i, thus t is attacker's best response. The last two constraints (8)-(9) ensure a feasible signaling scheme, where $\Theta_s = \{\bar{s}, s+, s-\}$ denote the subset of allocation states with a sensor.

3.4. Solution Method

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 & 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.

Specifically, we adopt a column generation technique for one LP w.r.t. a specific t to address the issue of the exponential size of set $\mathcal{E}$. At a high level, we start by solving the LP for a small subset $\mathcal{E}' \subset \mathcal{E}$, and then search for a pure strategy $e \in \mathcal{E} \setminus \mathcal{E}'$ such that adding e to $\mathcal{E}'$ improves the optimal objective value strictly. This procedure continues until convergence, i.e., no objective value improvement. The key component in this technique is an algorithm to search for the new pure strategy, which is a specially-crafted problem derived from LP duality and referred to as the *slave problem*.

Slave Problem: Given different weights $\alpha_i^\theta \in \mathbb{R}$ for $\theta \in \Theta$, for each target i, solve the *weight maximization problem*:

$$\max_{e \in \mathcal{E}} \sum_{\theta \in \Theta} \sum_{i: e_i = \theta} \alpha_i^\theta \tag{10}$$

Note that $\{\alpha_i^\theta\}_{\theta \in \Theta}$ are the optimal dual variables for the previous LP constraints. Despite the more complex structure than classic SSGs, in this section, we compactly represent this slave problem as a mixed integer linear program (MILP) by introducing binary vectors to encode for each target what state it is in.

Handling the Reaction Stage: Due to the reaction stage, we have to add constraints to specify (a) which vertices have a patroller at a neighboring target; (b) which patroller goes to which nearby vertex if both sensors and patrollers do not detect the attacker. Constraint (b) means that patrollers must be "re-matched" to new vertices in the reaction stage.

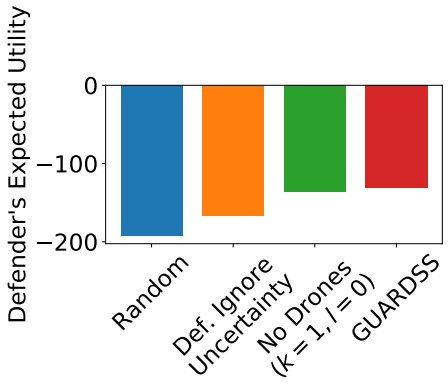

Figure 3. Results from the case study.

4. Conservation Drones

We have deployed a drone in South Africa, equipped with a thermal camera and detection system (Bondi et al., 2018). There are several challenges in using conservation drones. They must determine where to fly the drone, where to allocate human patrollers, and whether to signal while flying. Without a strategy, they may signal too frequently, thereby rendering the signals meaningless. Additionally, the detection system is not perfect. To ease these challenges, we apply GUARDSS and show that it provides positive results in a simulated scenario to support future potential deployment of the algorithm. In this scenario, we use the targets shown in Fig. 1. These areas were selected due to their proximity to the park border and rivers. Any targets within 5 km are connected via edges in a graph, as this is a distance the end-users could cover for response. We use $\gamma = 0.7$. There are 3 sensors and 1 patroller (other than the no drones scenario, in which there are 0 sensors and 1 patroller). Fig. 3 shows that we perform better with GUARDSS than using a random allocation and ignoring uncertainty. Testing with random graphs of various sizes, including Watts Strogatz and cycle graphs, shows similar relationships. Forgoing drones performs similarly to GUARDSS in this scenario. In general, we see that for high γ, drones are not as helpful as human patrollers, though they are more helpful for lower γ. Cost-benefit analysis would be necessary before deploying in the real world depending on the expected γ. However, the results emphasize the importance of correctly optimizing the defender's strategy, such as in GUARDSS, to get value from drones with detection uncertainty.

5. Acknowledgements

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

[1] Although we minimize this behavior, we still model it.

References

Air Shepherd. Air shepherd: The lindbergh foundation. http://airshepherd.org, 2018. Accessed: 2018-08-01.

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

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

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

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

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

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

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

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

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

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

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

Ma, X., He, Y., Luo, X., Li, J., Zhao, M., An, B., and Guan, X. Camera placement based on vehicle traffic for better city security surveillance. *IEEE Intelligent Systems*, 33 (4):49–61, 2018.

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

Mulero-Pázmány, M., Stolper, R., Van Essen, L., Negro, J. J., and Sassen, T. Remotely piloted aircraft systems as a rhinoceros anti-poaching tool in africa. *PloS one*, 9(1): e83873, 2014.

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

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

Rosenfeld, A. and Kraus, S. When security games hit traffic: Optimal traffic enforcement under one sided uncertainty. In *Proceedings of the 26th International Conference on Artificial Intelligence, IJCAI*, 2017.

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

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

van Gemert, J. C., Verschoor, C. R., Mettes, P., Epema, K., Koh, L. P., Wich, S., et al. Nature conservation drones for automatic localization and counting of animals. In *ECCV Workshops*, 2014.

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

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

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