To Signal or Not To Signal: Exploiting Uncertain Real-Time Information in Signaling Games for Security and Sustainability

Elizabeth Bondi,[1] **Hoon Oh,**[2] **Haifeng Xu,**[3] **Fei Fang,**[2] **Bistra Dilkina,**[4] **Milind Tambe**[1]
[1]Center for Research on Computation and Society, Harvard University, ebondi@g.harvard.edu, tambe@seas.harvard.edu
[2]Carnegie Mellon University, hooh@andrew.cmu.edu, feifang@cmu.edu
[3]University of Virginia, hx4ad@virginia.edu
[4]University of Southern California, dilkina@usc.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 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 uncertain real-time information. For example, even in the presence of uncertainty, the defender still has an informational advantage in knowing that she has or has not actually detected the attacker; and she can design a signaling scheme to "mislead" the attacker who is uncertain as to whether he has been detected. We provide theoretical results, 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 (Air Shepherd 2019). This requires a great deal of planning and coordination, as well as constant video monitoring. Rather than constant monitoring, we have recently worked with Air Shepherd to deploy an automatic detection system to locate humans and animals in these videos. Although an automatic detection system is 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.

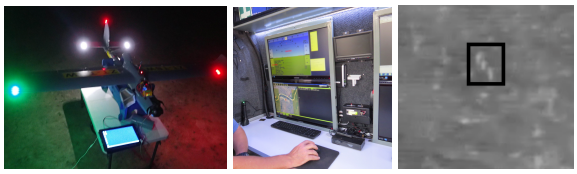


Figure 1: A drone and drone team member who are currently searching for poachers in a South African park at night.

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 (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 from sensors (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 (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 (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 at the end of this section.

We therefore focus on uncertainty in security games, in which real-time information comes from sensors that alert the defender when an attacker is detected and can also send warning signals to the attacker to deter the attack in real time. We consider both uncertainty in the sensor's detection of adversaries (henceforth detection uncertainty) and uncer-

tainty in the adversaries' observation of the sensor's signals (henceforth observational uncertainty), and show that ignoring uncertainty hurts the defender's expected utility. In our motivating domain of wildlife conservation with drones, automatic detection algorithms may make incorrect detections because humans in thermal infrared frames look similar to other objects (e.g., Fig. 1) and may even be occluded by other objects from the aerial perspective. The drone is also used to emit light to deter poachers, but such signals could sometimes be difficult for poachers to see in the wild, e.g., when trees block the sight.

We make contributions in (i) modeling, (ii) theoretical analysis, (iii) algorithmic design, and (iv) 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) We provide several theoretical results on the impact of uncertainties, e.g., the loss due to ignoring observational uncertainty can be arbitrarily large, illustrating the need to handle uncertainty. (iii) 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. (iv) 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) (Tambe 2011; Bucarey et al. 2017), SSGs with real-time information have been studied recently. Some recent work in deception for cybersecurity, such as (Cooney et al. 2019; Thakoor et al. 2019), considers strategic signaling with boundedly rational attackers and attackers with different objectives and abilities, but no sensing is required to identify attackers; rather, the systems may interact with both normal and adversarial users. Some other work relies on human patrollers for real-time information (Zhang et al. 2019; Wang et al. 2019), and others rely on sensors that can notify the patroller when an opponent is detected (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) (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 (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 (Černý, Boỳanskỳ, and Kiekintveld 2018; Cermak et al. 2016). However, GUARDSS is more scalable than the general EFG approach in this case (see Appendix).

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. 4). 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). The attacker is caught immediately at state p, so there is no signal. Therefore, we omit p and let $\Omega = \{\text{n}, \sigma_0, \sigma_1\}$ be the set of signaling states.

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 detection could be incorrect due to the inaccuracy of image detection techniques in the conservation domain (Bondi et al. 2020; 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 with 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. 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 use $\hat{\omega}$ to denote the attacker's observed signaling state, and use ω to denote the true signaling state based on the defender signaling scheme. We introduce uncertainty matrix Π to capture observational uncertainty. The uncertainty matrix Π will contain the conditional probability $\Pr[\hat{\omega}|\omega]$ for all $\hat{\omega}, \omega \in \Omega$ to describe how likely the attacker will observe a signaling state $\hat{\omega}$ given the true signaling state is ω.

$$\Pi = \begin{bmatrix} \Pr[\hat{\omega} = \text{n}|\text{n}] & \Pr[\hat{\omega} = \text{n}|\sigma_0] & \Pr[\hat{\omega} = \text{n}|\sigma_1] \\ \Pr[\hat{\omega} = \sigma_0|\text{n}] & \Pr[\hat{\omega} = \sigma_0|\sigma_0] & \Pr[\hat{\omega} = \sigma_0|\sigma_1] \\ \Pr[\hat{\omega} = \sigma_1|\text{n}] & \Pr[\hat{\omega} = \sigma_1|\sigma_0] & \Pr[\hat{\omega} = \sigma_1|\sigma_1] \end{bmatrix}$$

Considering an arbitrary uncertainty matrix may unnecessarily complicate the problem, since some uncertainties never happen. We thus focus on a restricted class of uncertainty matrices that are natural in our domain.[2] In our uncertainty model, we assume that a weak signal will never be observed as strong; moreover, n (the signaling state without any resource) will never be observed as strong or weak. As a result, the uncertainty matrix Π can be reduced to the following form, parameterized by κ, λ, μ, where $\kappa = \Pr[\hat{\omega} = \text{n}|\sigma_0]$, $\lambda = \Pr[\hat{\omega} = \text{n}|\sigma_1]$, $\mu = \Pr[\hat{\omega} = \sigma_0|\sigma_1]$:

$$\Pi_{\kappa\lambda\mu} = \begin{bmatrix} 1 & \kappa & \lambda \\ 0 & 1-\kappa & \mu \\ 0 & 0 & 1-\lambda-\mu \end{bmatrix}$$

As a result of this uncertainty, the attacker may not behave as expected. For example, if he knows that he has difficulty seeing the strong signal, he may decide to attack only when there is no drone, whereas typically we would expect him to attack on a weak signal. Therefore, let $\eta \in \{0,1\}^3$ be the vector that depicts attacker behavior for each observation $\{\text{n}, \sigma_0, \sigma_1\} \in \Omega$, where 1 represents attacking, and 0 represents running away. So, $\eta = \mathbf{1}$ means an attacker will attack no matter what signaling state is observed, and $\eta = \mathbf{0}$ means an attacker will never attack.

3.2 Reaction Stage

Uncertainty motivates us to 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. 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 the current model. 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.

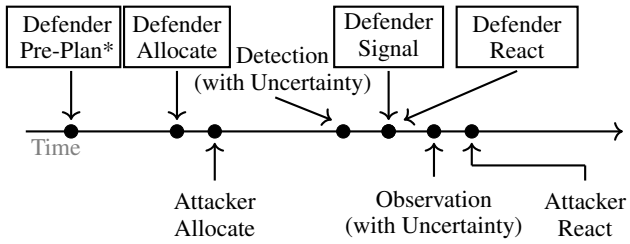


Figure 2: Game timing. Top and bottom are defender and attacker actions, respectively. *Defender fixes strategy offline.

3.3 Defender and Attacker Strategies

Defender Strategy: The strategy space consists of randomized resource allocation and re-allocation, and signaling. A deterministic resource allocation and re-allocation strategy (henceforth, a *defender pure strategy*) consists of allocating the patrollers to k targets, the sensors to l targets, and the neighboring target to which each patroller moves if no attackers are observed. Re-allocation can be equivalently thought of as matching each patroller's original target to a neighboring target. A patroller goes to the matched target only if the attacker is not observed, and may respond to any nearby sensor detection, regardless of matching.

As a result of this rich structure, a pure strategy in the model needs to represent not only if the target is assigned a patroller (p), nothing (n), or a sensor (s), but also the allocation in neighboring targets. We compactly encode this pure strategy via 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 1. 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").

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

[2] Most results can be extended to general uncertainty matrices.

	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. Note that not all vectors in Θ^N correspond to a feasible defender strategy due to the limited number of patrollers and sensors. 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 $\theta \in \Theta$. This is similar to the coverage vector used in basic SSGs with schedules (Jain et al. 2010). We introduce the constraints that x needs to satisfy to be a valid mixed strategy in Section 5.

The defender also deploys a signaling process w.r.t. each target i. The defender's signaling strategy can be specified by probabilities $\psi_i^{\text{s}-}$, $\psi_i^{\text{s}+}$, and $\psi_i^{\bar{\text{s}}}$. $\psi_i^{\text{s}-}$ is the *joint* probability of allocation state s$-$ and sending signal σ_0 together conditioned on the sensor detecting an attacker, i.e., $\Pr[\text{s}- \wedge \sigma_0|\text{detected}]$. To be a valid signaling strategy, $\psi_i^{\text{s}-} \in [0, x_i^{\text{s}-}]$. Note that $x_i^{\text{s}-} - \psi_i^{\text{s}-}$ will be the joint probability of realized state s$-$ and sending signal σ_1, together conditioned on detection. The conditional probability of sending σ_0 given the target is in state s$-$ and it is detected is $\psi_i^{\text{s}-}/x_i^{\text{s}-}$. We use the joint probability instead of the conditional probability as it results in linear terms for the optimal defender strategy. *Because of detection uncertainty, we add the option to signal without detecting the attacker.* Let $\varphi_i^\theta \in [0, x_i^\theta]$ be the joint probability of allocation state θ and sending signal σ_0 conditioned on the sensor not detecting an attacker, for all $\theta \in \{\bar{\text{s}}, \text{s}-, \text{s}+\}$. We use χ to denote the allocation, reaction, and signaling scheme, or *defender's deployment strategy*: $\chi = (x, \psi, \varphi)$.

Attacker Strategy: Recall the attacker has the allocation and reaction stages. In the allocation stage, the attacker chooses a target to attack based on the defender deployment strategy χ. He will be caught if the target is at state p. When the attacker is not caught, he may observe any of the signaling states $\hat{\omega} \in \Omega$. Based on his observation, the attacker then has a choice in the reaction stage to run away or continue the attack. The attacker knows the defender mixed strategy χ when choosing a target to attack, and he can observe the realization of the target (with uncertainty) when choosing to attack or run away. Since this is a Stackelberg game and the defender commits to allocation and signaling schemes, it suffices to consider only the attacker's pure responses.

4 Why Do We Need to Handle Uncertainty

In this section, we prove several theoretical properties regarding how uncertainties affect the defender's optimal strategy and utility. All formal proofs are deferred to the Appendix. Let $\chi^*(\gamma, \Pi)$ be the optimal allocation under detection uncertainty of γ and observational uncertainty Π. Let $\mathsf{DefEU}(\chi, \gamma, \Pi)$ be the defender expected utility when the actual uncertainties are γ, Π and the defender's deployment is χ. Let $\Pi_0 = \mathbf{I}$ denote no observational uncertainty. We assume in Propositions 1 and 2 and Theorem 1 that $\Pi = \Pi_0$ and analyze detection uncertainty, so omit for conciseness. We first show the loss due to ignoring detection uncertainty.

Proposition 1. *Let $\chi_0^* = \chi^*(0)$ be the defender optimal deployment when no uncertainties exist. There exist instances where $\mathsf{DefEU}(\chi_0^*, \gamma) < \mathsf{DefEU}(\chi^*(\gamma), \gamma)$ for some γ.*

In fact, $\mathsf{DefEU}(\chi^*(\gamma), \gamma) - \mathsf{DefEU}(\chi_0^*, \gamma) \geq \gamma \cdot \max_{i\in[N]} |U_-^d(i)|$

for some instance. If we ignore γ, we do not signal when we do not detect an attacker. Furthermore, the defender would never match a patroller to a target with a sensor (s+) in χ_0^*. Thus, if we ignore uncertainty, there can be a steep penalty; in contrast, with the optimal strategy considering uncertainty, if the false negative rate is high, we may match a patroller to a target to confirm the presence of an attacker. Given the attacker's knowledge of the defender mixed strategy, the attacker is therefore more likely to run away.

Our next result (Theorem 1) shows that the defender expected utility is non-increasing as detection uncertainty γ increases. As a byproduct of the proof for Theorem 1, we also show that the optimal solution may change as detection uncertainty changes. This illustrates the necessity of an algorithm for dealing with detection uncertainties.

Theorem 1. $\mathsf{DefEU}(\chi^*(\gamma), \gamma) \geq \mathsf{DefEU}(\chi^*(\gamma'), \gamma')$ *for any $\gamma' > \gamma$ in any problem instance.*

Proposition 2. *$\chi^*(\gamma)$ differs from $\chi^*(\gamma')$ for any $\gamma' > \gamma$ when x_t^{s-} is nonzero for $\chi^*(\gamma')$, where target t is the attacker best responding target in $\chi^*(\gamma')$.*

The intuition underlying the proof of Theorem 1 is that if we have a drone with a low false negative rate, then we can simulate a drone with a high false negative rate by ignoring some of its detections. The optimal solution for drones with a low false negative rate cannot be worse than that for drones with a high false negative rate.

We now show several results for observational uncertainty. First, we show that the loss due to observational uncertainty can be arbitrarily large.

Proposition 3. *There exists Π such that the loss due to ignoring observational uncertainty is arbitrarily large. In other words,* $\mathsf{DefEU}(\chi^*(\gamma_0, \Pi), \gamma_0, \Pi) - \mathsf{DefEU}(\chi^*(\gamma_0, \Pi_0), \gamma_0, \Pi) > M, \forall M > 0$.

The original signaling strategy tries to ensure the attacker only attacks when he observes the weak signal, σ_0, or nothing, n. However, with observational uncertainty, this may not be true because the true signal may be σ_1, but the attacker may have observed it mistakenly as σ_0. Therefore, we need

to enforce different attacker behaviors in order to obtain a better solution quality.

Now, we examine the attacker's behavior given a fixed deployment χ as observational uncertainty changes. Let (t, η) represent an attacker strategy of attacking target t and behaving according to η. Theorems 2 and 3 show that if we do not consider observational uncertainty, then the attacker behavior is more likely to converge to always attacking ($\eta = \mathbf{1}$) as observational uncertainty increases, where higher observational uncertainty means the attacker cannot distinguish between signaling states. Theorems 2 and 3 show that a deployment χ that does not consider observational uncertainty is more likely to result in this worst-case behavior of $\eta = \mathbf{1}$.

Theorem 2. *For any fixed deployment χ, if the attacker's best response is $(t, \mathbf{0})$ or $(t, \mathbf{1})$ at the Stackelberg equilibrium with Π_0, then it stays as an equilibrium for any Π'.*

Note that $\eta = \mathbf{0}$ and $\eta = \mathbf{1}$ result in an action that is independent of the attacker's observation. Thus, no matter what the attacker observes, the attacker can obtain the same utility with $\eta = \mathbf{0}$ or $\eta = \mathbf{1}$. It's only left to show that the attacker cannot get strictly better utility in Π' with a different attacker behavior. Intuitively, Π_0 implies a perfect observation, thus the attacker cannot get better utility than the perfect observation. So, if $(t, \mathbf{1})$ or $(t, \mathbf{0})$ is a Stackelberg equilibrium, the defender can safely deploy the same strategy for any uncertainty matrix Π', without any loss in her expected utility.

Even if $(t, \mathbf{0})$ or $(t, \mathbf{1})$ is not a best response with Π_0, $(t, \mathbf{1})$ may still be a best response at high levels of uncertainty. First, we say a target t is a *weak-signal-attack target* if $\mathsf{AttEU}(\sigma_0) \geq 0$ at t. Note that if $\mathsf{AttEU}(\sigma_0) \geq 0$, then the attacker will either always attack at $\hat{\omega} = \sigma_0$, or is indifferent between attacking and running away. We say χ is a *weak-signal-attack deployment* if all targets are weak-signal-attack targets.

Theorem 3. *If $(t, \mathbf{1})$ is a best response for $\Pi_{\kappa\lambda\mu}$ and χ is a weak-signal-attack deployment, then $(t, \mathbf{1})$ is a best response for $\Pi_{\kappa'\lambda'\mu'}$ and χ for all $\kappa' \geq \kappa$, $\lambda' \geq \lambda$, $\mu' \geq \mu$.*

In our model of observational uncertainty, more uncertainty means that the attacker sees a weak signal more often. Further, the attacker always attacks when he observes a weak signal. Thus, if the attacker is always attacking with less uncertainty, he will only attack more often with more uncertainty. However, in order to obtain predictable attacker behavior, we need to show that a weak-signal-attack deployment always exists as an optimal solution. In other words, Theorem 3 holds if there is weak-signal-attack deployment, so we now have to show that such a deployment exists.

Proposition 4. *There always exists an optimal solution that is a weak-signal-attack deployment with Π_0.*

The intuition behind the proof is that we can always decrease the probability of a weak signal such that we either do not send a weak signal, or the attacker attacks when he observes a weak signal. This holds optimally because when observational uncertainty is Π_0, signals are interchangeable. To summarize, if the attacker behavior is $\mathbf{0}$ or $\mathbf{1}$, then the attacker behavior is independent of observational uncertainty. We may see this behavior emerge as uncertainty increases.

5 How to Handle Uncertainty

We provide a solution approach based on the well-known multiple LPs approach from (Conitzer and Sandholm 2006). In particular, for each target $t \in [N]$, we compute the optimal defender strategy given that the attacker's best response is t. Then, the optimal defender strategy is the mixed strategy that leads to the maximum defender expected utility among all $t \in [N]$. The problem is NP-hard without uncertainty (Xu et al. 2018), thus our ultimate goal is to develop an efficient algorithm to solve the problem. For expository purposes, we first focus on presenting the LP for detection uncertainty.

5.1 Detection Uncertainty

Using notation from Section 3.3, we first formulate each player's utility function by breaking it into three parts according to signaling states: 1) no sensor is allocated (states $\mathrm{n}(+/-)$ and p, which we denote by $-s$); 2) sensor is allocated and sends σ_0; and 3) sensor is allocated and sends σ_1.

1. $U_{\text{-}s}^{d/a}(i) = x_i^{\mathrm{p}} \cdot U_+^{d/a}(i) + x_i^{\mathrm{n}+} \cdot U_+^{d/a}(i) + x_i^{\mathrm{n}-} \cdot U_-^{d/a}(i)$ is the expected defender/attacker utility of target i being attacked over states when i has no sensor ($\mathrm{p}, \mathrm{n}+, \mathrm{n}-$).
2. $U_{\sigma_0}^{d/a}(i) = (1-\gamma) \cdot [\psi_i^{\mathrm{s}+} \cdot U_+^{d/a}(i) + \psi_i^{\mathrm{s}-} \cdot U_+^{d/a}(i) + \psi_i^{\bar{\mathrm{s}}} \cdot U_-^{d/a}(i)] + \gamma \cdot [\varphi_i^{\mathrm{s}+} \cdot U_+^{d/a}(i) + \varphi_i^{\mathrm{s}-} \cdot U_-^{d/a}(i) + \varphi_i^{\bar{\mathrm{s}}} \cdot U_-^{d/a}(i)]$ is the defender/attacker expected utility when the attacker attacks target i and the defender signals σ_0.
3. $U_{\sigma_1}^{d/a}(i) = (1-\gamma) \cdot [(x_i^{\mathrm{s}+} - \psi_i^{\mathrm{s}+}) \cdot U_+^{d/a}(i) + (x_i^{\mathrm{s}-} - \psi_i^{\mathrm{s}-}) \cdot U_+^{d/a}(i) + (x_i^{\bar{\mathrm{s}}} - \psi_i^{\bar{\mathrm{s}}}) \cdot U_-^{d/a}(i)] + \gamma \cdot [(x_i^{\mathrm{s}+} - \varphi_i^{\mathrm{s}+}) \cdot U_+^{d/a}(i) + (x_i^{\mathrm{s}-} - \varphi_i^{\mathrm{s}-}) \cdot U_-^{d/a}(i) + (x_i^{\bar{\mathrm{s}}} - \varphi_i^{\bar{\mathrm{s}}}) \cdot U_-^{d/a}(i)]$

In words, 2) and 3) are the sum of expected utility on a detection and the sum of expected utility on no detection. In 3), in the no detection case, the defender exploits information asymmetry in signaling σ_1. In particular, the defender knows that there is no detection, but in sending σ_1 to indicate a detection, relies on the uncertainty the attacker faces in determining if there was a detection. We are now ready to describe an (exponentially-large) linear program (LP) formulation for computing the optimal defender strategy assuming best attacker response t (not (t, η) since only detection uncertainty):

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

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

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

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

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

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

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

$$0 \leq \psi_i^{\theta} \leq x_i^{\theta} \quad \forall\, \theta \in \{\bar{\mathrm{s}}, \mathrm{s}-, \mathrm{s}+\}, \forall\, i \in [N] \tag{8}$$

$$0 \leq \varphi_i^{\theta} \leq x_i^{\theta} \quad \forall\, \theta \in \{\bar{\mathrm{s}}, \mathrm{s}-, \mathrm{s}+\}, \forall\, i \in [N] \tag{9}$$

The objective function (1) maximizes defender expected utility. Since the attacker is running away when he observes

σ_1, $U^d_{\sigma_1} = 0$. Constraints (2)-(4) enforce that the randomized resource allocation is feasible ($\mathcal{E}$ has exponential number of elements); (5)-(6) guarantee that σ_1, σ_0 result in the attacker best responses of running away and attacking[3]; (7) ensures the attacker expected utility at target t is bigger than at any other target i, thus t is attacker's best response; (8)-(9) ensure a feasible signaling scheme.

5.2 Acceleration via Branch and Price

We now describe the branch-and-price solution framework, which can be used for both uncertainty scenarios. There are two main challenges in efficiently solving the LP (1)-(9). First, the total number of possible q_e is $O(6^N)$. Second, we will need to solve N LPs (for each $t \in [N]$). Solving many of these large LPs is a significant barrier for scaling up. We therefore introduce Games with Uncertainty And Response to Detection with Signaling Solver (GUARDSS), which employs the branch-and-price framework. This framework is well-known for solving large-scale optimization programs, but the main challenges of applying this framework are to (1) design the efficient subroutine called the *slave problem* for solving each LP, and to (2) carefully design an upper bound for pruning LPs.

First, for one LP w.r.t. a specific t, to address the issue of the exponential size of set $\mathcal{E}$, we adopt the column generation technique. 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 $\alpha^\theta_i \in \mathbb{R}$ for $\theta \in \Theta$, for each target i, solve the *weight maximization problem*:

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

Note that $\{\alpha^\theta_i\}_{\theta \in \Theta}$ are the optimal dual variables for the previous LP constraint (2). We want to solve this without enumerating all of the elements in $\mathcal{E}$. Despite the added complexity compared to classic SSGs, in this section, we compactly represent this slave problem as a mixed integer linear program (MILP). To formulate the MILP, we introduce six binary vectors $\mathbf{v}^{\mathrm{p}}$, $\mathbf{v}^{\mathrm{n}+}$, $\mathbf{v}^{\mathrm{n}-}$, $\mathbf{v}^{\bar{\mathrm{s}}}$, $\mathbf{v}^{\mathrm{s}-}$, $\mathbf{v}^{\mathrm{s}+} \in \{0,1\}^N$ to encode for each target whether it is in each allocation state. For example, target i is at allocation state $\bar{\mathrm{s}}$ if and only if $v^{\bar{\mathrm{s}}}_i = 1$. The main challenge then is to properly set up linear (in)equalities of these vectors to precisely capture their constraints and relations. The capacity for each resource type results in two constraints (number of patrollers and sensors):

$$\textstyle\sum_{i \in [N]} v^{\mathrm{p}}_i \le k \tag{11}$$

$$\textstyle\sum_{i \in [N]} (v^{\bar{\mathrm{s}}}_i + v^{\mathrm{s}-}_i + v^{\mathrm{s}+}_i) \le l \tag{12}$$

Moreover, each target must be at one of these states:

$$v^{\mathrm{p}}_i + v^{\mathrm{n}-}_i + v^{\mathrm{n}+}_i + v^{\bar{\mathrm{s}}}_i + v^{\mathrm{s}-}_i + v^{\mathrm{s}+}_i = 1 \quad \forall i \in [N] \tag{13}$$

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

Due to the reaction stage, we have to add constraints to specify (a) which targets have a patroller at a neighboring target; (b) which patroller goes to which nearby target if both sensors and patrollers do not detect the attacker. For (a), the non-zero entries of $A \cdot \mathbf{v}^{\mathrm{p}}$ specify the targets with a patroller nearby, where A is the adjacency matrix of the underlying graph. Since three vectors encode the states requiring a nearby patroller, we have this constraint:

$$A \cdot \mathbf{v}^{\mathrm{p}} \ge \mathbf{v}^{\mathrm{n}+} + \mathbf{v}^{\mathrm{s}-} + \mathbf{v}^{\mathrm{s}+} \tag{14}$$

We ensure that a vertex with a patroller nearby cannot be $\mathbf{v}^{\bar{\mathrm{s}}}$:

$$A \cdot \mathbf{v}^{\mathrm{p}} \le \mathbf{v}^{\mathrm{p}} + \mathbf{v}^{\mathrm{n}+} + \mathbf{v}^{\mathrm{n}-} + \mathbf{v}^{\mathrm{s}-} + \mathbf{v}^{\mathrm{s}+} \tag{15}$$

Constraint (b) means that patrollers must be "re-matched" to new vertices in the reaction stage. Specifically, targets in states $\mathrm{p}, \mathrm{n}+, \mathrm{s}+$ must form a matching. To enforce this constraint, let G' be the directed version of G, i.e. for all $(i,j) \in E$ we have $(i,j), (j,i) \in E'$. *We further introduce edge variables $y_{(i,j)} \in \{0,1\}$ indicating whether the directed edge (i,j) is in the matching or not.* The matching constraint can be expressed by the following linear constraints:

$$\textstyle\sum_{(i,j) \in E' : j \in [N]} y_{(i,j)} = v^{\mathrm{p}}_i \qquad \forall i \in [N] \tag{16}$$

$$v^{\mathrm{n}+}_j + v^{\mathrm{s}+}_j \ge y_{(i,j)} \qquad \forall (i,j) \in E' \tag{17}$$

The resulting MILP for the slave problem is as follows.

$$\begin{aligned} \max_{\mathbf{v}, y} \quad & \textstyle\sum_\theta \sum_i v^\theta_i \alpha^\theta_i \\ \text{s.t.} \quad & (11) - (17) && (18)\\ & \mathbf{v}^\theta \in \{0,1\}^N && \forall \theta \in \Theta \quad (19)\\ & y_{(i,j)} \in \{0,1\} && \forall (i,j) \in E' \quad (20) \end{aligned}$$

Second, to avoid solving LPs for all different targets $t \in [N]$, we use the branch and bound technique which finds an upper bound for each LP for pruning. The natural approach for finding an upper bound is to solve a relaxed LP corresponding to the original LP — in our case, essentially relax the original LP into its marginal space. As the set $\mathcal{E}$ is exponentially large, we relax variables and constraints corresponding to $\mathcal{E}$ in our LP. Concretely, we relax (2) - (4) into a polynomial number of variables and constraints. These variables and constraints are (18) - (20) with $\mathbf{v}^\theta$ replaced by $\mathbf{x}^\theta$. We first use the relaxed LP to efficiently compute an upper bound for each LP. After solving each relaxed LP exactly, we solve original LPs chosen according to some heuristic order (typically the descending order of the relaxed optimal objective) using the column generation techniques, and we can safely prune out those LPs whose optimal relaxed value is less than the current largest achievable objective value. This process continues until no LP is left to solve, in which case the current largest objective value is optimal.

5.3 Detection and Observational Uncertainty

Finally, we briefly discuss the case with both uncertainties, as it can be solved in a similar way. Constraints (2)-(4) and (8)-(9) are the same. However, the remaining constraints must now account for attacker behavior, η. For example, the

utility functions $U^{d/a}_{\sigma_0}$ and $U^{d/a}_{\sigma_1}$ must change to incorporate attacker behaviors, and the objective function becomes that in (21) since the attacker may not run away when he observes σ_1 in the presence of observational uncertainty. Also, we add a constraint to ensure the attacker utilities are aligned with the attacker behavior $\eta \in \{0,1\}^3$. These are primarily notational changes. We therefore provide the full LP for this case in the Appendix.

$$\max_{x,q,\psi,\varphi} \quad U^d_{\text{-}s}(t) + U^d_{\sigma_1}(t) + U^d_{\sigma_0}(t) \tag{21}$$

6 Experiments

We generate random Watts-Strogatz graphs, which have small-world properties to describe more complex environments, such as roads connecting far-away nodes. For all tests, we average over 20 random graphs and include p-values. Utilities are randomly generated with a maximum absolute value of 1090 and based on the idea that the losses from undetected attacks are higher than the utility of catching adversaries (similar to (Xu et al. 2015)). This is realistic to the situation of preventing poaching, as animals are worth more for ecotourism than for sale on the black market as discussed in the Appendix. Additionally, we see that if we test on a set of utilities that is slightly different from the original input, the defender's utility does not vary greatly. Fig. 3a-3b show timing tests run on a cluster with Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.1 GHz with at most 16 GB RAM. We set the number of patrollers to be $k = \sqrt{N/2}$ and the number of drones to be $l = 2N/3 - k$. As shown, the full LP scales up to graphs of $N = 14$ only and exceeds the cutoff time limit of 3600s for all $N = 16$ graphs. Branch and price scales up to $N = 80$ and runs out of time for larger games, and a warm-up enhancement that greedily select an initial set of $\mathcal{E}$ further improves scalability and solves 13/40 graphs within cutoff time at $N = 90$ and $N = 100$. This is sufficient for middle-scale real-world problems, with further scalability being an interesting direction for future work. The heuristics provide the same solution as the full LP in most of the instances tested.

Next, we show the loss due to ignoring uncertainty empirically. In Figs. 3c-3d we compare $\mathsf{DefEU}(\chi^*(\gamma,\Pi),\gamma,\Pi)$ computed by GUARDSS and $\mathsf{DefEU}(\chi^*(0,\Pi_0),\gamma,\Pi)$, the defender expected utility when ignoring uncertainty for graphs with $N = 10$, $k = 1$, $l = 3$. We consider only one type of uncertainty at a time (e.g., $\gamma = 0$ when varying observational uncertainty). For detection uncertainty, GUARDSS's defender expected utility only decreases by 12%, whereas ignoring uncertainty decreases by 210% when γ varies from 0 to 0.9 ($p \leq 1.421\text{e}{-03}$ for $\gamma \geq 0.2$ in Fig. 3c)[4]. Some initial analysis shows that it is robust in most of the cases when we slightly under- or overestimate γ (e.g., the differences in defender expected utility are typically within 5-6% when the estimate of gamma is off by 0.1 or 0.2), but further investigation on dealing with such uncertainty over uncertainty would be an interesting direction for future work. For observational uncertainty, GUARDSS's defender expected utility only decreases by 1%, whereas ignoring uncertainty decreases by 18% as the observational uncertainty, parameterized by κ ($\lambda = \frac{\kappa}{2}$, and $\mu = \frac{\kappa}{2}$) varies from 0 to 0.9 ($p \leq 0.058$ for $\kappa \geq 0.4$ in Fig. 3d).

We also observe that when ignoring detection uncertainty, the attacker's best response is typically a target with a sensor, which implies that the attacker is taking advantage of the defender's ignorance of uncertainty. In fact, there is a statistically significant ($p = 1.52\text{e}{-08}$) difference in the mean probability of a sensor at the attacker's best response when ignoring uncertainty (0.68) versus GUARDSS (0.19).

How does the defender avoid these challenges and achieve such a small performance drop with GUARDSS when facing uncertainty? Statistics of the resulting defender strategy as well as Fig. 3e indicate that the defender *exploits the uncertain real-time information and the information asymmetry*, including (a) frequently but not always sending patrollers to check important targets when there is no detection; (b) sending strong signals more frequently than the probability that the patroller will visit the target (either due to response to detection or planned reallocation in the case of no detection), leveraging the informational advantage in which the attacker does not know whether he is detected or whether a patroller is matched; (c) using different signaling schemes with and without detection, leveraging the information advantage that the attacker does not know whether he is detected. In the GUARDSS strategies in Figs. 3c-3d, the mean probability of the attacker's best response target being at state s$-$ (with sensor but without a matched patroller) is 0.04, versus 0.43 when ignoring uncertainty ($p = 2.70\text{e}{-09}$), indicating point (a). If we call the strong signal sent when there is no detection a *fake signal*, Fig. 3e shows that the probability of the strong signal an attacker observes is a fake signal is nonzero and increases in a non-linear fashion, indicating points (b) and (c). Also, note that the strong signal is used with nonzero probability on average on targets with a nonzero probability of having a drone present.

Despite considering uncertainty, sensors may be less valuable at a high level of uncertainty. In Fig. 3f, the defender expected utility is influenced by the number of drones and uncertainty in size $N = 15$ graphs. In Fig. 3g, drones are better than an extra patroller at $\gamma = 0.3$ ($p \leq 6.661\text{e}{-02}$), but at $\gamma = 0.8$, patrollers are better than drones ($p \leq 1.727\text{e}{-07}$).

7 Conservation Drones

We have deployed a drone in South Africa, equipped with a thermal camera and detection system (Air Shepherd 2019). 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 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

[4] % change once normalized by largest defender/attacker utility.

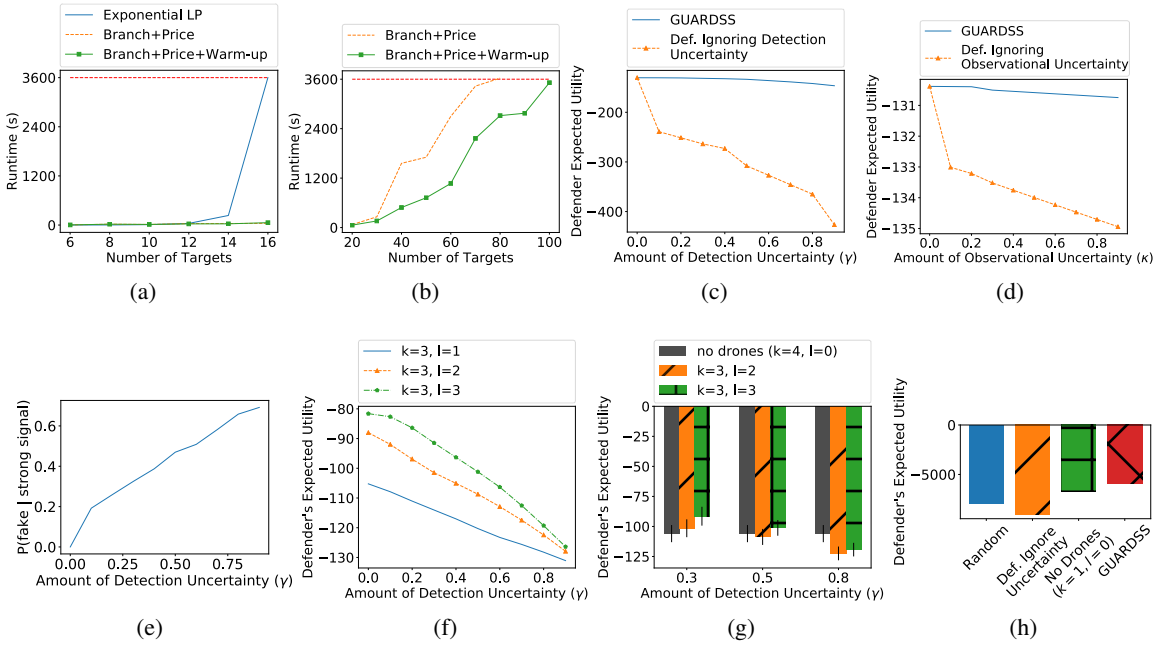


Figure 3: Experimental results. Figs. 3a-3b compare multiple LPs approach (Exponential LP) with GUARDSS branch-and-price and heuristic method. Figs. 3c-3d show defender expected utility when amount of detection uncertainty γ and observational uncertainty vary. Defender expected utility decreases much more when uncertainties are ignored. Fig. 3e shows the informational advantage of the defender as uncertainty increases. Figs. 3f-3g show that in the presence of a high false negative rate, extra patrollers may be more useful than drones. Fig. 3h contains the results from the case study, where GUARDSS performs best.

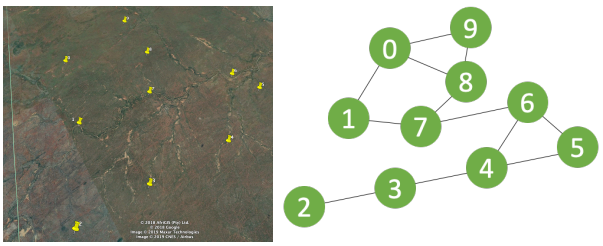


Figure 4: A park in Google Maps with potential poaching hotspots and the resulting graph (edges for < 5 km).

in (Gholami et al. 2018). The targets are shown in Fig. 4 (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. 4 (right). The utilities are included in the Appendix along with further details. 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 (see the Appendix for results with other γ). These details are directly input to GUARDSS, and then a mixed strategy is determined to cover the park. Fig. 3h 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.* For varying γ (see Appendix), the gap between ignoring detection uncertainty and GUARDSS increases as γ increases, and the gap between the no drones case and GUARDSS decreases as γ increases, showing a similar trend to Fig. 3g. However, in all cases, the results emphasize the importance of correctly optimizing to get value from drones even with uncertainty.

8 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. In this case, despite being aware of uncertainty, the attacker does not know whether he was detected, nor whether a patroller will respond in the reaction stage. Our results illustrate that the defender can exploit this informational advantage even with uncertain information. Thriving under this uncertainty makes real-world deployment of GUARDSS promising, as shown through simulation.

9 Acknowledgements

This was supported by Microsoft AI for Earth, NSF CCF-1522054 and IIS-1850477, and MURI W911NF-17-1-0370. We would also like to thank Air Shepherd for their valuable insights and collaboration.

References

Air Shepherd. 2019. http://airshepherd.org.

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.

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*.

Bondi, E.; Jain, R.; Aggrawal, P.; Anand, S.; Hannaford, R.; Kapoor, A.; Piavis, J.; Shah, S.; Joppa, L.; Dilkina, B.; and Tambe, M. 2020. Birdsai: A dataset for detection and tracking in aerial thermal infrared videos. In *WACV*.

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 *EC*.

Cooney, S.; Wang, K.; Bondi, E.; Nguyen, T.; Vayanos, P.; Winetrobe, H.; Cranford, E. A.; Gonzalez, C.; Lebiere, C.; and Tambe, M. 2019. Learning to signal in the goldilocks zone: Improving adversary compliance in security games. In *ECML PKDD (Research Track)*.

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*.

Jain, M.; Kardes, E.; Kiekintveld, C.; Ordónez, F.; and Tambe, M. 2010. Security games with arbitrary schedules: A branch and price approach. In *AAAI*.

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. *EC*.

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.

Thakoor, O.; Tambe, M.; Vayanos, P.; Xu, H.; Kiekintveld, C.; and Fang, F. 2019. Cyber camouflage games for strategic deception. In *GameSec*, 525–541. Springer.

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.; Rabinovich, Z.; Dughmi, S.; and Tambe, M. 2015. Exploring Information Asymmetry in Two-Stage Security Games. In *AAAI*, 1057–1063.

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*.

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*.

A EFG and POMDP

We first discuss in more detail how we exploit the structure of our game and provide a scalable algorithm by extending the use of coverage probabilities and multiple LPs, instead of using an EFG. The game tree is shown in Fig. 1. Our approach solves at most $8N$ LPs with $O(N + |\mathcal{E}|)$ constraints and $O(N+|\mathcal{E}|)$ variables, where $\mathcal{E}$ is the defender pure strategy set. Using the EFG approach, the size of the game tree is $O(N \cdot 4^l \cdot |\mathcal{E}|)$, where l is the number of drones. The EFG multiple LPs approach therefore solves exponentially more LPs, each with a much larger size than ours. One might also consider using a POMDP to model the movement of the defender from allocation to a new reaction target with the unobservable state being whether an attacker is present or not. However, a POMDP model does not capture all of the intricacies due to strategic game interactions. For example, it does not account for the fact that the attacker will choose a location to attack rationally.

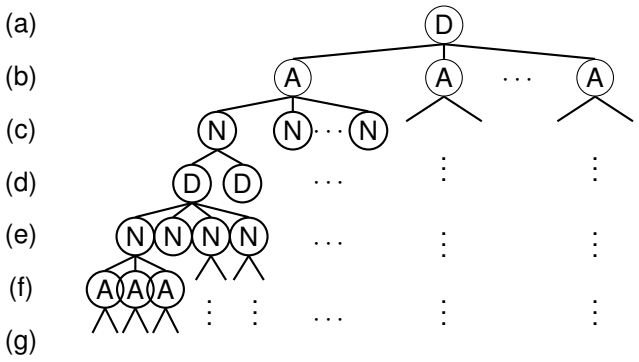


Figure 1: Game tree illustrating defenders' (D) allocation, signaling, and reaction steps, as well as those of the attacker (A) and nature (N) (uncertainty). Level (a) is the initial defender deployment, (b) is the attacker choice of target, (c) is the detection (with uncertainty), (d) is the defender signaling, (e) is the attacker's observation (with uncertainty), (f) is the attacker's decision to run away or not, (g) is when players receive payoffs.

B Omitted Proofs in Section 4

Proposition 1. *Let $\chi_0^* = \chi^*(0)$ be the defender optimal deployment when no uncertainties exist. There exist instances where $\mathsf{DefEU}(\chi_0^*, \gamma) < \mathsf{DefEU}(\chi^*(\gamma), \gamma)$ for some γ.*

Proof. We prove by constructing such an instance. Consider the graph in Fig. 2 with 4 targets, 1 human patroller, and 2 sensors. The attacker chooses one target to attack. A successful attack gives the attacker utility of $+2$ and defender utility of -5, whereas catching the attacker yields attacker utility -1 and defender utility 0. If the attacker chooses not to attack after observing a signal from the sensor, both the attacker and the defender receive 0.

If sensors have perfect detection ($\gamma = 0$), we can place the human patroller at t_2, place the two sensors at t_1 and t_3 respectively, and match the patroller to t_4. Thus, we cover all targets with probability 1. When we observe the attacker, we will always send the strong signal (σ_1) to ensure the attacker will run away. Therefore, the attacker is better off not attacking, yielding utility of 0 for both.

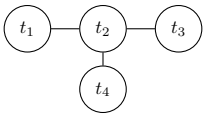


Figure 2: Diagram of detection uncertainty example.

To see how a false negative detection can affect the solution quality, we now consider the case with $\gamma = 0.5$. If we use the same strategy but with imperfect detection, the attacker can attack t_1 or t_3 successfully with probability $1/2$ (when the attacker does not observe a warning signal σ_1), and run away with probability $1/2$ (when the attacker observes the signal σ_1). The defender's expected utility is $\mathsf{DefEU}(\cdot) = \frac{1}{2}(0) + \frac{1}{2}(-5) = -\frac{5}{2} = -2.5$.

We want to show the optimal strategy when $\gamma = 0.5$. First, we will always have a patroller at t_2. We will always have one drone at t_1 and we will have another drone at t_3 with probability $6/9$ and at t_4 with probability $3/9$. We will match the patroller to t_1 with probability $2/9$; t_3 with probability $3/9$ ($1/9$ when there's sensor, $2/9$ when there's nothing); t_4 with probability 4/9 when there's nothing. We will first calculate defender expected utility at t_1, t_2 and t_4. At target t_1, the attacker always observe a sensor, with probability $7/9$ it's state s$-$ and with probability $2/9$ it's state s$+$. Thus $\mathsf{AttEU}(t_1) = 2/9 \cdot (-1) + 7/9 \cdot (1-\gamma)(-1) + 7/9 \cdot \gamma \cdot (2) = 1/6$, thus the attacker will attack. For target t_2 the attacker will always get caught, so he won't attack. For target t_4, if he observes nothing, he will get caught with probability $\frac{4/9}{4/9+2/9} = 2/3$, so he will not attack when he observes nothing. If he observes a drone, he will only get caught with probability $(1-\gamma)$, thus gains utility of $\gamma \cdot 1/3 = 1/6$.

Consider the following signaling scheme at t_3. If we detect an attacker or the state is matched, then we will send σ_0, if we don't detect and state is not matched we will send signal with marginal probability $\varphi^{s-} = 7/18$. If the attacker does not observe a drone, he will not attack. If he observes a drone with σ_0, then the attacker expected utility is $((1-\gamma) \cdot 2/3) \cdot (-1) + \gamma \cdot 1/9 \cdot (-1) + \gamma \cdot 7/18 \cdot (2) = 0$. Thus the attacker only attacks when he observes σ_1 this happens with probability $1/12$, thus the attacker gets expected utility of $2 \cdot 1/12 = 1/6$. Since we are getting attacked with probability $1/12$, the defender expected utility is now $-5 \cdot 1/12 = -0.416$ thus doing getting better expected utility by considering uncertainty.

With optimal deployment, we can get a defender expected utility of -0.416 when $\gamma = 0.5$. This example shows that when the detection uncertainty does not exist, a very simple deployment yields the optimal expected utility. However, this strategy is no longer optimal when detection uncertainty is present. Therefore, we need to consider detection uncertainty and compute the new optimal solution.

□

Theorem 1. $\mathsf{DefEU}(\chi^*(\gamma),\gamma) \geq \mathsf{DefEU}(\chi^*(\gamma'),\gamma')$ *for any* $\gamma' > \gamma$ *in any problem instance.*

Proof. Let $\chi^*_\gamma = \chi^*(\gamma)$. Throughout the proof, we assume no observational uncertainty. Assume for the contradiction that defender expected utility strictly increases as detection uncertainty increases, i.e., $\mathsf{DefEU}(\chi^*_\gamma,\gamma) < \mathsf{DefEU}(\chi^*_{\gamma+\epsilon},\gamma+\epsilon)$ for some ϵ. Let ψ and φ be the corresponding signaling variables from $\chi^*_{\gamma+\epsilon}$.

Consider the following new variables χ' and $\psi' = \frac{(1-\gamma-\epsilon)\psi+\epsilon\varphi}{1-\gamma}$, and let all other variables be the same as $\chi^*_{\gamma+\epsilon}$. First, note that $\psi'^\theta = \frac{(1-\gamma-\epsilon)\psi^\theta+\epsilon\varphi^\theta}{1-\gamma} \leq \frac{(1-\gamma-\epsilon)x^\theta+\epsilon x^\theta}{1-\gamma} = x^\theta$, for all $\theta \in \{\mathrm{s}+,\mathrm{s}-,\bar{\mathrm{s}}\}$; therefore, all variables are feasible. Thus, we have $\mathsf{DefEU}(\chi',\gamma) \leq \mathsf{DefEU}(\chi^*_\gamma,\gamma) < \mathsf{DefEU}(\chi^*_{\gamma+\epsilon},\gamma+\epsilon)$. Furthermore, consider an augmented strategy where when we observe an attacker at state $\psi'^{\mathrm{s}-}$, with marginal probability $\varphi^{\mathrm{s}-}$, we ignore the detection, thus make the target uncovered. Note that this strategy is still feasible, and makes our defender expected utility lower.

Now, we will calculate the defender expected utility when the defender allocates security resources according to χ'. We can decompose the expected utility by different signals, i.e. $\mathsf{DefEU}(\chi',\gamma) = \sum_{\omega\in\{\mathrm{n},\sigma_0,\sigma_1\}} \mathsf{DefEU}(\chi',\gamma|\omega)$, where $\mathsf{DefEU}(\chi',\gamma|\omega)$ is the defender expected utility given state ω. First, note that $\mathsf{DefEU}(\chi',\gamma|\mathrm{n})$ stays the same as detection uncertainty changes. Thus, we will only look at $\mathsf{DefEU}(\chi',\gamma|\sigma_0)$ and $\mathsf{DefEU}(\chi',\gamma|\sigma_1)$.

$$
\begin{aligned}
&\mathsf{DefEU}(\chi',\gamma|\sigma_0)\\
&\quad= (1-\gamma)\cdot(\psi'^{\mathrm{s}+}U^d_+ + \psi'^{\mathrm{s}-}U^d_+ + \psi'^{\bar{\mathrm{s}}}U^d_-)\\
&\qquad+\gamma\cdot(\varphi^{\mathrm{s}+}U^d_+ + \varphi^{\mathrm{s}-}U^d_- + \varphi^{\bar{\mathrm{s}}}U^d_-)\\
&\quad\geq (1-\gamma-\epsilon)\cdot(\psi^{\mathrm{s}+}U^d_+ + \psi^{\mathrm{s}-}U^d_+ + \psi^{\bar{\mathrm{s}}}U^d_-)\\
&\qquad+\epsilon(\varphi^{\mathrm{s}+}U^d_+ + \varphi^{\mathrm{s}-}U^d_- + \varphi^{\bar{\mathrm{s}}}U^d_-)\\
&\qquad\qquad\text{(by our augmented strategy)}\\
&\qquad+\gamma\cdot(\varphi^{\mathrm{s}+}U^d_+ + \varphi^{\mathrm{s}-}U^d_- + \varphi^{\bar{\mathrm{s}}}U^d_-)\\
&\quad= (1-\gamma-\epsilon)\cdot(\psi^{\mathrm{s}+}U^d_+ + \psi^{\mathrm{s}-}U^d_+ + \psi^{\bar{\mathrm{s}}}U^d_-)\\
&\qquad+(\gamma+\epsilon)\cdot(\varphi^{\mathrm{s}+}U^d_+ + \varphi^{\mathrm{s}-}U^d_- + \varphi^{\bar{\mathrm{s}}}U^d_-)\\
&\quad= \mathsf{DefEU}(\chi^*_{\gamma+\epsilon},\gamma+\epsilon|\sigma_0)
\end{aligned}
$$

We also want to show that $\mathsf{DefEU}(\chi',\gamma|\sigma_1) = \mathsf{DefEU}(\chi^*_{\gamma+\epsilon},\gamma+\epsilon|\sigma_1)$. Recall that they are both 0 because the attacker will run away. Then, we have $\mathsf{DefEU}(\chi',\gamma) \geq \mathsf{DefEU}(\chi^*_{\gamma+\epsilon},\gamma+\epsilon) > \mathsf{DefEU}(\chi^*_\gamma,\gamma)$. This contradicts χ^*_γ is an optimal solution.

□

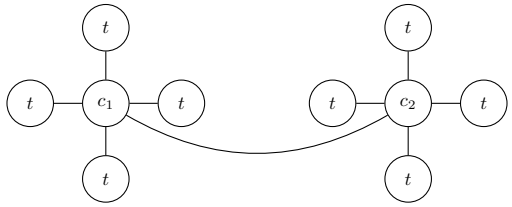
Proposition 2. $\chi^*(\gamma)$ *differs from* $\chi^*(\gamma')$ *for any* $\gamma' > \gamma$ *when* x^{s-}_t *is nonzero for* $\chi^*(\gamma')$*, where target* t *is the attacker best responding target in* $\chi^*(\gamma')$*.*

Proof. Suppose for contradiction that χ is an optimal solution for both $\mathsf{DefEU}(\chi,\gamma)$ and $\mathsf{DefEU}(\chi,\gamma+\epsilon)$. Consider χ' that is obtained same way as previous proof. i.e. $\psi' = \frac{(1-\gamma-\epsilon)\psi+\epsilon\varphi}{1-\gamma}$ and all other variables stays the same.

Note $\mathsf{DefEU}(\chi',\gamma)$ is strictly bigger than $\mathsf{DefEU}(\chi,\gamma+\epsilon)$ if $\psi^{\mathrm{s}-}$, $x^{\mathrm{s}-}-\psi^{\mathrm{s}-}$, $\varphi^{\mathrm{s}-}$ or $x^{\mathrm{s}-}-\varphi^{\mathrm{s}-}$ is non-zero. In other words, if $x^{\mathrm{s}-}$ is nonzero, the $\mathsf{DefEU}(\chi',\gamma) > \mathsf{DefEU}(\chi,\gamma+\epsilon)$, thus contradicts $\mathsf{DefEU}(\chi,\gamma)$ is an optimal solution. □

Proposition 3. *There exists* Π *such that the loss due to ignoring observational uncertainty is arbitrarily large. In other words,* $\mathsf{DefEU}(\chi^*(\gamma_0,\Pi),\gamma_0,\Pi)$ - $\mathsf{DefEU}(\chi^*(\gamma_0,\Pi_0),\gamma_0,\Pi) > M, \forall M > 0$.

Proof. We will show an example where the new signaling strategy is arbitrarily better than the naive signaling strategy. Consider the following example. We have 10 targets, 1 human patroller, and 8 sensors.



The optimal allocation strategy is to allocate sensors in all t's, allocate the human patroller in one of the center vertices (c_1,c_2) uniformly at random, and match to the other center vertex. For all targets, utility is defined as the following, for some arbitrarily big M. Note that the attacker's expected utilities for attacking c_1 and c_2 are both 0.

	covered	uncovered
Attacker	-1	$1+\epsilon$
Defender	0	$-M$

Consider the following uncertainty matrix. Let $r_0 = \frac{1-\epsilon}{1+\epsilon}\cdot(1-2\epsilon')$, for some $\epsilon' > 0$.

$\Pr[\hat{\omega}\|\omega]$	$\omega=\mathrm{n}$	$\omega=\sigma_0$	$\omega=\sigma_1$
$\hat{\omega}=\mathrm{n}$	1	$1-r_0$	ϵ'
$\hat{\omega}=\sigma_0$	0	r_0	$1-2\epsilon'$
$\hat{\omega}=\sigma_1$	0	0	ϵ'

Let r_{p} be the probability of state p, let r_{n} be the probability of state n+ and n−, and let $r_{\mathrm{s}+}$ and $r_{\bar{\mathrm{s}}}$ be the probability of the state s+ and s− and $\bar{\mathrm{s}}$, respectively. Let $r_o = \frac{1}{1+\epsilon}$ (the optimal signaling strategy ignoring uncertainty). In this strategy, the attacker attacks when he observes state n and σ_0. Therefore, $\bar{\eta} = [1,1,0]$ is the vector that depicts attacker behavior in this case.

$$
\begin{aligned}
&\mathsf{DefEU}(r_o)\\
&\quad= U^d_+ r_{\mathrm{p}} + U^d_- r_{\mathrm{n}} + U^d_- r_{\bar{\mathrm{s}}}(1-r_n)\bar{\eta}\cdot\Pr[\hat{\omega}|\omega=\sigma_0]\\
&\quad+ U^d_- r_{\bar{\mathrm{s}}}(r_n)\bar{\eta}\cdot\Pr[\hat{\omega}|\omega=\sigma_1] + U^d_+ r_{\mathrm{s}+}\bar{\eta}\cdot\Pr[\hat{\omega}|\omega=\sigma_1]\\
&\quad= -\frac{1}{2} - \frac{M-1}{2}(1-\epsilon') \leq -\frac{M-2}{2}(1-\epsilon')
\end{aligned}
$$

Let $r_n = \epsilon'$ be the new signaling strategy. Let η' be the new attacker's attacking vector. Observe the attacker will only attack when he observes state n. Therefore, $\eta' = [1, 0, 0]$.

$$\begin{aligned}
&\mathsf{DefEU}(r_n) \\
&\quad = U^d_+ r_{\mathrm{p}} + U^d_- r_{\mathrm{n}} + U^d_- r_{\bar{\mathrm{s}}}(1 - r_n)\eta' \cdot \Pr[\hat{\omega}|\omega = \sigma_0] \\
&\quad + U^d_- r_{\bar{\mathrm{s}}}(r_n)\eta' \cdot \Pr[\hat{\omega}|\omega = \sigma_1] + U^d_+ r_{\mathrm{s}+}\eta' \cdot \Pr[\hat{\omega}|\omega = \sigma_1] \\
&\quad = -\frac{M}{2}(1 - \epsilon')(1 - \frac{M-2}{M}(1 - 2\epsilon')) - \frac{M}{2}\epsilon'\epsilon' \\
&\quad \geq -(1 - \epsilon') - \frac{M}{2}\epsilon'\epsilon'
\end{aligned}$$

For $0 < \epsilon' < 1/M$, we get the gap of $O(M)$.

□

Theorem 2. *For any fixed deployment χ, if the attacker's best response is $(t, \mathbf{0})$ or $(t, \mathbf{1})$ at the Stackelberg equilibrium with Π_0, then it stays as an equilibrium for any Π'.*

Proof. The proof of Theorem 2 follows from this Lemma:

Lemma 1. *For any fixed χ, if $(t, \mathbf{1})$ (or $(t, \mathbf{0})$) is a best response for the attacker at Π_0, then $(t, \mathbf{1})$ (or $(t, \mathbf{0})$) is also a best response for all Π', for any $\Pi' \neq \Pi_0$.*

Proof. The proof of the Lemma follows from the following two claims. Let $\mathsf{DefEU}(\chi, t, \eta, \Pi)$ be the defender's expected utility when she plays the deployment χ. We add t and η to the typical notation to represent that the attacker's strategy is to attack t with behavior η, and the observational uncertainty matrix is Π. There is no detection uncertainty. We use a similar notation for $\mathsf{AttEU}(\chi, t, \eta, \Pi)$. Let us also index η with $\hat{\omega} \in \Omega$ as $\eta^{\hat{\omega}}$, and reference η for a target, i, as η_i, for a final notation of $\eta_i^{\hat{\omega}}$. For example, to reference the attacker behavior for $\hat{\omega} = \mathrm{n}$ and target i, we can write it as η_i^{n}.

Claim 1. *For any χ, t, η, $\mathsf{AttEU}(\chi, t, \eta, \Pi') \leq \mathsf{AttEU}(\chi, t, \eta^*_{0,t}, \Pi_0)$, where $\eta^*_{0,t}$ is the best response when $\Pi = \Pi_0$.*

Proof. Let $\mathsf{AttEU}(\omega) = \Pr[\omega]\mathsf{AttEU}(\chi, t, \mathbf{1})$ be the attacker's expected utility when true signaling state is ω and the attacker attacks the target t.

$$\begin{aligned}
&\mathsf{AttEU}(\chi, t, \eta, \Pi') = \\
&\qquad \sum_{\hat{\omega} \in \Omega} \eta^{\hat{\omega}} \cdot \left(\sum_{\omega \in \Omega} \Pr[\hat{\omega}|\omega] \mathsf{AttEU}(\omega) \right) \\
&\quad = \sum_{\omega \in \Omega} \mathsf{AttEU}(\omega) \left(\sum_{\hat{\omega} \in \Omega} \eta^{\hat{\omega}} \Pr[\hat{\omega}|\omega] \right) \\
&\quad \leq \sum_{\omega \in \Omega} \mathsf{AttEU}(\omega) \cdot \mathbb{1}\left(\mathsf{AttEU}(\omega) \geq 0\right) \\
&\quad \leq \mathsf{AttEU}(\chi, t, \eta^*_{0,t}, \Pi_0) \\
&\quad \text{(Note } \Pi_0 \text{ means } \Pr[\hat{\omega}|\omega] = 1 \text{ for all } \hat{\omega} = \omega)
\end{aligned}$$

□

Where $\mathbb{1}(\cdot)$ is an indicator function, $\mathbb{1}(\cdot) = 1$ if the corresponding expression is true, and $\mathbb{1}(\cdot) = 0$ otherwise. Note that $\mathsf{AttEU}(\chi, t, \eta, \Pi') \leq \mathsf{AttEU}(\chi, t, \eta, \Pi_0)$ is not true. Consider a Π_i which is some permutation matrix of $\mathbf{I}$.

Claim 2. *For any χ and t, the attacker's expected attacker utility stays the same for any Π if the attacker behavior is $\mathbf{1}$ or $\mathbf{0}$. In other words, $\mathsf{AttEU}(\chi, t, \mathbf{1}, \Pi') = \mathsf{AttEU}(\chi, t, \mathbf{1}, \Pi_0)$ and $\mathsf{AttEU}(\chi, t, \mathbf{0}, \Pi') = \mathsf{AttEU}(\chi, t, \mathbf{0}, \Pi_0)$.*

Corollary 1. *We also have $\mathsf{DefEU}(\chi, t, \mathbf{1}, \Pi') = \mathsf{DefEU}(\chi, t, \mathbf{1}, \Pi_0)$ and $\mathsf{DefEU}(\chi, t, \mathbf{0}, \Pi') = \mathsf{DefEU}(\chi, t, \mathbf{0}, \Pi_0)$.*

Proof. If $\eta = \mathbf{1}$ then $\sum_{\hat{\omega} \in \Omega} \eta^{\hat{\omega}} \Pr[\hat{\omega}|\omega] = 1$ for all $\omega \in \Omega$ independent of Π. Therefore, we get $\mathsf{AttEU}(\chi, t, a, \Pi) = \sum_{\omega \in \Omega} \mathsf{AttEU}(\omega)$ independent of Π, and the claim holds.

Similarly, if $\eta = \mathbf{0}$ then $\sum_{\hat{\omega} \in \Omega} \eta^{\hat{\omega}} \Pr[\hat{\omega}|\omega] = 0$ for all $\omega \in \Omega$ independent of Π.

Exactly the same argument holds for calculating $\mathsf{DefEU}(\cdot)$. □

By combining the two claims we get the following: $\mathsf{AttEU}(\chi, t, \eta^*_{0,t}, \Pi_0) = \mathsf{AttEU}(\chi, t, \mathbf{1}, \Pi_0) = \mathsf{AttEU}(\chi, t, \mathbf{1}, \Pi') \geq \mathsf{AttEU}(\chi, t, \eta, \Pi')$, thus we get $(t, \mathbf{1})$ as a best response for Π', for any Π' and χ. □

This shows if $(t, \mathbf{1})$ or $(t, \mathbf{0})$ is a Stackelberg equilibrium, the defender can safely deploy the same strategy for any uncertainty matrix Π', without any loss in her expected utility. □

Theorem 3. *If $(t, \mathbf{1})$ is a best response for $\Pi_{\kappa\lambda\mu}$ and χ is a weak-signal-attack deployment, then $(t, \mathbf{1})$ is a best response for $\Pi_{\kappa'\lambda'\mu'}$ and χ for all $\kappa' \geq \kappa$, $\lambda' \geq \lambda$, $\mu' \geq \mu$.*

Let $\mathsf{AttEU}(\hat{\omega}) = \Pr[\hat{\omega}]\mathsf{AttEU}(\chi, t, \mathbf{1})$ be the attacker's expected utility when observed signaling state is $\hat{\omega}$ and the attacker attacks the target t.

Proof. We have $\eta^{\sigma_1} = 1$, which implies $\mathsf{AttEU}(\hat{\omega} = \sigma_1) \geq 0$ because of our Π structure. Therefore, increasing λ or μ only increases $\mathsf{AttEU}(\hat{\omega} = \mathrm{n})$ and $\mathsf{AttEU}(\hat{\omega} = \sigma_0)$, respectively. This implies $\mathsf{AttEU}(\hat{\omega} = \mathrm{n})$ and $\mathsf{AttEU}(\hat{\omega} = \sigma_0)$ stays positive. Therefore, the attacker behavior also stays the same, when we increase λ or μ.

Since χ is a weak-signal-attack deployment, we know $\mathsf{AttEU}(\hat{\omega} = \sigma_0) \geq 0$. Therefore, increasing κ only makes $\mathsf{AttEU}(\hat{\omega} = \mathrm{n})$, and $\mathsf{AttEU}(\hat{\omega} = \sigma_0)$ more positive; therefore, the attacker behavior stays as $\mathbf{1}$. □

Proposition 4. *There always exists an optimal solution that is a weak-signal-attack deployment with Π_0.*

Proof. Suppose for contradiction there does not exist an optimal solution that is a weak-signal-attack deployment. Then, consider the optimal solution χ^* with the least number of non-weak-signal-attack targets. By the assumption, we know $\mathsf{AttEU}(\sigma_0) < 0$ for some target t. Fix an arbitrary target t that is a non-weak-signal-attack target.

Then, we know $\mathsf{AttEU}(\sigma_0) = U^a_+(t) \cdot (\psi^{\mathrm{s}+} + \psi^{\mathrm{s}-}) + U^a_-(t) \cdot (\psi^{\bar{\mathrm{s}}}) < 0$. Since $\mathsf{AttEU}(\sigma_0) < 0$ and $\Pi = \Pi_0$,

we know the attacker is not attacking when he observes σ_0. Also, since $\mathsf{AttEU}(\cdot)$ is strictly negative, we know ψ^{s+} or ψ^{s-} is strictly greater than 0.

Consider the new deployment χ where we can decrease ψ^{s+} and/or ψ^{s-} until $\mathsf{AttEU}(\sigma_0) = 0$. Since the attacker is still not attacking when he observes σ_0 (recall we break ties in favor of the defender), the defender expected utility stays the same. Furthermore, this change only increases $x^{s+} - \psi^{s+}$ and $x^{s-} - \psi^{s-}$. Thus, $\mathsf{DefEU}(\sigma_1)$ also only increases. Our new χ is therefore still an optimal solution and t is now a weak-signal-attack target. Thus, it contradicts the assumption that χ^* is the optimal solution with the least number of non-weak-signal-attack targets. □

B.1 Handling Observational Uncertainty

The problem uses a similar linear program as for the case without observational uncertainty:

$$\max_{x,\psi,\varphi} \quad U^d_{-s}(t) + U^d_{\sigma_1}(t) + U^d_{\sigma_0}(t) \tag{1}$$

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

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

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

$$U^a_{\hat\omega}(\psi_i,\varphi_i,x_i) \leq b_i^{\hat\omega} \qquad \forall\hat\omega\in\Omega, \forall i\neq t \tag{5}$$

$$0\leq b_i^{\mathrm{n}} \quad 0\leq b_i^{\sigma_0} \quad 0\leq b_i^{\sigma_1} \qquad \forall i\neq t \tag{6}$$

$$U^a_{-s}(t) + U^a_{\sigma_1}(t) + U^a_{\sigma_0}(t) \geq x_i^{\mathrm{p}}\cdot U^a_+(i) + b_i^{\mathrm{n}} + b_i^{\sigma_0} + b_i^{\sigma_1} \qquad \forall\, i\neq t \tag{7}$$

$$0\leq \psi_i^\theta \leq x_i^\theta \quad \forall\,\theta\in\{\bar{\mathrm{s}},\mathrm{s}-,\mathrm{s}+\},\forall\, i\in[N] \tag{8}$$

$$0\leq \varphi_i^\theta \leq x_i^\theta \quad \forall\,\theta\in\{\bar{\mathrm{s}},\mathrm{s}-,\mathrm{s}+\},\forall\, i\in[N] \tag{9}$$

$$(2\eta_t^{\hat\omega}-1)\cdot U^a_{\hat\omega}(\psi_t,\varphi_t,x_t)\leq 0 \qquad \forall\,\hat\omega\in\Omega \tag{10}$$

However, here the utility functions need to be redefined in order to take observational uncertainty into account. We define $p^a_\omega(i) = \sum_{\hat\omega\in\Omega}\eta_i^{\hat\omega}\cdot\Pr[\hat\omega|\omega]$ as the probability of the attacker attacking target i given the true signaling state is $\omega\in\Omega$.

1. $U^{d/a}_{-s}(i) = x_i^{\mathrm{p}}\cdot U^{d/a}_+(i) + x_i^{\mathrm{n}+}\cdot U^{d/a}_+(i) + x_i^{\mathrm{n}-}\cdot U^{d/a}_-(i)$ is the expected defender/attacker utility of target i being attacked over states when i has no sensor (p, n+, n−). This is the same as the version with only detection uncertainty.

2. $U^{d/a}_{\sigma_0}(i) = (1-\gamma)\cdot p^a_{\sigma_0}(i)\cdot[\psi_i^{s+}\cdot U^{d/a}_+(i) + \psi_i^{s-}\cdot U^{d/a}_+(i) + \psi_i^{\bar{s}}\cdot U^{d/a}_-(i)] + \gamma\cdot p^a_{\sigma_0}(i)\cdot[\varphi_i^{s+}\cdot U^{d/a}_+(i) + \varphi_i^{s-}\cdot U^{d/a}_-(i) + \varphi_i^{\bar{s}}\cdot U^{d/a}_-(i)]$ is the defender/attacker expected utility when the attacker attacks target i and the defender signals σ_0. This has the added $p^a_{\sigma_0}(i)$ compared to the version with only detection uncertainty.

3. $U^{d/a}_{\sigma_1}(i) = (1-\gamma)\cdot p^a_{\sigma_1}(i)\cdot[(x_i^{s+}-\psi_i^{s+})\cdot U^{d/a}_+(i) + (x_i^{s-}-\psi_i^{s-})\cdot U^{d/a}_+(i) + (x_i^{\bar{s}}-\psi_i^{\bar{s}})\cdot U^{d/a}_-(i)] + \gamma\cdot p^a_{\sigma_1}(i)\cdot[(x_i^{s+}-\varphi_i^{s+})\cdot U^{d/a}_+(i) + (x_i^{s-}-\varphi_i^{s-})\cdot U^{d/a}_-(i) + (x_i^{\bar{s}}-\varphi_i^{\bar{s}})\cdot U^{d/a}_-(i)]$

Now we will define the attacker observational expected utility of any signaling state $\hat\omega\in\Omega$. Let $U^{d/a}_{\sigma_j}(i,\eta_i^{\hat\omega}{=}1)$ be $U^{d/a}_{\sigma_j}(i)$ with $\eta_i^{\hat\omega}{=}1$ and $\eta_i^{\hat\omega'}=0\;\forall\hat\omega\neq\hat\omega'$ and $j\in\{0,1\}$.

4. $U^{d/a}_{\hat\omega}(\psi_i,\varphi_i,x_i) = \Pr[\hat\omega|\mathrm{n}]\cdot[x_i^{\mathrm{n}-}\cdot U^{d/a}_-(i) + x_i^{\mathrm{n}+}\cdot U^{d/a}_+(i)] + \Pr[\hat\omega|\sigma_0]\cdot U^{d/a}_{\sigma_0}(i,\eta_i^{\hat\omega}{=}1) + \Pr[\hat\omega|\sigma_1]\cdot U^{d/a}_{\sigma_1}(i,\eta_i^{\hat\omega}{=}1)$ is the attacker observational expected utility. This is used in (5) and (10).

The set of constraints (2) - (4) enforce the randomized allocation is feasible, as in the version with only detection uncertainty. The set of constraints (5) - (7) ensure target t is the attacker's best response. b variables ensure attacker's utilities are nonnegative. The set of constraints (8)-(9) ensure the marginal probabilities of signaling (ψ and φ) are valid. Lastly, constraint (10) ensures η is a valid attacker behavior. In other words, if $\eta_t^{\hat\omega}$ is zero, then the attacker observational expected utility should be negative, otherwise the attacker utility should be positive.

C Experimental Results

In Fig. 3e, we show the probability of a fake signal given that a warning signal is used. Equation 11 describes this fully. This is then averaged over all of the targets, and finally averaged over 20 random graphs, as summarized in Equation 12.

$$P(fakesignal|\sigma_1)(i) = \frac{\gamma\cdot[(x_i^{s+}-\varphi_i^{s+})+(x_i^{s-}-\varphi_i^{s-})+(x_i^{\bar{s}}-\varphi_i^{\bar{s}})]}{(1-\gamma)\cdot[(x_i^{s+}-\psi_i^{s+})+(x_i^{s-}-\psi_i^{s-})+(x_i^{\bar{s}}-\psi_i^{\bar{s}})]+\gamma\cdot[(x_i^{s+}-\varphi_i^{s+})+(x_i^{s-}-\varphi_i^{s-})+(x_i^{\bar{s}}-\varphi_i^{\bar{s}})]} \tag{11}$$

$$\frac{1}{G}\frac{1}{N}\sum_{j=1}^{G}\sum_{i=1}^{N} P(\text{fake signal}|\sigma_1)(i) \tag{12}$$

The p-values for the experimental results are summarized in Table 1. Fig. 3h does not have a p-value because it is based solely on the graph illustrated in Fig. 4, and the utilities described in Section D.

D Conservation Drones

D.1 Utilities

The utilities used for the experiment in Section 7 are included in Table 2. We construct this payoff matrix to reflect the fact that the reward and penalty of the attackers are impacted by the following features: number of animals, distance to various park features such as boundary, rivers, and roads (some of the features used in (Gholami et al. 2018) to predict poaching activity), and price.

To arrive at specific values, we first chose locations of interest near the park boundary, rivers, and roads in a region of the park known for the presence of animals. The park and specific coordinates are withheld to protect wildlife. We then measured the distances from the locations of interest to the

Plot	p-values
3a	$p \leq 3.457e{-}16$ for $N \geq 14$
3b	$p \leq 2.579e{-}3$ at $40 \leq N \leq 90$
3c	$p \leq 1.421e{-}03$ for $\gamma \geq 0.2$
3d	$p \leq 0.058$ for $\gamma \geq 0.4$
3e	$p = 2.167e{-}22$ when comparing $\gamma = 0$ and $\gamma = 0.9$
3f	1 vs. 2: $p \leq 1.371e{-}04$ for $\gamma \leq 0.7$ 2 vs. 3: $p \leq 2.852e{-}02$ for $\gamma \leq 0.7$ 1 vs. 3: $p \leq 1.984e{-}05$ for $\gamma \leq 0.8$
3g	$p \leq 6.661e{-}02$ at $\gamma = 0.3$ No difference at $\gamma = 0.5$ $p \leq 1.727e{-}07$ at $\gamma = 0.8$

Table 1: p-values for results in Fig. 3 in the main paper.

closest rivers and roads, and the park boundary. The locations were ranked for each of these distances (e.g., node 6 is closest to the river and node 9 is farthest from the river, whereas node 9 is closest to a road and node 4 is farthest from a road). Next, a weighted average of these ranks was taken for each node to estimate the attractiveness of the node to animals (e.g., elephants), with weights of 0.8, 0.1, and 0.1 for distance to river, boundary, and road, respectively, according to the intuition that water matters most to animals. This was ranked from 1 to 10, with 10 being the best node for animals, and this served as a proxy for the number of animals at that node (e.g., 10 animals at node 6). To determine the relative poaching attractiveness for the attacker, the weighted average rank was calculated from the number of animals, and the river, boundary, and road distances, with weights of 0.7, 0.05, 0.15, and 0.1, respectively. This is based on the intuition that animals are the most important factor, but ease of reaching the location and getting away quickly is also a factor (e.g., nodes 6 and 7 are most attractive, while node 9 is least attractive).

We finally take elephants as an example animal, and use the price of ivory (approximately $40,000 (Brito 2019)), the approximate monetary benefit of ecotourism ($1.6 million (Platt 2014)), and an elephant poaching fine ($20,000 (Siyabona Africa (Pty) Ltd 2017)) to assign values to each of the 10 nodes. The defender payoffs are related to the ecotourism benefits of elephants – if the node is uncovered, it is related to the full amount, whereas if it is covered, it is related to an amount for one day. The attacker payoffs are related to the price of ivory, the attractiveness of a target, and the fines associated with a covered target. Given historical data, it may be possible to learn these values in the future from historical data (Gholami et al. 2019), or possibly from park ranger knowledge (Gurumurthy et al. 2018).

D.2 Other False Negative Rates

We include results in Fig. 3h for a single $\gamma = 0.3$, but the relationship varies with γ as we have seen in the rest of Fig. 3. We include several other examples here for $\gamma = 0$, $\gamma = 0.1$, $\gamma = 0.5$, $\gamma = 0.7$, and $\gamma = 0.9$ in Figs. 3a, 3b, 3c, 3d, 3e, respectively. At low values of γ, there is a small gap between ignoring detection uncertainty and GUARDSS, as expected,

Node	U_{du}	U_{dc}	U_{au}	U_{ac}
0	-3200	29	120	-20
1	-12800	56	320	-20
2	-8000	42	240	-20
3	-6400	38	160	-20
4	-4800	33	80	-20
5	-11200	51	280	-20
6	-16000	64	400	-20
7	-14400	59	400	-20
8	-9600	46	200	-20
9	-1600	24	40	-20

Table 2: Utilities for Fig. 3h and Fig. 4 in the main paper.

which indicates that it may be acceptable to ignore detection uncertainty at that point. However, as γ increases, the gap becomes wider, and ignoring detection uncertainty even becomes worse than a random allocation. The gap between no drones and GUARDSS also decreases as γ increases, meaning it becomes less beneficial to use drones under higher uncertainty, as seen in Fig. 3g.

References

Brito, C. 2019. Ivory from hundreds of elephants found in $48 million seizure. *CBS News*.

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*.

Gholami, S.; Yadav, A.; Tran-Thanh, L.; Dilkina, B.; and Tambe, M. 2019. Don't put all your strategies in one basket: Playing green security games with imperfect prior knowledge. In *AAMAS*.

Gurumurthy, S.; Yu, L.; Zhang, C.; Jin, Y.; Li, W.; Zhang, X.; and Fang, F. 2018. Exploiting data and human knowledge for predicting wildlife poaching. In *ACM COMPASS*.

Platt, J. R. 2014. Elephants are worth 76 times more alive than dead: Report. *Scientific American*.

Siyabona Africa (Pty) Ltd. 2017. Stiffer penalties for poaching in zimbabwe. http://www.krugerpark.co.za/krugerpark-times-e-3-stiffer-penalties-for-poaching-in-zimbabwe-25062.html. Accessed: 2019-08-27.

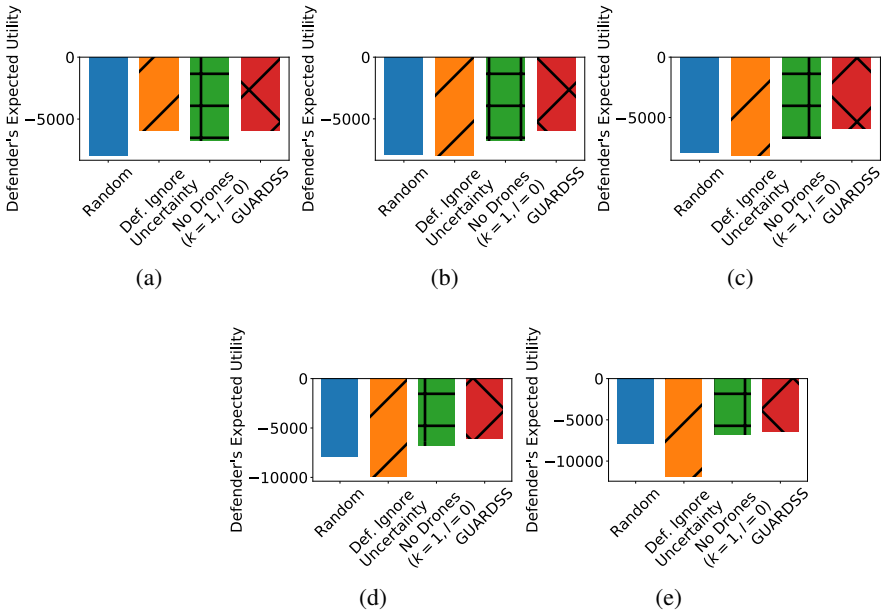


Figure 3: Case study results for multiple values of γ: Fig. 3a has $\gamma = 0$, Fig. 3b has $\gamma = 0.1$, Fig.3c has $\gamma = 0.5$, Fig. 3d has $\gamma = 0.7$, Fig. 3e has $\gamma = 0.9$.